Lecture Notes on

Introduction to Mathematical Economics

Walter Bossert

Département de Sciences Economiques
Université de Montréal
C.P. 6128, succursale Centre-ville
Montréal QC H3C 3J7
Canada
walter.bossert@umontreal.ca

Contents

Chapter 1

Basic Concepts

1.1 Elementary Logic

In all academic disciplines, systems of *logical statements* play a central role. To a large extent, scientific theories attempt to verify or falsify specific statements concerning the objects to be studied in the respective discipline. Statements that are of importance in economic theory include, for example, statements about commodity prices, interest rates, gross national product, quantities of goods bought and sold.

Statements are not restricted to academic considerations. For example, commonly used propositions such as

> There are 1000 students registered in this course

or

> The instructor of this course is less than 70 years old

are examples of statements.

A property common to all statements that we will consider here is that they are *either* true *or* false. For example, the first of the above examples can easily be shown to be false (we just have to consult the class list to see that the number of students registered in this course is *not* 1000), whereas the second statement is a true statement. Therefore, we will use the term "statement" in the following sense.

Definition 1.1.1 *A* statement *is a proposition which is* either *true* or *false.*

Note that, by using the formulation "either . . . or", we rule out statements that are neither true nor false, and we exclude statements with the property of being true *and* false. This restriction is imposed to avoid logical inconsistencies.

To put it simply, elementary logic is concerned with the analysis of statements as defined above, and with combinations of and relations among such statements. We will now introduce specific methods to derive new statements from given statements.

Definition 1.1.2 *Given a statement a, the* negation of *a is the statement "a is false". We denote the negation of a statement a by $\neg a$ (in words: "not a").*

For example, for the statements

a: There are 1000 students registered in this course,
b: The instructor of this course is less than 70 years old,
c: $2 \cdot 3 = 5$,

the corresponding negations can be formulated as

$\neg a$: The number of students registered in this course is not equal to 1000,
$\neg b$: The instructor of this course is at least 70 years old,
$\neg c$: $2 \cdot 3 \neq 5$.

Two statements can be combined in different ways to obtain further statements. The most important ways of formulating such *compound statements* are introduced in the following definitions.

a	b	$\neg a$	$\neg b$	$a \wedge b$	$a \vee b$	$\neg(a \wedge b)$	$\neg(a \vee b)$
T	T	F	F	T	T	F	F
T	F	F	T	F	T	T	F
F	T	T	F	F	T	T	F
F	F	T	T	F	F	T	T

Table 1.1: Negation, conjunction, and disjunction.

a	b	$a \Rightarrow b$	$b \Rightarrow a$	$a \Leftrightarrow b$	$\neg b \Rightarrow \neg a$	$\neg a \Rightarrow \neg b$
T	T	T	T	T	T	T
T	F	F	T	F	F	T
F	T	T	F	F	T	F
F	F	T	T	T	T	T

Table 1.2: Implication and equivalence.

Definition 1.1.3 *Given two statements a and b, the* conjunction *of a and b is the statement "a is true and b is true". We denote the conjunction of a and b by $a \wedge b$ (in words: "a and b").*

Definition 1.1.4 *Given two statements a and b, the* disjunction *of a and b is the statement "a is true or b is true". We denote the disjunction of a and b by $a \vee b$ (in words: "a or b").*

It is very important to note that "or" in Definition 1.1.4 is *not* an "exclusive or" as in "either ... or". The statement $a \vee b$ is true whenever *at least* one of the two statements a, b is true. In particular, $a \vee b$ is true if both a and b are true.

A convenient way of illustrating statements is to use a *truth table*. In Table 1.1, the *truth values* ("T" for "true" and "F" for "false") of the statements $\neg a$, $\neg b$, $a \wedge b$, $a \vee b$, $\neg(a \wedge b)$, $\neg(a \vee b)$ are illustrated for different combinations of truth values of a and b.

Other compound statements which are of importance are *implication* and *equivalence*.

Definition 1.1.5 *Given two statements a and b, the* implication *"a implies b" is the statement "If a is true, then b is true". We denote this implication by $a \Rightarrow b$ (in words: "a implies b").*

Definition 1.1.6 *Given two statements a and b, the* equivalence *of a and b is the statement "a is true if and only if b is true". We denote this equivalence by $a \Leftrightarrow b$ (in words: "a if and only if b").*

In a truth table, implication and equivalence can be illustrated as in Table 1.2. Note that $a \Leftrightarrow b$ is equivalent to $(a \Rightarrow b) \wedge (b \Rightarrow a)$ (*Exercise:* use a truth table to verify this). Furthermore, the statement $a \Rightarrow b$ is equivalent to the statement $\neg a \vee b$ (again, use a truth table to prove this equivalence), and, as is demonstrated in Table 1.2, $a \Rightarrow b$ is equivalent to $\neg b \Rightarrow \neg a$. Some other useful equivalences are sumarized in Table 1.3. In particular, note that $\neg(\neg a)$ is equivalent to a, $\neg(a \wedge b)$ is equivalent to $\neg a \vee \neg b$, and $\neg(a \vee b)$ is equivalent to $\neg a \wedge \neg b$.

Any compound statement involving negation, conjunction, disjunction, implication, and equivalence can be expressed equivalently as a statement involving negation and conjunction (or negation and disjunction) only.

a	b	$\neg(\neg a)$	$\neg(\neg b)$	$\neg(a \wedge b)$	$\neg a \vee \neg b$	$\neg(a \vee b)$	$\neg a \wedge \neg b$
T	T	T	T	F	F	F	F
T	F	T	F	T	T	F	F
F	T	F	T	T	T	F	F
F	F	F	F	T	T	T	T

Table 1.3: Negation.

The tools of elementary logic are useful in proving mathematical theorems. To illustrate that, we provide a discussion of some common proof techniques.

One possibility to prove that a statement is true is to use a *direct proof.* In the case of an implication, a direct proof of the statement $a \Rightarrow b$ proceeds by assuming that a is true, and then showing that b must necessarily be true as well.

Below is an example for a direct proof. Recall that a *natural number* (or *positive integer*) x is *even* if and only if there exists a natural number n such that $x = 2n$. A natural number x is *odd* if and only if there exists a natural number m such that $x = 2m - 1$. Consider the following statements.

a: x is an even natural number and y is an even natural number,
b: xy is an even natural number.

We now give a direct proof of the implication $a \Rightarrow b$. Assume a is true. Because x and y are even, there exist natural numbers n and m such that $x = 2n$ and $y = 2m$. Therefore, $xy = (2n)(2m) = 2(2nm) = 2r$, where $r := 2nm$ is a natural number (the notation $:=$ stands for "is defined by"). This means $xy = 2r$ for some natural number r, which proves that xy must be even. $\|$

(The symbol $\|$ is used to denote the end of a proof.)

Another possibility to prove that a statement is true is to show that its negation is false (it should be clear from the equivalence of $\neg(\neg a)$ and a—see Table 1.3—that this is indeed equivalent to a direct proof of a). This method of proof is called an *indirect proof* or a *proof by contradiction.*

For example, consider the statements

a: $x \neq 0$,
b: There exists exactly one real number y such that $xy = 1$.

We prove $a \Rightarrow b$ by contradiction, that is, we show that $\neg(a \Rightarrow b)$ must be false. Note that $\neg(a \Rightarrow b)$ is equivalent to $\neg(\neg a \vee b)$, which, in turn, is equivalent to $a \wedge \neg b$. Assume $a \wedge \neg b$ is true (that is, $a \Rightarrow b$ is false). We will lead this assumption to a contradiction, which will prove that $a \Rightarrow b$ is true.

Because a is true, $x \neq 0$. If b is false, there are two possible cases. The first possible case is that there exists *no* real number y such that $xy = 1$, and the second possibility is that there exist (at least) *two* different real numbers y and z such that $xy = 1$ and $xz = 1$. Consider the first case. Because $x \neq 0$, we can choose $y = 1/x$. Clearly, $xy = x(1/x) = 1$, which is a contradiction. In the second case, we have $xy = 1 \wedge xz = 1 \wedge y \neq z$. Because $x \neq 0$, we can divide the two equations by x to obtain $y = 1/x$ and $z = 1/x$. But this implies $y = z$, which is a contradiction to $y \neq z$. Hence, in all possible cases, the assumption $\neg(a \Rightarrow b)$ leads to a contradiction. Therefore, this assumption must be false, which means that $a \Rightarrow b$ is true. $\|$

Because, for any two statements a and b, $a \Leftrightarrow b$ is equivalent to $(a \Rightarrow b) \wedge (b \Rightarrow a)$, proving the equivalence $a \Leftrightarrow b$ can be accomplished by proving the implications $a \Rightarrow b$ and $b \Rightarrow a$.

We conclude this section with another example of a mathematical proof, namely, the proof of the *quadratic formula.* Consider the quadratic equation

$$x^2 + px + q = 0 \tag{1.1}$$

where p and q are given real numbers. The following theorem provides conditions under which real numbers x satisfying this equation exist, and shows how to find these solutions to (1.1).

Theorem 1.1.7 *(i) The equation (1.1) has a real solution if and only if* $(p/2)^2 \geq q$.
(ii) A real number x is a solution to (1.1) if and only if

$$\left[x = -\frac{p}{2} + \sqrt{\left(\frac{p}{2}\right)^2 - q}\right] \vee \left[x = -\frac{p}{2} - \sqrt{\left(\frac{p}{2}\right)^2 - q}\right]. \tag{1.2}$$

Proof. (i) Adding $(p/2)^2$ and subtracting q on both sides of (1.1), it follows that (1.1) is equivalent to

$$x^2 + px + \left(\frac{p}{2}\right)^2 = \left(\frac{p}{2}\right)^2 - q$$

which, in turn, is equivalent to

$$\left(x + \frac{p}{2}\right)^2 = \left(\frac{p}{2}\right)^2 - q. \tag{1.3}$$

The left side of (1.3) is nonnegative. Therefore, (1.3) has a solution if and only if the right side of (1.3) is nonnegative as well, that is, if and only if $(p/2)^2 \geq q$.

(ii) Because (1.3) is equivalent to (1.1), x is a solution to (1.1) if and only if x solves (1.3). Taking square roots on both sides, we obtain

$$\left[x + \frac{p}{2} = \sqrt{\left(\frac{p}{2}\right)^2 - q}\right] \vee \left[x + \frac{p}{2} = -\sqrt{\left(\frac{p}{2}\right)^2 - q}\right].$$

Subtracting $p/2$ from both sides, we obtain (1.2). ‖

For example, consider the equation

$$x^2 + 6x + 5 = 0.$$

We have $p = 6$ and $q = 5$. Note that $(p/2)^2 - q = 4 \geq 0$, and therefore, the equation has a real solution. According to Theorem 1.1.7, a solution x must be such that

$$\left[x = -3 + \sqrt{\left(\frac{6}{2}\right)^2 - 5}\right] \vee \left[x = -3 - \sqrt{\left(\frac{6}{2}\right)^2 - 5}\right],$$

that is, the solutions are $x = -1$ and $x = -5$.

As another example, consider

$$x^2 + 2x + 1 = 0.$$

We obtain $(p/2)^2 - q = 0$, and it follows that we have the unique solution $x = -1$.

Finally, consider

$$x^2 + 2x + 2 = 0.$$

We obtain $(p/2)^2 - q = -1 < 0$, and therefore, this equation does not have a real solution.

1.2 Sets

As is the case for logical statements, *sets* are encountered frequently in everyday life. A set is a collection of objects such as, for example, the set of all provinces of Canada, the set of all students registered in this course, or the set of all natural numbers. The precise formal definition of a set that we will be using is the following.

Definition 1.2.1 *A* set *is a collection of objects such that, for each object under consideration, the object is* either *in the set* or *not in the set, and each object appears* at most once *in a given set.*

Note that, according to Definition 1.2.1, it is ruled out that an object belongs to a set and, at the same time, does not belong to this set. Analogously to the assumption that a statement must be either true or false (see Definition 1.1.1), such situations must be excluded in order to avoid logical inconsistencies.

For a set A and an object x, we use the notation $x \in A$ for "Object x is an element (or a member) of A" (in the sense that x belongs to A). If x is not an element (a member) of A, we write $x \notin A$. Clearly, the statement $x \notin A$ is equivalent to $\neg(x \in A)$.

There are different possibilities of describing a set. Some sets can be described by enumerating their elements. For example, consider the sets

$$\begin{aligned} A &:= \{\text{Applied Health Sciences, Arts, Engineering, Environmental Studies,} \\ &\qquad \text{Mathematics, Science}\}, \\ B &:= \{2, 4, 6, 8, 10\}, \\ \mathbb{N} &:= \{1, 2, 3, 4, 5, \ldots\}, \\ \mathcal{Z} &:= \{0, 1, -1, 2, -2, 3, -3, \ldots\}, \\ \emptyset &:= \{\}. \end{aligned}$$

The set $\emptyset$ is called the *empty set* (the set which contains no elements).

There are sets that cannot be described by enumerating their elements, such as the set of real numbers. Therefore, another method must be used to describe these sets. The second commonly used way of describing a set is to enumerate the *properties* that are shared by its elements. For example, the sets A, B, $\mathbb{N}$, $\mathcal{Z}$ defined above can be described in terms of the properties of their members as

$A = \{x \mid x \text{ is a faculty of this university}\}$,
$B = \{x \mid x \text{ is an even natural number between 1 and 10}\}$,
$\mathbb{N} = \{x \mid x \text{ is a natural number}\}$,
$\mathcal{Z} = \{x \mid x \text{ is an integer}\}$.

The symbol $\mathbb{N}$ will be used throughout to denote the set of natural numbers. $\mathcal{Z}$ denotes the set of integers. Other important sets are

$\mathbb{N}_0 := \{x \mid x \in \mathcal{Z} \wedge x \geq 0\}$,
$\mathbb{R} := \{x \mid x \text{ is a real number}\}$,
$\mathbb{R}_+ := \{x \mid x \in \mathbb{R} \wedge x \geq 0\}$,
$\mathbb{R}_{++} := \{x \mid x \in \mathbb{R} \wedge x > 0\}$,
$\mathcal{Q} := \{x \mid x \in \mathbb{R} \wedge (\exists p \in \mathcal{Z}, q \in \mathbb{N} \text{ such that } x = p/q)\}$.

$\mathcal{Q}$ is the set of *rational* numbers. The symbol $\exists$ stands for "there exists". An example for a real number that is not a rational number is $\pi = 3.141593\ldots$.

The following definitions describe some important relationships between sets.

Definition 1.2.2 *For two sets A and B, A is a* subset *of B if and only if*

$$x \in A \Rightarrow x \in B.$$

We denote this subset relationship by $A \subseteq B$.

Therefore, A is a subset of B if and only if each element of A is also an element of B.

An alternative way of formulating the statement $x \in A \Rightarrow x \in B$ is

$$\forall x \in A,\ x \in B$$

where the symbol $\forall$ denotes "for all". In general, implications such as $x \in A \Rightarrow b$ where A is a set and b is a statement can equivalently be formulated as

$$\forall x \in A,\ b.$$

Sometimes, the notation $B \supseteq A$ is used instead of $A \subseteq B$, which means "B is a *superset* of A". The statements $A \subseteq B$ and $B \supseteq A$ are equivalent.

Two sets are equal if and only if they contain the same elements. We can define this property of two sets in terms of the subset relation.

Definition 1.2.3 *Two sets A and B are* equal *if and only if $(A \subseteq B) \wedge (B \subseteq A)$. In this case, we write $A = B$.*

Examples for subset relationships are

$$\mathbb{N} \subseteq \mathcal{Z},\ \mathcal{Z} \subseteq \mathcal{Q},\ \mathcal{Q} \subseteq \mathbb{R},\ \mathbb{R}_+ \subseteq \mathbb{R},\ \{1, 2, 4\} \subseteq \{1, 2, 3, 4\}.$$

Intervals are important subsets of $\mathbb{R}$. We distinguish between *non-degenerate* and *degenerate* intervals. Let $a, b \in \mathbb{R}$ be such that $a < b$. Then the following non-degenerate intervals can be defined.

$[a, b] := \{x \mid x \in \mathbb{R} \wedge (a \leq x \leq b)\}$ (*closed* interval),
$(a, b) := \{x \mid x \in \mathbb{R} \wedge (a < x < b)\}$ (*open* interval),
$[a, b) := \{x \mid x \in \mathbb{R} \wedge (a \leq x < b)\}$ (*half-open* interval),
$(a, b] := \{x \mid x \in \mathbb{R} \wedge (a < x \leq b)\}$ (*half-open* interval).

Using the symbols ∞ and $-\infty$ for "infinity" and "minus infinity", and letting $a \in \mathbb{R}$, the following sets are also non-degenerate intervals.

$(-\infty, a] := \{x \mid x \in \mathbb{R} \wedge x \leq a\}$,
$(-\infty, a) := \{x \mid x \in \mathbb{R} \wedge x < a\}$,
$[a, \infty) := \{x \mid x \in \mathbb{R} \wedge x \geq a\}$,
$(a, \infty) := \{x \mid x \in \mathbb{R} \wedge x > a\}$.

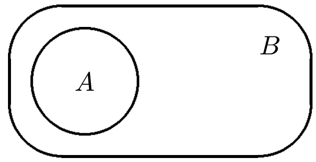


Figure 1.1: $A \subseteq B$.

In particular, $\mathbb{R}_+ = [0, \infty)$ and $\mathbb{R}_{++} = (0, \infty)$ are non-degenerate intervals. Furthermore, $\mathbb{R}$ is the interval $(-\infty, \infty)$. Degenerate intervals are either empty or contain one element only; that is, $\emptyset$ is a degenerate interval, and so are sets of the form $\{a\}$ with $a \in \mathbb{R}$.

If a set A is a subset of a set B and B, in turn, is a subset of a set C, then A must be a subset of C. (This is the *transitivity* property of the relation $\subseteq$.) Formally,

Theorem 1.2.4 *For any three sets A, B, C,*

$$(A \subseteq B) \wedge (B \subseteq C) \Rightarrow A \subseteq C.$$

Proof. Suppose $(A \subseteq B) \wedge (B \subseteq C)$. If $A = \emptyset$, A clearly is a subset of C (the empty set is a subset of *any* set). Now suppose $A \neq \emptyset$. We have to prove $x \in A \Rightarrow x \in C$. Let $x \in A$. Because $A \subseteq B$, $x \in B$. Because $B \subseteq C$, $x \in C$, which completes the proof. ‖

The following definitions introduce some important set operations.

Definition 1.2.5 *The* intersection *of two sets A and B is defined by*

$$A \cap B := \{x \mid x \in A \wedge x \in B\}.$$

Definition 1.2.6 *The* union *of two sets A and B is defined by*

$$A \cup B := \{x \mid x \in A \vee x \in B\}.$$

Two sets A and B are *disjoint* if and only if $A \cap B = \emptyset$, that is, two sets are disjoint if and only if they do not have any common elements.

Definition 1.2.7 *The* difference *between a set A and a set B is defined by*

$$A \setminus B := \{x \mid x \in A \wedge x \notin B\}.$$

The set $A \setminus B$ is called "A without B" or "A minus B".

Definition 1.2.8 *The* symmetric difference *of two sets A and B is defined by*

$$A \triangle B := (A \setminus B) \cup (B \setminus A).$$

Clearly, for any two sets A and B, we have $A \triangle B = B \triangle A$ (prove this as an exercise).

Sets can be illustrated diagramatically by using so-called *Venn diagrams*. For example, the subset relation $A \subseteq B$ can be illustrated as in Figure 1.1. The intersection $A \cap B$ and the union $A \cup B$ are illustrated in Figures 1.2 and 1.3. Finally, the difference $A \setminus B$ and the symmetric difference $A \triangle B$ are shown in Figures 1.4 and 1.5.

As an example, consider the sets $A = \{1, 2, 3, 6\}$ and $B = \{2, 3, 4, 5\}$. Then we obtain

$$A \cap B = \{2, 3\},\ A \cup B = \{1, 2, 3, 4, 5, 6\},\ A \setminus B = \{1, 6\},\ B \setminus A = \{4, 5\},\ A \triangle B = \{1, 4, 5, 6\}.$$

For the applications of set theory discussed in this course, *universal sets* can be defined, where, for a given universal set, all sets under consideration are subsets of this universal set. For example, we will frequently be concerned with subsets of $\mathbb{R}$, so that in these situations, $\mathbb{R}$ can be considered the universal set. For obvious reasons, we will always assume that the universal set under consideration is nonempty.

Given a universal set X and a set $A \subseteq X$, the *complement* of A in X can be defined.

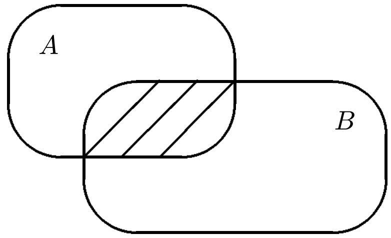


Figure 1.2: $A \cap B$.

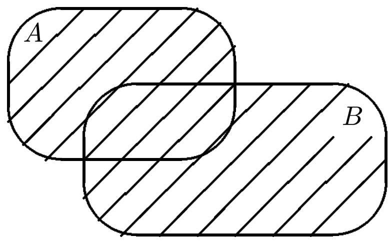


Figure 1.3: $A \cup B$.

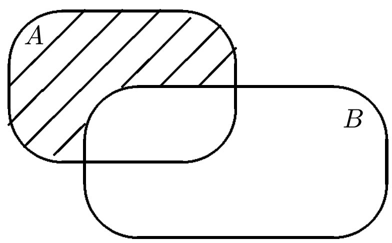


Figure 1.4: $A \setminus B$.

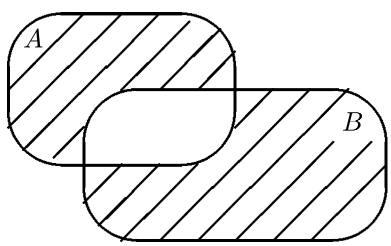


Figure 1.5: $A \triangle B$.

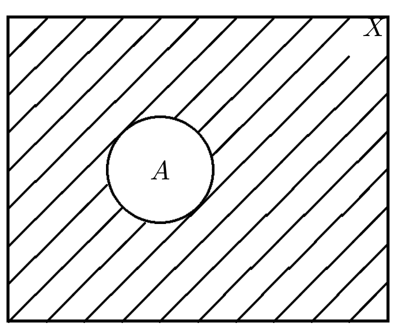


Figure 1.6: $\overline{A}$.

Definition 1.2.9 *Let X be a nonempty universal set, and let $A \subseteq X$. The* complement of A in X *is defined by* $\overline{A} := X \setminus A$.

In a Venn diagram, the complement of $A \subseteq X$ can be illustrated as in Figure 1.6.

For example, if $A = [a, b) \subseteq X = \mathbb{R}$, the complement of A in $\mathbb{R}$ is given by $\overline{A} = (-\infty, a) \cup [b, \infty)$. As another example, let $X = \mathbb{N}$ and $A = \{x \mid x \in \mathbb{N} \wedge (x \text{ is odd})\} \subseteq \mathbb{N}$. We have $\overline{A} = \{x \mid x \in \mathbb{N} \wedge (x \text{ is even})\}$.

The following theorem provides a few useful results concerning complements.

Theorem 1.2.10 *Let X be a nonempty universal set, and let $A \subseteq X$.*

(i) $\overline{\overline{A}} = A$,
(ii) $\overline{X} = \emptyset$,
(iii) $\overline{\emptyset} = X$.

Proof. (i) $\overline{\overline{A}} = X \setminus \overline{A} = \{x \mid x \in X \wedge x \notin \overline{A}\} = \{x \mid x \in X \wedge x \notin \{y \mid y \notin A\}\} = \{x \mid x \in X \wedge x \in A\} = A$.

(ii) We proceed by contradiction. Suppose $\overline{X} \neq \emptyset$. Then there exists $y \in \overline{X}$. By definition, $\overline{X} = \{x \mid x \in X \wedge x \notin X\}$. Therefore, $y \in X \wedge y \notin X$. But this is a contradiction, because no object can be a member of a set and, at the same time, not be a member of this set.

(iii) By way of contradiction, suppose $\overline{\emptyset} \neq X$. Then there exists $x \in X$ such that $x \notin \overline{\emptyset}$. But this implies $x \in \emptyset$, which is a contradiction, because the empty set has no elements. ‖

Part (i) of Theorem 1.2.10 states that, as one would expect, the complement of the complement of a set A is the set A itself.

Some important properties of set operations are summarized in the following theorem.

Theorem 1.2.11 *Let A, B, C be sets.*

(i.1) $A \cap B = B \cap A$,
(i.2) $A \cup B = B \cup A$,
(ii.1) $A \cap (B \cap C) = (A \cap B) \cap C$,
(ii.2) $A \cup (B \cup C) = (A \cup B) \cup C$,
(iii.1) $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$,
(iii.2) $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$.

The proof of Theorem 1.2.11 is left as an exercise. Properties (i) are the *commutative* laws, (ii) are the *associative* laws, and (iii) are the *distributive* laws of the set operations $\cup$ and $\cap$.

Next, we introduce the *Cartesian product* of sets.

Definition 1.2.12 *For two sets A and B, the* Cartesian product *of A and B is defined by*

$$A \times B := \{(x, y) \mid x \in A \wedge y \in B\}.$$

$A \times B$ is the set of all *ordered pairs* (x, y), the first component of which is a member of A, and the second component of which is an element of B. The term "ordered" is very important in the previous sentence. A pair $(x, y) \in A \times B$ is, in general, *different* from the pair (y, x). Note that nothing guarantees that (y, x) is even an element of $A \times B$. Some examples for Cartesian products are given below.

Let $A = \{1, 2, 4\}$ and $B = \{2, 3\}$. Then

$$A \times B = \{(1, 2), (1, 3), (2, 2), (2, 3), (4, 2), (4, 3)\}.$$

As another example, let $A = (1, 2)$ and $B = [0, 1]$. The Cartesian product of A and B is given by

$$A \times B = \{(x, y) \mid (1 < x < 2) \wedge (0 \leq y \leq 1)\}.$$

Finally, let $A = \{1\}$ and $B = [1, 2]$. Then

$$A \times B = \{(x, y) \mid x = 1 \wedge (1 \leq y \leq 2)\}.$$

If A and B are subsets of $\mathbb{R}$, the Cartesian product $A \times B$ can be illustrated in a diagram. The above examples are depicted in Figures 1.7 to 1.9.

We can also form Cartesian products of more than two sets. The following definition introduces the notion of an n-fold Cartesian product.

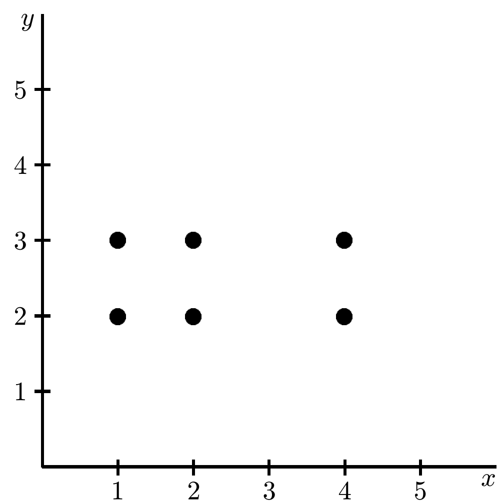


Figure 1.7: $A \times B$, first example.

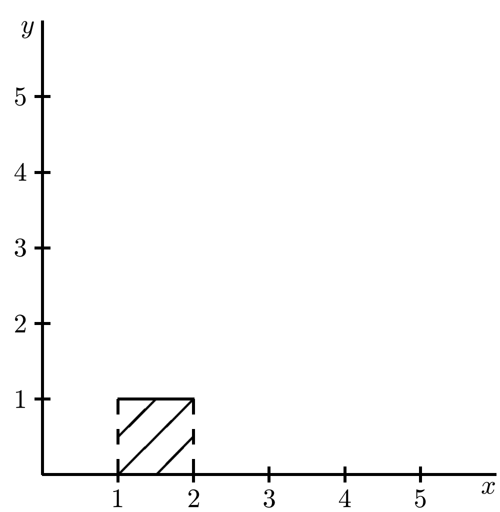


Figure 1.8: $A \times B$, second example.

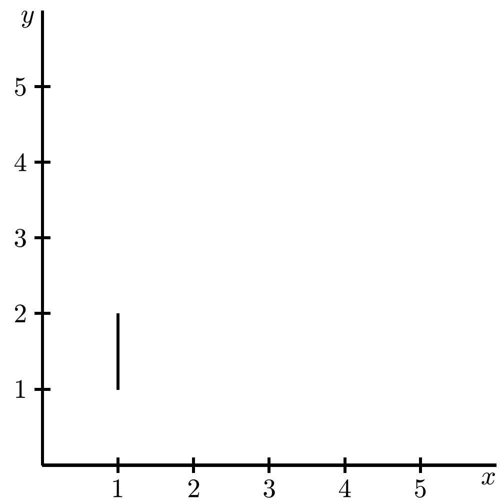


Figure 1.9: $A \times B$, third example.

Definition 1.2.13 *Let $n \in \mathbb{N}$. For n sets $A_1, A_2, \ldots, A_n$, the Cartesian product of $A_1, A_2, \ldots, A_n$ is defined by*

$$A_1 \times A_2 \times \ldots \times A_n := \{(x_1, x_2, \ldots, x_n) \mid x_i \in A_i \ \ \forall i = 1, \ldots, n\}.$$

The elements of an n-fold Cartesian product are called *ordered n-tuples* (again, note that the order of the components of an n-tuple is important).

For example, if $A_1 = \{1, 2\}$, $A_2 = \{0, 1\}$, $A_3 = \{1\}$, we obtain

$$A_1 \times A_2 \times A_3 = \{(1, 0, 1), (1, 1, 1), (2, 0, 1), (2, 1, 1)\}.$$

Of course, (some of) the sets $A_1, A_2, \ldots, A_n$ can be *equal*—Definitions 1.2.12 and 1.2.13 do *not* require the sets which define a Cartesian product to be distinct. For example, if $A = \{1, 2\}$, we can, for example, form the Cartesian products

$$A \times A = \{(1, 1), (1, 2), (2, 1), (2, 2)\}$$

and

$$A \times A \times A = \{(1, 1, 1), (1, 1, 2), (1, 2, 1), (1, 2, 2), (2, 1, 1), (2, 1, 2), (2, 2, 1), (2, 2, 2)\}.$$

For simplicity, the n-fold Cartesian product of a set A is denoted by A^n, that is,

$$A^n := \underbrace{A \times A \times \ldots \times A}_{n \text{ times}}.$$

The most important Cartesian product in this course is the n-fold Cartesian product of $\mathbb{R}$, defined by

$$\mathbb{R}^n := \{(x_1, x_2, \ldots, x_n) \mid x_i \in \mathbb{R} \ \ \forall i = 1, \ldots, n\}.$$

$\mathbb{R}^n$ is called the *n-dimensional Euclidean space*. (The term "space" is sometimes used for sets that have certain structural properties.) The elements of $\mathbb{R}^n$ (ordered n-tuples of real numbers) are usually referred to as *vectors*—details will follow in Chapter 2.

We conclude this section with some notation that will be used later on. For $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$, we define

$$\sum_{i=1}^{n} x_i := x_1 + x_2 + \ldots + x_n.$$

Therefore, $\sum_{i=1}^{n} x_i$ denotes the *sum* of the n numbers $x_1, x_2, \ldots, x_n$.

1.3 Sets of Real Numbers

This section discusses some properties of subsets of $\mathbb{R}$ that are of major importance in later chapters.

First, we define *neighborhoods* of points in $\mathbb{R}$. Intuitively, a neighborhood of a point $x_0 \in \mathbb{R}$ is a set of real numbers that are, in some sense, "close" to x_0. In order to introduce neighborhoods formally, the definition of *absolute values* of real numbers is needed. For $x \in \mathbb{R}$, the absolute value of x is defined as

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

The definition of a neighborhood in $\mathbb{R}$ is

Definition 1.3.1 *For $x_0 \in \mathbb{R}$ and $\varepsilon \in \mathbb{R}_{++}$, the* ε-neighborhood of x_0 *is defined by*

$$\mathcal{U}_\varepsilon(x_0) := \{x \in \mathbb{R} \mid |x - x_0| < \varepsilon\}.$$

Note that, in this definition, we used the formulation

$$\ldots x \in \mathbb{R} \mid \ldots$$

instead of

$$\ldots x \mid x \in \mathbb{R} \wedge \ldots$$

in order to simplify notation. This notation will be used at times if one of the properties defining the elements of a set is the membership in some given set.

$|x - x_0|$ is the *distance* between the points x and x_0 in $\mathbb{R}$. According Definition 1.3.1, an ε-neighborhood of $x_0 \in \mathbb{R}$ is the set of points $x \in \mathbb{R}$ such that the distance between x and x_0 is less than ε. An ε-neighborhood of x_0 is a specific open interval containing x_0—clearly, the neighborhood $\mathcal{U}_\varepsilon(x_0)$ can be written as

$$\mathcal{U}_\varepsilon(x_0) = (x_0 - \varepsilon, x_0 + \varepsilon).$$

Neighborhoods can be used to define *interior points* of a set $A \subseteq \mathbb{R}$.

Definition 1.3.2 *Let $A \subseteq \mathbb{R}$. $x_0 \in A$ is an* interior point *of A if and only if there exists $\varepsilon \in \mathbb{R}_{++}$ such that $\mathcal{U}_\varepsilon(x_0) \subseteq A$.*

According to this definition, a point $x_0 \in A$ is an interior point of A if and only if there exists a neighborhood of x_0 that is contained in A.

For example, consider the set $A = [0, 1)$. We will prove that all points $x \in (0, 1)$ are interior points of A, but the point 0 is not.

First, let $x_0 \in [1/2, 1)$. Let $\varepsilon := 1 - x_0$. This implies $x_0 - \varepsilon = 2x_0 - 1 \geq 0$ and $x_0 + \varepsilon = 1$, and hence, $\mathcal{U}_\varepsilon(x_0) = (x_0 - \varepsilon, x_0 + \varepsilon) \subseteq [0, 1) = A$. Therefore, all points in $[1/2, 1)$ are interior points of A.

Now let $x_0 \in (0, 1/2)$. Define $\varepsilon := x_0$. Then we have $x_0 - \varepsilon = 0$ and $x_0 + \varepsilon = 2x_0 < 1$. Again, $\mathcal{U}_\varepsilon(x_0) = (x_0 - \varepsilon, x_0 + \varepsilon) \subseteq [0, 1) = A$. Hence, all points in the interval $(0, 1/2)$ are interior points of A.

To show that 0 is not an interior point of A, we proceed by contradiction. Suppose 0 is an interior point of $A = [0, 1)$. Then there exists $\varepsilon \in \mathbb{R}_{++}$ such that $\mathcal{U}_\varepsilon(0) = (-\varepsilon, \varepsilon) \subseteq [0, 1) = A$. Let $\delta := \varepsilon/2$. Then it follows that $-\varepsilon < -\delta < 0$, and hence, $-\delta \in \mathcal{U}_\varepsilon(0)$. Because $\mathcal{U}_\varepsilon(0) \subseteq A$, this implies $-\delta \in A$. Because $-\delta$ is negative, this is a contradiction to the definition of A.

If all elements of a set $A \subseteq \mathbb{R}$ are interior points, then A is called an *open set* in $\mathbb{R}$. Furthermore, if the *complement* of a set $A \subseteq \mathbb{R}$ is open in $\mathbb{R}$, then A is called *closed* in $\mathbb{R}$. Formally,

Definition 1.3.3 *A set $A \subseteq \mathbb{R}$ is* open in $\mathbb{R}$ *if and only if*

$$x \in A \Rightarrow x \text{ is interior point of } A.$$

Definition 1.3.4 *A set $A \subseteq \mathbb{R}$ is* closed in $\mathbb{R}$ *if and only if $\overline{A}$ is open in $\mathbb{R}$.*

Openness and closedness can be defined in more abstract spaces than $\mathbb{R}$. However, for the purposes of this course, we can restrict attention to subsets of $\mathbb{R}$ (and $\mathbb{R}^n$, which will be discussed later on). If there is no ambiguity concerning the universal set under consideration (in this course, $\mathbb{R}$ or $\mathbb{R}^n$), we will sometimes simply write "open" (respectively "closed") instead of "open (respectively closed) in $\mathbb{R}$" (or in $\mathbb{R}^n$).

We have already seen that the set $A = [0, 1)$ is not open in $\mathbb{R}$, because 0 is an element of A which is not an interior point of A. To find out whether or not A is closed in $\mathbb{R}$, we have to consider the complement of A in $\mathbb{R}$. This complement is given by $\overline{A} = (-\infty, 0) \cup [1, \infty)$. $\overline{A}$ is not an open set in $\mathbb{R}$, because the point 1 is not an interior point of $\overline{A}$ (prove this as an exercise—the proof is analogous to the proof that 0 is not an interior point of A). Therefore, A is not closed in $\mathbb{R}$. This example establishes that there exist subsets of $\mathbb{R}$ which are neither open nor closed in $\mathbb{R}$.

Any open interval is an open set in $\mathbb{R}$ (which justifies the terminology *open* intervals for these sets), and all closed intervals are closed in $\mathbb{R}$. Furthermore, unions of disjoint open intervals are open, and unions of disjoint closed intervals are closed. $\mathbb{R}$ itself is an open set. To show this, let $x_0 \in \mathbb{R}$, and choose any $\varepsilon \in \mathbb{R}_{++}$. Clearly, $(x_0 - \varepsilon, x_0 + \varepsilon) \subseteq \mathbb{R}$, and therefore, all elements of $\mathbb{R}$ are interior points of $\mathbb{R}$.

The empty set is another example of an open set in $\mathbb{R}$. This is the case, because the empty set does not contain any elements. According to Definition 1.3.3, openness of $\emptyset$ requires

$$x \in \emptyset \Rightarrow x \text{ is interior point of } \emptyset.$$

From Section 1.1, we know that the implication $a \Rightarrow b$ is equivalent to $\neg a \vee b$. Therefore, if a is false, the implication is true. For any $x \in \mathbb{R}$, the statement $x \in \emptyset$ is false (because no object can be an element of the empty set). Consequently, the above implication is true for all $x \in \mathbb{R}$, which shows that $\emptyset$ is open in $\mathbb{R}$.

Note that the openness of $\mathbb{R}$ implies the closedness of $\emptyset$, and the openness of $\emptyset$ implies the closedness of $\mathbb{R}$. Therefore, $\mathbb{R}$ and $\emptyset$ are sets which are *both* open *and* closed in $\mathbb{R}$.

We now define *convex* subsets of $\mathbb{R}$.

Definition 1.3.5 *A set $A \subseteq \mathbb{R}$ is* convex *if and only if*

$$[\lambda x + (1 - \lambda)y] \in A \quad \forall x, y \in A,\ \forall \lambda \in [0, 1].$$

Geometrically, a set $A \subseteq \mathbb{R}$ is convex if, for any two points x and y in this set, all points on the line segment joining x and y belong to A as well. A point $\lambda x + (1 - \lambda)y$ where $\lambda \in [0, 1]$ is called a *convex combination* of x and y. A convex combination of two points is simply a *weighted average* of these points. For example, if we set $\lambda = 1/2$, the corresponding convex combination is

$$\frac{1}{2}x + \frac{1}{2}y,$$

and for $\lambda = 1/4$, we obtain the convex combination

$$\frac{1}{4}x + \frac{3}{4}y.$$

The convex subsets of $\mathbb{R}$ are easy to describe. All intervals (including $\mathbb{R}$ itself) are convex (no matter whether they are open, closed, or half-open), all sets consisting of a single point are convex, and the empty set is convex.

For example, let $A = [0, 1)$. To prove that A is convex, let $x, y \in A$. We have to show that any convex combination of x and y must be in A. Let $\lambda \in [0, 1]$. Without loss of generality, suppose $x \leq y$. Then it follows that

$$\lambda x + (1 - \lambda)y \geq \lambda x + (1 - \lambda)x = x$$

and

$$\lambda x + (1 - \lambda)y \leq \lambda y + (1 - \lambda)y = y.$$

Therefore, $x \leq \lambda x + (1 - \lambda)y \leq y$. Because x and y are elements of A, $x \geq 0$ and $y < 1$. Therefore, $0 \leq \lambda x + (1 - \lambda)y < 1$, which implies $[\lambda x + (1 - \lambda)y] \in A$.

An example of a subset of $\mathbb{R}$ which is not convex is $A = [0, 1] \cup \{2\}$. Let $x = 1$, $y = 2$, and $\lambda = 1/2$. Then $x \in A$ and $y \in A$ and $\lambda \in [0, 1]$, but $\lambda x + (1 - \lambda)y = 3/2 \notin A$.

The following definition introduces *upper and lower bounds* of subsets of $\mathbb{R}$.

Definition 1.3.6 *Let $A \subseteq \mathbb{R}$ be nonempty, and let $u, \ell \in \mathbb{R}$.*

(i) u is an upper bound *of A if and only if $x \leq u$ for all $x \in A$.*
(ii) ℓ is a lower bound *of A if and only if $x \geq \ell$ for all $x \in A$.*

A nonempty set $A \subseteq \mathbb{R}$ is *bounded from above* (resp. *bounded from below*) if and only if it has an upper (resp. lower) bound. A nonempty set $A \subseteq \mathbb{R}$ is *bounded* if and only if A has an upper bound and a lower bound.

Not all subsets of $\mathbb{R}$ have upper or lower bounds. For example, the set $\mathbb{R}_+ = [0, \infty)$ has no upper bound. To show this, we proceed by contradiction. Suppose $u \in \mathbb{R}$ is an upper bound of $\mathbb{R}_+$. But $\mathbb{R}_+$ contains elements x such that $x > u$, which contradicts the assumption that u is an upper bound of $\mathbb{R}_+$. On the other hand, $\mathbb{R}_+$ is bounded from below—any number in the interval $(-\infty, 0]$ is a lower bound of $\mathbb{R}_+$.

The above example shows that an upper bound or a lower bound need not be unique, if it exists. Specific upper and lower bounds are introduced in the next definition.

Definition 1.3.7 *Let $A \subseteq \mathbb{R}$ be nonempty, and let $u, \ell \in \mathbb{R}$.*
(i) u is the least upper bound *(the* supremum*) of A if and only if u is an upper bound of A and $u \leq u'$ for all $u' \in \mathbb{R}$ that are upper bounds of A.*
(ii) ℓ is the greatest lower bound *(the* infimum*) of A if and only if ℓ is a lower bound of A and $\ell \geq \ell'$ for all $\ell' \in \mathbb{R}$ that are lower bounds of A.*

Every nonempty subset of $\mathbb{R}$ which has an upper (resp. lower) bound has a supremum (resp. an infimum). This is *not* necessarily the case if $\mathbb{R}$ is replaced by some other universal set—for example, the set of rational numbers $\mathcal{Q}$ does not have this property.

If a set $A \subseteq \mathbb{R}$ has a supremum (resp. an infimum), the supremum (resp. infimum) is *unique*. Formally,

Theorem 1.3.8 *Let $A \subseteq \mathbb{R}$ be nonempty, and let $u, u', \ell, \ell' \in \mathbb{R}$.*

(i) u is a supremum of A and u' is a supremum of $A \Rightarrow u = u'$.
(ii) ℓ is an infimum of A and ℓ' is an infimum of $A \Rightarrow \ell = \ell'$.

Proof. (i) Let $A \subseteq \mathbb{R}$. Suppose $u \in \mathbb{R}$ is a supremum of A and $u' \in \mathbb{R}$ is a supremum of A. This implies that u and u' are upper bounds of A. By definition of a supremum, it follows that $u \leq u'$ and $u' \leq u$, and therefore, $u = u'$.

The proof of part (ii) is analogous. ‖

Note that the above result justifies the terms "the" supremum and "the" infimum used in Definition 1.3.7. We will denote the supremum (resp. infimum) of $A \subseteq \mathbb{R}$ by $\sup(A)$ (resp. $\inf(A)$).

Note that it is *not* required that the supremum (resp. infimum) of a set $A \subseteq \mathbb{R}$ is itself an element of A. For example, let $A = [0, 1)$. As can be shown easily, the supremum and the infimum of A exist and are given by $\sup(A) = 1$ and $\inf(A) = 0$. Therefore, $\inf(A) \in A$, but $\sup(A) \notin A$. If the supremum (resp. infimum) of $A \subseteq \mathbb{R}$ *is* an element of A, it is sometimes called the *maximum* (resp. *minimum*) of A, denoted by $\max(A)$ (resp. $\min(A)$). Therefore, we can define the maximum and the minimum of a set by

Definition 1.3.9 *Let $A \subseteq \mathbb{R}$ be nonempty, and let $u, \ell \in \mathbb{R}$.*

(i) $u = \max(A)$ if and only if $u = \sup(A) \wedge u \in A$.
(ii) $\ell = \min(A)$ if and only if $\ell = \inf(A) \wedge \ell \in A$.

1.4 Functions

Given two sets A and B, a *function that maps A into B* assigns one element in B to each element in A. The use of functions is widespread (but not always recognized and explicitly declared as such). For example, giving final grades for a course to students is an example of establishing a function from the set of students registered in a course to the set of possible course grades. Each student is assigned exactly one final grade. The formal definition of a function is

Definition 1.4.1 *Let A and B be nonempty sets. If there exists a mechanism f that assigns* exactly one *element in B to each element in A, then f is called a* function from A to B.

A function from A to B is denoted by

$$f : A \mapsto B,\ x \mapsto y = f(x)$$

where $y = f(x) \in B$ is the *image* of $x \in A$. A is the *domain* of the function f, B is the *range* of f. The set

$$f(A) := \{y \in B \mid \exists x \in A \text{ such that } y = f(x)\}$$

is the *image* of A under f (sometimes also called the image of f). More generally, the image of $S \subseteq A$ under f is defined as

$$f(S) := \{y \in B \mid \exists x \in S \text{ such that } y = f(x)\}.$$

Note that defining a function f involves two steps: First, the domain and the range of the function have to be specified, and then, for each element x in the domain of the function, it has to be stated which element in the range of f is assigned to x according to f.

Equality of functions is defined as

Definition 1.4.2 *Two functions $f_1 : A_1 \mapsto B_1$ and $f_2 : A_2 \mapsto B_2$ are* equal *if and only if*

$$(A_1 = A_2) \wedge (B_1 = B_2) \wedge (f_1(x) = f_2(x)\ \ \forall x \in A_1).$$

Note that equality of two functions requires that their domains *and* their ranges are equal.

To illustrate possible applications of functions, consider the above mentioned example. Suppose we have a course in which six students are registered. For simplicity, we number the students from 1 to 6. The possible course grades are $\{\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D}, \mathcal{F}\}$ (script letters are used in this example to avoid confusion with the domain A and the range B of the function considered). An assignment of grades to students can be expressed as a function $f : \{1, 2, 3, 4, 5, 6\} \mapsto \{\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D}, \mathcal{F}\}$. For example, if students 1 and 3 get an $\mathcal{A}$, student 2 gets a $\mathcal{B}$, student 4 fails, and students 5 and 6 get a $\mathcal{D}$, the function f is defined as

$$f : \{1, 2, 3, 4, 5, 6\} \mapsto \{\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D}, \mathcal{F}\},\ x \mapsto \begin{cases} \mathcal{A} & \text{if } x \in \{1, 3\} \\ \mathcal{B} & \text{if } x = 2 \\ \mathcal{D} & \text{if } x \in \{5, 6\} \\ \mathcal{F} & \text{if } x = 4. \end{cases} \tag{1.4}$$

The image of f is $f(A) = \{\mathcal{A}, \mathcal{B}, \mathcal{D}, \mathcal{F}\}$. As this example demonstrates, $f(A)$ is not necessarily equal to the range B—there may exist elements $y \in B$ such that there exists no $x \in A$ with $y = f(x)$ (as is the case for $\mathcal{C} \in B$ in the above example). Of course, by definition of $f(A)$, we always have the relationship $f(A) \subseteq B$.

As another example, consider the function defined by

$$f : \mathbb{R} \mapsto \mathbb{R},\ x \mapsto x^2. \tag{1.5}$$

The image of f is

$$f(\mathbb{R}) = \{y \in \mathbb{R} \mid \exists x \in \mathbb{R} \text{ such that } y = x^2\} = \{y \in \mathbb{R} \mid y \geq 0\} = \mathbb{R}_+.$$

Again, $B = \mathbb{R} \neq \mathbb{R}_+ = f(\mathbb{R}) = f(A)$.

Next, the *graph* of a function is defined.

Definition 1.4.3 *The* graph *G of a function $f : A \mapsto B$ is a subset of the Cartesian product $A \times B$, defined as*

$$G := \{(x, y) \mid x \in A \wedge y = f(x)\}.$$

In other words, the graph of a function $f : A \mapsto B$ is the set of all pairs $(x, f(x))$, where $x \in A$. For functions such that $A \subseteq \mathbb{R}$ and $B \subseteq \mathbb{R}$, the graph is a useful tool to give a diagrammatic illustration of the function. For example, the graph of the function f defined in (1.5) is

$$G = \{(x, y) \mid x \in \mathbb{R} \wedge y = x^2\}$$

and is illustrated in Figure 1.10.

As mentioned before, the range of a function is not necessarily equal to the image of this function. In the special case where $f(A) = B$, we say that the function f is *surjective* (or *onto*). We define

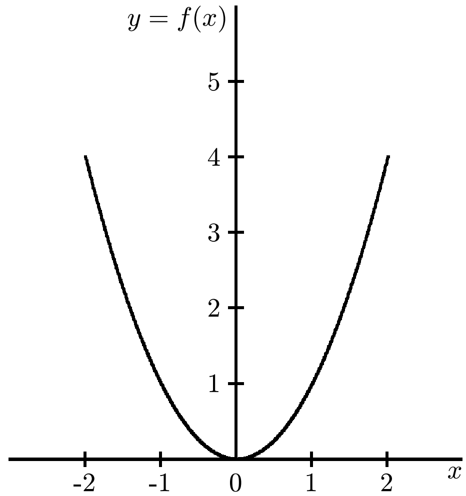


Figure 1.10: The graph of a function.

Definition 1.4.4 *A function $f : A \mapsto B$ is* surjective *(*onto*) if and only if $f(A) = B$.*

A function $f : A \mapsto B$ such that, for each $y \in f(A)$, there exists *exactly one* $x \in A$ with $y = f(x)$ is called *injective* (or *one-to-one*). Formally,

Definition 1.4.5 *A function $f : A \mapsto B$ is* injective *(*one-to-one*) if and only if*

$$\forall x, y \in A,\ x \neq y \Rightarrow f(x) \neq f(y).$$

If a function is onto *and* one-to-one, the function is called *bijective*.

Definition 1.4.6 *A function $f : A \mapsto B$ is* bijective *if and only if f is surjective and injective.*

As an example, consider the function f defined in (1.4). This function is not surjective, because there exists no $x \in \{1, 2, 3, 4, 5, 6\}$ such that $f(x) = \mathcal{C}$. Furthermore, this function is not injective, because $1 \neq 3$, but $f(1) = f(3) = \mathcal{A}$.

The function f defined in (1.5) is not surjective, because, for $y = -1 \in B$, there exists no $x \in \mathbb{R}$ such that $f(x) = x^2 = y$. f is not injective, because, for example, $f(1) = f(-1) = 1$.

As another example, define a function f by

$$f : \mathbb{R} \mapsto \mathbb{R}_+,\ x \mapsto x^2.$$

Note that this is *not* the same function as the one defined in (1.5), because it has a different range. That choosing a different range indeed gives us a different function can be seen by noting that this function *is* surjective, whereas the one in the previous example is *not*. The function is not injective, and therefore, not bijective.

Here is another example that illustrates the importance of defining the domain and the range properly. Let

$$f : \mathbb{R}_+ \mapsto \mathbb{R},\ x \mapsto x^2.$$

This function is not surjective (note that its range is $\mathbb{R}$, but its image is $\mathbb{R}_+$), but it is injective. To prove that, consider any $x, y \in \mathbb{R}_+ = A$ such that $x \neq y$. Note that the domain of f is $\mathbb{R}_+$, and therefore, x and y are nonnegative. Because $x \neq y$, we can, without loss of generality, assume $x > y$. Because both x and y are nonnegative, it follows that $x^2 > y^2$, and therefore, $f(x) = x^2 \neq y^2 = f(y)$, which proves that f is injective.

As a final example, let

$$f : \mathbb{R}_+ \mapsto \mathbb{R}_+,\ x \mapsto x^2.$$

Now the domain *and* the range of f are given by $\mathbb{R}_+$. For each $y \in \mathbb{R}_+ = B$, there exists $x \in \mathbb{R}_+ = A$ such that $y = f(x)$ (namely, $x = \sqrt{y}$), which proves that f is onto. Furthermore, as in the previous example, $f(x) \neq f(y)$ whenever $x, y \in \mathbb{R}_+ = A$ and $x \neq y$. Therefore, f is one-to-one, and hence, bijective.

Bijective functions allow us to find, for *each* $y \in B$, a *unique* element $x \in A$ such that $y = f(x)$. This suggests that a bijective function from A to B can be used to define another function with domain B and range A, which assigns each $x \in A$ to its image under f. This motivates the following definition of an *inverse* function.

Definition 1.4.7 *Let $f : A \mapsto B$ be bijective. The function defined by*

$$f^{-1} : B \mapsto A,\ f(x) \mapsto x$$

is called the inverse function *of f.*

Note that an inverse function is *not defined* if f is not bijective.

For example, consider the function

$$f : \mathbb{R} \mapsto \mathbb{R},\ x \mapsto x^3.$$

This function is bijective (Exercise: prove this), and consequently, its inverse function f^{-1} exists. By definition of the inverse, we have, for all $x \in \mathbb{R}$, $y \in \mathbb{R}$,

$$f^{-1}(y) = x \Leftrightarrow y = x^3 \Leftrightarrow y^{1/3} = x,$$

and therefore, the inverse of f is given by

$$f^{-1} : \mathbb{R} \mapsto \mathbb{R},\ y \mapsto y^{1/3}.$$

Two functions with appropriate domains and ranges can be combined to form a *composite* function. More precisely, composite functions are defined as

Definition 1.4.8 *Suppose two functions $f : A \mapsto B$ and $g : B \mapsto C$ are given. The function*

$$g \circ f : A \mapsto C,\ x \mapsto g(f(x))$$

is the composite function *of f and g.*

An important property of the inverse function f^{-1} of a bijective function f is that its inverse is given by f. Hence, for a bijective function $f : A \mapsto B$, we have

$$f^{-1}(f(x)) = x \ \ \forall x \in A$$

and

$$f(f^{-1}(y)) = y \ \ \forall y \in B.$$

The following definition introduces some important special cases of bijective functions, namely, *permutations*.

Definition 1.4.9 *Let A be a finite subset of $\mathbb{N}$. A* permutation of A *is a bijective function $\pi : A \mapsto A$.*

For example, a permutation of $A = \{1, 2, 3\}$ is given by

$$\pi : \{1,2,3\} \mapsto \{1,2,3\},\ x \mapsto \begin{cases} 1 & \text{if } x = 1 \\ 2 & \text{if } x = 3 \\ 3 & \text{if } x = 2. \end{cases}$$

Other permutations of A are

$$\pi : \{1,2,3\} \mapsto \{1,2,3\},\ x \mapsto \begin{cases} 1 & \text{if } x = 2 \\ 2 & \text{if } x = 3 \\ 3 & \text{if } x = 1 \end{cases} \tag{1.6}$$

and

$$\pi : \{1,2,3\} \mapsto \{1,2,3\},\ x \mapsto \begin{cases} 1 & \text{if } x = 3 \\ 2 & \text{if } x = 2 \\ 3 & \text{if } x = 1. \end{cases}$$

A permutation can be used to change the *numbering* of objects. For example, if

$$\pi : \{1, \ldots, n\} \mapsto \{1, \ldots, n\}$$

is a permutation of $\{1, \ldots, n\}$ and $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$, a renumbering of the components of x is obtained by applying the permutation π. The resulting vector is

$$x_\pi = (x_{\pi(1)}, \ldots, x_{\pi(n)}).$$

More specifically, if π is given by (1.6) and $x = (x_1, x_2, x_3) \in \mathbb{R}^3$, we obtain

$$x_\pi = (x_3, x_1, x_2).$$

1.5 Sequences of Real Numbers

In addition to having some economic applications in their own right, sequences of real numbers are very useful for the formulation of some properties of real-valued functions. This section provides a brief introduction to sequences of real numbers. We restrict attention to considerations that will be of importance for this course.

A sequence of real numbers is a special case of a function as defined in the previous section, namely, a function with the domain $\mathbb{N}$ and the range $\mathbb{R}$.

Definition 1.5.1 *A* sequence of real numbers *is a function $a : \mathbb{N} \mapsto \mathbb{R}$, $n \mapsto a(n)$. To simplify notation, we will write a_n instead of $a(n)$ for $n \in \mathbb{N}$, and use $\{a_n\}$ to denote such a sequence.*

More general sequences (not necessarily of real numbers) could be defined by allowing the range to be a set that is not necessarily equal to $\mathbb{R}$ in the above definition. However, all sequences encountered in this chapter will be sequences of real numbers, and we will, for simplicity of presentation, refer to them as "sequences" and omit "of real numbers". Sequences of elements of $\mathbb{R}^n$ will be discussed in Chapter 4.

Here is an example of a sequence. Define

$$a : \mathbb{N} \mapsto \mathbb{R},\ n \mapsto 1 - \frac{1}{n}. \tag{1.7}$$

The first few points in this sequence are

$$a_1 = 0,\ a_2 = 1/2,\ a_3 = 2/3,\ a_4 = 3/4, \ldots.$$

Sequences also appear in economic problems. For example, suppose a given amount of money $x \in \mathbb{R}_{++}$ is deposited to a bank account, and there is a fixed rate of interest $r \in \mathbb{R}_{++}$. After one year, the value of this investment is $x + rx = (1 + r)x$. Assuming this amount is reinvested and the rate of interest is unchanged, the value after two years is $(1 + r)x + r(1 + r)x = (1 + r)(1 + r)x = x(1 + r)^2$. In general, after $n \in \mathbb{N}$ years, the value of the investment is $x(1 + r)^n$. These values of the investment in different years can be expressed as a sequence, namely, the sequence $\{a_n\}$, where

$$a_n = x(1 + r)^n \ \ \forall n \in \mathbb{N}.$$

This is a special case of a *geometric* sequence.

Definition 1.5.2 *A sequence $\{a_n\}$ is a* geometric *sequence if and only if there exists $q \in \mathbb{R}$ such that*

$$a_{n+1} = qa_n \ \ \forall n \in \mathbb{N}.$$

A geometric sequence has a quite simple structure in the sense that all elements of the sequence can be derived from the *first* element of the sequence, given the number $q \in \mathbb{R}$. This is the case because, according to Definition 1.5.2, $a_2 = qa_1$, $a_3 = qa_2 = q^2 a_1$ and, for $n \in \mathbb{N}$ with $n \geq 2$, $a_n = q^{n-1} a_1$. The above example of interest accumulation is a geometric sequence, where $q = 1 + r$ and $a_1 = qx$.

It is often important to analyze the behaviour of a sequence as n becomes, loosely speaking, "large". To formalize this notion more precisely, we introduce the following definition.

Definition 1.5.3 *(i) A sequence* $\{a_n\}$ converges to $\alpha \in \mathbb{R}$ *if and only if*

$$\forall \varepsilon \in I\!R_{++}, \exists n_0 \in I\!N \text{ such that } a_n \in \mathcal{U}_\varepsilon(\alpha)\ \forall n \geq n_0.$$

(ii) If $\{a_n\}$ *converges to* $\alpha \in I\!R$*,* α *is the* limit *of* $\{a_n\}$*, and we write*

$$\lim_{n\to\infty} a_n = \alpha.$$

Recall that $\mathcal{U}_\varepsilon(\alpha)$ is the ε-neighborhood of $\alpha \in \mathbb{R}$, where $\varepsilon \in \mathbb{R}_{++}$. Therefore, the statement "$a_n \in \mathcal{U}_\varepsilon(\alpha)$" is equivalent to "$|a_n - \alpha| < \varepsilon$".

In words, $\{a_n\}$ converges to $\alpha \in \mathbb{R}$ if and only if, for any $\varepsilon \in \mathbb{R}_{++}$, at most a finite number of elements of $\{a_n\}$ are *outside* the ε-neighborhood of α.

To illustrate this definition, consider again the sequence defined in (1.7). We prove that this sequence converges to the limit 1. In order to do so, we show that, for all $\varepsilon \in \mathbb{R}_{++}$, there exists $n_0 \in \mathbb{N}$ such that $|a_n - 1| < \varepsilon$ for all $n \geq n_0$. For any $\varepsilon \in \mathbb{R}_{++}$, choose $n_0 \in \mathbb{N}$ such that $n_0 > 1/\varepsilon$. For $n \geq n_0$, it follows that $n > 1/\varepsilon$, and therefore,

$$\varepsilon > \frac{1}{n} = |1 - \frac{1}{n} - 1| = |a_n - 1|,$$

which shows that the sequence $\{a_n\}$ converges to the limit $\alpha = 1$.

The following terminology will be used.

Definition 1.5.4 *A sequence* $\{a_n\}$ *is* convergent *if and only if there exists* $\alpha \in I\!R$ *such that*

$$\lim_{n\to\infty} a_n = \alpha.$$

A sequence $\{a_n\}$ *is* divergent *if and only if* $\{a_n\}$ *is not convergent.*

Here is an example of a divergent sequence. Define $\{a_n\}$ by

$$a_n = n\ \ \forall n \in N. \tag{1.8}$$

That this sequence diverges can be shown by contradiction. Suppose $\{a_n\}$ is convergent. Then there exists $\alpha \in \mathbb{R}$ such that $\lim_{n\to\infty} a_n = \alpha$. Therefore, for any $\varepsilon \in \mathbb{R}_{++}$, there exists $n_0 \in \mathbb{N}$ such that

$$|n - \alpha| < \varepsilon\ \ \forall n \geq n_0. \tag{1.9}$$

Let $n_1 \in \mathbb{N}$ be such that $n_1 > \alpha + \varepsilon$. Then $n_1 - \alpha > \varepsilon$, and therefore,

$$|n - \alpha| = n - \alpha > \varepsilon\ \ \forall n \geq n_1. \tag{1.10}$$

Let $n_2 \in \mathbb{N}$ be such that $n_2 \geq n_0$ and $n_2 \geq n_1$. Then (1.9) implies $|n_2 - \alpha| < \varepsilon$, and (1.10) implies $|n_2 - \alpha| > \varepsilon$, which is a contradiction.

Two special cases of divergence, defined below, are of particular importance.

Definition 1.5.5 *A sequence* $\{a_n\}$ diverges to ∞ *if and only if*

$$\forall c \in I\!R, \exists n_0 \in I\!N \text{ such that } a_n \geq c\ \forall n \geq n_0.$$

Definition 1.5.6 *A sequence* $\{a_n\}$ diverges to $-\infty$ *if and only if*

$$\forall c \in I\!R, \exists n_0 \in I\!N \text{ such that } a_n \leq c\ \forall n \geq n_0.$$

For example, the sequence $\{a_n\}$ defined in (1.8) diverges to ∞ (Exercise: provide a proof).

Note that there are divergent sequences which do not diverge to ∞ or $-\infty$. A divergent sequence which does not diverge to ∞ or $-\infty$ is said to *oscillate*. For example, consider the sequence $\{a_n\}$ defined by

$$a_n = \begin{cases} 0 & \text{if } n \text{ is even} \\ 1 & \text{if } n \text{ is odd.} \end{cases} \tag{1.11}$$

As an exercise, prove that this sequence oscillates.

For a sequence $\{a_n\}$ and $\alpha \in \mathbb{R} \cup \{-\infty, \infty\}$, we will sometimes use the notation

$$a_n \longrightarrow \alpha$$

if $\alpha \in \mathbb{R}$ and $\{a_n\}$ converges to α, or if $\alpha \in \{-\infty, \infty\}$ and $\{a_n\}$ diverges to α.

In Definition 1.5.3, we referred to α as "the" limit of $\{a_n\}$, if $\{a_n\}$ converges to α. This terminology is justified, because a sequence cannot have more than one limit. That is, the limit of a convergent sequence is *unique*, as stated in the following theorem.

Theorem 1.5.7 *Let $\{a_n\}$ be a sequence, and let $\alpha, \beta \in \mathbb{R}$. If $\lim_{n\to\infty} a_n = \alpha$ and $\lim_{n\to\infty} a_n = \beta$, then $\alpha = \beta$.*

Proof. By way of contradiction, assume

$$(\lim_{n\to\infty} a_n = \alpha) \wedge (\lim_{n\to\infty} a_n = \beta) \wedge \alpha \neq \beta.$$

Without loss of generality, suppose $\alpha < \beta$. Define

$$\varepsilon := \frac{\beta - \alpha}{2} > 0.$$

Then it follows that

$$\mathcal{U}_\varepsilon(\alpha) = \left(\alpha - \beta + \frac{\alpha + \beta}{2}, \frac{\alpha + \beta}{2}\right)$$

and

$$\mathcal{U}_\varepsilon(\beta) = \left(\frac{\alpha + \beta}{2}, \frac{\alpha + \beta}{2} + \beta - \alpha\right).$$

Therefore,

$$\mathcal{U}_\varepsilon(\alpha) \cap \mathcal{U}_\varepsilon(\beta) = \emptyset. \tag{1.12}$$

Because $\lim_{n\to\infty} a_n = \alpha$, there exists $n_0 \in \mathbb{N}$ such that $a_n \in \mathcal{U}_\varepsilon(\alpha)$ for all $n \geq n_0$. Analogously, because $\lim_{n\to\infty} a_n = \beta$, there exists $n_1 \in \mathbb{N}$ such that $a_n \in \mathcal{U}_\varepsilon(\beta)$ for all $n \geq n_1$. Let $n_2 \in \mathbb{N}$ be such that $n_2 \geq n_0$ and $n_2 \geq n_1$. Then

$$a_n \in \mathcal{U}_\varepsilon(\alpha) \wedge a_n \in \mathcal{U}_\varepsilon(\beta) \quad \forall n \geq n_2,$$

which is equivalent to $a_n \in \mathcal{U}_\varepsilon(\alpha) \cap \mathcal{U}_\varepsilon(\beta)$ for all $n \geq n_2$, a contradiction to (1.12). ‖

The convergence of some sequences can be established by showing that they have certain properties. We define

Definition 1.5.8

(i) A sequence $\{a_n\}$ is monotone nondecreasing $\Leftrightarrow$ $a_{n+1} \geq a_n \;\; \forall n \in \mathbb{N}$,

(ii) A sequence $\{a_n\}$ is monotone nonincreasing $\Leftrightarrow$ $a_{n+1} \leq a_n \;\; \forall n \in \mathbb{N}$.

Definition 1.5.9

(i) A sequence $\{a_n\}$ is bounded from above $\Leftrightarrow \exists c \in \mathbb{R}$ *such that* $a_n \leq c \;\; \forall n \in \mathbb{N}$,

(ii) A sequence $\{a_n\}$ is bounded from below $\Leftrightarrow \exists c \in \mathbb{R}$ *such that* $a_n \geq c \;\; \forall n \in \mathbb{N}$,

(iii) A sequence $\{a_n\}$ is bounded $\Leftrightarrow \{a_n\}$ *is bounded from above and from below.*

For example, consider the sequence $\{a_n\}$ defined in (1.7). This sequence is monotone nondecreasing, because

$$a_{n+1} = 1 - \frac{1}{n+1} \geq 1 - \frac{1}{n} = a_n \quad \forall n \in \mathbb{N}.$$

Furthermore, $\{a_n\}$ is bounded, because $0 \leq a_n \leq 1$ for all $n \in \mathbb{N}$.

There are some important relationships between convergent, monotone, and bounded sequences. First, we show that a convergent sequence must be bounded.

Theorem 1.5.10 *Let $\{a_n\}$ be a sequence. If $\{a_n\}$ is convergent, then $\{a_n\}$ is bounded.*

Proof. Suppose $\{a_n\}$ is convergent with limit $\alpha \in \mathbb{R}$. Let $\varepsilon = 1$. Then there exists $n_0 \in \mathbb{N}$ such that

$$|a_n - \alpha| < 1 \;\; \forall n \geq n_0.$$

Define

$$N_1 := \{n \in \mathbb{N} \mid n > n_0 \wedge a_n \geq \alpha\},$$
$$N_2 := \{n \in \mathbb{N} \mid n > n_0 \wedge a_n < \alpha\}.$$

Then it follows that

$$a_n < \alpha + 1 \;\; \forall n \in N_1, \tag{1.13}$$
$$a_n > \alpha - 1 \;\; \forall n \in N_2. \tag{1.14}$$

Now define

$$c_1 := \max(\{a_1, \ldots, a_{n_0}, \alpha + 1\}),$$
$$c_2 := \min(\{a_1, \ldots, a_{n_0}, \alpha - 1\}).$$

(Clearly, c_1 and c_2 are well-defined, because a finite set of real numbers must have a maximum and a minimum.) This, together with (1.13) and (1.14), implies

$$c_2 \leq a_n \leq c_1 \;\; \forall n \in N,$$

which proves that $\{a_n\}$ is bounded. $\|$

Boundedness of a sequence does *not* imply the convergence of this sequence. For example, the sequence $\{a_n\}$ defined in (1.11) is bounded (because $0 \leq a_n \leq 1$ for all $n \in \mathbb{N}$), but it is not convergent. However, boundedness from above (resp. below) together with monotone nondecreasingness (resp. nonincreasingness) together imply that a sequence is convergent. This result provides a criterion that is often very useful when it is to be determined whether a sequence is convergent. The theorem below states this result formally.

Theorem 1.5.11 *Let $\{a_n\}$ be a sequence.*

(i) $\{a_n\}$ is monotone nondecreasing and bounded from above $\Rightarrow$
$\{a_n\}$ is convergent with limit $\sup(\{a_n \mid n \in \mathbb{N}\})$.
(ii) $\{a_n\}$ is monotone nonincreasing and bounded from below $\Rightarrow$
$\{a_n\}$ is convergent with limit $\inf(\{a_n \mid n \in \mathbb{N}\})$.

Proof. (i) Suppose $\{a_n\}$ is bounded from above and monotone nondecreasing. Boundedness from above of $\{a_n\}$ is equivalent to the boundedness from above of the set $\{a_n \mid n \in \mathbb{N}\}$. Therefore, $\sup(\{a_n \mid n \in \mathbb{N}\})$ exists. Letting $\alpha = \sup(\{a_n \mid n \in \mathbb{N}\})$, we have $a_n \leq \alpha$ for all $n \in \mathbb{N}$. Next, we show

$$\forall \varepsilon \in \mathbb{R}_{++}, \exists n_0 \in \mathbb{N} \text{ such that } \alpha - a_{n_0} < \varepsilon. \tag{1.15}$$

Suppose this is not the case. Then there exists $\varepsilon \in \mathbb{R}_{++}$ such that $\alpha - a_n \geq \varepsilon$ for all $n \in \mathbb{N}$. Equivalently, $\alpha - \varepsilon \geq a_n$ for all $n \in \mathbb{N}$, which means that $\alpha - \varepsilon$ is an upper bound of $\{a_n \mid n \in \mathbb{N}\}$. Because $\alpha - \varepsilon < \alpha$, this contradicts the definition of α as the supremum of this set. Therefore, (1.15) is true. This implies $\alpha - \varepsilon < a_{n_0}$. Because $\{a_n\}$ is monotone nondecreasing, $a_n \geq a_{n_0}$ for all $n \geq n_0$. Therefore, $\alpha - \varepsilon < a_{n_0} \leq a_n$ for all $n \geq n_0$, which implies

$$\alpha - a_n < \varepsilon \;\; \forall n \geq n_0. \tag{1.16}$$

Because $a_n \leq \alpha$ for all $n \in \mathbb{N}$, (1.16) is equivalent to

$$|a_n - \alpha| < \varepsilon \;\; \forall n \geq n_0,$$

which proves that α is the limit of $\{a_n\}$.

The proof of part (ii) of the theorem is analogous. $\|$

The results below show how the limits of sums, products, and ratios of convergent sequences can be obtained. The proofs of these theorems are left as exercises.

Theorem 1.5.12 *Let $\{a_n\}$, $\{b_n\}$ be sequences, and let $\alpha, \beta \in \mathbb{R}$. If $\lim_{n\to\infty} a_n = \alpha$ and $\lim_{n\to\infty} b_n = \beta$, then $\lim_{n\to\infty}(a_n + b_n) = \alpha + \beta$ and $\lim_{n\to\infty}(a_n b_n) = \alpha\beta$.*

Theorem 1.5.13 *Let $\{a_n\}$, $\{b_n\}$ be sequences, and let $\alpha, \beta \in \mathbb{R}$. Furthermore, let $b_n \neq 0$ for all $n \in \mathbb{N}$ and $\beta \neq 0$. If $\lim_{n\to\infty} a_n = \alpha$ and $\lim_{n\to\infty} b_n = \beta$, then $\lim_{n\to\infty}(a_n / b_n) = \alpha/\beta$.*

Chapter 2

Linear Algebra

2.1 Vectors

In Chapter 1, we introduced elements of the space $\mathbb{R}^n$ as ordered n-tuples $(x_1, \ldots, x_n)$, where $x_i \in \mathbb{R}$ for all $i = 1, \ldots, n$. Geometrically, we can think of $x \in \mathbb{R}^n$ as a *point* in the n-dimensional space. For example, for $n = 2$, the point $x = (1, 2) \in \mathbb{R}^2$ can be represented by a point in the two-dimensional coordinate system with the coordinates 1 and 2.

Another possibility is to think of $x \in \mathbb{R}^n$ as a parametrization of a *vector* in $\mathbb{R}^n$. We can visualize the vector $x \in \mathbb{R}^n$ as an "arrow" starting at the *origin* $\mathbf{0} = (0, \ldots, 0) \in \mathbb{R}^n$ which "points" at $x = (x_1, \ldots, x_n)$. For example, the vector $x = (1, 2) \in \mathbb{R}^2$ can be represented as in Figure 2.1.

For some of our applications, it is important whether the components of a vector in $\mathbb{R}^n$ are arranged in a *column* or in a *row*. A *column vector* in $\mathbb{R}^n$ is written as

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n,$$

and the corresponding row vector (the vector with the same elements as x, but arranged in a row) is

$$x' = (x_1, \ldots, x_n) \in \mathbb{R}^n.$$

The notation $'$ stands for "transpose". For example, the transpose of

$$x = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \in \mathbb{R}^2$$

is $x' = (1, 2)$. If x is a column vector, then its transpose is a row vector. The transpose of the transpose of a vector $x \in \mathbb{R}^n$ is the vector x itself, that is, $(x')' = x$. For simplicity, we will omit the $'$ for row vectors whenever it is of no importance whether $x \in \mathbb{R}^n$ is to be treated as a column vector or a row vector. It should be kept in mind, though, that for some applications, this distinction is of importance.

The following operations can be applied to vectors.

Definition 2.1.1 *Let $n \in \mathbb{N}$, and let $x, y \in \mathbb{R}^n$. The* sum of x and y *is defined by*

$$x + y := \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{pmatrix}.$$

Definition 2.1.2 *Let $n \in \mathbb{N}$, let $\alpha \in \mathbb{R}$, and let $x \in \mathbb{R}^n$. The* product of α and x *is defined by*

$$\alpha x := \begin{pmatrix} \alpha x_1 \\ \vdots \\ \alpha x_n \end{pmatrix}.$$

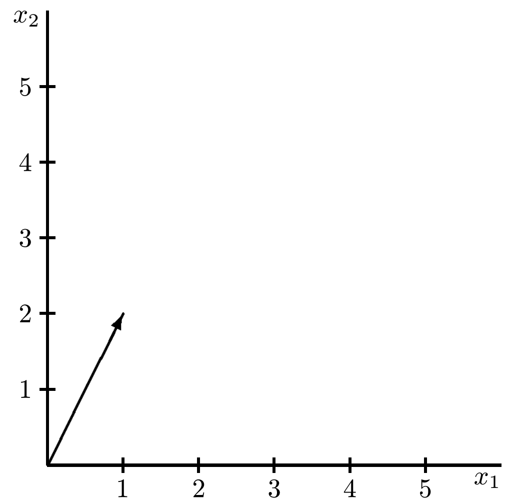


Figure 2.1: The vector $x = (1, 2)$.

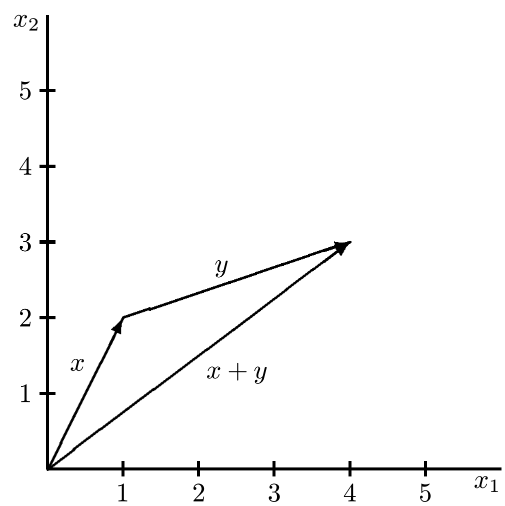


Figure 2.2: Vector addition.

The operation described in Definition 2.1.1 is called *vector addition*, the operation introduced in Definition 2.1.2 is called *scalar multiplication*.

Vector addition and scalar multiplication have very intuitive geometric interpretations. Consider, for example, $x = (1, 2) \in \mathbb{R}^2$ and $y = (3, 1) \in \mathbb{R}^2$. Then we obtain $x + y = (1 + 3, 2 + 1) = (4, 3)$. This vector addition is illustrated in Figure 2.2.

Now consider $x = (1, 2) \in \mathbb{R}^2$ and $\alpha = 2$. We obtain $\alpha x = (2 \cdot 1, 2 \cdot 2) = (2, 4)$, which is illustrated in Figure 2.3.

Geometrically, the vector $2x$ points in the same direction as $x \in \mathbb{R}^n$, but is twice as long (we will discuss the notion of the "length" of a vector in more detail below).

Using vector addition and scalar multiplication, we can define the operation *vector subtraction*.

Definition 2.1.3 *Let $n \in \mathbb{N}$, and let $x, y \in \mathbb{R}^n$. The* difference of x and y *is defined by*

$$x - y := x + (-1)y.$$

The following theorem summarizes some properties of vector addition and scalar multiplication.

Theorem 2.1.4 *Let $n \in \mathbb{N}$, let $\alpha, \beta \in \mathbb{R}$, and let $x, y, z \in \mathbb{R}^n$.*

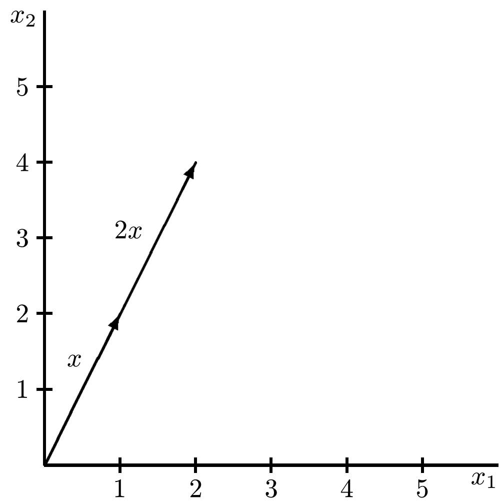


Figure 2.3: Scalar multiplication.

(i) $x + y = y + x$,
(ii) $(x + y) + z = x + (y + z)$,
(iii) $\alpha(\beta x) = (\alpha\beta)x$,
(iv) $(\alpha + \beta)x = \alpha x + \beta x$,
(v) $\alpha(x + y) = \alpha x + \alpha y$.

The proof of this theorem is left as an exercise.

Vector addition and scalar multiplication yield new vectors in $\mathbb{R}^n$. The operation introduced in the following definition assigns a *real number* to a pair of vectors.

Definition 2.1.5 *Let $n \in \mathbb{N}$, and let $x, y \in \mathbb{R}^n$. The* inner product *of* x *and* y *is defined by*

$$xy := \sum_{i=1}^{n} x_i y_i.$$

The inner product of two vectors is used frequently in economic models. For example, if $x \in \mathbb{R}^n_+$ represents a *commodity bundle* and $p \in \mathbb{R}^n_{++}$ denotes a *price vector*, the inner product $px = \sum_{i=1}^{n} p_i x_i$ is the *value* of the commodity bundle at these prices.

The inner product has the following properties.

Theorem 2.1.6 *Let $n \in \mathbb{N}$, let $\alpha \in \mathbb{R}$, and let $x, y, z \in \mathbb{R}^n$.*

(i) $xy = yx$,
(ii) $xx \geq 0$,
(iii) $xx = 0 \Leftrightarrow x = \mathbf{0}$,
(iv) $(x + y)z = xz + yz$,
(v) $(\alpha x)y = \alpha(xy)$.

Again, the proof of this theorem is left as an exercise.

Of special importance are the so-called *unit vectors* in $\mathbb{R}^n$. For $i \in \{1, \ldots, n\}$, the i^{th} *unit vector* is defined by

$$e^i_j := \begin{cases} 0 & \text{if } j \neq i \\ 1 & \text{if } j = i \end{cases} \quad \forall j = 1, \ldots, n.$$

For example, for $n = 2$, $e^1 = (1, 0)$ and $e^2 = (0, 1)$. See Figure 2.4.

We now define the *Euclidean norm* of a vector.

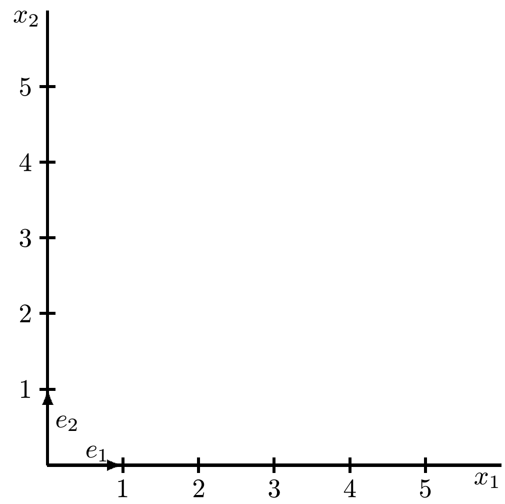


Figure 2.4: Unit vectors.

Definition 2.1.7 *Let $n \in I\!N$, and let $x \in I\!R^n$. The* Euclidean norm of x *is defined by*

$$\|x\| := \sqrt{xx} = \sqrt{\sum_{i=1}^{n} x_i^2}.$$

Because the Euclidean norm is the only norm considered in this chapter, we will refer to $\|x\|$ simply as the *norm* of $x \in \mathbb{R}^n$. Geometrically, the norm of a vector can be interpreted as a measure of its *length.* For example, let $x = (1,2) \in \mathbb{R}^2$. Then

$$\|x\| = \sqrt{1 \cdot 1 + 2 \cdot 2} = \sqrt{5}.$$

Clearly, $\|e^i\| = 1$ for the unit vector $e^i \in \mathbb{R}^n$.

Based on the Euclidean norm, we can define the *Euclidean distance* of two vectors $x, y \in \mathbb{R}^n$ as the norm of the difference $x - y$.

Definition 2.1.8 *Let $n \in I\!N$, and let $x, y \in I\!R^n$. The* Euclidean distance of x and y *is defined by*

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

Again, we will omit the term "Euclidean", because the above defined distance is the only notion of distance used in this chapter. See Chapter 4 for a more detailed discussion of distance functions for vectors. For $n = 1$, we obtain the usual norm and distance used for real numbers. For $x \in \mathbb{R}$, $\|x\| = \sqrt{x^2} = |x|$, and for $x, y \in \mathbb{R}$, $d(x,y) = \|x - y\| = |x - y|$.

For many applications discussed in this course, the question whether or not some vectors are *linearly independent* is of great importance. Before defining the terms "linear (in)dependence" precisely, we have to introduce *linear combinations* of vectors.

Definition 2.1.9 *Let $m, n \in I\!N$, and let $x^1, \ldots, x^m, y \in I\!R^n$. y is a* linear combination of $x^1, \ldots, x^m$ *if and only if there exist $\alpha_1, \ldots, \alpha_m \in I\!R$ such that*

$$y = \sum_{j=1}^{m} \alpha_j x^j.$$

For example, consider the vectors $x^1 = (1,2)$, $x^2 = (0,1)$, $x^3 = (5,-1)$, $y = (-9/2, 5)$ in $\mathbb{R}^2$. y is a linear combination of x^1, x^2, x^3, because, with $\alpha_1 = 1/2$, $\alpha_2 = 3$, $\alpha_3 = -1$, we obtain

$$y = \alpha_1 x^1 + \alpha_2 x^2 + \alpha_3 x^3.$$

Note that *any* vector $x \in \mathbb{R}^n$ can be expressed as a linear combination of the unit vectors $e^1, \ldots, e^n$— simply choose $\alpha_i = x_i$ for all $i = 1, \ldots, n$ to verify this.

Linear (in)dependence of vectors is defined as

Definition 2.1.10 *Let $m, n \in I\!N$, and let $x^1, \ldots x^m \in I\!R^n$. The vectors $x^1, \ldots, x^m$ are* linearly independent *if and only if*

$$\sum_{j=1}^{m} \alpha_j x^j = \mathbf{0} \Rightarrow \alpha_j = 0 \; \forall j = 1, \ldots, m.$$

The vectors $x^1, \ldots x^m$ are linearly dependent *if and only if they are not linearly independent.*

Linear independence of $x^1, \ldots, x^m$ requires that *the only* possibility to choose $\alpha_1, \ldots, \alpha_m \in \mathbb{R}$ such that $\sum_{j=1}^{m} \alpha_j x^j = \mathbf{0}$ is to choose $\alpha_1 = \ldots = \alpha_m = 0$. Clearly, the vectors $x^1, \ldots, x^m$ are linearly dependent if and only if there exist $\alpha_1, \ldots, \alpha_m \in \mathbb{R}$ such that

$$\sum_{j=1}^{m} \alpha_j x^j = \mathbf{0} \wedge (\exists k \in \{1, \ldots, m\} \text{ such that } \alpha_k \neq 0).$$

As an example, consider the vectors $x^1 = (2, 1)$, $x^2 = (1, 0)$. To check whether these vectors are linearly independent, we have to consider the equation

$$\alpha_1 \begin{pmatrix} 2 \\ 1 \end{pmatrix} + \alpha_2 \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{2.1}$$

(2.1) is equivalent to

$$2\alpha_1 + \alpha_2 = 0 \wedge \alpha_1 = 0. \tag{2.2}$$

By (2.2), we must have $\alpha_1 = 0$, and therefore, $\alpha_2 = 0$ in order to satisfy (2.1). Hence, the only possibility to satisfy (2.1) is to choose $\alpha_1 = \alpha_2 = 0$, which means that the two vectors are linearly independent.

Another example for a set of linearly independent vectors is the set of unit vectors $\{e^1, \ldots, e^n\}$ in $\mathbb{R}^n$. Clearly, the equation

$$\sum_{j=1}^{n} \alpha_j e^j = \mathbf{0}$$

is equivalent to

$$\alpha_1 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \ldots + \alpha_n \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix},$$

which requires $\alpha_j = 0$ for all $j = 1, \ldots, n$. For a *single* vector $x \in \mathbb{R}^n$, it is not very hard to find out whether the vector is linearly dependent or independent. This is shown in

Theorem 2.1.11 *Let $n \in I\!N$, and let $x \in I\!R^n$. x is linearly independent if and only if $x \neq \mathbf{0}$.*

Proof. Let $x = \mathbf{0}$. Then *any* $\alpha \in \mathbb{R}$ satisfies

$$\alpha x = \mathbf{0}. \tag{2.3}$$

Therefore, there exists $\alpha \neq 0$ satisfying (2.3), which shows that $\mathbf{0}$ is linearly dependent.

Now let $x \neq \mathbf{0}$. Then there exists $k \in \{1, \ldots, n\}$ such that $x_k \neq 0$. To satisfy

$$\alpha \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}, \tag{2.4}$$

we must, in particular, have $\alpha x_k = 0$. Because $x_k \neq 0$, the only possibility to satisfy (2.4) is to choose $\alpha = 0$. Therefore, x is linearly independent. ‖

The following theorem provides an alternative formulation of linear dependence for $m \geq 2$ vectors.

Theorem 2.1.12 *Let $n, m \in I\!N$ with $m \geq 2$, and let $x^1, \ldots, x^m \in I\!R^n$. The vectors $x^1, \ldots, x^m$ are linearly dependent if and only if (at least) one of these vectors is a linear combination of the remaining vectors.*

Proof. "Only if": Suppose $x^1, \ldots, x^m$ are linearly dependent. Then there exist $\alpha_1, \ldots \alpha_m \in \mathbb{R}$ such that

$$\sum_{j=1}^{m} \alpha_j x^j = \mathbf{0} \tag{2.5}$$

and there exists $k \in \{1, \ldots, m\}$ such that $\alpha_k \neq 0$. Without loss of generality, suppose $k = m$. Because $\alpha_m \neq 0$, we can divide (2.5) by α_m to obtain

$$\sum_{j=1}^{m} \frac{\alpha_j}{\alpha_m} x^j = \mathbf{0}$$

or, equivalently,

$$x^m = \sum_{j=1}^{m-1} \beta_j x^j,$$

where $\beta_j := -\alpha_j/\alpha_m$ for all $j = 1, \ldots, m-1$. But this means that x^m is a linear combination of the remaining vectors.

"If": Suppose one of the vectors $x^1, \ldots, x^m$ is a linear combination of the remaining vectors. Without loss of generality, suppose x^m is this vector. Then there exist $\alpha_1, \ldots, \alpha_{m-1} \in \mathbb{R}$ such that

$$x^m = \sum_{j=1}^{m-1} \alpha_j x^j.$$

Defining $\alpha_m := -1$, this is equivalent to

$$\sum_{j=1}^{m} \alpha_j x^j = \mathbf{0}.$$

Because $\alpha_m \neq 0$, this implies that the vectors $x^1, \ldots, x^m$ are linearly dependent. ‖

Further results concerning linear (in)dependence will follow later in this chapter, once matrices and the solution of systems of linear equations have been discussed.

2.2 Matrices

Matrices are *arrays* of real numbers. The formal definition is

Definition 2.2.1 *Let $m, n \in \mathbb{N}$. An $m \times n$* matrix *is an array*

$$A = (a_{ij}) = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix}$$

where $a_{ij} \in \mathbb{R}$ for all $i = 1, \ldots, m$, $j = 1, \ldots, n$.

Therefore, an $m \times n$ matrix is an array of real numbers with m rows and n columns. An $m \times n$ matrix such that $m = n$ is called a *square matrix*, that is, a square matrix has the same number of rows and columns.

Equality of two matrices is defined as

Definition 2.2.2 *Let $m, n, q, r \in \mathbb{N}$. Furthermore, let $A = (a_{ij})$ be an $m \times n$ matrix, and let $B = (b_{ij})$ be a $q \times r$ matrix. A and B are* equal *if and only if*

$$(m = q) \wedge (n = r) \wedge (a_{ij} = b_{ij} \; \forall i = 1, \ldots, m, \; \forall j = 1, \ldots, n).$$

Vectors are special cases of matrices. A column vector

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$$

is an $n \times 1$ matrix, and a row vector $x' = (x_1, \ldots, x_n) \in \mathbb{R}^n$ is a $1 \times n$ matrix. We can also think of an $m \times n$ matrix as an ordered n-tuple of m-dimensional column vectors

$$\left(\begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix} \cdots \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix} \right)$$

or as an ordered m-tuple of n-dimensional row vectors

$$\begin{matrix} (a_{11}, \ldots, a_{1n}) \\ \vdots \\ (a_{m1}, \ldots, a_{mn}). \end{matrix}$$

The *transpose* of a matrix is defined as

Definition 2.2.3 *Let $m, n \in \mathbb{N}$, and let A be an $m \times n$ matrix. The* transpose *of A is defined by*

$$A' := \begin{pmatrix} a_{11} & a_{21} & \ldots & a_{m1} \\ a_{12} & a_{22} & \ldots & a_{m2} \\ \vdots & \vdots & & \vdots \\ a_{1n} & a_{2n} & \ldots & a_{mn} \end{pmatrix}.$$

Clearly, if A is an $m \times n$ matrix, A' is an $n \times m$ matrix. The transpose of A is obtained by *interchanging* the roles of rows and columns. For $i = 1, \ldots, m$ and $j = 1, \ldots, n$, the i^{th} row of A is the i^{th} column of A', and the j^{th} column of A is the j^{th} row of A'. For example, the transpose of the 2×3 matrix

$$A = \begin{pmatrix} 3 & -1 & 0 \\ 2 & 0 & 1 \end{pmatrix}$$

is the 3×2 matrix

$$A' = \begin{pmatrix} 3 & 2 \\ -1 & 0 \\ 0 & 1 \end{pmatrix}.$$

The transpose of the transpose of a matrix A is the matrix A itself, that is, for any matrix A, $(A')' = A$.

Symmetric square matrices will play an important role in this course. We define

Definition 2.2.4 *Let $n \in \mathbb{N}$. An $n \times n$ matrix A is* symmetric *if and only if $A' = A$.*

Note that symmetry is defined only for square matrices. For example, the 2×2 matrix

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 0 \end{pmatrix}$$

is symmetric, because

$$A' = \begin{pmatrix} 1 & 2 \\ 2 & 0 \end{pmatrix} = A.$$

On the other hand, the 2×2 matrix

$$B = \begin{pmatrix} 1 & 2 \\ -2 & 0 \end{pmatrix}$$

is not symmetric, because

$$B' = \begin{pmatrix} 1 & -2 \\ 2 & 0 \end{pmatrix} \neq B.$$

Now we will introduce some important matrix operations. First, we define *matrix addition*.

Definition 2.2.5 *Let $m, n \in \mathbb{N}$, and let $A = (a_{ij})$ and $B = (b_{ij})$ be $m \times n$ matrices. The* sum of A and B *is defined by*

$$A + B := (a_{ij} + b_{ij}) = \begin{pmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \dots & a_{1n} + b_{1n} \\ a_{21} + b_{21} & a_{22} + b_{22} & \dots & a_{2n} + b_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} + b_{m1} & a_{m2} + b_{m2} & \dots & a_{mn} + b_{mn} \end{pmatrix}.$$

Note that the sum $A + B$ is defined only for matrices of the same type, that is, A and B must have the same number of rows and the same number of columns. As examples, consider the 2×3 matrices

$$A = \begin{pmatrix} 2 & 1 & 0 \\ 3 & 0 & -1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 2 \end{pmatrix}.$$

The sum of theses two matrices is

$$A + B = \begin{pmatrix} 3 & 1 & 0 \\ 3 & 2 & 1 \end{pmatrix}.$$

Next, *scalar multiplication* is defined.

Definition 2.2.6 *Let $m, n \in \mathbb{N}$, and let $A = (a_{ij})$ be an $m \times n$ matrix. Furthermore, let $\alpha \in \mathbb{R}$. The* product of α and A *is defined by*

$$\alpha A := (\alpha a_{ij}) = \begin{pmatrix} \alpha a_{11} & \alpha a_{12} & \dots & \alpha a_{1n} \\ \alpha a_{21} & \alpha a_{22} & \dots & \alpha a_{2n} \\ \vdots & \vdots & & \vdots \\ \alpha a_{m1} & \alpha a_{m2} & \dots & \alpha a_{mn} \end{pmatrix}.$$

For example, let $\alpha = -2$ and

$$A = \begin{pmatrix} 0 & -2 \\ 1 & 2 \\ 5 & 0 \end{pmatrix}.$$

Then

$$\alpha A = \begin{pmatrix} 0 & 4 \\ -2 & -4 \\ -10 & 0 \end{pmatrix}.$$

The following theorem summarizes some properties of matrix addition and scalar multiplication (note the similarity to Theorem 2.1.4).

Theorem 2.2.7 *Let $m, n \in \mathbb{N}$, and let A, B, C be $m \times n$ matrices. Furthermore, let $\alpha, \beta \in \mathbb{R}$.*

(i) $A + B = B + A$,
(ii) $(A + B) + C = A + (B + C)$,
(iii) $\alpha(\beta A) = (\alpha\beta)A$,
(iv) $(\alpha + \beta)A = \alpha A + \beta A$,
(v) $\alpha(A + B) = \alpha A + \alpha B$.

As an exercise, prove this theorem.

The *multiplication* of two matrices is only defined if the matrices satisfy a *conformability* condition for matrix multiplication. In particular, the matrix product AB is defined only if the number of *columns* in A is equal to the number of *rows* in B.

Definition 2.2.8 *Let $m, n, r \in \mathbb{N}$. Furthermore, let $A = (a_{ij})$ be an $m \times n$ matrix and let $B = (b_{ij})$ be an $n \times r$ matrix. The* matrix product AB *is defined by*

$$AB := \left(\sum_{k=1}^{n} a_{ik} b_{kj} \right) = \begin{pmatrix} \sum_{k=1}^{n} a_{1k} b_{k1} & \sum_{k=1}^{n} a_{1k} b_{k2} & \dots & \sum_{k=1}^{n} a_{1k} b_{kr} \\ \sum_{k=1}^{n} a_{2k} b_{k1} & \sum_{k=1}^{n} a_{2k} b_{k2} & \dots & \sum_{k=1}^{n} a_{2k} b_{kr} \\ \vdots & \vdots & & \vdots \\ \sum_{k=1}^{n} a_{mk} b_{k1} & \sum_{k=1}^{n} a_{mk} b_{k2} & \dots & \sum_{k=1}^{n} a_{mk} b_{kr} \end{pmatrix}.$$

If A is an $m \times n$ matrix and B is an $n \times r$ matrix, the matrix AB is an $m \times r$ matrix. If A is an $m \times n$ matrix, B is a $q \times r$ matrix, and $n \neq q$, the product AB is *not* defined.

The above definition is best illustrated with an example. Let

$$A = \begin{pmatrix} 2 & 1 \\ 0 & 2 \\ 1 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 0 & 3 & 0 \\ -1 & 2 & 2 & 1 \end{pmatrix}.$$

A is a 3×2 matrix, and B is a 2×4 matrix, and therefore, the conformability condition for the multiplication of these matrices is satisfied (because the number of columns in A is equal to the number of rows in B). The matrix AB is a 3×4 matrix, and it is obtained as follows.

$$\begin{aligned} AB &= \begin{pmatrix} 2 & 1 \\ 0 & 2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 3 & 0 \\ -1 & 2 & 2 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 2\cdot 1 + 1\cdot(-1) & 2\cdot 0 + 1\cdot 2 & 2\cdot 3 + 1\cdot 2 & 2\cdot 0 + 1\cdot 1 \\ 0\cdot 1 + 2\cdot(-1) & 0\cdot 0 + 2\cdot 2 & 0\cdot 3 + 2\cdot 2 & 0\cdot 0 + 2\cdot 1 \\ 1\cdot 1 + 1\cdot(-1) & 1\cdot 0 + 1\cdot 2 & 1\cdot 3 + 1\cdot 2 & 1\cdot 0 + 1\cdot 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 2 & 8 & 1 \\ -2 & 4 & 4 & 2 \\ 0 & 2 & 5 & 1 \end{pmatrix}. \end{aligned}$$

Some results on matrix multiplication are stated below.

Theorem 2.2.9 *Let $m, n, q, r \in \mathbb{N}$. If A is an $m \times n$ matrix, B is an $n \times q$ matrix, and C is a $q \times r$ matrix, then $(AB)C = A(BC)$.*

Proof. By definition, $AB = (\sum_{k=1}^{n} a_{ik} b_{kj})$. Therefore,

$$\begin{aligned} (AB)C &= \left(\sum_{l=1}^{q} \left(\sum_{k=1}^{n} a_{ik} b_{kl} \right) c_{lj} \right) \\ &= \left(\sum_{l=1}^{q} \sum_{k=1}^{n} a_{ik} b_{kl} c_{lj} \right) \\ &= \left(\sum_{k=1}^{n} \sum_{l=1}^{q} a_{ik} b_{kl} c_{lj} \right) \\ &= \left(\sum_{k=1}^{n} a_{ik} \left(\sum_{l=1}^{q} b_{kl} c_{lj} \right) \right) \\ &= A(BC). \quad \| \end{aligned}$$

Theorem 2.2.10 *Let $m, n, r \in \mathbb{N}$. If A is an $m \times n$ matrix and B and C are $n \times r$ matrices, then $A(B + C) = AB + AC$.*

Theorem 2.2.11 *Let $m, n, r \in \mathbb{N}$. If A and B are $m \times n$ matrices and C is an $n \times r$ matrix, then $(A + B)C = AC + BC$.*

The proofs of Theorems 2.2.10 and 2.2.11 are left as exercises.

Theorem 2.2.12 *Let $m, n, r \in \mathbb{N}$. If A is an $m \times n$ matrix and B is an $n \times r$ matrix, then $(AB)' = B'A'$.*

Proof. First, note that $(AB)'$ is an $r \times m$ matrix, and $B'A'$ is an $r \times m$ matrix. Because $AB = (\sum_{k=1}^{n} a_{ik} b_{kj})$, $(AB)' = (\sum_{k=1}^{n} a_{jk} b_{ki})$. Furthermore,

$$B'A' = \left(\sum_{k=1}^{n} b_{ki} a_{jk} \right) = \left(\sum_{k=1}^{n} a_{jk} b_{ki} \right),$$

and therefore, $(AB)' = B'A'$. $\|$

Theorem 2.2.9 describes the *associative* laws of matrix multiplication, and Theorems 2.2.10 and 2.2.11 contain the *distributive* laws for matrix addition and multiplication. Note that matrix multiplication is *not* commutative, that is, in general, AB is *not* equal to BA. If A is an $m \times n$ matrix, B is an $n \times r$ matrix, and $m \neq r$, the product BA is not even defined, even though AB *is* defined. If A is an $m \times n$ matrix, B is an $n \times m$ matrix, and $m \neq n$, both products AB and BA are defined, but these matrices are not of the same type—AB is an $m \times m$ matrix, whereas BA is an $n \times n$ matrix. Clearly, if $m \neq n$, these matrices cannot be equal. If A and B both are $n \times n$ matrices, AB and BA are defined and are of the same type (both are $n \times n$ matrices), but still, AB and BA are not necessarily equal. For example, consider

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & -1 \\ 6 & 7 \end{pmatrix}.$$

Then we obtain

$$AB = \begin{pmatrix} 12 & 13 \\ 24 & 25 \end{pmatrix}, \quad BA = \begin{pmatrix} -3 & -4 \\ 27 & 40 \end{pmatrix}.$$

Clearly, $AB \neq BA$.

Therefore, the *order* in which we form a matrix product is important. For a matrix product AB, we use the terminology "A is *post*multiplied by B" or "B is *pre*multiplied by A".

The *rank* of a matrix will be of importance in solving systems of linear equations later on in this chapter. First, we define the *row rank* and the *column rank* of a matrix.

Definition 2.2.13 *Let $m, n \in \mathbb{N}$, and let A be an $m \times n$ matrix.*

(i) The row rank *of A, $R_r(A)$, is the maximal number of linearly independent row vectors in A.*
(ii) The column rank *of A, $R_c(A)$, is the maximal number of linearly independent column vectors in A.*

It can be shown that, for *any* matrix A, the row rank of A is equal to the column rank of A (see the next section for more details). Therefore, we can define the *rank* of an $m \times n$ matrix A as

$$R(A) := R_r(A) = R_c(A).$$

For example, consider the matrix

$$A = \begin{pmatrix} 2 & 0 \\ 1 & 1 \end{pmatrix}.$$

The column vectors of A,

$$\begin{pmatrix} 2 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

are linearly independent, and so are the row vectors $(2, 0)$ and $(1, 1)$ (show this as an exercise). Therefore, the maximal number of linearly independent row (column) vectors in A is 2, which implies $R(A) = 2$.

Now consider the matrix

$$B = \begin{pmatrix} 2 & -4 \\ 1 & -2 \end{pmatrix}.$$

We have

$$\begin{pmatrix} -4 \\ -2 \end{pmatrix} = -2 \begin{pmatrix} 2 \\ 1 \end{pmatrix},$$

and therefore, the column vectors of B are linearly dependent. The maximal number of linearly independent column vectors is one (because we can find a column vector that is not equal to $\mathbf{0}$), and therefore, $R(B) = R_c(B) = 1$. (As an exercise, show that the row rank of B is equal to one.)

Special matrices that are of importance are *null matrices* and *identity matrices*.

Definition 2.2.14 *Let $m, n \in \mathbb{N}$. The $m \times n$* null matrix *is defined by*

$$\mathbf{0} := \begin{pmatrix} 0 & 0 & \ldots & 0 \\ 0 & 0 & \ldots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & 0 \end{pmatrix}.$$

Definition 2.2.15 *Let $n \in \mathbb{N}$. The $n \times n$* identity matrix *$E = (e_{ij})$ is defined by*

$$e_{ij} := \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j \end{cases} \quad \forall i, j = 1, \ldots, n.$$

Hence, all entries of a null matrix are equal to zero, and an identity matrix has *ones* along the main diagonal, and all other entries are equal to zero. Note that only *square* identity matrices are defined.

If the $m \times n$ null matrix is added to *any* $m \times n$ matrix A, the resulting matrix is the matrix A itself. Therefore, null matrices play a role analogous to the role played by the number zero for the addition of real numbers (the role of the *neutral element* for addition). Analogously, the $n \times n$ identity matrix has the property $AE = A$ for all $m \times n$ matrices A and $EB = B$ for all $n \times m$ matrices B. Therefore, the identity matrix is the neutral element for postmultiplication and premultiplication of matrices.

2.3 Systems of Linear Equations

Solving systems of equations is a frequently occuring problem in economic theory. For example, *equilibrium* conditions can be formulated as equations, and if we want to look for equilibria in several markets simultaneously, we obtain a whole set (or *system*) of equations. In this section, we deal with the special case where these equations are *linear*. For $m, n \in \mathbb{N}$, a *system of m linear equations in n variables* can be written as

$$\begin{array}{rcl} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n & = & b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n & = & b_2 \\ \vdots & & \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n & = & b_m \end{array} \tag{2.6}$$

where the a_{ij} and the b_i, $i = 1, \ldots, m$, $j = 1, \ldots, n$, are given real numbers. Clearly, (2.6) can be written in matrix notation as

$$Ax = b$$

where A is an $m \times n$ matrix and $b \in \mathbb{R}^m$. An alternative way of formulating (2.6) (which will be convenient in employing a specific solution method) is

$$\left(\begin{array}{cccc|c} x_1 & x_2 & \ldots & x_n & \\ \hline a_{11} & a_{12} & \ldots & a_{1n} & b_1 \\ a_{21} & a_{22} & \ldots & a_{2n} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} & b_m \end{array} \right).$$

To solve (2.6) means to find a *solution vector* $x^* \in \mathbb{R}^n$ such that all m equations in (2.6) are satisfied. A solution to (2.6) need not exist, and if a solution exists, it need not be unique.

Clearly, the solvability of a system of linear equations will depend on the properties of A and b. We will derive a general method of solving systems of linear equations and provide necessary and sufficient conditions on A and b for the existence of a solution.

The basic idea underlying the method of solution described below—the *Gaussian elimination method*—is that certain *transformations* of a system of linear equations do not affect the solution of the system (by a solution, we mean the set of solution vectors, which may, of course, be empty), such that the solution of the transformed system is easy to find.

We say that two systems of equations are *equivalent* if they have the same solution. Because *renumbering* the variables involved does not change the property of a vector solving a system of linear equations, this possibility is included by allowing for *permutations* of the columns of the system. Formally,

Definition 2.3.1 *Let $m, n \in \mathbb{N}$, and let π be a permutation of $\{1, \ldots, n\}$. Furthermore, let A and C be $m \times n$ matrices, and let $b, d \in \mathbb{R}^m$. The systems of linear equations $Ax = b$ and $Cx_\pi = d$ are* equivalent *if and only if, for all $x \in \mathbb{R}^n$,*

$$x \text{ solves } Ax = b \Leftrightarrow x_\pi \text{ solves } Cx_\pi = d.$$

We denote the equivalence of two systems of equations by $(Ax = b) \sim (Cx_\pi = d)$.

The following theorem provides a list of transformations that lead to equivalent systems of equations.

Theorem 2.3.2 *Let $m, n \in \mathbb{N}$, and let π be a permutation of $\{1, \ldots, n\}$. Furthermore, let A be an $m \times n$ matrix, and let $b \in \mathbb{R}^m$.*

(i) If two equations in the system of linear equations $Ax = b$ are interchanged, the resulting system is equivalent to $Ax = b$.
(ii) If an equation in the system $Ax = b$ is multiplied by a real number $\alpha \neq 0$, the resulting system of equations is equivalent to $Ax = b$.
(iii) If an equation in the system $Ax = b$ is replaced by the sum of this equation and a multiple of another *equation in $Ax = b$, the resulting system is equivalent to $Ax = b$.*
(iv) If A_π is the $m \times n$ matrix obtained by applying the permutation π to the n column vectors of A, the system $A_\pi x_\pi = b$ is equivalent to $Ax = b$.

The proof of Theorem 2.3.2 is left as an exercise (that the transformations defined in this theorem lead to equivalent systems of equations can be verified by simple substitution).

To illustrate the use of Theorem 2.3.2, consider the following example. Let

$$A = \begin{pmatrix} 2 & -1 \\ -4 & 0 \end{pmatrix}, \quad b = \begin{pmatrix} 2 \\ 4 \end{pmatrix}.$$

Then the system of linear equations $Ax = b$ is equivalent to

$$\left(\begin{array}{cc|c} x_1 & x_2 & \\ \hline 2 & -1 & 2 \\ -4 & 0 & 4 \end{array}\right) \sim \left(\begin{array}{cc|c} x_2 & x_1 & \\ \hline -1 & 2 & 2 \\ 0 & -4 & 4 \end{array}\right) \sim \left(\begin{array}{cc|c} x_2 & x_1 & \\ \hline 1 & -2 & -2 \\ 0 & -4 & 4 \end{array}\right)$$

(application of (iv)—exchange columns 1 and 2; application of (ii)—multiply Equation 1 by -1)

$$\sim \left(\begin{array}{cc|c} x_2 & x_1 & \\ \hline 1 & -2 & -2 \\ 0 & 1 & -1 \end{array}\right) \sim \left(\begin{array}{cc|c} x_2 & x_1 & \\ \hline 1 & 0 & -4 \\ 0 & 1 & -1 \end{array}\right)$$

(application of (ii)—multiply Equation 2 by $-1/4$; application of (iii)—add two times Equation 2 to Equation 1). The solution of this system of linear equations (and therefore, the solution of the equivalent system $Ax = b$) can be obtained easily. It is $x_2^* = -4$ and $x_1^* = -1$, which means we obtain the unique solution vector $x^* = (-1, -4)$.

This procedure can be generalized. The transformations mentioned in Theorem 2.3.2 allow us to find, for *any* system of linear equations, an equivalent system with a simple structure. This procedure—the Gaussian elimination method—is described below.

Let $m, n \in \mathbb{N}$, let A be an $m \times n$ matrix, and let $b \in \mathbb{R}^m$. By repeated application of Theorem 2.3.2, we can find a permutation π of $\{1, \ldots, n\}$, a number $k \in \{1, \ldots, n\}$, an $m \times n$ matrix C, and a vector $d \in \mathbb{R}^m$ such that $(Cx_\pi = d) \sim (Ax = b)$ and

$$\left(\begin{array}{ccccccc|c} x_{\pi(1)} & x_{\pi(2)} & \ldots & x_{\pi(k)} & x_{\pi(k+1)} & \ldots & x_{\pi(n)} & \\ \hline 1 & 0 & \ldots & 0 & c_{1(k+1)} & \ldots & c_{1n} & d_1 \\ 0 & 1 & \ldots & 0 & c_{2(k+1)} & \ldots & c_{2n} & d_2 \\ \vdots & & & \vdots & \vdots & & \vdots & \vdots \\ 0 & & \ldots & 1 & c_{k(k+1)} & \ldots & c_{kn} & d_k \\ 0 & & \ldots & 0 & 0 & \ldots & 0 & d_{k+1} \\ \vdots & & & \vdots & \vdots & & \vdots & \vdots \\ 0 & & \ldots & 0 & 0 & \ldots & 0 & d_m \end{array}\right). \tag{2.7}$$

The following example illustrates the application of the Gaussian elimination method. Consider the following system of linear equations. Applying the elimination method, we obtain

$$\left(\begin{array}{ccccc|c} x_1 & x_2 & x_3 & x_4 & x_5 & \\ \hline 1 & 0 & 2 & 0 & 1 & 1 \\ 2 & -2 & 1 & 3 & 0 & 2 \\ 0 & -1 & 0 & 1 & 0 & 2 \\ 0 & 0 & -3 & 1 & -2 & -4 \end{array}\right) \sim \left(\begin{array}{ccccc|c} x_1 & x_2 & x_3 & x_4 & x_5 & \\ \hline 1 & 0 & 2 & 0 & 1 & 1 \\ 0 & -2 & -3 & 3 & -2 & 0 \\ 0 & -1 & 0 & 1 & 0 & 2 \\ 0 & 0 & -3 & 1 & -2 & -4 \end{array}\right)$$

(add -2 times Equation 1 to Equation 2)

$$\sim \left(\begin{array}{ccccc|c} x_1 & x_2 & x_3 & x_4 & x_5 & \\ \hline 1 & 0 & 2 & 0 & 1 & 1 \\ 0 & -1 & 0 & 1 & 0 & 2 \\ 0 & -2 & -3 & 3 & -2 & 0 \\ 0 & 0 & -3 & 1 & -2 & -4 \end{array}\right) \sim \left(\begin{array}{ccccc|c} x_1 & x_2 & x_3 & x_4 & x_5 & \\ \hline 1 & 0 & 2 & 0 & 1 & 1 \\ 0 & 1 & 0 & -1 & 0 & -2 \\ 0 & -2 & -3 & 3 & -2 & 0 \\ 0 & 0 & -3 & 1 & -2 & -4 \end{array}\right)$$

(interchange Equations 2 and 3; multiply Equation 2 by -1)

$$\sim \left(\begin{array}{ccccc|c} x_1 & x_2 & x_3 & x_4 & x_5 & \\ \hline 1 & 0 & 2 & 0 & 1 & 1 \\ 0 & 1 & 0 & -1 & 0 & -2 \\ 0 & 0 & -3 & 1 & -2 & -4 \\ 0 & 0 & -3 & 1 & -2 & -4 \end{array}\right) \sim \left(\begin{array}{ccccc|c} x_1 & x_2 & x_3 & x_4 & x_5 & \\ \hline 1 & 0 & 2 & 0 & 1 & 1 \\ 0 & 1 & 0 & -1 & 0 & -2 \\ 0 & 0 & -3 & 1 & -2 & -4 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right) \tag{2.8}$$

(add 2 times Equation 2 to Equation 3; add -1 times Equation 3 to Equation 4)

$$\sim \left(\begin{array}{ccccc|c} x_1 & x_2 & x_4 & x_3 & x_5 & \\ \hline 1 & 0 & 0 & 2 & 1 & 1 \\ 0 & 1 & -1 & 0 & 0 & -2 \\ 0 & 0 & 1 & -3 & -2 & -4 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right) \sim \left(\begin{array}{ccccc|c} x_1 & x_2 & x_4 & x_3 & x_5 & \\ \hline 1 & 0 & 0 & 2 & 1 & 1 \\ 0 & 1 & 0 & -3 & -2 & -6 \\ 0 & 0 & 1 & -3 & -2 & -4 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

(interchange columns 3 and 4; add Equation 3 to Equation 2). We have now transformed the original system of linear equations into an equivalent system of the form (2.7). It is now easy to see that any $x \in \mathbb{R}^5$ satisfying

$$\begin{aligned} x_1 &= 1 - 2x_3 - x_5 \\ x_2 &= -6 + 3x_3 + 2x_5 \\ x_4 &= -4 + 3x_3 + 2x_5 \end{aligned}$$

is a solution. Therefore, we can choose $\alpha_1 := x_3$ and $\alpha_2 := x_5$ *arbitrarily*, and any solution x^* can be written as

$$x^* = \begin{pmatrix} x_1^* \\ x_2^* \\ x_3^* \\ x_4^* \\ x_5^* \end{pmatrix} = \begin{pmatrix} 1 \\ -6 \\ 0 \\ -4 \\ 0 \end{pmatrix} + \alpha_1 \begin{pmatrix} -2 \\ 3 \\ 1 \\ 3 \\ 0 \end{pmatrix} + \alpha_2 \begin{pmatrix} -1 \\ 2 \\ 0 \\ 2 \\ 1 \end{pmatrix}.$$

Now let us modify this example by changing b_4 from -4 to 0. The resulting system of linear equations is

$$\left(\begin{array}{ccccc|c} x_1 & x_2 & x_3 & x_4 & x_5 & \\ \hline 1 & 0 & 2 & 0 & 1 & 1 \\ 2 & -2 & 1 & 3 & 0 & 2 \\ 0 & -1 & 0 & 1 & 0 & 2 \\ 0 & 0 & -3 & 1 & -2 & 0 \end{array}\right).$$

Using the same transformations that led to (2.8) in the previous example, it follows that this system is equivalent to

$$\left(\begin{array}{ccccc|c} x_1 & x_2 & x_3 & x_4 & x_5 & \\ \hline 1 & 0 & 2 & 0 & 1 & 1 \\ 0 & 1 & 0 & -1 & 0 & -2 \\ 0 & 0 & -3 & 1 & -2 & -4 \\ 0 & 0 & -3 & 1 & -2 & 0 \end{array}\right) \sim \left(\begin{array}{ccccc|c} x_1 & x_2 & x_3 & x_4 & x_5 & \\ \hline 1 & 0 & 2 & 0 & 1 & 1 \\ 0 & 1 & 0 & -1 & 0 & -2 \\ 0 & 0 & -3 & 1 & -2 & -4 \\ 0 & 0 & 0 & 0 & 0 & 4 \end{array}\right)$$

(add -1 times Equation 3 to Equation 4). Clearly, this system of equations *cannot* have a solution, because the last equation requires $0 = 4$, which is, of course, impossible.

In general, whenever *any* of the numbers $d_{k+1}, \ldots, d_m$ in (2.7) is different from zero, we see that the corresponding system of linear equations cannot have a solution.

Theorem 2.3.2 and (2.7) can also be used to show that the column rank of a matrix must be equal to the row rank of a matrix (see Section 2.2). This follows from the observation that the transformations mentioned in Theorem 2.3.2, if applied to a matrix, leave the column and row rank of this matrix unchanged.

Clearly, $Ax = b$ has a solution if and only if $d_{k+1} = \ldots = d_m = 0$ in (2.7). Because the transformations used in the elimination method do not change the rank of A and of (A, b), where

$$(A,b) := \begin{pmatrix} a_{11} & \ldots & a_{1n} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{m1} & \ldots & a_{mn} & b_m \end{pmatrix},$$

we obtain

Theorem 2.3.3 *Let $m, n \in \mathbb{N}$, let A be an $m \times n$ matrix, and let $b \in \mathbb{R}^m$. The system $Ax = b$ has at least one solution if and only if $R(A) = R(A, b)$.*

Theorem 2.3.3 gives a precise answer to the question under which conditions solutions to a system of linear equations exist.

The question whether vectors are linearly independent can be formulated in terms of a system of linear equations. The vectors $x^1, \ldots, x^m \in \mathbb{R}^n$, where $m, n \in \mathbb{N}$, are linearly independent if and only if the only solution to the system of linear equations

$$\alpha_1 x^1 + \ldots + \alpha_m x^m = \mathbf{0}$$

is $\alpha_1^* = \ldots = \alpha_m^* = 0$. We now show that *at most* n vectors of dimension n can be linearly independent.

Theorem 2.3.4 *Let $m, n \in \mathbb{N}$, and let $x^1, \ldots, x^m \in \mathbb{R}^n$. If $m > n$, then $x^1, \ldots x^m$ are linearly dependent.*

Proof. Consider the system of linear equations

$$\left(\begin{array}{ccc|c} \alpha_1 & \ldots & \alpha_m & \\ \hline x_1^1 & \ldots & x_1^m & 0 \\ \vdots & & \vdots & \vdots \\ x_n^1 & \ldots & x_n^m & 0 \end{array}\right). \tag{2.9}$$

Using the Gaussian elimination method, we can find an equivalent system of the form

$$\left(\begin{array}{ccccccc|c} \alpha_{\pi(1)} & \alpha_{\pi(2)} & \ldots & \alpha_{\pi(k)} & \alpha_{\pi(k+1)} & \ldots & \alpha_{\pi(m)} & \\ \hline 1 & 0 & \ldots & 0 & c_{1(k+1)} & \ldots & c_{1m} & 0 \\ 0 & 1 & \ldots & 0 & c_{2(k+1)} & \ldots & c_{2m} & 0 \\ \vdots & & & \vdots & \vdots & & \vdots & \vdots \\ 0 & & \ldots & 1 & c_{k(k+1)} & \ldots & c_{km} & 0 \\ 0 & & \ldots & 0 & 0 & \ldots & 0 & 0 \\ \vdots & & & \vdots & \vdots & & \vdots & \vdots \\ 0 & & \ldots & 0 & 0 & \ldots & 0 & 0 \end{array}\right)$$

(note that, because the right sides in (2.9) are equal to zero, the right sides of the transformed system must be equal to zero—the transformations that are used to obtain this equivalent system leave these zeroes unchanged). Because $k \leq n < m$, the set $\{\alpha_{\pi(k+1)}, \ldots, \alpha_{\pi(m)}\}$ is nonempty. Therefore, we can choose $\alpha_{\pi(k+1)}, \ldots, \alpha_{\pi(m)}$ *arbitrarily* (in particular, different from zero) and $\alpha_{\pi(1)}, \ldots, \alpha_{\pi(k)}$ so that the equations in the above system are satisfied. But this means there exist $\alpha_1, \ldots, \alpha_m \in \mathbb{R}$ satisfying (2.9) such that at least one α_j is different from zero. This means that the vectors $x^1, \ldots x^m$ must be linearly dependent. ‖

We conclude this section with a special case of systems of linear equations, where the number of equations is equal to the number of variables. These systems are important in many applications.

The following result deals with the existence *and* uniqueness of a solution to a system of n equations in n variables, where $n \in \mathbb{N}$. In many economic models, not only existence, but also uniqueness of a solution is of importance, because multiple solutions can pose a serious selection problem and can complicate the analysis of the behaviour of economic variables considerably. Again, the *rank* of the matrix of coefficients is the crucial factor in determining whether or not a unique solution to a system of linear equations exists.

Theorem 2.3.5 *Let $n \in \mathbb{N}$, and let A be an $n \times n$ matrix. $Ax = b$ has a unique solution for all $b \in \mathbb{R}^n$ if and only if $R(A) = n$.*

Proof. "Only if": By way of contradiction, suppose the system $Ax = b$ has a unique solution $x^* \in \mathbb{R}^n$ for all $b \in \mathbb{R}^n$, and $R(A) < n$. Then one of the row vectors in A can be written as a linear combination of the remaining row vectors. Without loss of generality, suppose

$$(a_{n1}, \ldots, a_{nn}) = \alpha_1(a_{11}, \ldots, a_{1n}) + \ldots + \alpha_{n-1}(a_{(n-1)1}, \ldots, a_{(n-1)n}) \tag{2.10}$$

with $\alpha_1, \ldots, \alpha_{n-1} \in \mathbb{R}$. Let

$$b = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} \in \mathbb{R}^n,$$

and consider the system of equations

$$Ax = b. \tag{2.11}$$

Now successively add $-\alpha_1$ times the first, $-\alpha_2$ times the second,..., $-\alpha_{n-1}$ times the $(n-1)^{st}$ equation to the n^{th} equation. By (2.10) and the definition of b, the n^{th} equation becomes

$$0 = 1,$$

which means that (2.11) cannot have a solution for $b \in \mathbb{R}^n$ as chosen above. This is a contradiction.

"If": Suppose $R(A) = n$. Let $b \in \mathbb{R}^n$. We have to prove

(i) the existence,
(ii) the uniqueness

of a solution to the system $Ax = b$.

(i) *Existence.* By Theorem 2.3.4, the vectors

$$\begin{pmatrix} a_{11} \\ \vdots \\ a_{n1} \end{pmatrix}, \ldots, \begin{pmatrix} a_{1n} \\ \vdots \\ a_{nn} \end{pmatrix}, \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$$

must be linearly dependent ($n+1$ vectors in $\mathbb{R}^n$ cannot be linearly independent). Therefore, there exist $\alpha_1, \ldots, \alpha_n, \beta \in \mathbb{R}$ such that

$$\alpha_1 \begin{pmatrix} a_{11} \\ \vdots \\ a_{n1} \end{pmatrix} + \ldots + \alpha_n \begin{pmatrix} a_{1n} \\ \vdots \\ a_{nn} \end{pmatrix} + \beta \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix},$$

where at least one of the real numbers $\alpha_1, \ldots, \alpha_n, \beta$ must be different from zero.

If $\beta = 0$,

$$\alpha_1 \begin{pmatrix} a_{11} \\ \vdots \\ a_{n1} \end{pmatrix} + \ldots + \alpha_n \begin{pmatrix} a_{1n} \\ \vdots \\ a_{nn} \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix},$$

and there exists $k \in \{1, \ldots, n\}$ such that $\alpha_k \neq 0$. But this contradicts the assumption $R(A) = n$. Therefore, $\beta \neq 0$. This implies

$$\left(-\frac{\alpha_1}{\beta}\right) \begin{pmatrix} a_{11} \\ \vdots \\ a_{n1} \end{pmatrix} + \ldots + \left(-\frac{\alpha_n}{\beta}\right) \begin{pmatrix} a_{1n} \\ \vdots \\ a_{nn} \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix},$$

which means that

$$x^* = \left(-\frac{\alpha_1}{\beta}, \ldots, -\frac{\alpha_n}{\beta}\right)$$

solves $Ax = b$. This proves (i).

(ii) *Uniqueness.* Suppose $Ax = b$ and $Ay = b$, where $x, y \in \mathbb{R}^n$. We have to show that $x = y$. Subtracting the second of the above systems from the first, we obtain $A(x - y) = \mathbf{0}$, which can be written as

$$(x_1 - y_1)\begin{pmatrix} a_{11} \\ \vdots \\ a_{n1} \end{pmatrix} + \ldots + (x_n - y_n)\begin{pmatrix} a_{1n} \\ \vdots \\ a_{nn} \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}. \tag{2.12}$$

Because $R(A) = n$, the column vectors in A are linearly independent, and therefore, (2.12) can be satisfied only if $x_i = y_i$ for all $i = 1, \ldots, n$, that is, $x = y$. ‖

2.4 The Inverse of a Matrix

Recall that for any real number $x \neq 0$, there exists a number $y \in \mathbb{R}$ such that $xy = 1$, namely, $y = 1/x$. Analogously, we could ask ourselves whether, for an $n \times n$ matrix A, we can find a matrix B such that $AB = E$. As we will see, this is not the case for all $n \times n$ matrices—there are square matrices A (and the null matrix is not the only one) such that there exists no such matrix B.

We will use the following terminology.

Definition 2.4.1 *Let $n \in \mathbb{N}$, and let A be an $n \times n$ matrix.*

(i) A is nonsingular *if and only if there exists a matrix B such that $AB = E$.*
(ii) A is singular *if and only if A is not nonsingular.*

For example, consider the matrix

$$A = \begin{pmatrix} 2 & 0 \\ 1 & 1 \end{pmatrix}.$$

For

$$B = \begin{pmatrix} 1/2 & 0 \\ -1/2 & 1 \end{pmatrix},$$

we obtain $AB = E$, and therefore, the matrix A is nonsingular.

Now let

$$A = \begin{pmatrix} 2 & -4 \\ 1 & -2 \end{pmatrix}.$$

A is nonsingular if and only if we can find a 2×2 matrix $B = (b_{ij})$ such that

$$\begin{pmatrix} 2 & -4 \\ 1 & -2 \end{pmatrix}\begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

This implies that B must satisfy

$$(2b_{12} - 4b_{22} = 0) \wedge (b_{12} - 2b_{22} = 1).$$

But this is equivalent to

$$(b_{12} - 2b_{22} = 0) \wedge (b_{12} - 2b_{22} = 1),$$

and clearly, there exists no such matrix B. Therefore, A is singular.

Nonsingularity of a square matrix A is *equivalent* to A having *full rank*.

Theorem 2.4.2 *Let $n \in \mathbb{N}$, and let A be an $n \times n$ matrix. A is nonsingular if and only if $R(A) = n$.*

Proof. "Only if": Suppose A is nonsingular. Then there exists a matrix B such that $AB = E$. Let $b \in \mathbb{R}^n$. Then

$$b = Eb = (AB)b = A(Bb).$$

This means that the system of linear equations $Ax = b$ has a solution x^* (namely, $x^* = Bb$) for *all* $b \in \mathbb{R}^n$. By the argument used in the proof of Theorem 2.3.5, this implies $R(A) = n$.

"If": Suppose $R(A) = n$. By Theorem 2.3.5, this implies that $Ax = b$ has a unique solution for all $b \in \mathbb{R}^n$. In particular, for $b = e^i$, where $i \in \{1, \ldots, n\}$, there exists a unique $x^{*i} \in \mathbb{R}^n$ such that $Ax^{*i} = e^i$. Because the matrix composed of the column vectors $(e^1, \ldots, e^n)$ is the identity matrix E, it follows that $AB = E$, where $B := (x^{*1}, \ldots, x^{*n})$. ‖

In the second part of the above proof, we *constructed* a matrix B such that $AB = E$ for a given square matrix A with full rank. Note that, because of the uniqueness of the solution x^{*i}, $i = 1, \ldots, n$, the resulting matrix $B = (x^{*1}, \ldots, x^{*n})$ is uniquely determined. Therfore, for any nonsingular square matrix A, there exists *exactly one* matrix B such that $AB = E$. Therefore, the above proof also has shown

Theorem 2.4.3 *Let $n \in \mathbb{N}$, and let A be an $n \times n$ matrix. If $AB = E$ and $AC = E$, then $B = C$.*

For a nonsingular square matrix A, we call the unique matrix B such that $AB = E$ the *inverse of* A, and we denote it by A^{-1}.

Definition 2.4.4 *Let $n \in \mathbb{N}$, and let A be a nonsingular $n \times n$ matrix. The unique matrix A^{-1} such that*

$$AA^{-1} = E$$

is called the inverse matrix *of A.*

Some important properties of inverse matrices are summarized below.

Theorem 2.4.5 *Let $n \in \mathbb{N}$. If A is a nonsingular $n \times n$ matrix, then A^{-1} is nonsingular and $(A^{-1})^{-1} = A$.*

Proof. Because nonsingularity is equivalent to $R(A) = n$, and because $R(A') = R(A)$, it follows that A' is nonsingular. Therefore, there exists a unique matrix B such that $A'B = E$. Clearly, $E' = E$, and therefore,

$$E = E' = (A'B)' = B'A,$$

and we obtain

$$B' = B'E = B'(AA^{-1}) = (B'A)A^{-1} = EA^{-1} = A^{-1}.$$

This implies $A^{-1}A = B'A = E$, and hence, A is the inverse of A^{-1}. ‖

Theorem 2.4.6 *Let $n \in \mathbb{N}$. If A is a nonsingular $n \times n$ matrix, then A' is nonsingular and $(A')^{-1} = (A^{-1})'$.*

Proof. That A' is nonsingular follows from $R(A') = R(A)$ and the nonsingularity of A. Using Theorem 2.4.5, we obtain

$$E = E' = (A^{-1}A)' = A'(A^{-1})'.$$

Therefore, $(A')^{-1} = (A^{-1})'$. ‖

Theorem 2.4.7 *Let $n \in \mathbb{N}$. If A and B are nonsingular $n \times n$ matrices, then AB is nonsingular and $(AB)^{-1} = B^{-1}A^{-1}$.*

Proof. By the rules of matrix multiplication,

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AEA^{-1} = AA^{-1} = E.$$

Therefore, $(AB)^{-1} = B^{-1}A^{-1}$. ‖

2.5 Determinants

The *determinant* of a square matrix A is a real number associated with A that is useful for several purposes. We first consider a geometric interpretation of the determinant of a 2×2 matrix.

Consider two vectors $a^1, a^2 \in \mathbb{R}^2$. Suppose we want to find the *area* of the parallelogram spanned by the two vectors a^1 and a^2. Figure 2.5 illustrates this area V for the vectors $a^1 = (3, 1)$ and $a^2 = (1, 2)$.

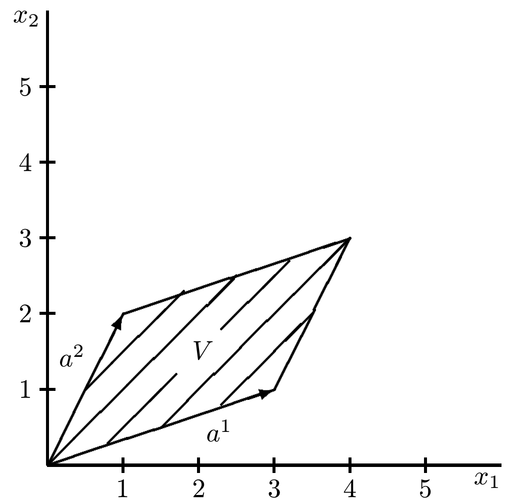


Figure 2.5: Geometric interpretation of determinants.

Of course, we would like to be able to assign a real number representing this area to *any* pair of vectors $a^1, a^2 \in \mathbb{R}^2$. Letting $\mathcal{A}^2$ denote the set of all 2×2 matrices, this means that we want to find a function $f : \mathcal{A}^2 \mapsto \mathbb{R}$ such that, for any $A \in \mathcal{A}^2$, $f(A)$ is interpreted as the area of the parallelogram spanned by the column vectors (or the row vectors) of A.

It can be shown that the *only* function that has certain plausible properties (properties that we would expect from a function assigning the area of such a parallelogram to each matrix) is the function that assigns the absolute value of the determinant of A to each matrix $A \in \mathcal{A}^2$. More generally, the only function having such properties in *any* dimension $n \in \mathbb{N}$ is the function that assigns the absolute value of A's determinant to each square matrix A.

In addition to this geometrical interpretation, the determinant of a matrix has several very important properties, as we will see later. In order to introduce the determinant of a square matrix formally, we need some further definitions involving permutations.

Definition 2.5.1 *Let $n \in \mathbb{N}$, and let $\pi : \{1, \ldots, n\} \mapsto \{1, \ldots, n\}$ be a permutation of $\{1, \ldots, n\}$. The numbers $\pi(i)$ and $\pi(j)$ form an* inversion *(in π) if and only if*

$$i < j \wedge \pi(i) > \pi(j).$$

For example, let $n = 3$, and define the permutation

$$\pi : \{1, 2, 3\} \mapsto \{1, 2, 3\},\ x \mapsto \begin{cases} 1 & \text{if } x = 2 \\ 2 & \text{if } x = 3 \\ 3 & \text{if } x = 1 \end{cases} \tag{2.13}$$

Then $\pi(1)$ and $\pi(2)$ form an inversion, because $\pi(1) > \pi(2)$, and $\pi(1)$ and $\pi(3)$ form an inversion, because $\pi(1) > \pi(3)$. Therefore, there are two inversions in the above permutation. We define

Definition 2.5.2 *Let $n \in \mathbb{N}$, and let $\pi : \{1, \ldots, n\} \mapsto \{1, \ldots, n\}$ be a permutation of $\{1, \ldots, n\}$. The number of inversions in π is denoted by $N(\pi)$.*

Definition 2.5.3 *Let $n \in \mathbb{N}$, and let $\pi : \{1, \ldots, n\} \mapsto \{1, \ldots, n\}$ be a permutation of $\{1, \ldots, n\}$.*

(i) π is an odd *permutation of $\{1, \ldots, n\} \Leftrightarrow N(\pi)$ is odd.*
(ii) π is an even *permutation of $\{1, \ldots, n\} \Leftrightarrow N(\pi)$ is even $\vee\ N(\pi) = 0$.*

The number of inversions in the permutation π defined in (2.13) is $N(\pi) = 2$, and therefore, π is an even permutation.

Letting Π denote the set of all permutations of $\{1, \ldots, n\}$, the determinant of an $n \times n$ matrix can now be defined.

Definition 2.5.4 *Let $n \in \mathbb{N}$, and let A be an $n \times n$ matrix. The* determinant of A *is defined by*

$$|A| := \sum_{\pi \in \Pi} a_{\pi(1)1} a_{\pi(2)2} \ldots a_{\pi(n)n} (-1)^{N(\pi)}.$$

To illustrate Definition 2.5.4, consider first the case $n = 2$. There are two permutations π^1 and π^2 of $\{1, 2\}$, namely,

$$\pi^1(1) = 1 \wedge \pi^1(2) = 2$$

and

$$\pi^2(1) = 2 \wedge \pi^2(2) = 1.$$

Clearly, $N(\pi^1) = 0$ and $N(\pi^2) = 1$ (that is, π^1 is even and π^2 is odd). Therefore, the determinant of a 2×2 matrix A is given by

$$|A| = a_{11}a_{22}(-1)^0 + a_{21}a_{12}(-1)^1 = a_{11}a_{22} - a_{21}a_{12}.$$

For example, the determinant of

$$A = \begin{pmatrix} 1 & 2 \\ -1 & 1 \end{pmatrix}$$

is

$$|A| = 1 \cdot 1 - (-1) \cdot 2 = 1 + 2 = 3.$$

In general, the determinant of a 2×2 matrix is obtained by subtracting the product of the off-diagonal elements from the product of the elements on the main diagonal.

For $n = 3$, there are six possible permutations, namely,

$$\begin{aligned}
\pi^1(1) = 1 \wedge \pi^1(2) = 2 \wedge \pi^1(3) = 3, \\
\pi^2(1) = 1 \wedge \pi^2(2) = 3 \wedge \pi^2(3) = 2, \\
\pi^3(1) = 2 \wedge \pi^3(2) = 1 \wedge \pi^3(3) = 3, \\
\pi^4(1) = 2 \wedge \pi^4(2) = 3 \wedge \pi^4(3) = 1, \\
\pi^5(1) = 3 \wedge \pi^5(2) = 1 \wedge \pi^5(3) = 2, \\
\pi^6(1) = 3 \wedge \pi^6(2) = 2 \wedge \pi^6(3) = 1.
\end{aligned}$$

We have $N(\pi^1) = 0$, $N(\pi^2) = 1$, $N(\pi^3) = 1$, $N(\pi^4) = 2$, $N(\pi^5) = 2$, and $N(\pi^6) = 3$ (verify this as an exercise). Therefore, the determinant of a 3×3 matrix A is given by

$$\begin{aligned}
|A| &= a_{11}a_{22}a_{33}(-1)^0 + a_{11}a_{32}a_{23}(-1)^1 + a_{21}a_{12}a_{33}(-1)^1 \\
&+ a_{21}a_{32}a_{13}(-1)^2 + a_{31}a_{12}a_{23}(-1)^2 + a_{31}a_{22}a_{13}(-1)^3 \\
&= a_{11}a_{22}a_{33} - a_{11}a_{32}a_{23} - a_{21}a_{12}a_{33} + a_{21}a_{32}a_{13} + a_{31}a_{12}a_{23} - a_{31}a_{22}a_{13}.
\end{aligned}$$

For example, the determinant of

$$A = \begin{pmatrix} 1 & 0 & 4 \\ 2 & 1 & -1 \\ 0 & 1 & 0 \end{pmatrix}$$

is

$$\begin{aligned}
|A| &= 1 \cdot 1 \cdot 0 - 1 \cdot 1 \cdot (-1) - 2 \cdot 0 \cdot 0 + 2 \cdot 1 \cdot 4 + 0 \cdot 0 \cdot (-1) - 0 \cdot 1 \cdot 4 \\
&= 0 + 1 + 0 + 8 + 0 + 0 = 9.
\end{aligned}$$

For large n, the calculation of determinants can become computationally quite involved. In general, there are $n!$ (in words: "n factorial") permutations of $\{1, \ldots, n\}$, where, for $n \in \mathbb{N}$,

$$n! := 1 \cdot 2 \cdot \ldots \cdot n.$$

For example, if A is a 5×5 matrix, the calculation of $|A|$ involves the summation of $5! = 120$ terms. However, it is possible to simplify the calculation of determinants of matrices of higher dimensions. We introduce some more definitions in order to do so.

Definition 2.5.5 *Let $n \in I\!N$, and let A be an $n \times n$ matrix. For $i, j \in \{1, \ldots, n\}$, let A_{ij} denote the $(n-1) \times (n-1)$ matrix that is obtained from A by removing row i and column j. The* cofactor of a_{ij} *is defined by*

$$|C_{ij}| := (-1)^{i+j} |A_{ij}|.$$

Definition 2.5.6 *Let $n \in I\!N$, and let A be an $n \times n$ matrix. The* adjoint of A *is defined by*

$$adj(A) := \begin{pmatrix} |C_{11}| & \ldots & |C_{n1}| \\ \vdots & & \vdots \\ |C_{1n}| & \ldots & |C_{nn}| \end{pmatrix}.$$

Therefore, the adjoint of A is the *transpose* of the matrix of cofactors of A.

The cofactors of a square matrix A can be used to *expand* the determinant of A along a row or a column of A. This expansion procedure is described in the following theorem, which is stated without a proof.

Theorem 2.5.7 *Let $n \in I\!N$, and let A be an $n \times n$ matrix.*

(i) $|A| = \sum_{j=1}^{n} a_{ij} |C_{ij}| \quad \forall i = 1, \ldots, n.$
(ii) $|A| = \sum_{i=1}^{n} a_{ij} |C_{ij}| \quad \forall j = 1, \ldots, n.$

Expanding the determinant of an $n \times n$ matrix A along a row or column of A is a method that proceeds by calculating n determinants of $(n-1) \times (n-1)$ matrices, which can simplify the calculation of a determinant considerably—especially if a matrix has a row or column with many zero entries. For example, consider

$$A = \begin{pmatrix} 1 & 0 & 4 \\ 2 & 1 & -1 \\ 0 & 1 & 0 \end{pmatrix}.$$

If we expand $|A|$ along the third row (which is the natural choice, because this row has many zeroes), we obtain

$$|A| = 0 \cdot (-1)^{3+1} \cdot \begin{vmatrix} 0 & 4 \\ 1 & -1 \end{vmatrix} + 1 \cdot (-1)^{3+2} \cdot \begin{vmatrix} 1 & 4 \\ 2 & -1 \end{vmatrix} + 0 \cdot (-1)^{3+3} \cdot \begin{vmatrix} 1 & 0 \\ 2 & 1 \end{vmatrix} = 0 + (-1) \cdot (-1-8) + 0 = 9.$$

Next, some important properties of determinants are summarized (the proof of the following theorem is omitted).

Theorem 2.5.8 *Let $n \in I\!N$, and let A and B be $n \times n$ matrices.*

(i) $|A'| = |A|$.
(ii) If B is obtained from A by adding a multiple of one row (column) of A to another row (column) of A, then $|B| = |A|$.
(iii) If B is obtained from A by multiplying one row (column) of A by $\alpha \in I\!R$, then $|B| = \alpha|A|$.
(iv) If B is obtained from A by interchanging two rows (columns) of A, then $|B| = -|A|$.
(v) $|AB| = |A||B|$.

Part (v) of the above theorem gives us a convenient way to find the determinant of the inverse of a nonsingular matrix A. First, note that $|E| = 1$ (Exercise: show this). For a nonsingular square matrix A, we have $AA^{-1} = E$, and therefore, $|AA^{-1}| = |E| = 1$. By part (v) of Theorem 2.5.8, $|AA^{-1}| = |A||A^{-1}| = 1$, and therefore,

$$|A^{-1}| = \frac{1}{|A|}.$$

Note that $|A||A^{-1}| = 1$ implies that the determinant of a nonsingular matrix (and the determinant of its inverse) must be different from zero. In fact, a nonzero determinant is *equivalent* to the nonsingularity of a square matrix. Formally,

Theorem 2.5.9 *Let $n \in I\!N$, and let A be an $n \times n$ matrix. A is nonsingular if and only if $|A| \neq 0$.*

For a system of linear equations $Ax = b$, where A is a nonsingular $n \times n$ matrix and $b \in \mathbb{R}^n$, determinants can be used to find the unique solution to this system. The following theorem describes this method, which is known as *Cramer's rule.*

Theorem 2.5.10 *Let $n \in \mathbb{N}$. Furthermore, let A be a nonsingular $n \times n$ matrix, and let $b \in \mathbb{R}^n$. The unique solution $x^* \in \mathbb{R}^n$ to the system of linear equations $Ax = b$ satisfies*

$$x_j^* = \frac{\begin{vmatrix} a_{11} & \ldots & a_{1(j-1)} & b_1 & a_{1(j+1)} & \ldots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \ldots & a_{n(j-1)} & b_n & a_{n(j+1)} & \ldots & a_{nn} \end{vmatrix}}{|A|} \quad \forall j = 1, \ldots, n.$$

Proof. Let $x^* \in \mathbb{R}^n$ be the unique solution to $Ax = b$ (existence and uniqueness of x^* follow from the nonsingularity of A). By part (iii) of Theorem 2.5.8,

$$x_j^*|A| = \begin{vmatrix} a_{11} & \ldots & a_{1(j-1)} & x_j^* a_{1j} & a_{1(j+1)} & \ldots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \ldots & a_{n(j-1)} & x_j^* a_{nj} & a_{n(j+1)} & \ldots & a_{nn} \end{vmatrix} \quad \forall j = 1, \ldots, n.$$

By part (ii) of Theorem 2.5.8, the determinant of a matrix is unchanged if we add a multiple of a column to another column. Repeated application of this property yields

$$x_j^*|A| = \begin{vmatrix} a_{11} & \ldots & a_{1(j-1)} & \sum_{k=1}^n x_k^* a_{1k} & a_{1(j+1)} & \ldots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \ldots & a_{n(j-1)} & \sum_{k=1}^n x_k^* a_{nk} & a_{n(j+1)} & \ldots & a_{nn} \end{vmatrix} \quad \forall j = 1, \ldots, n.$$

Because x^* solves $Ax = b$, $\sum_{k=1}^n x_k^* a_{ik} = b_i$ for all $i = 1, \ldots, n$. Therefore,

$$x_j^*|A| = \begin{vmatrix} a_{11} & \ldots & a_{1(j-1)} & b_1 & a_{1(j+1)} & \ldots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \ldots & a_{n(j-1)} & b_n & a_{n(j+1)} & \ldots & a_{nn} \end{vmatrix} \quad \forall j = 1, \ldots, n. \tag{2.14}$$

Because A is nonsingular, $|A| \neq 0$, and therefore, we can divide both sides of (2.14) by $|A|$ to complete the proof. ‖

As an example, consider the system of equations $Ax = b$, where

$$A = \begin{pmatrix} 2 & 0 & 1 \\ 1 & -1 & 0 \\ 1 & 0 & 3 \end{pmatrix}, \quad b = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}.$$

First, we calculate $|A|$ to determine whether A is nonsingular. Expanding $|A|$ along the second column yields

$$|A| = (-1) \begin{vmatrix} 2 & 1 \\ 1 & 3 \end{vmatrix} = -5 \neq 0.$$

By Cramer's rule,

$$x_1^* = \frac{\begin{vmatrix} 1 & 0 & 1 \\ 0 & -1 & 0 \\ 1 & 0 & 3 \end{vmatrix}}{|A|} = \frac{(-1)\begin{vmatrix} 1 & 1 \\ 1 & 3 \end{vmatrix}}{-5} = \frac{2}{5},$$

$$x_2^* = \frac{\begin{vmatrix} 2 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 3 \end{vmatrix}}{|A|} = \frac{(-1)\begin{vmatrix} 1 & 1 \\ 1 & 3 \end{vmatrix}}{-5} = \frac{2}{5},$$

$$x_3^* = \frac{\begin{vmatrix} 2 & 0 & 1 \\ 1 & -1 & 0 \\ 1 & 0 & 1 \end{vmatrix}}{|A|} = \frac{(-1)\begin{vmatrix} 2 & 1 \\ 1 & 1 \end{vmatrix}}{-5} = \frac{1}{5}.$$

Determinants can also be used to provide an alternative way of finding the inverse of a nonsingular square matrix.

Theorem 2.5.11 *Let $n \in \mathbb{N}$. If A is a nonsingular $n \times n$ matrix, then*

$$A^{-1} = \frac{1}{|A|} adj(A).$$

Proof. By definition, $AA^{-1} = E$. Letting a_{ij}^{-1} denote the element in the i^{th} row and j^{th} column of A^{-1}, we can write this matrix equation as

$$A \begin{pmatrix} a_{1j}^{-1} \\ \vdots \\ a_{nj}^{-1} \end{pmatrix} = e^j \quad \forall j = 1, \ldots, n.$$

By Cramer's rule,

$$a_{ij}^{-1} = \frac{\begin{vmatrix} a_{11} & \ldots & a_{1(i-1)} & e_1^j & a_{1(i+1)} & \ldots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \ldots & a_{n(i-1)} & e_n^j & a_{n(i+1)} & \ldots & a_{nn} \end{vmatrix}}{|A|} \quad \forall i, j = 1, \ldots, n.$$

Expanding along the i^{th} column, we obtain (recall that $e_j^j = 1$ and $e_i^j = 0$ for $i \neq j$)

$$a_{ij}^{-1} = \frac{(-1)^{i+j} A_{ji}}{|A|} = \frac{|C_{ji}|}{|A|} \quad \forall i, j = 1, \ldots, n,$$

that is,

$$A^{-1} = \frac{1}{|A|} adj(A). \quad \|$$

2.6 Quadratic Forms

Quadratic forms are expressions in several variables such that each variable appears either as a square or in a product with another variable. Formally, we can write a quadratic form in n variables $x_1, \ldots, x_n$ in the following way.

Definition 2.6.1 *Let $n \in \mathbb{N}$. Furthermore, let $\alpha_{ij} \in \mathbb{R}$ for all $i = 1, \ldots, n$, $j = 1, \ldots, i$, where $\alpha_{ij} \neq 0$ for at least one α_{ij}. The expression*

$$\sum_{i=1}^{n} \sum_{j=1}^{i} \alpha_{ij} x_i x_j$$

is a quadratic form *in the variables $x_1, \ldots, x_n$.*

Every quadratic form in n variables can be expressed as a matrix product

$$x'Ax$$

where A is the *symmetric* $n \times n$ matrix given by

$$A = \begin{pmatrix} \alpha_{11} & \frac{1}{2}\alpha_{12} & \ldots & \frac{1}{2}\alpha_{1n} \\ \frac{1}{2}\alpha_{12} & \alpha_{22} & \ldots & \frac{1}{2}\alpha_{2n} \\ \vdots & \vdots & & \vdots \\ \frac{1}{2}\alpha_{1n} & \frac{1}{2}\alpha_{2n} & \ldots & \alpha_{nn} \end{pmatrix}. \tag{2.15}$$

It can be shown that, for given α_{ij}, the matrix A defined in (2.15) is *the only* symmetric $n \times n$ matrix such that

$$x'Ax = \sum_{i=1}^{n} \sum_{j=1}^{i} \alpha_{ij} x_i x_j.$$

Therefore, *any* symmetric square matrix uniquely defines a quadratic form.

For example, consider the symmetric matrix

$$A = \begin{pmatrix} 1 & 2 \\ 2 & -1 \end{pmatrix}.$$

The quadratic form defined by this matrix is

$$x'Ax = x_1^2 + 2x_1x_2 + 2x_1x_2 - x_2^2 = x_1^2 + 4x_1x_2 - x_2^2.$$

The properties of a quadratic form are determined by the properties of the symmetric square matrix defining this quadratic form. Of particular importance are the following *definiteness* properties of symmetric square matrices.

Definition 2.6.2 *Let $n \in \mathbb{N}$. A symmetric $n \times n$ matrix A is*

(i) positive definite $\Leftrightarrow x'Ax > 0 \ \ \forall x \in \mathbb{R}^n \setminus \{\mathbf{0}\}$,
(ii) negative definite $\Leftrightarrow x'Ax < 0 \ \ \forall x \in \mathbb{R}^n \setminus \{\mathbf{0}\}$,
(iii) positive semidefinite $\Leftrightarrow x'Ax \geq 0 \ \ \forall x \in \mathbb{R}^n$,
(iv) negative semidefinite $\Leftrightarrow x'Ax \leq 0 \ \ \forall x \in \mathbb{R}^n$.

These properties will be of importance later on when we discuss second-order conditions for the optimization of functions of several variables.

Here are some examples. Let

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}, \ B = \begin{pmatrix} -2 & 1 \\ 1 & -1 \end{pmatrix}, \ C = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}.$$

The quadratic form defined by A is

$$x'Ax = 2x_1^2 + 2x_1x_2 + x_2^2 = x_1^2 + (x_1 + x_2)^2,$$

which is positive for all $x \in \mathbb{R}^2 \setminus \{\mathbf{0}\}$. Therefore, A is positive definite. Similarly, we obtain

$$x'Bx = -2x_1^2 + 2x_1x_2 - x_2^2 = (-1)(x_1^2 + (x_1 - x_2)^2),$$

which is negative for all $x \in \mathbb{R}^2 \setminus \{\mathbf{0}\}$, and hence, B is negative definite. Finally, the quadratic form defined by C is

$$x'Cx = x_1^2 + 4x_1x_2 + x_2^2 = 2x_1x_2 + (x_1 + x_2)^2.$$

For $x' = (1, 1)$, we obtain $x'Cx = 6 > 0$, and substituting $(-1, 1)$ for x' yields $x'Cx = -2 < 0$. Therefore, C is neither positive (semi)definite nor negative (semi)definite. A matrix which is neither positive semidefinite nor negative semidefinite is called *indefinite*.

For matrices of higher dimensions, checking definiteness properties can be a quite involved task. There are some useful criteria for definiteness properties in terms of the determinants of a matrix and some of its submatrices. Some further definitions are needed in order to introduce these criteria.

Definition 2.6.3 *Let $n \in \mathbb{N}$, and let A be an $n \times n$ matrix. Furthermore, let $k \in \{1, \ldots, n\}$.*

(i) A principal submatrix of order k *of A is a $k \times k$ matrix that is obtained by removing $(n-k)$ rows and the $(n-k)$ columns* with the same numbers *from A.*
(ii) The leading principal submatrix of order k *of A is the $k \times k$ matrix that is obtained by removing* the last $(n-k)$ *rows and columns from A. The leading principal submatrix of order k of A is denoted by A_k.*

For example, the principal submatrices of order 2 of

$$A = \begin{pmatrix} 1 & 2 & 0 \\ 0 & -1 & 0 \\ 2 & 1 & -2 \end{pmatrix}$$

are

$$\begin{pmatrix} 1 & 2 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 2 & -2 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ 1 & -2 \end{pmatrix}.$$

The leading principal submatrices of orders $k = 1, 2, 3$ of A are

$$(1), \begin{pmatrix} 1 & 2 \\ 0 & -1 \end{pmatrix}, A,$$

respectively. Note that principal submatrices are obtained by removing rows and columns *with the same numbers.* For the above matrix A, the submatrix

$$\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$$

which is obtained by removing row 2 and column 3 from A is *not* a principal submatrix of A.

The *determinant* of a (leading) principal submatrix is called a *(leading) principal minor.* Hence, the leading principal minor of order k of an $n \times n$ matrix A is

$$|A_k| = \begin{vmatrix} a_{11} & \dots & a_{1k} \\ \vdots & & \vdots \\ a_{k1} & \dots & a_{kk} \end{vmatrix}.$$

We now can state

Theorem 2.6.4 *Let $n \in \mathbb{N}$, and let A be a symmetric $n \times n$ matrix. A is*

(i) positive definite $\Leftrightarrow$ $|M_k| > 0$ *for all principal submatrices M_k of order k, for all $k = 1, \dots, n$,*
(ii) negative definite $\Leftrightarrow$ $(-1)^k |M_k| > 0$ *for all principal submatrices M_k of order k, for all $k = 1, \dots, n$,*
(iii) positive semidefinite $\Leftrightarrow$ $|M_k| \geq 0$ *for all principal submatrices M_k of order k, for all $k = 1, \dots, n$,*
(iv) negative semidefinite $\Leftrightarrow$ $(-1)^k |M_k| \geq 0$ *for all principal submatrices M_k of order k, for all $k = 1, \dots, n$.*

Proof (for $n = 2$). Theorem 2.6.4 is true for *any* $n \in \mathbb{N}$, but we will only prove the case $n = 2$ here. See more specialized literature for the general proof.

(i) "$\Rightarrow$": By way of contradiction, suppose A is positive definite, but there exists $k \in \{1, 2\}$ and a principal submatrix M_k of order k such that $|M_k| \leq 0$. For $n = 2$, there are three possible cases.

I: $a_{11} \leq 0$;
II: $a_{22} \leq 0$;
III: $|A| \leq 0$.

In case I, let $x' = (1, 0)$. Then $x'Ax = a_{11} \leq 0$, which contradicts the positive definiteness of A.

In case II, let $x' = (0, 1)$. Then $x'Ax = a_{22} \leq 0$, which again contradicts the positive definiteness of A.

Finally, consider case III. If $a_{22} \leq 0$, we can use the same reasoning as in case II to obtain a contradiction. If $a_{22} > 0$, let

$$x_1 = \sqrt{a_{22}}, \quad x_2 = -\frac{a_{12}}{\sqrt{a_{22}}}.$$

Then $x'Ax = a_{11}a_{22} - a_{12}^2 = |A| \leq 0$, contradicting the positive definiteness of A.

"$\Leftarrow$": Suppose $a_{11} > 0$, $a_{22} > 0$, and $|A| = a_{11}a_{22} - a_{12}^2 > 0$. For any $x \in \mathbb{R}^n \setminus \{\mathbf{0}\}$, we obtain

$$\begin{aligned}
x'Ax &= a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2 \\
&= a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2 + \frac{a_{12}^2}{a_{11}}x_2^2 - \frac{a_{12}^2}{a_{11}}x_2^2 \\
&= a_{11}\left(x_1^2 + 2\frac{a_{12}}{a_{11}}x_1x_2 + \frac{a_{12}^2}{a_{11}^2}x_2^2\right) + \left(a_{22} - \frac{a_{12}^2}{a_{11}}\right)x_2^2 \\
&= a_{11}\left(x_1 + \frac{a_{12}}{a_{11}}x_2\right)^2 + \frac{a_{11}a_{22} - a_{12}^2}{a_{11}}x_2^2 \\
&= a_{11}\left(x_1 + \frac{a_{12}}{a_{11}}x_2\right)^2 + \frac{|A|}{a_{11}}x_2^2 > 0,
\end{aligned}$$

and therefore, A is positive definite.

The proof of part (ii) is analogous and left as an exercise.

(iii) "$\Rightarrow$": Suppose A is positive semidefinite, but there exists $k \in \{1,2\}$ and a principal submatrix M_k of order k such that $|M_k| < 0$. Again, there are three possible cases.

I: $a_{11} < 0$;
II: $a_{22} < 0$;
III: $|A| < 0$.

In case I, let $x' = (1,0)$. Then $x'Ax = a_{11} < 0$, which contradicts the positive semidefiniteness of A.

In case II, let $x' = (0,1)$. Then $x'Ax = a_{22} < 0$, which again contradicts the positive semidefiniteness of A.

Now consider case III. If $a_{22} < 0$, we can use the same reasoning as in case II to obtain a contradiction. If $a_{22} = 0$, let $x' = (a_{12}, -(a_{11}+1)/2)$. Then $x'Ax = -a_{12}^2 = |A| < 0$, contradicting the positive semidefiniteness of A. If $a_{22} > 0$, let

$$x_1 = \sqrt{a_{22}}, \quad x_2 = -\frac{a_{12}}{\sqrt{a_{22}}}.$$

Again, we obtain $x'Ax = a_{11}a_{22} - a_{12}^2 = |A| < 0$, contradicting the positive semidefiniteness of A.

"$\Leftarrow$": Suppose $a_{11} \geq 0$, $a_{22} \geq 0$, and $|A| = a_{11}a_{22} - a_{12}^2 \geq 0$. Let $x \in \mathbb{R}^n$. If $a_{11} = 0$, $|A| \geq 0$ implies $a_{12} = 0$, and we obtain

$$x'Ax = a_{22}x_2^2 \geq 0.$$

If $a_{11} > 0$, we obtain (analogously to the proof of part (i))

$$x'Ax = a_{11}\left(x_1 + \frac{a_{12}}{a_{11}}x_2\right)^2 + \frac{|A|}{a_{11}}x_2^2 \geq 0,$$

which proves that A is positive semidefinite.

The proof of (iv) is analogous and left as an exercise. ‖

For positive and negative definiteness, it is sufficient to check the *leading* principal minors of A.

Theorem 2.6.5 *Let $n \in \mathbb{N}$, and let A be a symmetric $n \times n$ matrix. A is*

(i) positive definite $\Leftrightarrow |A_k| > 0 \;\; \forall k = 1, \ldots, n$,
(ii) negative definite $\Leftrightarrow (-1)^k|A_k| > 0 \;\; \forall k = 1, \ldots, n$.

Proof. Again, we give a proof for $n = 2$.

(i) By Theorem 2.6.4, all that needs to be shown is that the n *leading* principal minors of A are positive if and only if *all* principal minors of order $k = 1, \ldots, n$ of A are positive.

Clearly, if all principal minors of order k are positive, then, in particular, the *leading* principal minors of order k are positive for all $k = 1, \ldots, n$.

Conversely, suppose the leading principal minors of A are positive. Therefore, in the case $n = 2$, we have $a_{11} > 0$ and $|A| > 0$. Because A is symmetric, $|A| > 0$ implies $a_{11}a_{22} - a_{12}^2 > 0$. Therefore,

$$a_{11}a_{22} > a_{12}^2. \tag{2.16}$$

Because $a_{11} > 0$, dividing (2.16) by a_{11} yields $a_{22} > a_{12}^2/a_{11} \geq 0$, which completes the proof of (i).

Part (ii) is proven analogously. ‖

Therefore, a symmetric $n \times n$ matrix A is positive definite if and only if the leading principal minors of A are positive, and A is negative definite if and only if the signs of the leading principal minors alternate, starting with $|A_1| = a_{11} < 0$.

For positive and negative semidefiniteness, a result analogous to Theorem 2.6.5 can *not* be obtained. Checking the leading principal minors is, in general, *not* sufficient to determine whether a matrix is positive or negative semidefinite. For example, consider the matrix

$$A = \begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix}.$$

We have $|A_1| = 0$ and $|A_2| = |A| = 0$. Therefore, $|A_k| \geq 0$ for all $k = 1, 2$, but A is *not* positive semidefinite, because, for example, choosing $x' = (0, 1)$ leads to $x'Ax = -1 < 0$.

As an example for the application of Theorems 2.6.4 and 2.6.5, consider again

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}, \; B = \begin{pmatrix} -2 & 1 \\ 1 & -1 \end{pmatrix}, \; C = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}.$$

We obtain $|A_1| = 2 > 0$ and $|A_2| = |A| = 1 > 0$, and therefore, according to Theorem 2.6.5, A is positive definite. Furthermore, $|B_1| = -2 < 0$ and $|B_2| = |B| = 1 > 0$, which means that B is negative definite. Finally, $|C_1| = 1 > 0$ and $|C_2| = |C| = -3 < 0$. According to Theorem 2.6.4, C is neither positive semidefinite nor negative semidefinite, and therefore, C is indefinite.

Chapter 3

Functions of One Variable

3.1 Continuity

In this chapter, we discuss functions the domain and range of which are subsets of $\mathbb{R}$. Hence, a *real-valued function of one variable* is a function $f : A \mapsto B$ where $A \subseteq \mathbb{R}$ and $B \subseteq \mathbb{R}$. To simplify matters, we assume that A is an *interval* of the form $[a, b]$ with $a, b \in \mathbb{R}$ or (a, b) with $a \in \mathbb{R} \cup \{-\infty\}$, $b \in \mathbb{R} \cup \{\infty\}$ or $[a, b)$ with $a \in \mathbb{R}$, $b \in \mathbb{R} \cup \{\infty\}$ or $(a, b]$ with $a \in \mathbb{R} \cup \{-\infty\}$, $b \in \mathbb{R}$, where $a < b$. It should be noted that some, but not all, of the results stated in this chapter can be generalized to domains that are not necessarily intervals. The range of f will, in most cases, simply be the set $\mathbb{R}$ itself.

In many economic models (and in other applications), real-valued functions are assumed to be *continuous*. Loosely speaking, continuity ensures that "small" changes in the argument of a function do not lead to "large" changes in the value of the function. The formal definition of continuity is

Definition 3.1.1 *Let $A \subseteq \mathbb{R}$, and let $f : A \mapsto \mathbb{R}$ be a function. Furthermore, let $x_0 \in A$.*

(i) The function f is continuous at x_0 *if and only if*

$$\forall \delta \in \mathbb{R}_{++}, \exists \varepsilon \in \mathbb{R}_{++} \text{ such that } f(x) \in \mathcal{U}_\delta(f(x_0)) \;\forall x \in \mathcal{U}_\varepsilon(x_0) \cap A.$$

(ii) The function f is continuous on $A_0 \subseteq A$ *if and only if f is continuous at each $x_0 \in A_0$. If f is continuous on A, we will often simply say that f is continuous.*

According to this definition, a function f is continuous at a point x_0 in its domain if, for each neighborhood of $f(x_0)$, there exists a neighborhood of x_0 such that $f(x)$ is in this neighborhood of $f(x_0)$ for all x in the domain of f that are in the neighborhood of x_0.

Consider the following example. Let

$$f : \mathbb{R} \mapsto \mathbb{R},\; x \mapsto 2x,$$

and let $x_0 = 1$. Then $f(x_0) = f(1) = 2$. To show that the function f is continuous at $x_0 = 1$, we have to show that, for *any* $\delta \in \mathbb{R}_{++}$, we can find an $\varepsilon \in \mathbb{R}_{++}$ such that

$$f(x) \in (2 - \delta, 2 + \delta) \quad \forall x \in (1 - \varepsilon, 1 + \varepsilon).$$

For $\delta \in \mathbb{R}_{++}$, let $\varepsilon := \delta/2$. Then, for all $x \in \mathcal{U}_\varepsilon(1) = (1 - \delta/2, 1 + \delta/2)$,

$$f(x) > 2(1 - \varepsilon) = 2(1 - \delta/2) = 2 - \delta \text{ and } f(x) < 2(1 + \varepsilon) = 2(1 + \delta/2) = 2 + \delta.$$

Therefore, $f(x) \in \mathcal{U}_\delta(2)$ for all $x \in \mathcal{U}_\varepsilon(1)$. Similarly, it can be shown that f is continuous at any point $x_0 \in \mathbb{R}$ (Exercise: provide a proof), and therefore, f is continuous on its domain $\mathbb{R}$.

As another example, consider the function

$$f : \mathbb{R} \mapsto \mathbb{R},\; x \mapsto \begin{cases} 0 & \text{if } x \leq 0 \\ 1 & \text{if } x > 0. \end{cases}$$

This function is not continuous at $x_0 = 0$. To show this, let $\delta = 1/2$. Continuity of f at $x_0 = 0$ requires that there exists $\varepsilon \in \mathbb{R}_{++}$ such that $f(x) \in (-1/2, 1/2)$ for all $x \in \mathcal{U}_\varepsilon(0) = (-\varepsilon, \varepsilon)$. For *any* $\varepsilon \in \mathbb{R}_{++}$,

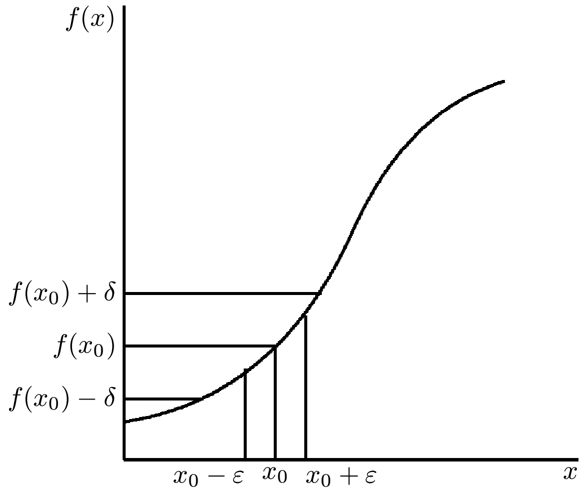


Figure 3.1: A continuous function.

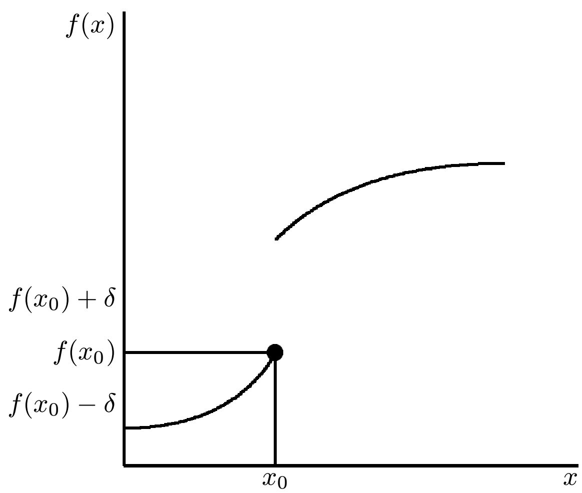


Figure 3.2: A discontinuous function.

$\mathcal{U}_\varepsilon(0)$ contains points $x > 0$, and therefore, any neighborhood of $x_0 = 0$ contains points x such that $f(x) = 1 \notin \mathcal{U}_\delta(f(x_0)) = (-1/2, 1/2)$. Therefore, f is not continuous at $x_0 = 0$.

Using the graph of a function, continuity can be illustrated in a diagram. Figure 3.1 gives an example of the graph of a function f that is continuous at x_0, whereas the function with the graph illustrated in Figure 3.2 is *not* continuous at the point x_0.

An alternative definition of continuity can be given in terms of *sequences* of real numbers.

Theorem 3.1.2 *Let $A \subseteq \mathbb{R}$ be an interval, and let $f : A \mapsto \mathbb{R}$ be a function. Furthermore, let $x_0 \in A$. f is continuous at x_0 if and only if, for all sequences $\{x_n\}$ such that $x_n \in A$ for all $n \in \mathbb{N}$,*

$$\lim_{n\to\infty} x_n = x_0 \Rightarrow \lim_{n\to\infty} f(x_n) = f(x_0).$$

Proof. "Only if": Suppose f is continuous at $x_0 \in A$. Then, for any $\delta \in \mathbb{R}_{++}$, there exists $\varepsilon \in \mathbb{R}_{++}$ such that

$$f(x) \in \mathcal{U}_\delta(f(x_0)) \; \forall x \in \mathcal{U}_\varepsilon(x_0) \cap A.$$

Let $\{x_n\}$ be a sequence such that $x_n \in A$ for all $n \in \mathbb{N}$ and $\{x_n\}$ converges to x_0. Because $\lim_{n\to\infty} x_n = x_0$, we can find $n_0 \in \mathbb{N}$ such that

$$|x_n - x_0| < \varepsilon \quad \forall n \geq n_0.$$

Therefore, $|f(x_n) - f(x_0)| < \delta$ for all $n \geq n_0$, which implies that $\{f(x_n)\}$ converges to $f(x_0)$.

"If": Suppose f is not continuous at $x_0 \in A$. Then there exists $\delta \in \mathbb{R}_{++}$ such that, for all $\varepsilon \in \mathbb{R}_{++}$, there exists $x \in \mathcal{U}_\varepsilon(x_0) \cap A$ with

$$f(x) \notin \mathcal{U}_\delta(f(x_0)).$$

Let $\varepsilon_n := 2/n$ for all $n \in \mathbb{N}$. By assumption, for any $n \in \mathbb{N}$, there exists $x_n \in \mathcal{U}_{\varepsilon_n}(x_0) \cap A$ such that

$$|f(x_n) - f(x_0)| \geq \delta. \tag{3.1}$$

Because $x_n \in \mathcal{U}_{\varepsilon_n}(x_0) \cap A$ for all $n \in \mathbb{N}$, the sequence $\{x_n\}$ converges to x_0. By (3.1), $\{f(x_n)\}$ does not converge to $f(x_0)$, which completes the proof. $\|$

Theorem 3.1.2 is particularly useful in proving that a function is not continuous at a point x_0 in its domain. One only has to find *one* sequence $\{x_n\}$ that converges to x_0 such that the sequence $\{f(x_n)\}$ does not converge to $f(x_0)$.

We now introduce *limits of a function* which are closely related to limits of sequences.

Definition 3.1.3 *Let $A \subseteq \mathbb{R}$, and let $f : A \mapsto \mathbb{R}$ be a function. Furthermore, let $\alpha \in \mathbb{R} \cup \{-\infty, \infty\}$.*

(i) Let $x_0 \in \mathbb{R} \cup \{-\infty\}$, and suppose $(x_0, x_0 + h) \subseteq A$ for some $h \in \mathbb{R}_{++}$. The right-side limit of f at x_0 exists and is equal to α *if and only if, for all monotone nonincreasing sequences $\{x_n\}$ such that $x_n \in A$ for all $n \in \mathbb{N}$,*

$$x_n \longrightarrow x_0 \Rightarrow f(x_n) \longrightarrow \alpha.$$

If the right-side limit of f at x_0 exists and is equal to α, we write $\lim_{x\downarrow x_0} f(x) = \alpha$.

(ii) Let $x_0 \in \mathbb{R} \cup \{\infty\}$, and suppose $(x_0 - h, x_0) \subseteq A$ for some $h \in \mathbb{R}_{++}$. The left-side limit of f at x_0 exists and is equal to α *if and only if, for all monotone nondecreasing sequences $\{x_n\}$ such that $x_n \in A$ for all $n \in \mathbb{N}$,*

$$x_n \longrightarrow x_0 \Rightarrow f(x_n) \longrightarrow \alpha.$$

If the left-side limit of f at x_0 exists and is equal to α, we write $\lim_{x\uparrow x_0} f(x) = \alpha$.

(iii) Let $x_0 \in \mathbb{R}$, and suppose $(x_0 - h, x_0) \cup (x_0, x_0 + h) \subseteq A$ for some $h \in \mathbb{R}_{++}$. The limit of f at x_0 exists and is equal to α *if and only if the right-side and left-side limits of f at x_0 exist and are equal to α. If the limit of f at x_0 exists and is equal to α, we write $\lim_{x\to x_0} f(x) = \alpha$.*

For finite values of x_0 and α, an equivalent definition of a limit can be given directly (that is, without using sequences). We obtain

Theorem 3.1.4 *Let $A \subseteq \mathbb{R}$, and let $f : A \mapsto \mathbb{R}$ be a function. Furthermore, let $x_0 \in \mathbb{R}$, $\alpha \in \mathbb{R}$, and suppose $(x_0 - h, x_0) \cup (x_0, x_0 + h) \subseteq A$ for some $h \in \mathbb{R}_{++}$. The limit of f at x_0 exists and is equal to α if and only if*

$$\forall \delta \in \mathbb{R}_{++}, \exists \varepsilon \in \mathbb{R}_{++} \text{ such that } f(x) \in \mathcal{U}_\delta(\alpha)\ \forall x \in \mathcal{U}_\varepsilon(x_0).$$

Analogous results are valid for one-sided limits.

For *interior* points of the domain of a function, continuity at these points can be formulated in terms of limits. We obtain

Theorem 3.1.5 *Let $A \subseteq \mathbb{R}$ be an interval, and let $f : A \mapsto \mathbb{R}$ be a function. Furthermore, let x_0 be an interior point of A. f is continuous at x_0 if and only if $\lim_{x\to x_0} f(x)$ exists and $\lim_{x\to x_0} f(x) = f(x_0)$.*

Theorem 3.1.5 can be proven by combining Theorem 3.1.2 and Definition 3.1.3.

Analogously, for boundary points of A, we obtain

Theorem 3.1.6 *Let $A = [a, b]$ with $a, b \in \mathbb{R}$ and $a < b$. Furthermore, let $f : A \mapsto \mathbb{R}$ be a function.*

(i) f is continuous at a if and only if $\lim_{x\downarrow a} f(x)$ exists and $\lim_{x\downarrow a} f(x) = f(a)$.
(ii) f is continuous at b if and only if $\lim_{x\uparrow b} f(x)$ exists and $\lim_{x\uparrow b} f(x) = f(b)$.

To illustrate the application of Theorem 3.1.5, consider the following examples. Let

$$f : \mathbb{R} \mapsto \mathbb{R},\ x \mapsto 2x.$$

We use Theorem 3.1.5 to prove that f is continuous at $x_0 = 1$. We obtain

$$\lim_{x \uparrow 1} f(x) = \lim_{x \uparrow 1} 2x = 2$$

and

$$\lim_{x \downarrow 1} f(x) = \lim_{x \downarrow 1} 2x = 2,$$

and therefore, $\lim_{x \to 1} f(x)$ exists and is equal to 2. Furthermore, $f(1) = 2$, and therefore, according to Theorem 3.1.5, f is continuous at $x_0 = 1$.

Now let

$$f : \mathbb{R} \mapsto \mathbb{R},\ x \mapsto \begin{cases} 0 & \text{if } x \leq 0 \\ 1 & \text{if } x > 0. \end{cases}$$

Let $x_0 = 0$. We have

$$\lim_{x \uparrow 0} f(x) = \lim_{x \uparrow 0} 0 = 0$$

and

$$\lim_{x \downarrow 0} f(x) = \lim_{x \downarrow 0} 1 = 1,$$

and therefore, $\lim_{x \to 0} f(x)$ does not exist. By Theorem 3.1.5, f is not continuous at $x_0 = 0$.

Finally, consider the function given by

$$f : \mathbb{R} \mapsto \mathbb{R},\ x \mapsto \begin{cases} 0 & \text{if } x = 0 \\ 1 & \text{if } x \neq 0. \end{cases}$$

For $x_0 = 0$, we obtain

$$\lim_{x \uparrow 0} f(x) = \lim_{x \uparrow 0} 1 = 1$$

and

$$\lim_{x \downarrow 0} f(x) = \lim_{x \downarrow 0} 1 = 1.$$

Therefore, $\lim_{x \to 0} f(x)$ exists and is equal to 1. But we have $f(0) = 0 \neq 1$, and hence, f is not continuous at $x_0 = 0$.

The following definition introduces some notation for functions that can be defined using other functions.

Definition 3.1.7 *Let $A \subseteq \mathbb{R}$ be an interval, and let $f : A \mapsto \mathbb{R}$ and $g : A \mapsto \mathbb{R}$ be functions. Furthermore, let $\alpha \in \mathbb{R}$.*

(i) The sum *of f and g is the function defined by*

$$f + g : A \mapsto \mathbb{R},\ x \mapsto f(x) + g(x).$$

(ii) The function αf is defined by

$$\alpha f : A \mapsto \mathbb{R},\ x \mapsto \alpha f(x).$$

(iii) The product *of f and g is the function defined by*

$$fg : A \mapsto \mathbb{R},\ x \mapsto f(x)g(x).$$

(iv) If $g(x) \neq 0$ for all $x \in A$, the ratio *of f and g is the function defined by*

$$\frac{f}{g} : A \mapsto \mathbb{R},\ x \mapsto \frac{f(x)}{g(x)}.$$

Some useful results concerning the continuity of functions are summarized in

Theorem 3.1.8 *Let $A \subseteq I\!R$ be an interval, and let $f : A \mapsto I\!R$ and $g : A \mapsto I\!R$ be functions. Furthermore, let $\alpha \in I\!R$. If f and g are continuous at $x_0 \in A$, then*

(i) $f + g$ is continuous at x_0.
(ii) αf is continuous at x_0.
(iii) fg is continuous at x_0.
(iv) If $g(x) \neq 0$ for all $x \in A$, f/g is continuous at x_0.

The proof of this theorem follows from the corresponding properties of sequences and Theorem 3.1.2.

Furthermore, if two functions $f : A \mapsto \mathbb{R}$ and $g : f(A) \mapsto \mathbb{R}$ are continuous, then the composite function $g \circ f$ is continuous. We state this result without a proof.

Theorem 3.1.9 *Let $A \subseteq I\!R$ be an interval. Furthermore, let $f : A \mapsto I\!R$ and $g : f(A) \mapsto I\!R$ be functions, and let $x_0 \in A$. If f is continuous at x_0 and g is continuous at $y_0 = f(x_0)$, then $g \circ f$ is continuous at x_0.*

If the domain of a continuous function is an interval, the inverse of this function is continuous.

Theorem 3.1.10 *Let $A \subseteq I\!R$ be an interval, let $B \subseteq I\!R$, and let $f : A \mapsto B$ be bijective. If f is continuous on A, then f^{-1} is continuous on $B = f(A)$.*

The assumption that A is an interval is essential in this theorem. For some bijective functions f with more general domains $A \subseteq \mathbb{R}$, continuity of f does *not* imply continuity of f^{-1}. To illustrate that, consider the following example. Let

$$f : [0,1] \cup (2,3] \mapsto [0,2],\ x \mapsto \begin{cases} x^2 & \text{if } x \in [0,1] \\ x-1 & \text{if } x \in (2,3]. \end{cases}$$

The domain of this function is $A = [0,1] \cup (2,3]$, which is not an interval. The function f is continuous on A and bijective (Exercise: show this). The inverse of f exists and is given by

$$f^{-1} : [0,2] \mapsto [0,1] \cup (2,3],\ y \mapsto \begin{cases} \sqrt{y} & \text{if } y \in [0,1] \\ y+1 & \text{if } y \in (1,2]. \end{cases}$$

Clearly, f^{-1} is not continuous at $y_0 = 1$.

Some important examples for continuous functions are *polynomials*.

Definition 3.1.11 *Let $A \subseteq I\!R$ be an interval, and let $f : A \mapsto I\!R$ be a function. Furthermore, let $n \in I\!N \cup \{0\}$. f is a* polynomial of degree n *if and only if there exist $\alpha_0, \ldots, \alpha_n \in I\!R$ with $\alpha_n \neq 0$ if $n \in I\!N$ such that*

$$f(x) = \sum_{i=0}^{n} \alpha_i x^i \quad \forall x \in A.$$

For example, the function

$$f : \mathbb{R} \mapsto \mathbb{R},\ x \mapsto -2 + x - 2x^3$$

is a polynomial of degree three. Other examples for polynomials are *constant* functions, defined by

$$f : \mathbb{R} \mapsto \mathbb{R},\ x \mapsto \alpha$$

where $\alpha \in \mathbb{R}$ is a constant, *affine* functions

$$f : \mathbb{R} \mapsto \mathbb{R},\ x \mapsto \alpha + \beta x$$

where $\alpha, \beta \in \mathbb{R}$, $\beta \neq 0$, and *quadratic* functions

$$f : \mathbb{R} \mapsto \mathbb{R},\ x \mapsto \alpha + \beta x + \gamma x^2$$

where $\alpha, \beta, \gamma \in \mathbb{R}$, $\gamma \neq 0$. Constant functions are polynomials of degree zero, affine functions are polynomials of degree one, and quadratic functions are polynomials of degree two.

Because the function

$$f : \mathbb{R} \mapsto \mathbb{R},\ x \mapsto x^n$$

is continuous for any $n \in \mathbb{N}$ (Exercise: prove this), Theorem 3.1.8 implies that all polynomials are continuous.

Analogously to monotonic sequences, monotonicity properties of functions can be defined. We conclude this section with the definitions of these properties.

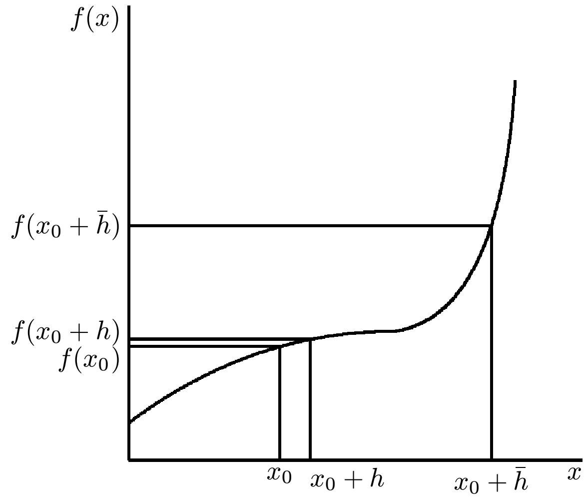


Figure 3.3: The difference quotient.

Definition 3.1.12 *Let $A \subseteq \mathbb{R}$ be an interval, and let $B \subseteq A$. Furthermore, let $f : A \mapsto \mathbb{R}$ be a function.*

(i) f is (monotone) nondecreasing on *B if and only if*

$$x > y \Rightarrow f(x) \geq f(y) \quad \forall x, y \in B.$$

(ii) f is (monotone) increasing on *B if and only if*

$$x > y \Rightarrow f(x) > f(y) \quad \forall x, y \in B.$$

(iii) f is (monotone) nonincreasing on *B if and only if*

$$x > y \Rightarrow f(x) \leq f(y) \quad \forall x, y \in B.$$

(iv) f is (monotone) decreasing on *B if and only if*

$$x > y \Rightarrow f(x) < f(y) \quad \forall x, y \in B.$$

If f is nondecreasing (respectively increasing, nonincreasing, decreasing) on its domain A, we will sometimes simply say that f is nondecreasing (respectively increasing, nonincreasing, decreasing).

3.2 Differentiation

An important issue concerning real-valued functions of one real variable is the question how the value of a function changes as a consequence of a change in its argument. We first give a diagrammatic illustration. Consider the graph of a function $f : \mathbb{R} \mapsto \mathbb{R}$ as illustrated in Figure 3.3.

Suppose we want to find the *rate of change* in the value of f as a consequence of a change in the variable x from x_0 to $x_0 + h$. A natural way to do this is to use the *slope* of the *secant* through the points $(x_0, f(x_0))$ and $(x_0 + h, f(x_0 + h))$. Clearly, this slope is given by the ratio

$$\frac{f(x_0 + h) - f(x_0)}{h}. \tag{3.2}$$

The quotient in (3.2) is called the *difference quotient* of f for x_0 and h. This slope depends on the number $h \in \mathbb{R}$ we add to x_0 (h can be positive or negative). Depending on f, the values of the difference quotient for given x_0 can differ substantially for different values of h—consider, for example, the difference quotients that are obtained with h and with $\bar{h}$ in Figure 3.3.

In order to obtain some information about the behaviour of f "close" to a point x_0, it is desirable to find an indicator that is *independent* of the number h which represents the deviation from x_0. The natural way to proceed is to use the *limit* of the difference quotient as h approaches zero as an indicator of the change in f at x_0. If this limit exists and is finite, we say that the function is *differentiable* at the point x_0. Formally,

Definition 3.2.1 *Let $A \subseteq I\!R$, and let $f : A \mapsto I\!R$ be a function. Furthermore, let x_0 be an interior point of A. The function f is* differentiable at x_0 *if and only if*

$$\lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h} \tag{3.3}$$

exists and is finite. If f is differentiable at x_0, we call the limit in (3.3) the derivative of f at x_0 *and denote it by $f'(x_0)$.*

Definition 3.2.1 applies to interior points of A only. We can also define "one-sided" derivatives at the endpoints of the interval A, if these points are elements of A.

Definition 3.2.2 *Let $f : [a, b] \mapsto I\!R$ with $a, b \in I\!R$ and $a < b$.*

(i) f is (right-side) differentiable at a *if and only if*

$$\lim_{h \downarrow 0} \frac{f(a + h) - f(a)}{h} \tag{3.4}$$

exists and is finite. If f is (right-side) differentiable at a, we call the limit in (3.4) the (right-side) derivative of f at a *and denote it by $f'(a)$.*

(ii) f is (left-side) differentiable at b *if and only if*

$$\lim_{h \uparrow 0} \frac{f(b + h) - f(b)}{h} \tag{3.5}$$

exists and is finite. If f is (left-side) differentiable at b, we call the limit in (3.5) the (left-side) derivative of f at b *and denote it by $f'(b)$.*

A function $f : A \mapsto I\!R$ is *differentiable on an interval* $B \subseteq A$ if $f'(x_0)$ exists for all interior points x_0 of B, and the right-side and left-side derivatives of f at the endpoints of B exist whenever these endpoints belong to B. If $f : A \mapsto I\!R$ is differentiable on A, the function

$$f' : A \mapsto I\!R,\ x \mapsto f'(x)$$

is called the *derivative* of f. If the function f' is differentiable at $x_0 \in A$, we can find the *second derivative* of f at x_0, which is just the derivative of f' at x_0. Formally, if $f' : A \mapsto I\!R$ is differentiable at $x_0 \in A$, we call

$$f''(x_0) = \lim_{h \to 0} \frac{f'(x_0 + h) - f'(x_0)}{h}$$

the *second derivative* of f at x_0. Analogously, we can define higher-order derivatives of f, and we write $f^{(n)}(x_0)$ for the n^{th} derivative of f at x_0 (where $n \in I\!N$), if this derivative exists.

For a function f which is differentiable at a point x_0 in its domain, $f'(x_0)$ is the *slope* of the tangent to the graph of f at x_0. The equation of this tangent for a differentiable function $f : A \mapsto I\!R$ at $x_0 \in A$ is

$$y = f(x_0) + f'(x_0)(x - x_0) \quad \forall x \in A.$$

As an example, consider the following function

$$f : I\!R \mapsto I\!R,\ x \mapsto 2 + x + 4x^2.$$

To determine whether or not this function is differentiable at a point x_0 in its domain, we have to find out whether the limit of the difference quotient as h approaches zero exists and is finite. For the above function f and $x_0 \in \mathbb{R}$, we obtain

$$\begin{aligned} \frac{f(x_0+h)-f(x_0)}{h} &= \frac{2+x_0+h+4(x_0+h)^2-2-x_0-4(x_0)^2}{h} \\ &= \frac{h+4((x_0)^2+2x_0h+h^2)-4(x_0)^2}{h} \\ &= \frac{h+8x_0h+4h^2}{h} = 1+8x_0+4h. \end{aligned}$$

Clearly,

$$\lim_{h\to 0}(1+8x_0+4h) = 1+8x_0,$$

and therefore, f is differentiable at any point $x_0 \in \mathbb{R}$ with $f'(x_0) = 1+8x_0$.

Differentiability is a *stronger* requirement than continuity. This is shown in the following theorems.

Theorem 3.2.3 *Let $A \subseteq \mathbb{R}$ be an interval, and let $f : A \mapsto \mathbb{R}$ be a function. Furthermore, let x_0 be an interior point of A. If f is differentiable at x_0, then f is continuous at x_0.*

Proof. For $h \in \mathbb{R}$ such that $(x_0+h) \in A$ and $h \neq 0$, we obtain

$$f(x_0+h)-f(x_0) = \left(\frac{f(x_0+h)-f(x_0)}{h}\right)h.$$

Because f is differentiable at x_0,

$$\lim_{h\to 0}(f(x_0+h)-f(x_0)) = f'(x_0)\lim_{h\to 0} h = 0.$$

This implies

$$\lim_{h\to 0} f(x_0+h) = f(x_0).$$

Defining $x := x_0+h$, this can be written as

$$\lim_{x\to x_0} f(x) = f(x_0),$$

which proves that f is continuous at x_0. ‖

Again, we can extend the result of Theorem 3.2.3 to boundary points of A.

Theorem 3.2.4 *Let $f : [a,b] \mapsto \mathbb{R}$ with $a, b \in \mathbb{R}$ and $a < b$.*

(i) If f is right-side differentiable at a, then f is continuous at a.
(ii) If f is left-side differentiable at b, then f is continuous at b.

The proof of this theorem is analogous to the proof of Theorem 3.2.3 and is left as an exercise.

Continuity does *not* imply differentiability. Consider, for example, the function

$$f : \mathbb{R} \mapsto \mathbb{R},\ x \mapsto |x|.$$

This function is continuous at $x_0 = 0$, but it is not differentiable at $x_0 = 0$. To show this, note that

$$\lim_{x\downarrow 0} f(x) = \lim_{x\downarrow 0} x = 0,\ \lim_{x\uparrow 0} f(x) = \lim_{x\uparrow 0} -x = 0,$$

and $f(0) = 0$. Therefore, $\lim_{x\to 0} f(x)$ exists and is equal to $f(0) = 0$, which shows that f is continuous at $x_0 = 0$.

Now note that

$$\lim_{h\downarrow 0}\frac{f(0+h)-f(0)}{h} = \frac{h}{h} = 1$$

and

$$\lim_{h\uparrow 0}\frac{f(0+h)-f(0)}{h} = \frac{-h}{h} = -1.$$

Clearly,

$$\lim_{h\downarrow 0} \frac{f(0+h)-f(0)}{h} \neq \lim_{h\uparrow 0} \frac{f(0+h)-f(0)}{h},$$

and therefore,

$$\lim_{h\to 0} \frac{f(0+h)-f(0)}{h}$$

does not exist, which implies that f is not differentiable at $x_0 = 0$.

The following theorem introduces some rules of differentiation which simplify the task of finding the derivatives of certain differentiable functions. It states that sums, multiples, products, and ratios of differentiable functions are differentiable, and shows how these derivatives are obtained.

Theorem 3.2.5 *Let $A \subseteq \mathbb{R}$ be an interval, and let $f : A \mapsto \mathbb{R}$, $g : A \mapsto \mathbb{R}$ be functions. Furthermore, let $\alpha \in \mathbb{R}$. Let x_0 be an interior point of A, and suppose f and g are differentiable at x_0.*

(i) $f + g$ is differentiable at x_0, and $(f+g)'(x_0) = f'(x_0) + g'(x_0)$.
(ii) αf is differentiable at x_0, and $(\alpha f)'(x_0) = \alpha f'(x_0)$.
(iii) fg is differentiable at x_0, and $(fg)'(x_0) = f'(x_0)g(x_0) + f(x_0)g'(x_0)$.
(iv) If $g(x_0) \neq 0$, then f/g is differentiable at x_0, and

$$\left(\frac{f}{g}\right)'(x_0) = \frac{f'(x_0)g(x_0) - f(x_0)g'(x_0)}{[g(x_0)]^2}.$$

Proof. (i) Because f and g are differentiable at x_0, $f'(x_0)$ and $g'(x_0)$ exist. Now it follows that

$$\begin{aligned}\lim_{h\to 0} \frac{f(x_0+h)+g(x_0+h)-f(x_0)-g(x_0)}{h} &= \lim_{h\to 0} \frac{f(x_0+h)-f(x_0)}{h} + \lim_{h\to 0} \frac{g(x_0+h)-g(x_0)}{h} \\ &= f'(x_0)+g'(x_0).\end{aligned}$$

$$\text{(ii)} \quad \lim_{h\to 0} \frac{\alpha f(x_0+h)-\alpha f(x_0)}{h} = \alpha \lim_{h\to 0} \frac{f(x_0+h)-f(x_0)}{h} = \alpha f'(x_0).$$

$$\begin{aligned}\text{(iii)} \quad \lim_{h\to 0} \frac{f(x_0+h)g(x_0+h)-f(x_0)g(x_0)}{h} &= \lim_{h\to 0} \frac{f(x_0+h)g(x_0+h)-f(x_0)g(x_0+h)}{h} \\ &+ \lim_{h\to 0} \frac{f(x_0)g(x_0+h)-f(x_0)g(x_0)}{h} \\ &= \lim_{h\to 0} \frac{f(x_0+h)-f(x_0)}{h} g(x_0+h) \\ &+ \lim_{h\to 0} f(x_0)\frac{g(x_0+h)-g(x_0)}{h} \\ &= f'(x_0)g(x_0)+f(x_0)g'(x_0).\end{aligned}$$

(iv) Because g is differentiable at x_0, g is continuous at x_0. Therefore, $g(x_0) \neq 0$ implies that there exists a neighborhood of x_0 such that $g(x) \neq 0$ for all $x \in A$ in this neighborhood. Let $h \in \mathbb{R}$ be such that $(x_0 + h)$ is in this neighborhood. Then we obtain

$$\lim_{h\to 0} \frac{f(x_0+h)/g(x_0+h)-f(x_0)/g(x_0)}{h} = \lim_{h\to 0} \frac{f(x_0+h)g(x_0)-f(x_0)g(x_0+h)}{hg(x_0+h)g(x_0)} =$$

$$\lim_{h\to 0} \frac{g(x_0)[f(x_0+h)-f(x_0)]-f(x_0)[g(x_0+h)-g(x_0)]}{hg(x_0+h)g(x_0)} = \frac{f'(x_0)g(x_0)-f(x_0)g'(x_0)}{[g(x_0)]^2}. \quad \|$$

Replacing the limits in the above proof by the corresponding one-sided limits, we obtain analogous results for boundary points of A. The formulation of these results is left as an exercise.

We now find the derivatives of some important functions. First, consider *affine* functions of the form

$$f : \mathbb{R} \mapsto \mathbb{R},\ x \mapsto \alpha + \beta x$$

with $\alpha, \beta \in \mathbb{R}$ and $\beta \neq 0$. Let $x \in \mathbb{R}$. Then

$$\lim_{h \to 0} \frac{f(x+h) - f(x)}{h} = \lim_{h \to 0} \frac{\alpha + \beta(x+h) - \alpha - \beta x}{h} = \lim_{h \to 0} \frac{\beta h}{h} = \beta.$$

Therefore, $f'(x) = \beta$ for all $x \in \mathbb{R}$.

Now consider the *linear* function

$$f : \mathbb{R} \mapsto \mathbb{R},\ x \mapsto \beta x$$

with $\beta \in \mathbb{R}$, $\beta \neq 0$. For any $x \in \mathbb{R}$, we obtain

$$\lim_{h \to 0} \frac{f(x+h) - f(x)}{h} = \lim_{h \to 0} \frac{\beta(x+h) - \beta x}{h} = \lim_{h \to 0} \beta = \beta,$$

and therefore, $f'(x) = \beta$ for all $x \in \mathbb{R}$.

If f is a constant function

$$f : \mathbb{R} \mapsto \mathbb{R},\ x \mapsto \alpha$$

with $\alpha \in \mathbb{R}$, we can write f as

$$f : \mathbb{R} \mapsto \mathbb{R},\ x \mapsto \alpha + \beta x - \beta x$$

with $\beta \in \mathbb{R}$, $\beta \neq 0$. Parts (i) and (ii) of Theorem 3.2.5 imply, together with the above results for affine and linear functions,

$$f'(x) = 0 \quad \forall x \in \mathbb{R}. \tag{3.6}$$

Next, let

$$f : \mathbb{R} \mapsto \mathbb{R},\ x \mapsto x^n$$

where $n \in \mathbb{N}$. For $x \in \mathbb{R}$, we obtain the difference quotient

$$\frac{f(x+h) - f(x)}{h} = \frac{(x+h)^n - x^n}{h}.$$

The *binomial formula* (which we will not prove here) says that, for $x, y \in \mathbb{R}$ and $n \in \mathbb{N}$,

$$(x+y)^n = \sum_{k=0}^{n} \frac{n!}{k!(n-k)!} x^{n-k} y^k.$$

Therefore, we obtain

$$\begin{aligned}
\lim_{h \to 0} \frac{f(x+h) - f(x)}{h} &= \lim_{h \to 0} \frac{\sum_{k=0}^{n} \frac{n!}{k!(n-k)!} x^{n-k} h^k - x^n}{h} \\
&= \lim_{h \to 0} \frac{\sum_{k=1}^{n} \frac{n!}{k!(n-k)!} x^{n-k} h^k}{h} \\
&= \lim_{h \to 0} \frac{n x^{n-1} h + \sum_{k=2}^{n} \frac{n!}{k!(n-k)!} x^{n-k} h^k}{h} \\
&= n x^{n-1} + \lim_{h \to 0} \sum_{k=2}^{n} \frac{n!}{k!(n-k)!} x^{n-k} h^{k-1} \\
&= n x^{n-1}.
\end{aligned}$$

Therefore,

$$f'(x) = n x^{n-1} \quad \forall x \in \mathbb{R}. \tag{3.7}$$

If we have a polynomial

$$f : \mathbb{R} \mapsto \mathbb{R},\ x \mapsto \sum_{i=0}^{n} \alpha_i x^i$$

with $n \in \mathbb{N}$, Theorem 3.2.5, (3.6), and (3.7) can be applied to conclude

$$f'(x) = \sum_{i=1}^{n} i \alpha_i x^{i-1} \quad \forall x \in \mathbb{R}.$$

A very useful rule for differentiation that applies to composite functions is the *chain rule*, which is described in the following theorem.

Theorem 3.2.6 *Let $A \subseteq I\!R$ be an interval, and let $f : A \mapsto I\!R$, $g : f(A) \mapsto I\!R$ be functions. Let x_0 be an interior point of A. If f is differentiable at x_0 and g is differentiable at $y_0 = f(x_0)$, then $g \circ f$ is differentiable at x_0, and*

$$(g \circ f)'(x_0) = g'(f(x_0))f'(x_0).$$

Proof. Let $y_0 = f(x_0)$. For $r \in \mathbb{R}$ such that $(y_0 + r) \in f(A)$, define

$$k(y_0 + r) := \begin{cases} \frac{g(y_0+r)-g(y_0)}{r} & \text{if } r \neq 0 \\ g'(y_0) & \text{if } r = 0. \end{cases}$$

Because g is differentiable at y_0,

$$\lim_{r \to 0} k(y_0 + r) = g'(y_0).$$

Furthermore, by definition,

$$g(y_0 + r) - g(y_0) = rk(y_0 + r).$$

Therefore, for $h \in \mathbb{R} \setminus \{0\}$ such that $(x_0 + h) \in A$,

$$g(f(x_0 + h)) - g(f(x_0)) = [f(x_0 + h) - f(x_0)]k(f(x_0 + h)),$$

and hence,

$$\frac{g(f(x_0 + h)) - g(f(x_0))}{h} = \frac{f(x_0 + h) - f(x_0)}{h} k(f(x_0 + h)).$$

Taking limits, we obtain

$$\begin{aligned} (g \circ f)'(x_0) &= \lim_{h \to 0} \frac{g(f(x_0 + h)) - g(f(x_0))}{h} = \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h} k(f(x_0 + h)) \\ &= g'(f(x_0))f'(x_0). \quad \| \end{aligned}$$

Again, it is straightforward to obtain analogous theorems for boundary points of A by considering one-sided limits.

As an example for the application of the chain rule, let

$$f : \mathbb{R} \mapsto \mathbb{R},\ x \mapsto 2x - 3x^2$$

and

$$g : \mathbb{R} \mapsto \mathbb{R},\ y \mapsto y^3.$$

Then

$$g \circ f : \mathbb{R} \mapsto \mathbb{R},\ x \mapsto (2x - 3x^2)^3.$$

We obtain

$$f'(x) = 2 - 6x \quad \forall x \in \mathbb{R}$$

and

$$g'(y) = 3y^2 \quad \forall y \in \mathbb{R}.$$

According to the chain rule,

$$(g \circ f)'(x) = g'(f(x))f'(x) = 3(2x - 3x^2)^2(2 - 6x) \quad \forall x \in \mathbb{R}.$$

The following theorem (which we state without a proof) provides a relationship between the derivative of a bijective function and its inverse.

Theorem 3.2.7 *Let $A, B \subseteq I\!R$, where $A = (a, b)$ with $a, b \in I\!R$, $a < b$. Furthermore, let $f : A \mapsto B$ be bijective and continuous. If f is differentiable at $x_0 \in A = (a, b)$ and $f'(x_0) \neq 0$, then f^{-1} is differentiable at $y_0 = f(x_0)$, and*

$$(f^{-1})'(y_0) = \frac{1}{f'(x_0)}.$$

To conclude this section, we introduce some important functions and their properties. First, we define the *exponential function*.

Definition 3.2.8 *The function*

$$E : \mathbb{R} \mapsto \mathbb{R}_{++},\ x \mapsto \sum_{n=0}^{\infty} \frac{x^n}{n!}$$

is called the exponential function.

The definition of E involves an *infinite sum*, but it can be shown that this sum converges for all $x \in \mathbb{R}$, and therefore, E is a well-defined function. Another (equivalent) definition of E is given by the following theorem, which is stated without a proof.

Theorem 3.2.9 *For all $x \in \mathbb{R}$,*

$$E(x) = \lim_{n\to\infty} \left(1 + \frac{x}{n}\right)^n.$$

The exponential function is used, for example, to describe certain *growth processes*. Before discussing an application, we introduce some important properties of E and related functions. We only prove part (i) of the following theorem.

Theorem 3.2.10 *(i) The function E is differentiable on $\mathbb{R}$, and $E'(x) = E(x) \quad \forall x \in \mathbb{R}$.*
(ii) E is increasing, $\lim_{x\uparrow\infty} E(x) = \infty$, $\lim_{x\downarrow-\infty} E(x) = 0$, *and* $E(\mathbb{R}) = \mathbb{R}_{++}$.
(iii) $E(0) = 1$.
(iv) $E(1) = e = 2.71828\ldots$.
(v) $E(x+y) = E(x)E(y) \quad \forall x, y \in \mathbb{R}$.

Proof of (i). For $x \in \mathbb{R}$, we obtain (using the rules of differentiation for polynomials)

$$E'(x) = \sum_{n=1}^{\infty} \frac{nx^{n-1}}{n!} = \sum_{n=1}^{\infty} \frac{x^{n-1}}{(n-1)!} = \sum_{m=0}^{\infty} \frac{x^m}{m!} = E(x). \ \|$$

Part (i) of this theorem is quite remarkable. It states that for *any* $x \in \mathbb{R}$, the derivative of E at x is *equal* to the value of E at x.

Parts (iv) and (v) of Theorem 3.2.10 imply

$$E(2) = E(1+1) = E(1)E(1) = e^2,$$

and, in general, for $n \in \mathbb{N}$,

$$E(n) = e^n.$$

Furthermore, we obtain

$$E(-n) = \frac{1}{E(n)} = e^{-n} \quad \forall n \in \mathbb{N}.$$

Now let $y \in \mathbb{R}$. For a rational number $x = p/q$ with $p \in \mathcal{Z}$ and $q \in \mathbb{N}$, we obtain

$$(E(xy))^q = E(qxy) = E(py) = (E(y))^p,$$

and therefore,

$$E(xy) = \sqrt[q]{(E(y))^p} = (E(y))^{p/q} = (E(y))^x.$$

For $y = 1$, this implies

$$E(x) = (E(1))^x = e^x \quad \forall x \in \mathcal{Q}.$$

More generally, we can define

$$e^x := E(x) \quad \forall x \in \mathbb{R},$$

that is, e^x is defined by $E(x)$, even if x is an *irrational* number.

The graph of the exponential function can be illustrated as in Figure 3.4.

Part (ii) of Theorem 3.2.10 implies that E is bijective, and therefore, E has an inverse. We call the inverse of the exponential function the *natural logarithm*.

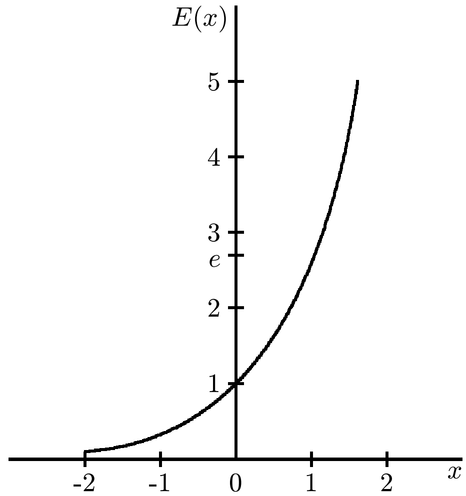


Figure 3.4: The exponential function.

Definition 3.2.11 *The* natural logarithm $\ln : I\!R_{++} \mapsto I\!R$ *is defined as the inverse function of the exponential function, that is,* $\ln = E^{-1}$.

By definition of an inverse, we have

$$E(\ln(x)) = e^{\ln(x)} = x \quad \forall x \in \mathbb{R}_{++}$$

and

$$\ln(E(x)) = \ln(e^x) = x \quad \forall x \in \mathbb{R}.$$

From the properties of E, it follows that ln must have certain properties. Some of them are summarized in the following theorem.

Theorem 3.2.12 *(i) The function* ln *is differentiable on* $I\!R_{++}$, *and* $\ln'(x) = 1/x \quad \forall x \in I\!R_{++}$.
(ii) ln *is increasing,* $\lim_{x\uparrow\infty} \ln(x) = \infty$, $\lim_{x\downarrow 0} \ln(x) = -\infty$, *and* $\ln(I\!R_{++}) = I\!R$.
(iii) $\ln(1) = 0$.
(iv) $\ln(e) = 1$.
(v) $\ln(xy) = \ln(x) + \ln(y) \quad \forall x, y \in I\!R_{++}$.

The natural logarithm can also be written as an infinite sum, namely,

$$\ln(x) = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n}(x-1)^n \quad \forall x \in \mathbb{R}_{++}.$$

An important implication of part (v) of Theorem 3.2.12 is

$$\ln(x^n) = n\ln(x) \quad \forall x \in \mathbb{R}_{++}, \ \forall n \in \mathbb{N}.$$

Because $\ln = E^{-1}$, we obtain

$$x^n = E(\ln(x^n)) = E(n\ln(x)) \quad \forall x \in \mathbb{R}_{++}, \ \forall n \in \mathbb{N}.$$

Therefore, the power function x^n can be expressed in terms of exponential and logarithmic functions. More generally, we can define *generalized powers* in the following way. For $x \in \mathbb{R}_{++}$ and $y \in \mathbb{R}$, we define

$$x^y := E(y\ln(x)).$$

This allows us to define generalizations of power functions such as

$$f : \mathbb{R}_{++} \mapsto \mathbb{R}, \ x \mapsto x^\alpha$$

where $\alpha \in \mathbb{R}$ (note that α is *not* restricted to natural or rational numbers). Functions of this type are called *generalized power functions.* Similarly, we can define functions such as

$$f : \mathbb{R} \mapsto \mathbb{R}_{++}, \; x \mapsto \alpha^x$$

where $\alpha \in \mathbb{R}_{++}$. Note that the variable x appears in the *exponent.* These functions are called *generalized exponential functions.* Clearly, a special case is the function E, which is obtained by setting $\alpha = e$.

Using the properties of E and ln, it is now easy to find the derivatives of generalized power functions and generalized exponential functions. Let

$$f : \mathbb{R}_{++} \mapsto \mathbb{R}, \; x \mapsto x^\alpha$$

with $\alpha \in \mathbb{R}$. By definition,

$$f(x) = E(\alpha \ln(x)) \quad \forall x \in \mathbb{R}_{++}.$$

Using the chain rule, we obtain, for $x \in \mathbb{R}_{++}$,

$$f'(x) = E'(\alpha \ln(x))\alpha\frac{1}{x} = E(\alpha \ln(x))\alpha\frac{1}{x} = x^\alpha\alpha\frac{1}{x} = \alpha x^{\alpha-1},$$

and therefore, the differentiation rule for power functions generalizes naturally.

For the function

$$f : \mathbb{R} \mapsto \mathbb{R}_{++}, \; x \mapsto \alpha^x$$

with $\alpha \in \mathbb{R}_{++}$, we have

$$f(x) = E(x \ln(\alpha)) \quad \forall x \in \mathbb{R}.$$

Applying the chain rule, it follows that

$$f'(x) = E'(x \ln(\alpha)) \ln(\alpha) = E(x \ln(\alpha)) \ln(\alpha) = \alpha^x \ln(\alpha)$$

for all $x \in \mathbb{R}$.

The *inverse* of a generalized exponential function is a *logarithmic function*, which generalizes the natural logarithm.

Definition 3.2.13 *Let $\alpha \in \mathbb{R}_{++} \setminus \{1\}$, and let $E_\alpha : \mathbb{R} \mapsto \mathbb{R}_{++}$ be a generalized exponential function. The* logarithm to the base α, $\log_\alpha : \mathbb{R}_{++} \mapsto \mathbb{R}$, *is the inverse of E_α, that is,* $\log_\alpha = E_\alpha^{-1}$.

By definition of the generalized exponential functions, we obtain, for all $x \in \mathbb{R}_{++}$,

$$x = E(\ln(x)) = E_\alpha(\log_\alpha(x)) = \alpha^{\log_\alpha(x)} = E(\log_\alpha(x) \ln(\alpha)).$$

Taking logarithms on both sides, we obtain

$$\ln(x) = \ln(E(\log_\alpha(x) \ln(\alpha))) = \log_\alpha(x) \ln(\alpha),$$

and therefore,

$$\log_\alpha(x) = \frac{\ln(x)}{\ln(\alpha)} \quad \forall x \in \mathbb{R}_{++}.$$

Therefore, the logarithm to *any* base $\alpha \in \mathbb{R}_{++} \setminus \{1\}$ can be expressed in terms of the natural logarithm.

As was mentioned earlier, exponential functions play an important role in growth models. Suppose, for example, the real national income in an economy is described by a function

$$Y : \mathbb{N} \cup \{0\} \mapsto \mathbb{R}_{++}, \; t \mapsto Y(t)$$

where t is a variable that represents time, and hence, $Y(t)$ is the national income in period $t = 0, 1, 2, \ldots$. A standard formulation of a model describing economic growth is to use functions involving *exponential* growth. For example, define

$$Y : \mathbb{N} \cup \{0\} \mapsto \mathbb{R}_{++}, \; t \mapsto y_0(1 + r)^t$$

where $y_0 \in \mathbb{R}_{++}$, and $r \in \mathbb{R}_{++}$ is the *growth rate.* By setting $t = 0$, we see that $y_0 = Y(0)$, that is, y_0 is the *initial value* for the growth process described by Y.

In discussing economic growth, one should keep in mind that *exponential* growth (that is, growth involving a growth rate at or above a given positive rate) cannot be expected to be sustainable in the long run. Consider the following example. Suppose the annual growth rate is 5%. How long will it take, given this growth rate, to *double* the initial national income? To answer this question, we have to find $t \in \mathbb{N}$ such that

$$Y(t) \geq 2Y(0). \tag{3.8}$$

For an exponential growth process with $r = 0.05$, we obtain

$$Y(t) = Y(0)(1.05)^t,$$

and therefore, (3.8) is equivalent to

$$(1.05)^t \geq 2.$$

Taking logarithms on both sides, we obtain

$$t \ln(1.05) \geq \ln(2),$$

or

$$t \geq \frac{\ln(2)}{\ln(1.05)} = 14.2067.$$

Therefore, if a growth process is described by an exponential function with a growth rate of 5%, it takes only 14 years to *double* the existing national income. If no further assumptions on *how* and in which sectors this growth comes about are made, this clearly could lead to rather undesirable results (such as irreparable environmental damage), and therefore, the growth rate of real national income *alone* is *not* a useful indicator for the societal well-being in an economy.

We conclude this section with a discussion of the two basic *trigonometric* functions.

Definition 3.2.14 *The function*

$$\sin : \mathbb{R} \mapsto \mathbb{R},\ x \mapsto \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!}$$

is called the sine function.

Definition 3.2.15 *The function*

$$\cos : \mathbb{R} \mapsto \mathbb{R},\ x \mapsto \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!}$$

is called the cosine function.

A geometric interpretation of those two functions is provided in Figure 3.5. Consider a circle around the origin of $\mathbb{R}^2$ with radius $r = 1$. Starting at the point $(1, 0)$, let x be the distance travelled counter-clockwise along the circle. The first coordinate of the resulting point is the value of the cosine function at x and its second coordinate is the value of the sine function at x.

The following theorem states some important properties of the sine and cosine functions.

Theorem 3.2.16 *(i) The function* sin *is differentiable on* $\mathbb{R}$, *and* $\sin'(x) = \cos(x) \quad \forall x \in \mathbb{R}$.
(ii) The function cos *is differentiable on* $\mathbb{R}$, *and* $\cos'(x) = -\sin(x) \quad \forall x \in \mathbb{R}$.
(iii) $(\sin(x))^2 + (\cos(x))^2 = 1 \quad \forall x \in \mathbb{R}$.
(iv) $\sin(x + y) = \sin(x)\cos(y) + \cos(x)\sin(y) \quad \forall x, y \in \mathbb{R}$.
(v) $\cos(x + y) = \cos(x)\cos(y) - \sin(x)\sin(y) \quad \forall x, y \in \mathbb{R}$.

Proof. (i) and (ii) follow immediately from applying the rules of differentiation for polynomials to the definitions of sin and cos.

(iii) Define

$$f : \mathbb{R} \mapsto \mathbb{R},\ x \mapsto (\sin(x))^2 + (\cos(x))^2.$$

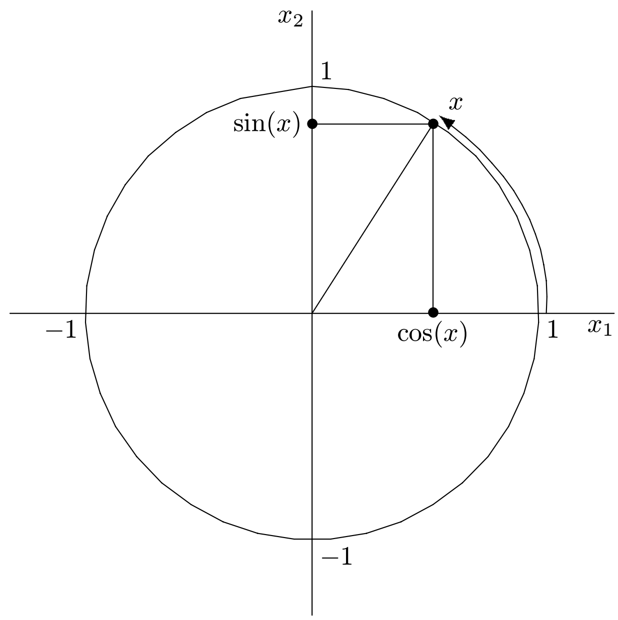


Figure 3.5: Trigonometric functions.

Using parts (i) and (ii), differentiating f yields

$$f'(x) = 2\sin(x)\cos(x) - 2\cos(x)\sin(x) = 0 \quad \forall x \in \mathbb{R}.$$

This means that the function f is constant, and we obtain

$$f(x) = f(0) = (\sin(0))^2 + (\cos(0))^2 = 0 + 1 = 1 \quad \forall x \in \mathbb{R}.$$

To prove (iv) and (v), an argument analogous to the one employed in the proof of (iii) can be employed; as an exercise, provide the details. ‖

Given the geometric interpretation illustrated in Figure 3.5, part (iii) of the above theorem is a variant of Pythagoras's theorem.

In order to provide an illustration of the graphs of sin and cos, we state (without proof) the following theorem which summarizes how some important values of those functions are obtained.

Theorem 3.2.17 *(i)* $\sin(\pi/2) = 1$ *and* $\cos(\pi/2) = 0$.
(ii) $\sin((n + 1/2)\pi) = \cos(n\pi) = (-1)^n$ *and* $\sin(n\pi) = \cos((n + 1/2)\pi) = 0 \quad \forall n \in \mathbb{N}$.
(iii) $\sin(x + \pi/2) = \cos(x)$ *and* $\cos(x + \pi/2) = -\sin(x) \quad \forall x \in \mathbb{R}$.
(iv) $\sin(x + n\pi) = (-1)^n \sin(x)$ *and* $\cos(x + n\pi) = (-1)^n \cos(x) \quad \forall x \in \mathbb{R}, \forall n \in \mathbb{N}$.
(v) $\sin(-x) = -\sin(x)$ *and* $\cos(-x) = \cos(x) \quad \forall x \in \mathbb{R}$.
(vi) $\sin(x) \in [-1, 1]$ *and* $\cos(x) \in [-1, 1] \quad \forall x \in \mathbb{R}$.

Using Theorem 3.2.17, the graphs of sin and cos are illustrated in Figures 3.6 and 3.7.

3.3 Optimization

The maximization and minimization of functions plays an important role in economic models. We will discuss an economic example at the end of the next section.

To introduce general optimization methods, we first define what is meant by *global* and *local* maxima and minima of real-valued functions.

Definition 3.3.1 *Let $A \subseteq \mathbb{R}$ be an interval, let $f : A \mapsto \mathbb{R}$, and let $x_0 \in A$.*

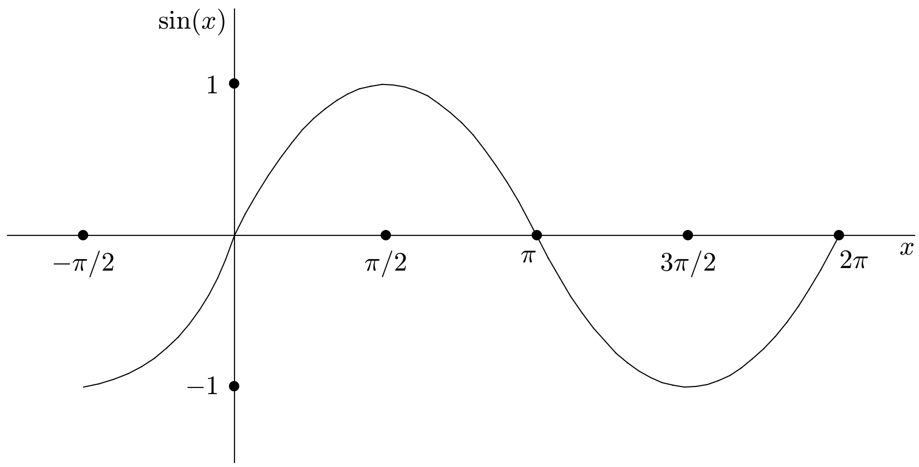


Figure 3.6: The graph of sin.

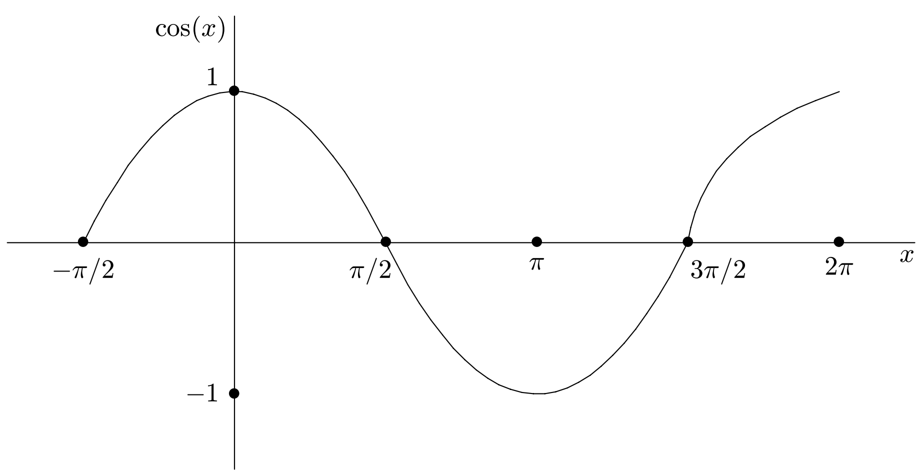


Figure 3.7: The graph of cos.

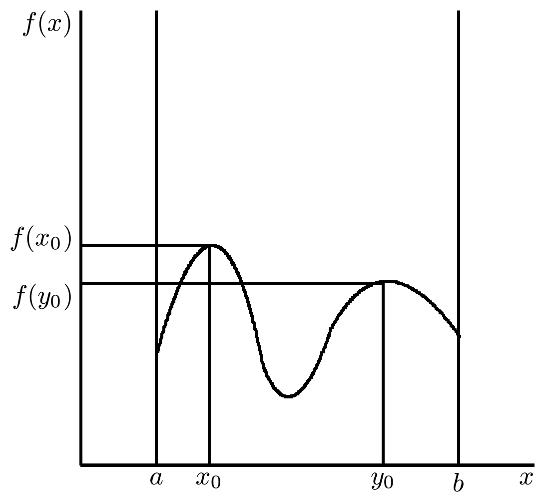


Figure 3.8: Local and global maxima.

(i) f has a global maximum *at* $x_0 \Leftrightarrow f(x_0) \geq f(x) \;\; \forall x \in A$.
(ii) f has a local maximum *at* $x_0 \Leftrightarrow \exists \varepsilon \in \mathbb{R}_{++}$ *such that* $f(x_0) \geq f(x) \;\; \forall x \in \mathcal{U}_\varepsilon(x_0) \cap A$.
(iii) f has a global minimum *at* $x_0 \Leftrightarrow f(x_0) \leq f(x) \;\; \forall x \in A$.
(iv) f has a local minimum *at* $x_0 \Leftrightarrow \exists \varepsilon \in \mathbb{R}_{++}$ *such that* $f(x_0) \leq f(x) \;\; \forall x \in \mathcal{U}_\varepsilon(x_0) \cap A$.

If a function has a global maximum (minimum) at a point x_0 in its domain, the value of f *cannot* be larger (smaller) than $f(x_0)$ at *any* point in its domain. For a local maximum (minimum), x_0 leads to a maximal (minimal) value of f *within a neighborhood* of x_0. Clearly, if f has a global maximum (minimum) at x_0, then f has a local maximum (minimum) at x_0, but a local maximum (minimum) need not be a global maximum (minimum).

As an example, consider the function $f : [a, b] \mapsto \mathbb{R}$ with a graph as illustrated in Figure 3.8.

The function f has a global maximum at x_0, because $f(x_0) \geq f(x)$ for all $x \in A = [a, b]$. Clearly, f also has a local maximum at x_0—if $f(x_0) \geq f(x)$ for all $x \in A$, it must be true that $f(x_0) \geq f(x)$ for all $x \in A$ that are in a neighborhood of x_0. Furthermore, f has a local maximum at y_0, but this maximum is *not* a global maximum, because $f(x_0) > f(y_0)$.

It should be noted that a global (local) maximum (minimum) *need not exist*, and if a local (global) maximum (minimum) exists, it *need not be unique.* For example, consider the function

$$f : \mathbb{R} \mapsto \mathbb{R}, \; x \mapsto x.$$

This function has no local maximum and no local minimum (and therefore, no global maximum and no global minimum). To show that f cannot have a local maximum, suppose, by way of contradiction, there exists $x_0 \in \mathbb{R}$ and $\varepsilon \in \mathbb{R}_{++}$ such that

$$f(x_0) \geq f(x) \quad \forall x \in \mathcal{U}_\varepsilon(x_0). \tag{3.9}$$

Let $\bar{x} := x_0 + \varepsilon/2$. Clearly, $\bar{x} \in \mathcal{U}_\varepsilon(x_0)$ and $\bar{x} > x_0$. By definition of f, $f(\bar{x}) = \bar{x} > x_0 = f(x_0)$, and therefore, we obtain a contradiction to (3.9). That f cannot have a local minimum is shown analogously.

Now consider the function

$$f : \mathbb{R} \mapsto \mathbb{R}, \; x \mapsto 1.$$

This function is *constant*, and it has infinitely many maxima and minima—f has a maximum and a minimum at *all* points $x_0 \in \mathbb{R}$, because

$$f(x_0) = 1 \geq 1 = f(x) \quad \forall x \in \mathbb{R} \;\text{ and }\; f(x_0) = 1 \leq 1 = f(x) \quad \forall x \in \mathbb{R}.$$

Therefore, f has a maximum and a minimum, but *not* a unique maximum or a unique minimum.

A very important theorem concerning maxima and minima states that a *continuous* function f *must* have a global maximum and a global minimum, if the domain of f is *closed and bounded.* To prove this theorem, we first state a result concerning *bounded* functions. Bounded functions are defined in

Definition 3.3.2 *Let $A \subseteq I\!R$ be an interval, and let $B \subseteq A$. Furthermore, let $f : A \mapsto I\!R$. f is* bounded on B *if and only if the set*

$$\{f(x) \mid x \in B\} \tag{3.10}$$

is bounded.

Note that boundedness of f on B implies that the set given by (3.10) has an infimum and a supremum.

We obtain

Theorem 3.3.3 *Let $A = [a, b]$ with $a, b \in I\!R$ and $a < b$, and let $f : A \mapsto I\!R$. If f is continuous on $[a, b]$, then f is bounded on $[a, b]$.*

Proof. Because f is continuous on $A = [a, b]$, for any $\delta \in \mathbb{R}_{++}$, there exists $\varepsilon \in \mathbb{R}_{++}$ such that $f(x) \in \mathcal{U}_\delta(f(a))$ for all $x \in [a, a + \varepsilon)$. Letting $\xi \in (a, a + \varepsilon)$, it follows that f is bounded on $[a, \xi]$. Now define

$$X := \{\gamma \in (a, b] \mid f \text{ is bounded on } [a, \gamma]\}.$$

We have already shown that there exists $\xi \in (a, b]$ such that f is bounded on $[a, \xi]$, and therefore, X is nonempty. Clearly, if $\xi \in X$, it follows that $\gamma \in X$ for all $\gamma \in (a, \xi]$. We now show that $b \in X$, and therefore, f is bounded on $[a, b]$.

By way of contradiction, suppose $b \notin X$. Let $\beta := \sup(X)$. This implies that $\xi \in X$ for all $\xi \in (a, \beta)$. First, suppose $\beta < b$. Because f is continuous on $[a, b]$, for any $\delta \in \mathbb{R}_{++}$, we can find $\varepsilon \in \mathbb{R}_{++}$ such that $f(x) \in \mathcal{U}_\delta(f(\beta))$ for all $x \in (\beta - \varepsilon, \beta + \varepsilon)$. This means that f is bounded on $[\beta - \xi, \beta + \xi]$ for some $\xi \in (0, \varepsilon)$. Because $\beta - \xi < \beta$, it follows that f is bounded on $[a, \beta - \xi)$ and on $[\beta - \xi, \beta + \xi]$, and therefore, on $[a, \beta + \xi]$. Therefore, $(\beta + \xi) \in X$. But we assumed $\beta = \sup(X)$, and therefore, $(\beta + \xi) \notin X$, a contradiction. If $\beta = b$, the above argument can be applied with $\beta + \xi$ replaced by b. $\|$

Now we can prove

Theorem 3.3.4 *Let $A = [a, b]$ with $a, b \in I\!R$ and $a < b$, and let $f : A \mapsto I\!R$. If f is continuous on $[a, b]$, then there exist $x_0, y_0 \in A$ such that f has a global maximum at x_0 and f has a global minimum at y_0.*

Proof. By Theorem 3.3.3, f is bounded on $[a, b]$. Let $M = \sup(\{f(x) \mid x \in [a, b]\})$ and $m = \inf(\{f(x) \mid x \in [a, b]\})$. We have to show that there exist $x^0, y^0 \in [a, b]$ such that $f(x^0) = M$ and $f(y^0) = m$.

We proceed by contradiction. Suppose there exists no $x^0 \in [a, b]$ such that $f(x^0) = M$. By definition of M, this implies

$$f(x) < M \quad \forall x \in [a, b]. \tag{3.11}$$

Because f is continuous, for any $\delta \in \mathbb{R}_{++}$, there exists $\alpha \in [a, b]$ such that $M - f(\alpha) < \delta$. Because $f(\alpha) < M$, this is equivalent to

$$\frac{1}{M - f(\alpha)} > \frac{1}{\delta}.$$

Therefore, the function

$$g : [a, b] \mapsto \mathbb{R}_{++},\ x \mapsto \frac{1}{M - f(x)}$$

is not bounded (because we can choose δ arbitrarily close to zero). Note that (3.11) guarantees that g is well-defined, because $M - f(x) > 0$ for all $x \in [a, b]$. It follows that g cannot be continuous (because continuity of g would, by Theorem 3.3.3, imply boundedness of g). But continuity of f implies continuity of g, a contradiction. Therefore, there must exist $x^0 \in [a, b]$ such that $f(x^0) = M$.

The existence of $y^0 \in [a, b]$ such that $f(y^0) = m$ is proven analogously. $\|$

The assumption that A is *closed and bounded* is essential for Theorem 3.3.4. For example, consider the function

$$f : (0, 1) \mapsto \mathbb{R},\ x \mapsto x.$$

f is continuous on $A = (0, 1)$, but A is not closed. It is easy to show that f has no local maximum and no local minimum (Exercise: prove this), and therefore, f has no global maximum and no global minimum. That boundedeness is important in Theorem 3.3.4 can be illustrated by the example

$$f : \mathbb{R} \mapsto \mathbb{R},\ x \mapsto x.$$

We have already shown earlier that this continuous function has no maximum and no minimum. The domain of f is closed, but not bounded.

If the domain of a function is closed and bounded but f is not continuous, maxima and minima need not exist. Consider the function

$$f : [-1, 1] \mapsto \mathbb{R},\ x \mapsto \begin{cases} 1/x & \text{if } x \neq 0 \\ 0 & \text{if } x = 0. \end{cases}$$

Clearly, this function has no global maximum and no global minimum (Exercise: provide a proof).

Theorem 3.3.4 allows us to show that the image of $[a, b]$ under a function $f : [a, b] \mapsto \mathbb{R}$ must be a closed and bounded interval. This result is called the *intermediate value theorem.*

Theorem 3.3.5 *Let $A = [a, b]$ with $a, b \in \mathbb{R}$ and $a < b$, and let $f : A \mapsto \mathbb{R}$. If f is continuous on A, then*

$$f(A) = [\alpha, \beta],$$

where $\alpha := \min(f(A))$ and $\beta := \max(f(A))$.

Proof. Note that, by Theorem 3.3.4, f has a global minimum and a global maximum, and therefore, α and β are well-defined. Let $c_0 \in [\alpha, \beta]$. We have to show that there exists $x_0 \in [a, b]$ such that $f(x_0) = c_0$. For $c_0 = \alpha$ or $c_0 = \beta$, we are done by the definition of α and β. Let $c_0 \in (\alpha, \beta)$. Let $x, y \in [a, b]$ be such that $f(x) = \alpha$ and $f(y) = \beta$. Define the sequences $\{x_n\}$, $\{y_n\}$, $\{z_n\}$ by

$$x_1 := x,\ \ y_1 := y,\ \ z_1 := \frac{x_1 + y_1}{2},$$

and, for $n \geq 2$,

$$x_n := \begin{cases} x_{n-1} & \text{if } f(z_{n-1}) \geq c_0 \\ z_{n-1} & \text{if } f(z_{n-1}) < c_0, \end{cases}$$

$$y_n := \begin{cases} z_{n-1} & \text{if } f(z_{n-1}) \geq c_0 \\ y_{n-1} & \text{if } f(z_{n-1}) < c_0, \end{cases}$$

$$z_n := \frac{x_n + y_n}{2}.$$

By definition, $\{x_n\}$ is monotone nondecreasing, and $\{y_n\}$ is monotone nonincreasing. Furthermore,

$$f(x_n) < c_0 \leq f(y_n) \quad \forall n \in \mathbb{N} \tag{3.12}$$

and

$$|x_n - y_n| = \frac{|x_1 - y_1|}{2^{n-1}} \quad \forall n \in \mathbb{N}.$$

This implies $|x_n - y_n| \longrightarrow 0$, and therefore, the sequences $\{x_n\}$ and $\{y_n\}$ converge to the same limit $x_0 \in [a, b]$. Because f is continuous, $f(x_0) = \lim_{n\to\infty} f(x_n) = \lim_{n\to\infty} f(y_n)$. By (3.12), $\lim_{n\to\infty} f(x_n) \leq c_0 \leq \lim_{n\to\infty} f(y_n)$, and therefore, $f(x_0) = c_0$. ‖

The intermediate value theorem can be generalized. In particular, if A is an interval (not necessarily closed and bounded) and f is continuous, then $f(A)$ must be an interval.

If a function is differentiable, the task of finding maxima and minima of this function can become much easier. The following result provides a *necessary* condition for a local maximum (minimum) at an *interior* point of A.

Theorem 3.3.6 *Let $A \subseteq \mathbb{R}$ be an interval, and let $f : A \mapsto \mathbb{R}$. Furthermore, let x_0 be an interior point of A, and suppose f is differentiable at x_0.*

(i) f has a local maximum at $x_0 \Rightarrow f'(x_0) = 0$.
(ii) f has a local minimum at $x_0 \Rightarrow f'(x_0) = 0$.

Proof. (i) Suppose f has a local maximum at an interior point $x_0 \in A$. Then there exists $\varepsilon \in \mathbb{R}_{++}$ such that $f(x_0) \geq f(x)$ for all $x \in \mathcal{U}_\varepsilon(x_0)$. Therefore,

$$f(x_0 + h) - f(x_0) \leq 0 \tag{3.13}$$

for all $h \in \mathbb{R}$ such that $(x_0 + h) \in \mathcal{U}_\varepsilon(x_0)$.

First, let $h > 0$. Then (3.13) implies

$$\frac{f(x_0 + h) - f(x_0)}{h} \leq 0.$$

Because f is differentiable at x_0, we obtain

$$\lim_{h \downarrow 0} \frac{f(x_0 + h) - f(x_0)}{h} \leq 0 \tag{3.14}$$

(note that differentiability of f at x_0 ensures that this limit exists).

Now let $h < 0$. By (3.13), we obtain

$$\frac{f(x_0 + h) - f(x_0)}{h} \geq 0,$$

and hence,

$$\lim_{h \uparrow 0} \frac{f(x_0 + h) - f(x_0)}{h} \geq 0. \tag{3.15}$$

Because f is differentiable at x_0,

$$\lim_{h \downarrow 0} \frac{f(x_0 + h) - f(x_0)}{h} = \lim_{h \uparrow 0} \frac{f(x_0 + h) - f(x_0)}{h} = \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h} = f'(x_0).$$

Therefore, (3.14) implies $f'(x_0) \leq 0$, and (3.15) implies $f'(x_0) \geq 0$, which is only possible if $f'(x_0) = 0$.

The proof of (ii) is analogous and is left as an exercise. ‖

Theorem 3.3.6 says that *if* f has a local maximum or minimum at an interior point x_0, *then* we must have $f'(x_0) = 0$. Therefore, this theorem provides a *necessary* condition for a local maximum or minimum, but this condition is, in general, *not sufficient*. This means that if we find an interior point $x_0 \in A$ such that $f'(x_0) = 0$, this does *not* imply that f has a local maximum or a local minimum at x_0. To illustrate that, consider the function

$$f : \mathbb{R} \mapsto \mathbb{R},\ x \mapsto x^3.$$

We have $f'(x) = 3x^2$ for all $x \in \mathbb{R}$, and therefore, $f'(0) = 0$. But f does *not* have a local maximum or minimum at $x_0 = 0$, because, for *any* $\varepsilon \in \mathbb{R}_{++}$, we have $\varepsilon/2 \in \mathcal{U}_\varepsilon(0)$ and $-\varepsilon/2 \in \mathcal{U}_\varepsilon(0)$, and furthermore,

$$f\left(\frac{\varepsilon}{2}\right) = \frac{\varepsilon^3}{8} > 0 = f(0) \tag{3.16}$$

and

$$f\left(\frac{-\varepsilon}{2}\right) = -\frac{\varepsilon^3}{8} < 0 = f(0). \tag{3.17}$$

(3.16) implies that f cannot have a local maximum at $x_0 = 0$, and (3.17) establishes that f cannot have a local minimum at $x_0 = 0$.

Theorem 3.3.6 applies to *interior* points of A only. If f is defined on an interval including one or both of its endpoints, a zero derivative is *not* necessary for a maximum at an endpoint. For example, consider a function $f : [a, b] \mapsto \mathbb{R}$ with a graph as illustrated in Figure 3.9.

Clearly, f has local maxima at a and at b. Because a and b are not interior points, Theorem 3.3.6 cannot be applied. However, if the right-side derivative of f exists at a, we can conclude that this derivative *cannot* be positive if f has a local maximum at a. Analogously, if f has a local maximum at b and the left-side derivative of f exists at b, this derivative *must* be nonnegative. Similar considerations apply to minima at the endpoints of $[a, b]$. This is summarized in

Theorem 3.3.7 *Let $A = [a, b]$, where $a, b \in \mathbb{R}$ and $a < b$. Furthermore, let $f : A \mapsto \mathbb{R}$.*

(i) If f has a local maximum at a and the right-side derivative of f at a exists, then $f'(a) \leq 0$.
(ii) If f has a local maximum at b and the left-side derivative of f at b exists, then $f'(b) \geq 0$.
(iii) If f has a local minimum at a and the right-side derivative of f at a exists, then $f'(a) \geq 0$.
(iv) If f has a local minimum at b and the left-side derivative of f at b exists, then $f'(b) \leq 0$.

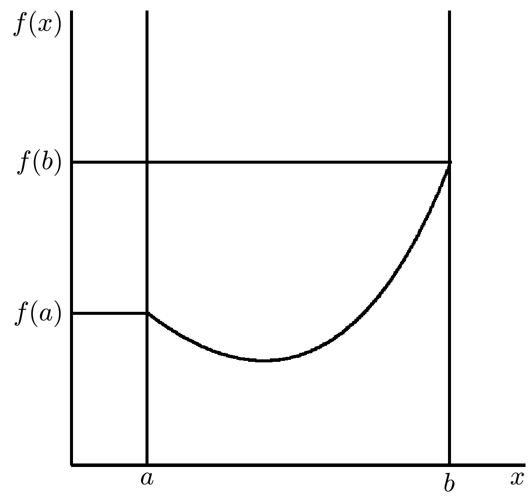


Figure 3.9: Maxima at endpoints.

The proof of Theorem 3.3.7 is analogous to the proof of Theorem 3.3.6 and is left as an exercise.

The necessary conditions for local maxima and minima presented in Theorems 3.3.6 and 3.3.7 are sometimes called *first-order conditions* for local maxima and minima, because they involve the first-order derivatives of the function f. Points $x_0 \in A$ satisfying these first-order conditions are called *critical points*.

As was mentioned earlier, the first-order conditions only provide *necessary*, but not sufficient conditions for local maxima and minima. Once we have found a critical point satisfying the first-order conditions, we have to do some more work to find out whether this point is a local maximum, minimum, or neither a maximum nor a minimum.

Next, we formulate some results that will be useful in determining the nature of critical points. The following result is usually referred to as *Rolle's theorem*.

Theorem 3.3.8 *Let $[a,b] \subseteq A$, where $A \subseteq \mathbb{R}$ is an interval, $a, b \in \mathbb{R}$, and $a < b$. Furthermore, let $f : A \mapsto \mathbb{R}$ be continuous on $[a,b]$ and differentiable on (a,b). If $f(a) = f(b)$, then there exists $x_0 \in (a,b)$ such that $f'(x_0) = 0$.*

Proof. Suppose $f(a) = f(b)$. By Theorem 3.3.4, there exist $x_0, y_0 \in [a,b]$ such that

$$f(x_0) \geq f(x) \geq f(y_0) \quad \forall x \in [a,b]. \tag{3.18}$$

If $f(x_0) = f(y_0)$, (3.18) implies that f has a local maximum and a local minimum at *any* point $x \in (a,b)$. By Theorem 3.3.6, it follows that $f'(x) = 0$ for all $x \in (a,b)$.

Now suppose $f(x_0) > f(y_0)$. If $f(x_0) > f(a)$, it follows that $x_0 \neq a$ and $x_0 \neq b$ (because, by assumption, $f(a) = f(b)$). Therefore, $x_0 \in (a,b)$, and by Theorem 3.3.6, $f'(x_0) = 0$.

If $f(x_0) = f(a)$, it follows that $f(y_0) < f(a)$, and therefore, $y_0 \neq a$ and $y_0 \neq b$. This implies that y_0 is an interior point, that is, $y_0 \in (a,b)$. Because f has a local minimum at y_0, Theorem 3.3.6 implies $f'(y_0) = 0$, which completes the proof. ‖

The next theorem is the *mean-value theorem*.

Theorem 3.3.9 *Let $[a,b] \subseteq A$, where $A \subseteq \mathbb{R}$ is an interval, $a, b \in \mathbb{R}$, and $a < b$. Furthermore, let $f : A \mapsto \mathbb{R}$ be continuous on $[a,b]$ and differentiable on (a,b). There exists $x_0 \in (a,b)$ such that*

$$f'(x_0) = \frac{f(b) - f(a)}{b - a}.$$

The mean-value theorem can be generalized. The following result is the *generalized mean-value theorem*.

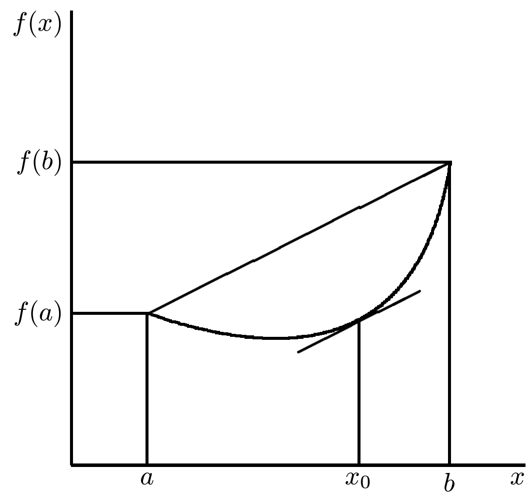


Figure 3.10: The mean-value theorem.

Theorem 3.3.10 *Let $[a,b] \subseteq A$, where $A \subseteq I\!R$ is an interval, $a, b \in I\!R$, and $a < b$. Furthermore, let $f : A \mapsto I\!R$ and $g : A \mapsto I\!R$ be continuous on $[a,b]$ and differentiable on (a,b), and suppose $g'(x) \neq 0$ for all $x \in (a,b)$. There exists $x_0 \in (a,b)$ such that*

$$\frac{f'(x_0)}{g'(x_0)} = \frac{f(b) - f(a)}{g(b) - g(a)}.$$

Proof. Because $g'(x) \neq 0$ for all $x \in (a,b)$, we must have $g(a) \neq g(b)$—otherwise, we would obtain a contradiction to Rolle's theorem. Therefore, $g(b) - g(a) \neq 0$, and we can define the function

$$k : [a,b] \mapsto I\!R,\ x \mapsto f(x) - \frac{f(b) - f(a)}{g(b) - g(a)}(g(x) - g(a)).$$

Because f and g are continuous on $[a,b]$ and differentiable on (a,b), k is continuous on $[a,b]$ and differentiable on (a,b). Furthermore, $k(a) = k(b) = f(a)$. By Theorem 3.3.8, there exists $x_0 \in (a,b)$ such that

$$k'(x_0) = 0. \tag{3.19}$$

By definition of k,

$$k'(x) = f'(x) - \frac{f(b) - f(a)}{g(b) - g(a)} g'(x) \quad \forall x \in (a,b),$$

and therefore, (3.19) implies

$$\frac{f'(x_0)}{g'(x_0)} = \frac{f(b) - f(a)}{g(b) - g(a)}. \quad \|$$

The mean-value theorem is obtained as a special case of Theorem 3.3.10, where $g(x) = x$ for all $x \in A$. See Figure 3.10 for an illustration of the mean-value theorem. The mean-value theorem says that there must exist a point $x_0 \in (a,b)$ such that the slope of the tangent to the graph of f at x_0 is equal to the slope of the secant through $(a, f(a))$ and $(b, f(b))$, given that f is continuous on $[a,b]$ and differentiable on (a,b).

The mean-value theorem can be used to provide criteria for the monotonicity properties introduced in Definition 3.1.12.

Theorem 3.3.11 *Let $A, B \subseteq I\!R$ be intervals with $B \subseteq A$, and let $f : A \mapsto I\!R$ be differentiable on B.*

(i) $f'(x) \geq 0\ \forall x \in B \Leftrightarrow f$ is nondecreasing on B.
(ii) $f'(x) > 0\ \forall x \in B \Rightarrow f$ is increasing on B.
(iii) $f'(x) \leq 0\ \forall x \in B \Leftrightarrow f$ is nonincreasing on B.
(iv) $f'(x) < 0\ \forall x \in B \Rightarrow f$ is decreasing on B.

Proof. (i) "$\Rightarrow$": Suppose $f'(x) \geq 0$ for all $x \in B$. Let $x, y \in B$ with $x > y$. By the mean-value theorem, there exists $x_0 \in (y, x)$ such that

$$f'(x_0) = \frac{f(x) - f(y)}{x - y}.$$

Because $f'(x_0) \geq 0$ and $x > y$, this implies $f(x) \geq f(y)$.

"$\Leftarrow$": Suppose f is nondecreasing on B. Let $x_0 \in B$ be such that $(x_0, x_0 + \varepsilon) \subseteq B$ for some $\varepsilon \in \mathbb{R}_{++}$. For $h \in (0, \varepsilon)$, nondecreasingness implies $f(x_0 + h) - f(x_0) \geq 0$, and therefore,

$$\frac{f(x_0 + h) - f(x_0)}{h} \geq 0.$$

This implies

$$\lim_{h \downarrow 0} \frac{f(x_0 + h) - f(x_0)}{h} \geq 0.$$

Because f is differentiable at x_0,

$$f'(x_0) = \lim_{h \downarrow 0} \frac{f(x_0 + h) - f(x_0)}{h} \geq 0.$$

If $x_0 \in B$ is the right-side endpoint of B, there exists $\varepsilon \in \mathbb{R}_{++}$ such that $(x_0 - \varepsilon, x_0) \subseteq B$, and the same argument as above can be applied with $h \in (-\varepsilon, 0)$.

(ii) Suppose $f'(x) > 0$ for all $x \in B$. Let $x, y \in B$ with $x > y$. As in part (i), the mean-value theorem implies that there exists $x_0 \in (y, x)$ such that

$$f'(x_0) = \frac{f(x) - f(y)}{x - y}.$$

Because $f'(x_0) > 0$ and $x > y$, we obtain $f(x) > f(y)$.

The proofs of parts (iii) and (iv) are analogous. ‖

Note that the reverse implications of (ii) and (iv) in Theorem 3.3.11 are *not* true. For example, consider the function

$$f : \mathbb{R} \mapsto \mathbb{R},\ x \mapsto x^3.$$

This function is increasing on its domain, but $f'(0) = 0$.

Furthermore, the assumption that B is an interval is essential in Theorem 3.3.11. To see this, consider the function

$$f : (0,1) \cup (1,2) \mapsto \mathbb{R},\ x \mapsto \begin{cases} x & \text{if } x \in (0,1) \\ x - 1 & \text{if } x \in (1,2). \end{cases}$$

This function is differentiable on its domain $A = (0,1) \cup (1,2)$. Furthermore, $f'(x) = 1 > 0$ for all $x \in A$, but f is *not* increasing (not even nondecreasing) on A, because, for example, $f(3/4) = 3/4 > 1/4 = f(5/4)$.

The next result can be very useful in finding limits of functions. It is called *l'Hôpital's rule.*

Theorem 3.3.12 *Let $A \subseteq \mathbb{R}$, and let $f : A \mapsto \mathbb{R}$ and $g : A \mapsto \mathbb{R}$ be functions. Let $\alpha \in \mathbb{R} \cup \{-\infty, \infty\}$.*

(i) Let $x_0 \in \mathbb{R} \cup \{-\infty\}$, and suppose $(x_0, x_0 + h) \subseteq A$ for some $h \in \mathbb{R}_{++}$. Furthermore, suppose f and g are differentiable on $(x_0, x_0 + h)$, and $g(x) \neq 0$, $g'(x) \neq 0$ for all $x \in (x_0, x_0 + h)$. If

$$\lim_{x \downarrow x_0} f(x) = \lim_{x \downarrow x_0} g(x) = 0 \vee \lim_{x \downarrow x_0} g(x) = -\infty \vee \lim_{x \downarrow x_0} g(x) = \infty,$$

then

$$\lim_{x \downarrow x_0} \frac{f'(x)}{g'(x)} = \alpha \Rightarrow \lim_{x \downarrow x_0} \frac{f(x)}{g(x)} = \alpha.$$

(ii) Let $x_0 \in \mathbb{R} \cup \{\infty\}$, and suppose $(x_0 - h, x_0) \subseteq A$ for some $h \in \mathbb{R}_{++}$. Furthermore, suppose f and g are differentiable on $(x_0 - h, x_0)$, and $g(x) \neq 0$, $g'(x) \neq 0$ for all $x \in (x_0 - h, x_0)$. If

$$\lim_{x \uparrow x_0} f(x) = \lim_{x \uparrow x_0} g(x) = 0 \vee \lim_{x \uparrow x_0} g(x) = -\infty \vee \lim_{x \uparrow x_0} g(x) = \infty,$$

then

$$\lim_{x \uparrow x_0} \frac{f'(x)}{g'(x)} = \alpha \Rightarrow \lim_{x \uparrow x_0} \frac{f(x)}{g(x)} = \alpha.$$

Proof. We provide a proof of (i) in the case where $x_0 \in \mathbb{R}$, $\alpha \in \mathbb{R}$, and $\lim_{x \downarrow x_0} f(x) = \lim_{x \downarrow x_0} g(x) = 0$; the other cases are similar. Suppose

$$\lim_{x \downarrow x_0} f(x) = \lim_{x \downarrow x_0} g(x) = 0. \tag{3.20}$$

Let $\delta \in \mathbb{R}_{++}$. Because

$$\lim_{x \downarrow x_0} \frac{f'(x)}{g'(x)} = \alpha,$$

there exists $\varepsilon \in \mathbb{R}_{++}$ such that

$$\frac{f'(x)}{g'(x)} \in \mathcal{U}_{\delta/2}(\alpha) \quad \forall x \in (x_0, x_0 + \varepsilon).$$

Let $y \in (x_0, x_0 + \varepsilon)$. By (3.20),

$$\frac{f(y)}{g(y)} = \lim_{x \downarrow x_0} \frac{f(y) - f(x)}{g(y) - g(x)}.$$

By the generalized mean-value theorem, for any $x \in (x_0, y)$, there exists $y_0 \in (x, y)$—which may depend on x—such that

$$\frac{f'(y_0)}{g'(y_0)} = \frac{f(y) - f(x)}{g(y) - g(x)}.$$

Therefore,

$$\frac{f(y)}{g(y)} = \lim_{x \downarrow x_0} \frac{f'(y_0)}{g'(y_0)}.$$

Because

$$\frac{f'(y_0)}{g'(y_0)} \in \mathcal{U}_{\delta/2}(\alpha),$$

it follows that

$$\frac{f(y)}{g(y)} \in \mathcal{U}_{\delta}(\alpha),$$

and therefore,

$$\lim_{x \downarrow x_0} \frac{f(x)}{g(x)} = \alpha. \quad \|$$

To illustrate the application of l'Hôpital's rule, consider the following example. Let

$$f : (1, 2) \mapsto \mathbb{R},\ x \mapsto x^2 - 1$$

and

$$g : (1, 2) \mapsto \mathbb{R},\ x \mapsto x - 1.$$

We obtain $\lim_{x \downarrow 1} f(x) = \lim_{x \downarrow 1} g(x) = 0$. Furthermore, $f'(x) = 2x$ and $g'(x) = 1$ for all $x \in (1, 2)$. Therefore,

$$\lim_{x \downarrow 1} \frac{f'(x)}{g'(x)} = \lim_{x \downarrow 1} 2x = 2.$$

According to l'Hôpital's rule,

$$\lim_{x \downarrow 1} \frac{f(x)}{g(x)} = 2.$$

Next, we present a result that can be used to *approximate* the value of a function by means of a *polynomial.* For functions with a complex structure, this can be a substantial simplification, because the values of polynomials are relatively easy to compute. We obtain

Theorem 3.3.13 *Let $A \subseteq \mathbb{R}$, and let $f : A \mapsto \mathbb{R}$. Let $n \in \mathbb{N}$, and let x_0 be an interior point of A. If f is $n+1$ times differentiable on $\mathcal{U}_\varepsilon(x_0)$ for some $\varepsilon \in \mathbb{R}_{++}$, then*

$$f(x) = f(x_0) + \sum_{k=1}^{n} \frac{f^{(k)}(x_0)(x - x_0)^k}{k!} + R(x)$$

for all $x \in \mathcal{U}_\varepsilon(x_0)$, where

$$R(x) = \frac{f^{(n+1)}(\xi)(x - x_0)^{n+1}}{(n+1)!} \tag{3.21}$$

for some ξ between x and x_0.

Proof. For $x \in \mathcal{U}_\varepsilon(x_0)$, let

$$R(x) := f(x) - f(x_0) - \sum_{k=1}^{n} \frac{f^{(k)}(x_0)(x - x_0)^k}{k!}.$$

We have to show that $R(x)$ must be of the form (3.21). By definition, $R(x_0) = 0$, and

$$R'(x_0) = f'(x_0) - \sum_{k=1}^{n} f^{(k)}(x_0)\frac{k(x - x_0)^{k-1}}{k!} = 0,$$

and analogously, $R^{(k)}(x_0) = 0$ for all $k = 1, \ldots, n$. Furthermore,

$$R^{(n+1)}(x) = f^{(n+1)}(x) \quad \forall x \in \mathcal{U}_\varepsilon(x_0). \tag{3.22}$$

We obtain

$$\frac{R(x)}{(x - x_0)^{n+1}} = \frac{R(x) - R(x_0)}{(x - x_0)^{n+1} - 0}.$$

By the generalized mean-value theorem, it follows that there exists ξ_1 between x_0 and x such that

$$\frac{R(x) - R(x_0)}{(x - x_0)^{n+1} - 0} = \frac{R'(\xi_1)}{(n+1)(\xi_1 - x_0)^n}.$$

Again, by definition of R,

$$\frac{R'(\xi_1)}{(n+1)(\xi_1 - x_0)^n} = \frac{R'(\xi_1) - R'(x_0)}{(n+1)(\xi_1 - x_0)^n - 0},$$

and by the generalized mean-value theorem, there exists ξ_2 between x_0 and ξ_1 such that

$$\frac{R'(\xi_1) - R'(x_0)}{(n+1)(\xi_1 - x_0)^n - 0} = \frac{R''(\xi_2)}{n(n+1)(\xi_2 - x_0)^{n-1}}.$$

By repeated application of this argument, we eventually obtain

$$\frac{R^{(n)}(\xi_n) - R^{(n)}(x_0)}{(n+1)!(\xi_n - x_0) - 0} = \frac{R^{(n+1)}(\xi)}{(n+1)!}$$

where ξ is between x_0 and ξ_n. Combining all these equalities, we obtain

$$\frac{R(x)}{(x - x_0)^{n+1}} = \frac{R^{(n+1)}(\xi)}{(n+1)!}$$

and, using (3.22),

$$R(x) = \frac{f^{(n+1)}(\xi)(x - x_0)^{n+1}}{(n+1)!}. \quad \|$$

Theorem 3.3.13 says that, in a neighborhood of a point x_0, the value of $f(x)$ can be approximated with a polynomial, where $R(x)$—the *remainder*—describes the error that is made in this approximation. The polynomial

$$f(x_0) + \sum_{k=1}^{n} \frac{f^{(k)}(x_0)(x - x_0)^k}{k!}$$

is called a *Taylor polynomial of order n around x_0*. The approximation is a "good" approximation if the error term $R(x)$ is "small". If $\lim_{n\to\infty} R(x) = 0$, the approximation approaches the true value of $f(x)$ as

n approaches infinity. If $R(x)$ approaches zero as n approaches infinity and the n^{th}-order derivative of f at x_0 exists for all $n \in \mathbb{N}$, it follows that

$$f(x) = f(x_0) + \sum_{k=1}^{\infty} \frac{f^{(k)}(x_0)(x - x_0)^k}{k!} \tag{3.23}$$

for all $x \in \mathcal{U}_\varepsilon(x_0)$. (3.23) is a *Taylor series expansion of f around x_0*.

Consider the following example. Let

$$f : \mathbb{R} \mapsto \mathbb{R},\ x \mapsto x^3 + 3x^2 + 2.$$

To find Taylor polynomials of f around $x_0 = 1$, note first that we have $f'(x) = 3x^2 + 6x$, $f''(x) = 6x + 6$, $f^{(3)}(x) = 6$, and $f^{(n)}(x) = 0$ for all $n \geq 4$, for all $x \in \mathbb{R}$. Therefore, $f(1) = 6$, $f'(1) = 9$, $f''(1) = 12$, $f^{(3)}(1) = 6$, and $f^{(n)}(1) = 0$ for all $n \geq 4$. Furthermore, $R(x) = 0$ for all $x \in \mathbb{R}$ and all $n \geq 3$. The Taylor polynomial of order one around $x_0 = 1$ is

$$f(x_0) + f'(x_0)(x - x_0) = 6 + 9(x - 1),$$

and the Taylor polynomial of order two is

$$f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)(x - x_0)^2}{2} = 6 + 9(x - 1) + 6(x - 1)^2.$$

For $n \geq 3$, the Taylor polynomial of order n gives the exact value of f at x because, in this case, $R(x) = 0$.

We now return to the problem of finding maxima and minima of functions. The following theorem provides *sufficient* conditions for maxima and minima of differentiable functions.

Theorem 3.3.14 *Let $A \subseteq \mathbb{R}$ be an interval, let $f : A \mapsto \mathbb{R}$, and let $x_0 \in A$. Suppose there exists $\varepsilon \in \mathbb{R}_{++}$ such that f is continuous on $\mathcal{U}_\varepsilon(x_0) \cap A$ and differentiable on $\mathcal{U}_\varepsilon(x_0) \cap A \setminus \{x_0\}$.*

(i) If $f'(x) \geq 0$ for all $x \in (x_0 - \varepsilon, x_0) \cap A$ and $f'(x) \leq 0$ for all $x \in (x_0, x_0 + \varepsilon) \cap A$, then f has a local maximum at x_0.
(ii) If $f'(x) \leq 0$ for all $x \in (x_0 - \varepsilon, x_0) \cap A$ and $f'(x) \geq 0$ for all $x \in (x_0, x_0 + \varepsilon) \cap A$, then f has a local minimum at x_0.

Proof. (i) First, we show

$$f(x_0) \geq f(x) \quad \forall x \in (x_0 - \varepsilon, x_0) \cap A. \tag{3.24}$$

If $(x_0 - \varepsilon, x_0) \cap A = \emptyset$, (3.24) is trivially true. Now let $x \in (x_0 - \varepsilon, x_0) \cap A$. By the mean-value theorem, there exists $y \in (x, x_0)$ such that

$$\frac{f(x_0) - f(x)}{x_0 - x} = f'(y) \geq 0,$$

and therefore, because $x_0 > x$, $f(x_0) - f(x) \geq 0$, which completes the proof of (3.24).

Next, we prove

$$f(x_0) \geq f(x) \quad \forall x \in (x_0, x_0 + \varepsilon) \cap A. \tag{3.25}$$

If $(x_0, x_0 + \varepsilon) \cap A = \emptyset$, (3.25) is trivially true. Now let $x \in (x_0, x_0 + \varepsilon) \cap A$. By the mean-value theorem, there exists $y \in (x_0, x)$ such that

$$\frac{f(x) - f(x_0)}{x - x_0} = f'(y) \leq 0,$$

and therefore, because $x > x_0$, $f(x) - f(x_0) \leq 0$. This proves (3.25). Combining (3.24) and (3.25), the proof of (i) is complete.

Part (ii) is proven analogously. $\|$

Note that Theorem 3.3.14 only requires f to be continuous at x_0—f need not be differentiable at x_0. Therefore, this result can be applied to functions such as

$$f : \mathbb{R} \mapsto \mathbb{R},\ x \mapsto |x|.$$

This function is not differentiable at $x_0 = 0$, but it is continuous at $x_0 = 0$ and differentiable at all points in $\mathbb{R} \setminus \{0\}$. We obtain

$$f'(x) = \begin{cases} 1 & \text{if } x > 0 \\ -1 & \text{if } x < 0, \end{cases}$$

and therefore, according to Theorem 3.3.14, f has a local minimum at $x_0 = 0$.

For boundary points of A, the following theorem provides further sufficient conditions for maxima and minima.

Theorem 3.3.15 *Let $A = [a, b]$, where $a, b \in \mathbb{R}$ and $a < b$. Furthermore, let $f : A \mapsto \mathbb{R}$.*

(i) f is right-side differentiable at a and $f'(a) < 0 \Rightarrow f$ has a local maximum at a.
(ii) f is right-side differentiable at a and $f'(a) > 0 \Rightarrow f$ has a local minimum at a.
(iii) f is left-side differentiable at b and $f'(b) > 0 \Rightarrow f$ has a local maximum at b.
(iv) f is left-side differentiable at b and $f'(b) < 0 \Rightarrow f$ has a local minimum at b.

Proof. (i) Suppose

$$f'(a) = \lim_{h \downarrow 0} \frac{f(a+h) - f(a)}{h} < 0.$$

Then there exists $\varepsilon \in \mathbb{R}_{++}$ such that

$$\frac{f(a+h) - f(a)}{h} < 0 \quad \forall h \in (0, \varepsilon).$$

Because $h > 0$, this implies $f(a+h) - f(a) < 0$ for all $h \in (0, \varepsilon)$, and therefore,

$$f(x) < f(a) \quad \forall x \in (a, a+\varepsilon),$$

which implies that f has a local maximum at a.

The proofs of parts (ii) to (iv) are analogous. ‖

If f is twice differentiable at a point x_0, the following conditions for local maxima and minima at interior points can be established. We already know that at a local interior maximum (minimum) of a differentiable function, the derivative of this function must be equal to zero. If we combine this condition with a condition involving the *second* derivative of f, we obtain another set of *sufficient* conditions for local maxima and minima.

Theorem 3.3.16 *Let $A \subseteq \mathbb{R}$ be an interval, let $f : A \mapsto \mathbb{R}$, and let x_0 be an interior point of A. Suppose f is twice differentiable at x_0, and there exists $\varepsilon \in \mathbb{R}_{++}$ such that f is differentiable on $\mathcal{U}_\varepsilon(x_0)$.*

(i) $f'(x_0) = 0 \wedge f''(x_0) < 0 \Rightarrow f$ has a local maximum at x_0.
(ii) $f'(x_0) = 0 \wedge f''(x_0) > 0 \Rightarrow f$ has a local minimum at x_0.

Proof. (i) Let $f'(x_0) = 0$ and $f''(x_0) < 0$. Then

$$0 > f''(x_0) = \lim_{h \to 0} \frac{f'(x_0+h) - f'(x_0)}{h} = \lim_{h \to 0} \frac{f'(x_0+h)}{h}.$$

Therefore, there exists $\delta \in (0, \varepsilon)$ such that, with $x = x_0 + h$,

$$f'(x) > 0 \quad \forall x \in (x_0 - \delta, x_0)$$

and

$$f'(x) < 0 \quad \forall x \in (x_0, x_0 + \delta).$$

By Theorem 3.3.14, this implies that f has a local maximum at x_0.

The proof of part (ii) is analogous. ‖

Analogous sufficient conditions for maxima and minima at boundary points can be formulated.

Theorem 3.3.17 *Let $A = [a, b]$, where $a, b \in \mathbb{R}$ and $a < b$. Furthermore, let $f : A \mapsto \mathbb{R}$.*

(i) Suppose f is twice right-side differentiable at a, and there exists $\varepsilon \in \mathbb{R}_{++}$ such that f is differentiable on $(a, a + \varepsilon)$. If $f'(a) \leq 0$ and $f''(a) < 0$, then f has a local maximum at a.

(ii) Suppose f is twice left-side differentiable at b, and there exists $\varepsilon \in \mathbb{R}_{++}$ such that f is differentiable on $(b - \varepsilon, b)$. If $f'(b) \geq 0$ and $f''(b) < 0$, then f has a local maximum at b.

(iii) Suppose f is twice right-side differentiable at a, and there exists $\varepsilon \in \mathbb{R}_{++}$ such that f is differentiable on $(a, a + \varepsilon)$. If $f'(a) \geq 0$ and $f''(a) > 0$, then f has a local minimum at a.

(iv) Suppose f is twice left-side differentiable at b, and there exists $\varepsilon \in \mathbb{R}_{++}$ such that f is differentiable on $(b - \varepsilon, b)$. If $f'(b) \leq 0$ and $f''(b) > 0$, then f has a local minimum at b.

The proof of Theorem 3.3.17 is analogous to the proof of Theorem 3.3.16 and is left as an exercise.

The above theorems provide *sufficient* conditions for local maxima and minima. These conditions are *not* necessary. For example, consider

$$f : \mathbb{R} \mapsto \mathbb{R},\ x \mapsto x^4.$$

The function f has a local minimum at $x_0 = 0$, but $f''(0) = 0$.

The conditions formulated in Theorems 3.3.16 and 3.3.17 involving the *second* derivatives of f are called sufficient *second-order* conditions.

For maxima and minima at *interior* points, we can also formulate *necessary* second-order conditions.

Theorem 3.3.18 *Let $A \subseteq \mathbb{R}$ be an interval, let $f : A \mapsto \mathbb{R}$, and let x_0 be an interior point of A. Suppose f is twice differentiable at x_0, and there exists $\varepsilon \in \mathbb{R}_{++}$ such that f is differentiable on $\mathcal{U}_\varepsilon(x_0)$.*

(i) f has a local maximum at $x_0 \Rightarrow f''(x_0) \leq 0$.
(ii) f has a local minimum at $x_0 \Rightarrow f''(x_0) \geq 0$.

Proof. (i) Suppose f has a local maximum at an interior point x_0. Then there exists $\varepsilon \in \mathbb{R}_{++}$ such that

$$f(x_0) \geq f(x) \quad \forall x \in \mathcal{U}_\varepsilon(x_0).$$

By way of contradiction, suppose $f''(x_0) > 0$. Therefore,

$$\lim_{h \downarrow 0} \frac{f'(x_0 + h) - f'(x_0)}{h} > 0.$$

This implies that we can find $\delta \in (0, \varepsilon)$ such that

$$f'(x) > f'(x_0) \quad \forall x \in (x_0, x_0 + \delta).$$

Because f has a local maximum at the interior point x_0, it follows that $f'(x_0) = 0$, and therefore, we obtain

$$f'(x) > 0 \quad \forall x \in (x_0, x_0 + \delta). \tag{3.26}$$

Let $x \in (x_0, x_0 + \delta)$. By the mean-value theorem, there exists $y \in (x_0, x)$ such that

$$\frac{f(x) - f(x_0)}{x - x_0} = f'(y).$$

By (3.26), $f'(y) > 0$, and because $x > x_0$, we obtain $f(x) > f(x_0)$, contradicting the assumption that f has a local maximum at x_0.

The proof of (ii) is analogous. ‖

Theorem 3.3.18 applies to *interior* points of A only. For boundary points, the second derivative being nonpositive (nonnegative) is *not* necessary for a maximum (minimum). For example, consider the function

$$f : [0, 1] \mapsto \mathbb{R},\ x \mapsto x^2.$$

This function has a local maximum at $x_0 = 1$, but $f''(1) = 2 > 0$.

To conclude this section, we work through an example for finding local maxima and minima. Define

$$f : [0, 3] \mapsto \mathbb{R},\ x \mapsto (x - 1)^2.$$

f is twice differentiable on its domain $A = [0, 3]$, and we obtain

$$f'(x) = 2(x-1),\ f''(x) = 2 \quad \forall x \in A.$$

Because $f'(0) = -2 < 0$, f has a local maximum at the boundary point 0 (see Theorem 3.3.15). Similarly, because $f'(3) = 4 > 0$, f has a local maximum at the boundary point 3. Furthermore, the only possibility for a local maximum or minimum at an interior point is at $x_0 = 1$, because this is the only interior point x_0 such that $f'(x_0) = 0$. The function f has a local minimum at $x_0 = 1$, because $f''(1) = 2 > 0$.

Because the domain of f is a closed and bounded interval, f must have a global maximum and a global minimum (see Theorem 3.3.4). The global minimum must be at $x_0 = 1$, because this is the only local minimum, and a global minimum must be a local minimum. There are two local maxima of f. To find out at which of these f has a global maximum, we calculate the values of f at the corresponding points. We obtain $f(0) = 1$ and $f(3) = 4$, and therefore, the function f has a global maximum at $x_0 = 3$.

3.4 Concave and Convex Functions

The conditions introduced in the previous section provide criteria for *local* maxima and minima of functions. It would, of course, be desirable to have conditions that allow us to determine whether a point is a *global* maximum or minimum. This can be done with the help of *concavity* and *convexity* of functions. The definition of these properties is

Definition 3.4.1 *Let $A \subseteq I\!R$ be an interval, and let $f : A \mapsto I\!R$.*

(i) f is concave *if and only if*

$$f(\lambda x + (1-\lambda)y) \geq \lambda f(x) + (1-\lambda)f(y) \quad \forall x, y \in A \text{ such that } x \neq y,\ \forall \lambda \in (0,1).$$

(ii) f is strictly concave *if and only if*

$$f(\lambda x + (1-\lambda)y) > \lambda f(x) + (1-\lambda)f(y) \quad \forall x, y \in A \text{ such that } x \neq y,\ \forall \lambda \in (0,1).$$

(iii) f is convex *if and only if*

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y) \quad \forall x, y \in A \text{ such that } x \neq y,\ \forall \lambda \in (0,1).$$

(iv) f is strictly convex *if and only if*

$$f(\lambda x + (1-\lambda)y) < \lambda f(x) + (1-\lambda)f(y) \quad \forall x, y \in A \text{ such that } x \neq y,\ \forall \lambda \in (0,1).$$

Be careful to distinguish between the convexity of a *set* (see Chapter 1) and the convexity of a *function* as defined above. Note that the assumption that A is an interval is very important in this definition. Because intervals are convex sets, this assumption ensures that f is defined at points $\lambda x + (1-\lambda)y$ for $x, y \in A$ and $\lambda \in (0, 1)$.

Figure 3.11 gives a diagrammatic illustration of the graph of a concave function.

Concavity requires that the value of f *at a convex combination* of two points $x, y \in A$ is greater than or equal to the corresponding *convex combination of the values of* f at these two points. Geometrically, concavity says that the line segment joining the points $(x, f(x))$ and $(y, f(y))$ can never be *above* the graph of the function.

Convexity has an analogous interpretation—see Figure 3.12 for an illustration of the graph of a convex function.

Clearly, a function f is concave if and only if the function $-f$ is convex (Exercise: prove this).

Concavity and convexity are *global* properties of a function, because they apply to the whole domain A rather than a neighborhood of a point only.

There are several necessary and sufficient conditions for concavity and convexity that will be useful for our purposes. First, we show

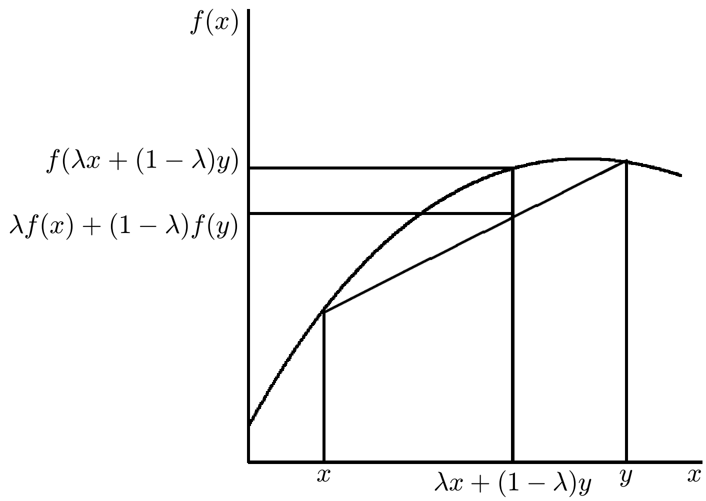


Figure 3.11: A concave function.

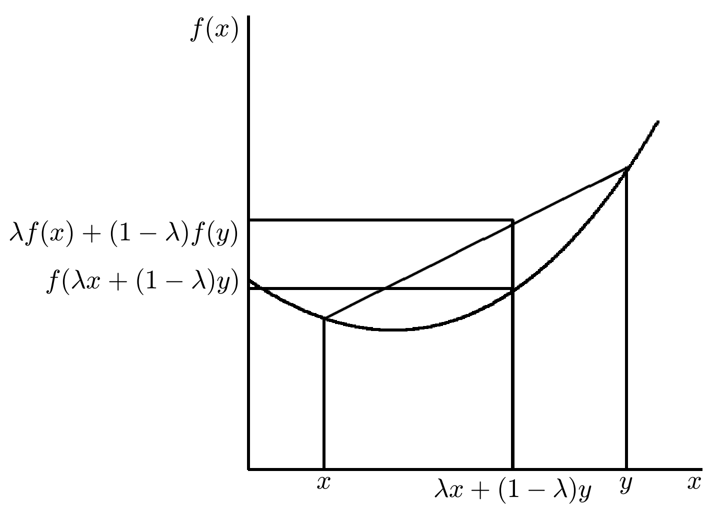


Figure 3.12: A convex function.

Theorem 3.4.2 *Let $A \subseteq I\!R$ be an interval, and let $f : A \mapsto I\!R$.*

(i) f is concave if and only if

$$\frac{f(y)-f(x)}{y-x} \geq \frac{f(z)-f(y)}{z-y} \quad \forall x,y,z \in A \text{ such that } x<y<z.$$

(ii) f is strictly concave if and only if

$$\frac{f(y)-f(x)}{y-x} > \frac{f(z)-f(y)}{z-y} \quad \forall x,y,z \in A \text{ such that } x<y<z.$$

(iii) f is convex if and only if

$$\frac{f(y)-f(x)}{y-x} \leq \frac{f(z)-f(y)}{z-y} \quad \forall x,y,z \in A \text{ such that } x<y<z.$$

(iv) f is strictly convex if and only if

$$\frac{f(y)-f(x)}{y-x} < \frac{f(z)-f(y)}{z-y} \quad \forall x,y,z \in A \text{ such that } x<y<z.$$

Proof. (i) "Only if": Suppose f is concave. Let $x,y,z \in A$ be such that $x<y<z$. Define

$$\lambda := \frac{z-y}{z-x}.$$

Then it follows that

$$1-\lambda = \frac{y-x}{z-x},$$

and $y = \lambda x + (1-\lambda)z$. Because f is concave,

$$f(y) \geq \frac{z-y}{z-x} f(x) + \frac{y-x}{z-x} f(z). \tag{3.27}$$

This is equivalent to

$$\begin{aligned} & (z-x)f(y) \geq (z-y)f(x) + (y-x)f(z) \\ \Leftrightarrow\ & (z-y)(f(y)-f(x)) \geq (y-x)(f(z)-f(y)) \\ \Leftrightarrow\ & \frac{f(y)-f(x)}{y-x} \geq \frac{f(z)-f(y)}{z-y}. \end{aligned} \tag{3.28}$$

"If": Let $\lambda \in (0,1)$ and $x,z \in A$ be such that $x \neq z$. Define $y := \lambda x + (1-\lambda)z$. Then we have

$$\frac{z-y}{z-x} = \lambda \ \wedge \ \frac{y-x}{z-x} = 1-\lambda.$$

As was shown above, (3.27) is equivalent to (3.28), and therefore,

$$f(\lambda x + (1-\lambda)z) \geq \lambda f(x) + (1-\lambda) f(z).$$

The proofs of (ii) to (iv) are analogous. ‖

A diagrammatic illustration of Theorem 3.4.2 is given in Figure 3.13.

The slope of the line segment joining $(x, f(x))$ and $(y, f(y))$ is greater than or equal to the slope of the line segment joining $(y, f(y))$ and $(z, f(z))$ whenever $x<y<z$ for a concave function f. The interpretation of parts (ii) to (iv) of the theorem is analogous.

For differentiable functions, there are some convenient criteria for concavity and convexity.

Theorem 3.4.3 *Let $A \subseteq I\!R$ be an interval, and let $f : A \mapsto I\!R$ be differentiable.*

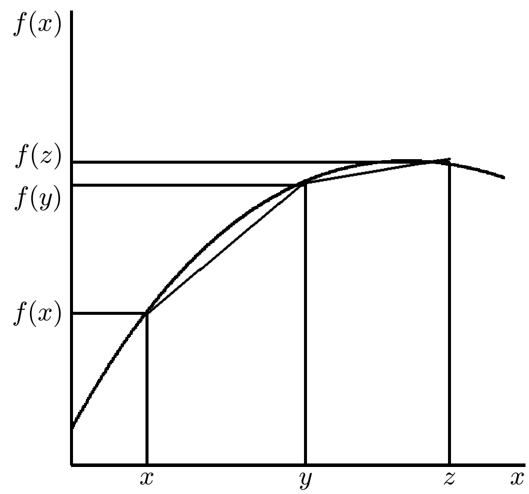


Figure 3.13: The secant inequality.

(i) f is concave $\Leftrightarrow$ f' is nonincreasing.
(ii) f is strictly concave $\Leftrightarrow$ f' is decreasing.
(iii) f is convex $\Leftrightarrow$ f' is nondecreasing.
(iv) f is strictly convex $\Leftrightarrow$ f' is increasing.

Proof. (i) "$\Rightarrow$": Suppose f is concave. Let $x, y \in A$ be such that $x < y$. Let $x_0 \in (x, y)$. For $k \in (0, x_0 - x)$ and $h \in (x_0 - y, 0)$, we have

$$x < x + k < x_0 < y + h < y.$$

By Theorem 3.4.2,

$$\frac{f(x+k) - f(x)}{k} \geq \frac{f(x_0) - f(x+k)}{x_0 - x - k} \geq \frac{f(y+h) - f(x_0)}{y + h - x_0} \geq \frac{f(y+h) - f(y)}{h}. \tag{3.29}$$

Because f is differentiable, taking limits as k and h approach zero yields

$$f'(x) = \lim_{k \downarrow 0} \frac{f(x+k) - f(x)}{k} \geq \frac{f(x_0) - f(x)}{x_0 - x} \geq \frac{f(y) - f(x_0)}{y - x_0} \geq \lim_{h \uparrow 0} \frac{f(y+h) - f(y)}{h} = f'(y), \tag{3.30}$$

and therefore, $f'(x) \geq f'(y)$, which proves that f' is nonincreasing.

"$\Leftarrow$": Suppose f' is nonincreasing. Let $x, y, z \in A$ be such that $x < y < z$. By the mean-value theorem, there exist $x_0 \in (x, y)$ such that

$$f'(x_0) = \frac{f(y) - f(x)}{y - x}$$

and $y_0 \in (y, z)$ such that

$$f'(y_0) = \frac{f(z) - f(y)}{z - y}.$$

Because $x_0 < y_0$, nonincreasingness of f' implies

$$f'(x_0) \geq f'(y_0), \tag{3.31}$$

and therefore,

$$\frac{f(y) - f(x)}{y - x} \geq \frac{f(z) - f(y)}{z - y}, \tag{3.32}$$

which, by Theorem 3.4.2, shows that f is concave.

(ii) The proof of (ii) is similar to the proof of part (i). To prove "$\Rightarrow$", note that in the case of strict concavity, the inequalities in (3.29) are strict, and therefore, (3.30) becomes

$$f'(x) = \lim_{k \downarrow 0} \frac{f(x+k) - f(x)}{k} \geq \frac{f(x_0) - f(x)}{x_0 - x} > \frac{f(y) - f(x_0)}{y - x_0} \geq \lim_{h \uparrow 0} \frac{f(y+h) - f(y)}{h} = f'(y),$$

which implies $f'(x) > f'(y)$.

To establish "$\Leftarrow$", note that decreasingness of f' implies that (3.31) is replaced by

$$f'(x_0) > f'(y_0),$$

and hence, the weak inequality in (3.32) is replaced by a strict inequality, proving that f is strictly concave.

The proofs of (iii) and (iv) are analogous. $\|$

If f is *twice* differentiable, Theorem 3.4.3 can be combined with Theorem 3.3.11 to obtain

Theorem 3.4.4 *Let $A \subseteq I\!R$ be an interval, and let $f : A \mapsto I\!R$ be twice differentiable.*

(i) $f''(x) \leq 0 \;\; \forall x \in A \;\Leftrightarrow\; f$ is concave.
(ii) $f''(x) < 0 \;\; \forall x \in A \;\Rightarrow\; f$ is strictly concave.
(iii) $f''(x) \geq 0 \;\; \forall x \in A \;\Leftrightarrow\; f$ is convex.
(iv) $f''(x) > 0 \;\; \forall x \in A \;\Rightarrow\; f$ is strictly convex.

As in Theorem 3.3.11, it is important to note that the reverse implications in parts (ii) and (iv) of Theorem 3.4.4 are *not* true.

For concave functions, any local maximum must also be a global maximum. Similarly, a local minimum of a convex function must be a global minimum.

Theorem 3.4.5 *Let $A \subseteq I\!R$ be an interval, and let $f : A \mapsto I\!R$. Furthermore, let $x_0 \in A$.*

(i) If f is concave and f has a local maximum at x_0, then f has a global maximum at x_0.
(ii) If f is convex and f has a local minimum at x_0, then f has a global minimum at x_0.

Proof. (i) Suppose f is concave and has a local maximum at $x_0 \in A$. By way of contradiction, suppose f does not have a global maximum at x_0. Then there exists $y \in A$ such that

$$f(y) > f(x_0). \tag{3.33}$$

Because f has a local maximum at x_0, there exists $\varepsilon \in I\!R_{++}$ such that

$$f(x_0) \geq f(x) \quad \forall x \in \mathcal{U}_\varepsilon(x_0) \cap A. \tag{3.34}$$

Clearly, there exists $\lambda \in (0,1)$ such that

$$z := \lambda y + (1-\lambda)x_0 \in \mathcal{U}_\varepsilon(x_0) \cap A \tag{3.35}$$

(just choose λ sufficiently close to zero for given ε). By concavity of f,

$$f(z) \geq \lambda f(y) + (1-\lambda)f(x_0),$$

and by (3.33),

$$\lambda f(y) + (1-\lambda)f(x_0) > f(x_0),$$

and therefore, $f(z) > f(x_0)$. Because of (3.35), this is a contradiction to (3.34).

The proof of (ii) is analogous. $\|$

If a function f is *strictly* concave and has a (local and global) maximum, this maximum is *unique.* Similarly, a strictly convex function has at most one (local and global) minimum.

Theorem 3.4.6 *Let $A \subseteq I\!R$ be an interval, and let $f : A \mapsto I\!R$.*

(i) If f is strictly concave, then f has at most one local (and global) maximum.
(ii) If f is strictly convex, then f has at most one local (and global) minimum.

Proof. (i) Suppose f is strictly concave, and therefore concave. By Theorem 3.4.5, all local maxima of f are global maxima. By way of contradiction, suppose f has two local, and therefore, global maxima at $x \in A$ and at $y \in A$, where $x \neq y$. By definition of a global maximum, we must have $f(x) \geq f(y)$ and $f(y) \geq f(x)$, and therefore, $f(x) = f(y)$. Let $z := \lambda x + (1-\lambda)y$ for some $\lambda \in (0,1)$. Clearly, $z \in A$. Because f is strictly concave, it follows that $f(z) > \lambda f(x) + (1-\lambda)f(y) = f(x)$, contradicting the assumption that f has a maximum at x.

The proof of (ii) is analogous. ‖

For differentiable functions, we can use concavity and convexity to state sufficient conditions for *global* maxima and minima.

Theorem 3.4.7 *Let $A \subseteq \mathbb{R}$ be an interval, and let $f : A \mapsto \mathbb{R}$. Furthermore, let x_0 be an interior point of A, and suppose f is differentiable at x_0.*

(i) $f'(x_0) = 0 \wedge f$ is concave $\Rightarrow f$ has a global maximum at x_0.
(ii) $f'(x_0) = 0 \wedge f$ is convex $\Rightarrow f$ has a global minimum at x_0.

Proof. (i) Suppose $f'(x_0) = 0$ and f is concave. Let $y \in A$, $y \neq x_0$. We have to show $f(x_0) \geq f(y)$.

First, suppose $y > x_0$. Because f is concave,

$$f(\lambda y + (1-\lambda)x_0) \geq \lambda f(y) + (1-\lambda)f(x_0) \quad \forall \lambda \in (0,1).$$

This is equivalent to

$$f(x_0 + \lambda(y - x_0)) - f(x_0) \geq \lambda(f(y) - f(x_0)) \quad \forall \lambda \in (0,1).$$

Because $y > x_0$,

$$\frac{f(x_0 + \lambda(y - x_0)) - f(x_0)}{\lambda(y - x_0)} \geq \frac{f(y) - f(x_0)}{y - x_0} \quad \forall \lambda \in (0,1).$$

Defining $h := \lambda(y - x_0)$, this can be written as

$$\frac{f(x_0 + h) - f(x_0)}{h} \geq \frac{f(y) - f(x_0)}{y - x_0} \quad \forall h \in (0, y - x_0).$$

Because f is differentiable at x_0, we obtain

$$f'(x_0) = \lim_{h \downarrow 0} \frac{f(x_0 + h) - f(x_0)}{h} \geq \frac{f(y) - f(x_0)}{y - x_0}.$$

Because $f'(x_0) = 0$ and $y > x_0$, this implies $f(y) - f(x_0) \leq 0$, which proves that $f(x_0) \geq f(y)$ for all $y \in A$ such that $y > x_0$. The case $y < x_0$ is proven analogously (prove that as an exercise).

The proof of (ii) is analogous. ‖

Theorem 3.4.7 applies to interior points only, but it is easy to generalize the result to boundary points. We obtain

Theorem 3.4.8 *Let $A = [a, b]$, where $a, b \in \mathbb{R}$ and $a < b$, and let $f : A \mapsto \mathbb{R}$.*

(i) If f is right-side differentiable at a, $f'(a) \leq 0$, and f is concave, then f has a global maximum at a.
(ii) If f is left-side differentiable at b, $f'(b) \geq 0$, and f is concave, then f has a global maximum at b.
(iii) If f is right-side differentiable at a, $f'(a) \geq 0$, and f is convex, then f has a global minimum at a.
(iv) If f is left-side differentiable at b, $f'(b) \leq 0$, and f is convex, then f has a global minimum at b.

The proof of Theorem 3.4.8 is analogous to the proof of Theorem 3.4.7 and is left as an exercise.

As an economic example for the application of optimization techniques, consider the following *profit maximization* problem of a *competitive firm*. Suppose a firm produces a good which is sold in a competitive market. The *cost function* of this firm is a function

$$C : \mathbb{R}_+ \mapsto \mathbb{R},\ y \mapsto C(y),$$

where $C(y)$ is the cost of producing $y \in \mathbb{R}_+$ units of output. The market price for the good produced by the firm is $p \in \mathbb{R}_{++}$. The assumption that the market is a *perfectly competitive* market implies that the firm has to take the market price as *given* and can only choose the quantity of output to be produced and sold.

Suppose the firm wants to maximize *profit*, that is, the difference between *revenue* and *cost*. If the firm produces and sells $y \in \mathbb{R}_+$ units of output, the revenue of the firm is py, and the cost is $C(y)$. Therefore, profit maximization means that the firm wants to maximize the function

$$\pi : \mathbb{R}_+ \mapsto \mathbb{R},\ y \mapsto py - C(y).$$

We denote this maximization problem by

$$\max_y \{py - C(y)\},$$

where the subscript indicates that y is the *choice variable* of the firm. Note that the price p is a *parameter* which cannot be chosen by the firm.

Suppose C is differentiable, which implies that π is differentiable. If C is convex, π is concave, and therefore, if a critical point can be found, the function π must have a global maximum at this point.

For example, let $p = 1$ and consider the cost function $C : \mathbb{R}_+ \mapsto \mathbb{R},\ y \mapsto y^2$. We obtain the profit maximization problem $\max_y\{y - y^2\}$. Because $\pi''(y) = -2 < 0$ for all $y \in \mathbb{R}_+$, π is strictly concave. If π has an interior maximum at a point $y_0 \in \mathbb{R}_{++}$, we must have $\pi'(y_0) = 0$, which is equivalent to $1 - 2y_0 = 0$, and therefore, $y_0 = 1/2$. By Theorem 3.4.6, π has a unique global maximum at $y_0 = 1/2$, and the maximal value of π is $\pi(1/2) = 1/4$.

More generally, consider the maximization problem $\max_y\{py - y^2\}$. The same steps as in the previous example lead to the conclusion that π has a unique global maximum at $y_0 = p/2$ for any $p \in \mathbb{R}_{++}$. Note that this defines the solution y_0 as a *function*

$$\bar{y} : \mathbb{R}_{++} \mapsto \mathbb{R},\ p \mapsto \frac{p}{2}$$

of the market price. We call this function a *supply function*, because it answers the question how much of the good a firm with the above cost function would want to sell at the market price $p \in \mathbb{R}_{++}$. If we substitute $y_0 = \bar{y}(p)$ into the objective function, we obtain the *maximal possible profit* of the firm as a function of the market price. In this example, we obtain the maximal possible profit

$$\pi_0 = \pi(\bar{y}(p)) = \pi\left(\frac{p}{2}\right) = p\frac{p}{2} - C\left(\frac{p}{2}\right) = \frac{p^2}{2} - \left(\frac{p}{2}\right)^2 = \left(\frac{p}{2}\right)^2.$$

Therefore, the maximal possible profit of the firm is given by the function

$$\bar{\pi} : \mathbb{R}_{++} \mapsto \mathbb{R},\ p \mapsto \left(\frac{p}{2}\right)^2.$$

We call the function $\bar{\pi}$ a *profit function*.

Now suppose the cost function is $C : \mathbb{R}_+ \mapsto \mathbb{R},\ y \mapsto y$. We obtain the profit maximization problem $\max_y\{py - y\}$. The function to be maximized is concave, because $\pi''(y) = 0$ for all $y \in \mathbb{R}_{++}$. For an interior solution, the necessary first-order condition is

$$p - 1 = 0,$$

which means that we have an interior maximum at *any* $y_0 \in \mathbb{R}_{++}$ if $p = 1$. To see under which circumstances we can have a *boundary solution* at $y_0 = 0$, we have to use the condition

$$\pi'(0) \leq 0,$$

which, in this example, leads to $p - 1 \leq 0$. Therefore, π has a global maximum at *any* $y_0 \in \mathbb{R}_+$ if $p = 1$, π has a unique maximum at $y_0 = 0$ if $p < 1$, and π has no maximum if $p > 1$.

Chapter 4

Functions of Several Variables

4.1 Sequences of Vectors

As is the case for functions of one variable, limits of functions of several variables can be related to limits of sequences, where the elements of these sequences are *vectors*. This section provides an introduction to sequences in $\mathbb{R}^n$.

Definition 4.1.1 *Let $n \in I\!N$. A* sequence of n-dimensional vectors *is a function $a : I\!N \mapsto I\!R^n$, $m \mapsto a(m)$. To simplify notation, we will write a^m instead of $a(m)$ for $m \in I\!N$, and use $\{a^m\}$ to denote such a sequence.*

Note that we use *superscripts* to denote elements of a sequence rather than *subscripts*. This is done to avoid confusion with *components* of vectors.

The notion of convergence of a sequence of vectors can be defined analogously to the convergence of sequences of real numbers, once we have a definition of *distance* in $\mathbb{R}^n$, which, in turn, can be used to define the notion of a *neighborhood* in $\mathbb{R}^n$.

There are several possibilities to define distances in $\mathbb{R}^n$ which could be used for our applications. One example is the *Euclidean* distance defined in Chapter 2. For our purposes, it will be convenient to use the following definition. For $n \in \mathbb{N}$ and $x, y \in \mathbb{R}^n$, define the distance between x and y as $d(x, y) := \max(\{|x_i - y_i| \mid i \in \{1, \ldots, n\}\})$. Clearly, for $n = 1$, we obtain the usual distance between real numbers. This distance function leads to the following definition of a neighborhood.

Definition 4.1.2 *Let $n \in I\!N$. For $x^0 \in I\!R^n$ and $\varepsilon \in I\!R_{++}$, the* ε-neighborhood of x^0 *is defined by*

$$\mathcal{U}_\varepsilon(x^0) := \left\{ x \in I\!R^n \,\middle|\, |x_i - x_i^0| < \varepsilon \;\; \forall i \in \{1, \ldots, n\} \right\}. \tag{4.1}$$

Figure 4.1 illustrates the ε-neighborhood of $x^0 = (1, 1)$ in $\mathbb{R}^2$ for $\varepsilon = 1/2$.

Once neighborhoods are defined, interior points, openness in $\mathbb{R}^n$, and closedness in $\mathbb{R}^n$ can be defined analogously to the corresponding definitions in Chapter 1—all that needs to be done is to replace $\mathbb{R}$ with $\mathbb{R}^n$ in these definitions. The corresponding formulations of these definitions are given below.

Definition 4.1.3 *Let $n \in I\!N$, and let $A \subseteq I\!R^n$. $x^0 \in A$ is an* interior point *of A if and only if there exists $\varepsilon \in I\!R_{++}$ such that $\mathcal{U}_\varepsilon(x^0) \subseteq A$.*

Definition 4.1.4 *Let $n \in I\!N$. A set $A \subseteq I\!R^n$ is* open in $\mathbb{R}^n$ *if and only if*

$$x \in A \Rightarrow x \text{ is interior point of } A.$$

Definition 4.1.5 *Let $n \in I\!N$. A set $A \subseteq I\!R^n$ is* closed in $\mathbb{R}^n$ *if and only if $\overline{A}$ is open in $I\!R^n$.*

Note that, as an alternative to Definition 4.1.2, we could define neighborhoods in terms of the Euclidean distance. This would lead to the following definition of an ε-neighborhood in $\mathbb{R}^n$.

$$\mathcal{U}_\varepsilon(x^0) := \left\{ x \in \mathbb{R}^n \,\middle|\, \sqrt{\sum_{i=1}^n (x_i - x_i^0)^2} < \varepsilon \right\}. \tag{4.2}$$

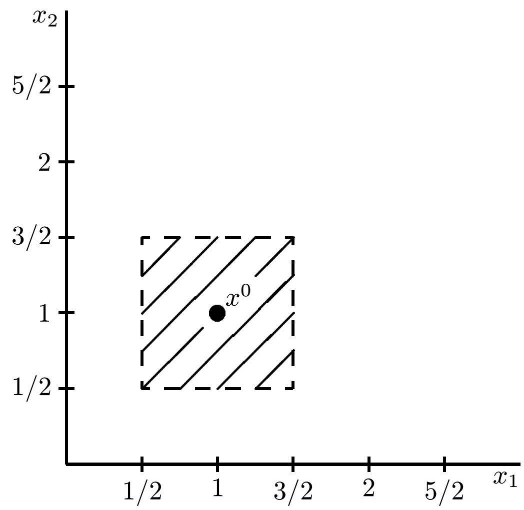


Figure 4.1: A neighborhood in $\mathbb{R}^2$.

Geometrically, a neighborhood $\mathcal{U}_\varepsilon(x^0)$ as defined in (4.2) can be represented as an open disc with center x^0 and radius ε for $n = 2$.

For our purposes, it does not matter whether (4.1) or (4.2) is used. This is the case because any set $A \subseteq \mathbb{R}^n$ is open according to the notion of a neighborhood defined in (4.1) if and only if A is open if we use (4.2) to define neighborhoods. That we use (4.1) rather than (4.2) is merely a matter of convenience, because some proofs are simplified by this choice.

We can now define *convergence* of sequences in $\mathbb{R}^n$.

Definition 4.1.6 *Let $n \in \mathbb{N}$.*

(i) A sequence of vectors $\{a^m\}$ converges to $\alpha \in \mathbb{R}^n$ *if and only if*

$$\forall \varepsilon \in \mathbb{R}_{++}, \exists m^0 \in \mathbb{N} \text{ such that } a^m \in \mathcal{U}_\varepsilon(\alpha)\ \forall m \geq m^0.$$

(ii) If $\{a^m\}$ converges to $\alpha \in \mathbb{R}^n$, α is the limit *of $\{a^m\}$, and we write*

$$\lim_{m\to\infty} a^m = \alpha.$$

The convergence of a sequence $\{a^m\}$ to $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{R}^n$ is *equivalent* to the convergence of the sequences of the *components* $\{a_i^m\}$ of $\{a^m\}$ to α_i for all $i \in \{1, \ldots, n\}$. This result—which is stated in the following theorem—considerably simplifies the task of checking sequences of vectors for convergence, because the problem can be reduced to the convergence of sequences of *real numbers*.

Theorem 4.1.7 *Let $n \in \mathbb{N}$, and let $\{a^m\}$ be a sequence in $\mathbb{R}^n$. If $\alpha \in \mathbb{R}^n$, then*

$$\lim_{m\to\infty} a^m = \alpha \Leftrightarrow \lim_{m\to\infty} a_i^m = \alpha_i\ \forall i = 1, \ldots, n.$$

Proof. "$\Rightarrow$": Suppose $\{a^m\}$ converges to $\alpha \in \mathbb{R}^n$. Let $\varepsilon \in \mathbb{R}_{++}$. By the definition of convergence, there exists $m^0 \in \mathbb{N}$ such that $a^m \in \mathcal{U}_\varepsilon(\alpha)$ for all $m \geq m^0$. By Definition 4.1.2, this is equivalent to

$$|a_i^m - \alpha_i| < \varepsilon \quad \forall m \geq m^0,\ \forall i = 1, \ldots, n,$$

which implies $a_i^m \in \mathcal{U}_\varepsilon(\alpha_i)$ for all $m \geq m^0$, for all $i = 1, \ldots, n$. This means that $\{a_i^m\}$ converges to α_i for all $i = 1, \ldots, n$.

"$\Leftarrow$": Suppose $\lim_{m\to\infty} a_i^m = \alpha_i$ for all $i = 1, \ldots, n$. This means that, for any $\varepsilon \in \mathbb{R}_{++}$, there exist $m_1^0, \ldots, m_n^0 \in \mathbb{N}$ such that

$$|a_i^m - \alpha_i| < \varepsilon \quad \forall m \geq m_i^0,\ \forall i = 1, \ldots, n.$$

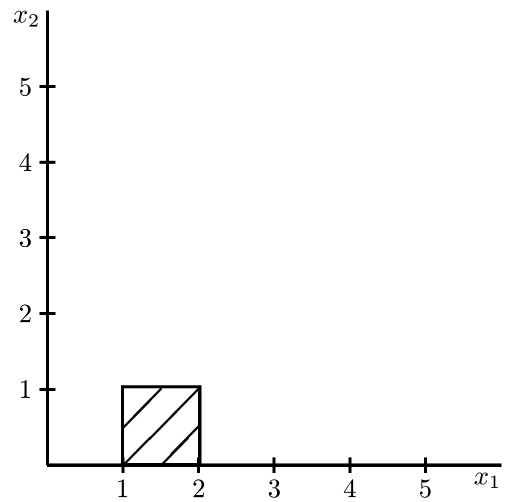


Figure 4.2: The set A.

Let $m^0 := \max(\{m_i^0 \mid i \in \{1, \ldots, n\}\})$. The we have

$$|a_i^m - \alpha_i| < \varepsilon \quad \forall m \geq m^0,\ \forall i = 1, \ldots, n,$$

which means that $\{a^m\}$ converges to α. $\|$

As an example, consider the sequence defined by

$$a^m = \left(\frac{1}{m}, 1 - \frac{1}{m}\right) \quad \forall m \in \mathbb{N}.$$

We obtain

$$\lim_{m\to\infty} a_1^m = \lim_{m\to\infty} \frac{1}{m} = 0$$

and

$$\lim_{m\to\infty} a_2^m = \lim_{m\to\infty} \left(1 - \frac{1}{m}\right) = 1,$$

and therefore, $\lim_{m\to\infty} a^m = (0, 1)$.

4.2 Continuity

A real-valued function of several variables is a function $f : A \mapsto B$, where $A \subseteq \mathbb{R}^n$ with $n \in \mathbb{N}$ and $B \subseteq \mathbb{R}$, that is, the domain of f is a set of n-dimensional vectors. Throughout this chapter, we assume that the domain A is a *convex* set, and the range B will usually be the set $\mathbb{R}$.

Convexity of a subset of $\mathbb{R}^n$ is defined analogously to the convexity of subsets of $\mathbb{R}$. We define

Definition 4.2.1 *Let $n \in \mathbb{N}$. A set $A \subseteq \mathbb{R}^n$ is* convex *if and only if*

$$[\lambda x + (1 - \lambda)y] \in A \quad \forall x, y \in A,\ \forall \lambda \in [0, 1].$$

Here are some examples for $n = 2$. Let $A = \{x \in \mathbb{R}^2 \mid (1 \leq x_1 \leq 2) \wedge (0 \leq x_2 \leq 1)\}$, $B = \{x \in \mathbb{R}^2 \mid (x_1)^2 + (x_2)^2 \leq 1\}$, $C = A \cup \{x \in \mathbb{R}^2 \mid (0 \leq x_1 \leq 2) \wedge (1 \leq x_2 \leq 2)\}$. The sets A, B, and C are illustrated in Figures 4.2, 4.3, and 4.4, respectively. As an exercise, show that A and B are convex, but C is not.

For functions of two variables, it is often convenient to give a graphical illustration by means of their *level sets.*

Definition 4.2.2 *Let $n \in \mathbb{N}$, let $A \subseteq \mathbb{R}^n$ be convex, and let $f : A \mapsto \mathbb{R}$ be a function. Furthermore, let $c \in f(A)$. The* level set *of f for c is the set $\{x \in A \mid f(x) = c\}$.*

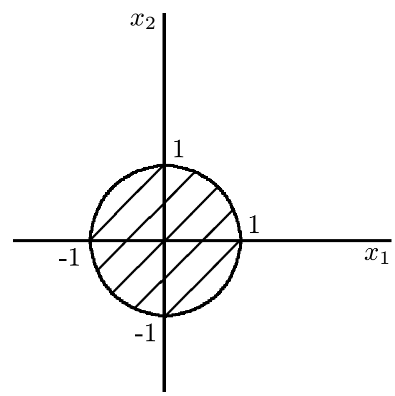


Figure 4.3: The set B.

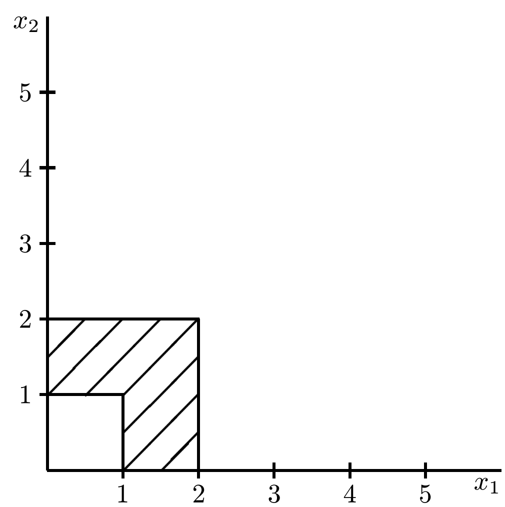


Figure 4.4: The set C.

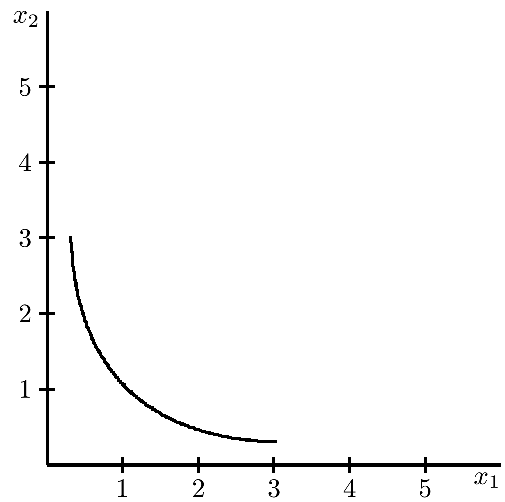


Figure 4.5: A level set.

Therefore, the level set of f for $c \in f(A)$ is the set of all points x in the domain of f such that the value of f at x is equal to c.

For example, let

$$f : \mathbb{R}^2_{++} \mapsto \mathbb{R},\ x \mapsto x_1 x_2.$$

We obtain $f(A) = f(\mathbb{R}^2_{++}) = \mathbb{R}_{++}$, and, for any $c \in \mathbb{R}_{++}$, the level set of f for c is given by $\{x \in \mathbb{R}^2_{++} \mid x_1 x_2 = c\}$. Figure 4.5 illustrates the level set of f for $c = 1$.

The definition of *continuity* of a function of several variables is analogous to the definition of continuity of a function of one variable.

Definition 4.2.3 *Let $n \in \mathbb{N}$, let $A \subseteq \mathbb{R}^n$ be convex, and let $f : A \mapsto \mathbb{R}$ be a function. Furthermore, let $x^0 \in A$.*

(i) The function f is continuous at x^0 *if and only if*

$$\forall \delta \in \mathbb{R}_{++}, \exists \varepsilon \in \mathbb{R}_{++} \text{ such that } f(x) \in \mathcal{U}_\delta(f(x^0))\ \forall x \in \mathcal{U}_\varepsilon(x^0) \cap A.$$

(ii) The function f is continuous on $B \subseteq A$ *if and only if f is continuous at each $x^0 \in B$. If f is continuous on A, we will often simply say that f is continuous.*

A useful criterion for the continuity of a function of several variables can be given in terms of convergent sequences.

Theorem 4.2.4 *Let $n \in \mathbb{N}$, let $A \subseteq \mathbb{R}^n$ be convex, and let $f : A \mapsto \mathbb{R}$ be a function. Furthermore, let $x^0 \in A$. f is continuous at x^0 if and only if, for all sequences $\{x^m\}$ such that $x^m \in A$ for all $m \in \mathbb{N}$,*

$$\lim_{m\to\infty} x^m = x^0 \Rightarrow \lim_{m\to\infty} f(x^m) = f(x^0).$$

The proof of Theorem 4.2.4 is analogous to the proof of Theorem 3.1.2.

Analogously to Theorems 3.1.8 and 3.1.9, we obtain

Theorem 4.2.5 *Let $n \in \mathbb{N}$, let $A \subseteq \mathbb{R}^n$ be convex, and let $f : A \mapsto \mathbb{R}$ and $g : A \mapsto \mathbb{R}$ be functions. Furthermore, let $\alpha \in \mathbb{R}$. If f and g are continuous at $x^0 \in A$, then*

(i) $f + g$ is continuous at x^0.
(ii) αf is continuous at x^0.
(iii) fg is continuous at x^0.
(iv) If $g(x) \neq 0$ for all $x \in A$, f/g is continuous at x^0.

Theorem 4.2.6 *Let $n \in \mathbb{N}$, and let $A \subseteq \mathbb{R}^n$ be convex. Furthermore, let $f : A \mapsto \mathbb{R}$ and $g : f(A) \mapsto \mathbb{R}$ be functions, and let $x^0 \in A$. If f is continuous at x^0 and g is continuous at $y^0 = f(x^0)$, then $g \circ f$ is continuous at x^0.*

Suppose f is a function of $n \geq 2$ variables. If we *fix* $n-1$ components of the vector of variables x, we can define a function of *one* variable. This function of one variable is continuous if the original function f is continuous.

Theorem 4.2.7 *Let $n \in \mathbb{N}$, let $A \subseteq \mathbb{R}^n$ be convex, and let $f : A \mapsto \mathbb{R}$ be a function. Furthermore, let $x^0 \in A$. For $i \in \{1, \ldots, n\}$, let*

$$A_i := \{y \in \mathbb{R} \mid \exists x \in A \text{ such that } x_i = y \ \wedge \ x_j = x_j^0 \ \forall j \in \{1, \ldots, n\} \setminus \{i\}\},$$

and define a function

$$f^i : A_i \mapsto \mathbb{R}, \ x_i \mapsto f(x_1^0, \ldots, x_{i-1}^0, x_i, x_{i+1}^0, \ldots, x_n^0). \tag{4.3}$$

If f is continuous at x^0, then f^i is continuous at x_i^0.

Proof. Suppose f is continuous at $x^0 \in A$. Let $i \in \{1, \ldots, n\}$, and let $\{x_i^m\}$ be a sequence of real numbers such that $x_i^m \in A_i$ for all $m \in \mathbb{N}$ and $\lim_{m\to\infty} x_i^m = x_i^0$. Furthermore, for $j \in \{1, \ldots, n\} \setminus \{i\}$, let $x_j^m = x_j^0$ for all $m \in \mathbb{N}$. Therefore, $\lim_{m\to\infty} x^m = x^0$, and, because f is continuous at x^0,

$$\lim_{m\to\infty} f(x^m) = f(x^0). \tag{4.4}$$

By definition, $f(x^m) = f^i(x_i^m)$ for all $m \in \mathbb{N}$, and $f(x^0) = f^i(x_i^0)$. Therefore, equation (4.4) implies $\lim_{m\to\infty} f^i(x_i^m) = f^i(x_i^0)$, which proves that f^i is continuous at x_i^0. $\|$

It is important to note that continuity of the restrictions of f to its i^{th} component as defined in (4.3) does *not* imply that f is continuous. It is possible that, for $x^0 \in A$, f^i is continuous at x_i^0 for all $i = 1, \ldots, n$, but f is *not* continuous at x^0. For example, consider the function

$$f : \mathbb{R}^2 \mapsto \mathbb{R}, \ x \mapsto \begin{cases} \frac{x_1 x_2}{(x_1)^2 + (x_2)^2} & \text{if } x \neq (0,0) \\ 0 & \text{if } x = (0,0). \end{cases}$$

Let $x^0 = (0,0)$. We obtain $f^1(x_1) = f(x_1, 0) = 0$ for all $x_1 \in \mathbb{R}$ and $f^2(x_2) = f(0, x_2) = 0$ for all $x_2 \in \mathbb{R}$. Therefore, f^1 is continuous at $x_1^0 = 0$ and f^2 is continuous at $x_2^0 = 0$. However, f is not continuous at $x^0 = (0,0)$. To see why this is the case, consider the sequence $\{x^m\}$ defined by

$$x^m = \left(\frac{1}{m}, \frac{1}{m}\right) \quad \forall m \in \mathbb{N}.$$

Clearly, $\lim_{m\to\infty} x^m = (0,0)$. We obtain

$$f(x^m) = \frac{\frac{1}{m}\frac{1}{m}}{\left(\frac{1}{m}\right)^2 + \left(\frac{1}{m}\right)^2} = \frac{1}{2} \quad \forall m \in \mathbb{N},$$

which implies

$$\lim_{m\to\infty} f(x^m) = \frac{1}{2} \neq 0 = f(0,0),$$

and therefore, f is not continuous at $x^0 = (0,0)$.

4.3 Differentiation

In Chapter 3, we introduced the derivative of a function of one variable in order to make statements about the change in the value of the function caused by a change in the argument of the function. If we consider functions of several variables, there are basically two possibilities of generalizing this approach.

The first possibility is to consider the effect of a change in *one* variable, keeping all other variables at a constant value. This leads to the concept of *partial* differentiation. Second, we could allow *all* variables to change and examine the effect on the value of the function, which is done by means of *total* differentiation.

Partial differentiation is very similar to the differentiation of functions of one variable.

Definition 4.3.1 *Let $n \in I\!N$, let $A \subseteq I\!R^n$ be convex, and let $f : A \mapsto I\!R$ be a function. Furthermore, let x^0 be an interior point of A, and let $i \in \{1, \ldots, n\}$. The function f is* partially differentiable with respect to x_i at x^0 *if and only if*

$$\lim_{h \to 0} \frac{f(x_1^0, \ldots, x_{i-1}^0, x_i^0 + h, x_{i+1}^0, \ldots, x_n^0) - f(x^0)}{h} \tag{4.5}$$

exists and is finite. If f is partially differentiable with respect to x_i at x^0, we call the limit in (4.5) the partial derivative of f with respect to x_i at x_0 *and denote it by $f_{x_i}(x^0)$ or*

$$\frac{\partial f(x^0)}{\partial x_i}.$$

For simplicity, we restrict attention to interior points of A in this chapter. Let $B \subseteq A$ be an open set. We say that f is *partially differentiable with respect to x_i on B* if and only if f is partially differentiable with respect to x_i at all points in B. If f is partially differentiable with respect to x_i on A, the function

$$f_{x_i} : A \mapsto \mathbb{R},\ x \mapsto f_{x_i}(x)$$

is called the *partial derivative of f with respect to x_i*.

Note that partial differentiation is very much like differentiation of a function of one variable, because the remaining $n-1$ variables are held *fixed*. Therefore, the same differentiation rules as introduced in Chapter 3 apply—we only have to treat all variables other than x_i as constants when differentiating with respect to x_i. For example, consider the function

$$f : \mathbb{R}^2 \mapsto \mathbb{R},\ x \mapsto x_1(x_2)^2 + 2(x_1)^3. \tag{4.6}$$

f is partially differentiable with respect to both arguments, and we obtain the partial derivatives

$$f_{x_1} : \mathbb{R}^2 \mapsto \mathbb{R},\ x \mapsto (x_2)^2 + 6(x_1)^2$$

and

$$f_{x_2} : \mathbb{R}^2 \mapsto \mathbb{R},\ x \mapsto 2x_1x_2.$$

Again, we can find higher-order partial derivatives of functions with partially differentiable partial derivatives. For a function $f : A \mapsto \mathbb{R}$ of $n \in \mathbb{N}$ variables, there are n^2 second-order partial derivatives at a point $x^0 \in A$ (assuming that these derivatives exist), namely,

$$\frac{\partial^2 f(x^0)}{\partial x_i \partial x_j} = f_{x_i x_j}(x^0) := \frac{\partial f_{x_i}(x^0)}{\partial x_j} \quad \forall i, j \in \{1, \ldots, n\}.$$

If we arrange these second-order partial derivatives in an $n \times n$ matrix, we obtain the so-called *Hessian matrix* of second-order partial derivatives of f at x^0, which we denote by $H(f(x^0))$, that is,

$$H(f(x^0)) := \begin{pmatrix} f_{x_1x_1}(x^0) & f_{x_1x_2}(x^0) & \ldots & f_{x_1x_n}(x^0) \\ f_{x_2x_1}(x^0) & f_{x_2x_2}(x^0) & \ldots & f_{x_2x_n}(x^0) \\ \vdots & \vdots & & \vdots \\ f_{x_nx_1}(x^0) & f_{x_nx_2}(x^0) & \ldots & f_{x_nx_n}(x^0) \end{pmatrix}.$$

As an example, consider the function f defined in (4.6). For any $x \in \mathbb{R}^2$, we obtain the Hessian matrix

$$H(f(x)) = \begin{pmatrix} 12x_1 & 2x_2 \\ 2x_2 & 2x_1 \end{pmatrix}.$$

Note that in this example, we have $f_{x_1x_2}(x) = f_{x_2x_1}(x)$ for all $x \in \mathbb{R}^2$, that is, the *order of differentiation* is irrelevant in finding the mixed second-order partial derivatives $f_{x_1x_2}$ and $f_{x_2x_1}$. This is not the case for all functions. For example, consider

$$f : \mathbb{R}^2 \mapsto \mathbb{R},\ x \mapsto \begin{cases} x_1x_2\frac{(x_1)^2-(x_2)^2}{(x_1)^2+(x_2)^2} & \text{if } x \neq (0,0) \\ 0 & \text{if } x = (0,0). \end{cases}$$

For $x_2 \in \mathbb{R}$, we obtain

$$f_{x_1}(0, x_2) = \lim_{h \to 0} \frac{hx_2 \frac{h^2 - (x_2)^2}{h^2 + (x_2)^2}}{h} = -\frac{(x_2)^3}{(x_2)^2} = -x_2.$$

For $x_1 \in \mathbb{R}$, it follows that

$$f_{x_2}(x_1, 0) = \lim_{h \to 0} \frac{x_1 h \frac{(x_1)^2 - h^2}{(x_1)^2 + h^2}}{h} = \frac{(x_1)^3}{(x_1)^2} = x_1.$$

We now obtain $f_{x_1 x_2}(0,0) = -1 \neq 1 = f_{x_2 x_1}(0,0)$, and therefore, in this example, the order of differentiation *does* matter. However, this can only happen if the second-order mixed partial derivatives are *not continuous*. The following theorem (which is often referred to as *Young's theorem*) states conditions under which the order of differentiation is irrelevant for finding second-order mixed partial derivatives. We will state this result without a proof.

Theorem 4.3.2 *Let $n \in \mathbb{N}$, let $A \subseteq \mathbb{R}^n$ be convex, and let $f : A \mapsto \mathbb{R}$ be a function. Furthermore, let x^0 be an interior point of A, and let $i, j \in \{1, \ldots, n\}$. If there exists $\varepsilon \in \mathbb{R}_{++}$ such that $f_{x_i}(x)$, $f_{x_j}(x)$, and $f_{x_i x_j}(x)$ exist for all $x \in \mathcal{U}_\varepsilon(x^0)$ and these partial derivatives are continuous at x^0, then $f_{x_j x_i}(x^0)$ exists, and $f_{x_j x_i}(x^0) = f_{x_i x_j}(x^0)$.*

Therefore, if all first-order and second-order partial derivatives of a function $f : A \mapsto \mathbb{R}$ are continuous at $x^0 \in A$, it follows that the Hessian matrix of f at x^0 is a *symmetric* matrix.

We now define *total* differentiability of a function of several variables. Recall that $\mathbf{0}$ denotes the origin $(0, \ldots, 0)$ of $\mathbb{R}^n$.

Definition 4.3.3 *Let $n \in \mathbb{N}$, let $A \subseteq \mathbb{R}^n$ be convex, and let $f : A \mapsto \mathbb{R}$ be a function. Furthermore, let x^0 be an interior point of A. f is* totally differentiable *at x^0 if and only if there exist $\varepsilon \in \mathbb{R}_{++}$ and functions $\varepsilon_i : \mathbb{R}^n \mapsto \mathbb{R}$ for $i = 1, \ldots, n$ such that, for all $h = (h_1, \ldots, h_n) \in \mathcal{U}_\varepsilon(\mathbf{0})$,*

$$f(x^0 + h) - f(x^0) = \sum_{i=1}^n f_{x_i}(x^0) h_i + \sum_{i=1}^n \varepsilon_i(h) h_i$$

and

$$\lim_{h \to \mathbf{0}} \varepsilon_i(h) = 0 \quad \forall i = 1, \ldots, n.$$

For example, consider the function

$$f : \mathbb{R}^2 \mapsto \mathbb{R},\ x \mapsto (x_1)^2 + (x_2)^2.$$

For $x^0 \in \mathbb{R}^2$ and $h \in \mathbb{R}^2$, we obtain

$$\begin{aligned} f(x^0 + h) - f(x^0) &= (x_1^0 + h_1)^2 + (x_2^0 + h_2)^2 - (x_1^0)^2 - (x_2^0)^2 \\ &= 2x_1^0 h_1 + (h_1)^2 + 2x_2^0 h_2 + (h_2)^2 \\ &= f_{x_1}(x^0) h_1 + f_{x_2}(x^0) h_2 + h_1 h_1 + h_2 h_2 \\ &= f_{x_1}(x^0) h_1 + f_{x_2}(x^0) h_2 + \varepsilon_1(h) h_1 + \varepsilon_2(h) h_2 \end{aligned}$$

where $\varepsilon_1(h) := h_1$ and $\varepsilon_2(h) := h_2$. Because $\lim_{h \to \mathbf{0}} \varepsilon_i(h) = 0$ for $i = 1, 2$, it follows that f is totally differentiable at any point x^0 in its domain.

In general, the partial differentiability of f with respect to all variables at an interior point x^0 in the domain of f is *not* sufficient for the total differentiability of f at x^0. However, if all partial derivatives exist in a neighborhood of x^0 and are continuous at x^0, it follows that f is totally differentiable at x^0. We state this result—which is very helpful in checking functions for total differentiability—without a proof.

Theorem 4.3.4 *Let $n \in \mathbb{N}$, let $A \subseteq \mathbb{R}^n$ be convex, and let $f : A \mapsto \mathbb{R}$ be a function. Furthermore, let x^0 be an interior point of A. If there exists $\varepsilon \in \mathbb{R}_{++}$ such that, for all $i \in \{1, \ldots, n\}$, $f_{x_i}(x)$ exists for all $x \in \mathcal{U}_\varepsilon(x^0)$ and f_{x_i} is continuous at x^0, then f is totally differentiable at x^0.*

The *total differential* is used to *approximate* the change in the value of a function caused by changes in its arguments. We define

Definition 4.3.5 *Let $n \in I\!N$, let $A \subseteq I\!R^n$ be convex, and let $f : A \mapsto I\!R$ be a function. Furthermore, let x^0 be an interior point of A, let f be totally differentiable at x^0, and let $h = (h_1, \ldots, h_n) \in I\!R^n$ be such that $(x^0 + h) \in A$. The* total differential of f at x^0 *and h is defined by*

$$df(x^0, h) := \sum_{i=1}^{n} f_{x_i}(x^0) h_i.$$

If f is totally differentiable, the total differential of f at x^0 and h can be used as an approximation of the change in f, $f(x^0 + h) - f(x^0)$, for *small* $h_1, \ldots, h_n$, because, in this case, the values of the functions ε_i (see Definition 4.3.3) approach zero, and therefore, $df(x^0, h)$ is "close" to the true value of this difference.

The following theorem generalizes the *chain rule* to functions of several variables.

Theorem 4.3.6 *Let $m, n \in I\!N$, and let $A \subseteq I\!R^n$ and $B \subseteq I\!R^m$. Furthermore, let $f^j : A \mapsto I\!R$, $j \in \{1, \ldots, m\}$, be such that $(f^1(x), \ldots, f^m(x)) \in B$ for all $x \in A$, let $g : B \mapsto I\!R$, and let $x^0 \in A$. Define the function k by*

$$k : A \mapsto I\!R, \; x \mapsto g(f^1(x), \ldots, f^m(x)).$$

Let $i \in \{1, \ldots, n\}$. If g is totally differentiable at $y^0 = (f^1(x^0), \ldots, f^m(x^0))$ and, for all $j \in \{1, \ldots, m\}$, f^j is partially differentiable with respect to x_i at x^0, then k is partially differentiable with respect to x_i, and

$$k_{x_i}(x^0) = \sum_{j=1}^{m} g_{y_j}(f^1(x^0), \ldots, f^m(x^0)) f^j_{x_i}(x^0).$$

For example, consider the functions

$$\begin{aligned} f^1 &: \mathbb{R}^2 \mapsto \mathbb{R}, \; x \mapsto x_1 x_2, \\ f^2 &: \mathbb{R}^2 \mapsto \mathbb{R}, \; x \mapsto e^{x_1} + x_2, \\ g &: \mathbb{R}^2 \mapsto \mathbb{R}, \; y \mapsto (y_1 + y_2)^2. \end{aligned}$$

We obtain

$$f^1_{x_1}(x) = x_2, \; f^1_{x_2}(x) = x_1, \; f^2_{x_1}(x) = e^{x_1}, \; f^2_{x_2}(x) = 1 \quad \forall x \in \mathbb{R}^2$$

and

$$g_{y_1}(y) = g_{y_2}(y) = 2(y_1 + y_2) \quad \forall y \in \mathbb{R}^2.$$

The function

$$k : \mathbb{R}^2 \mapsto \mathbb{R}, \; x \mapsto g(f^1(x), f^2(x))$$

is partially differentiable with respect to both variables, and its partial derivatives are given by

$$\begin{aligned} k_{x_1}(x) &= g_{y_1}(f^1(x), f^2(x)) f^1_{x_1}(x) + g_{y_2}(f^1(x), f^2(x)) f^2_{x_1}(x) \\ &= 2(x_1 x_2 + e^{x_1} + x_2) x_2 + 2(x_1 x_2 + e^{x_1} + x_2) e^{x_1} \\ &= 2(x_1 x_2 + e^{x_1} + x_2)(e^{x_1} + x_2) \end{aligned}$$

and

$$\begin{aligned} k_{x_2}(x) &= g_{y_1}(f^1(x), f^2(x)) f^1_{x_2}(x) + g_{y_2}(f^1(x), f^2(x)) f^2_{x_2}(x) \\ &= 2(x_1 x_2 + e^{x_1} + x_2) x_1 + 2(x_1 x_2 + e^{x_1} + x_2) \\ &= 2(x_1 x_2 + e^{x_1} + x_2)(x_1 + 1) \end{aligned}$$

for all $x \in \mathbb{R}^2$.

Functions of several variables can be used to describe relationships between a *dependent variable* and *independent variables*. If the relationship between a dependent variable y and n independent variables $x_1, \ldots, x_n$ is *explicitly* given, we can write

$$y = f(x_1, \ldots, x_n) \tag{4.7}$$

with a function $f : A \mapsto \mathbb{R}$, where $A \subseteq \mathbb{R}^n$. For example, the equation

$$-(x_1)^2 + ((x_2)^2 + 1)y = 0$$

can be solved for y as a function of $x = (x_1, x_2)$, namely, $y = f(x)$ with

$$f : \mathbb{R}^2 \mapsto \mathbb{R},\ x \mapsto \frac{(x_1)^2}{(x_2)^2 + 1}.$$

However, it is not always the case that a relationship between a dependent variable and several independent variables can be solved explicitly as in (4.7). Consider, for example, the relationship described by the equation

$$-y + 1 + (x_1)^2 + (x_2)^2 - e^y = 0. \tag{4.8}$$

This equation cannot easily be solved for y as a function of $x = (x_1, x_2)$—in fact, it is not even clear whether this equation defines y as a function of x at all.

Under certain circumstances, it is possible to find the partial derivatives of a function that is *implicitly* defined by an equation as in the above example, even if the function is not known explicitly. We first define *implicit functions.*

Definition 4.3.7 *Let $n \in \mathbb{N}$, let $B \subseteq \mathbb{R}^{n+1}$, $A \subseteq \mathbb{R}^n$, and let $F : B \mapsto \mathbb{R}$. The equation*

$$F(x_1, \ldots, x_n, y) = 0 \tag{4.9}$$

defines an implicit function *$f : A \mapsto \mathbb{R}$ if and only if for all $x^0 \in A$, there exists a* unique *$y^0 \in \mathbb{R}$ such that $(x_1^0, \ldots, x_n^0, y^0) \in B$ and $F(x_1^0, \ldots, x_n^0, y^0) = 0$.*

The *implicit function theorem* (which we state without a proof) provides us with conditions under which equations of the form (4.9) define an implicit function and a method of finding partial derivatives of such an implicit function.

Theorem 4.3.8 *Let $n \in \mathbb{N}$, let $B \subseteq \mathbb{R}^{n+1}$, and let $F : B \mapsto \mathbb{R}$. Let $C \subseteq \mathbb{R}^n$ and $D \subseteq \mathbb{R}$ be such that $C \times D \subseteq B$. Let x^0 be an interior point of C and let y^0 be an interior point of D such that $F(x^0, y^0) = 0$. Let $i \in \{1, \ldots, n\}$. Suppose the partial derivatives $F_{x_i}(x, y)$ and $F_y(x, y)$ exist and are continuous for all $x \in C$ and all $y \in D$. If $F_y(x^0, y^0) \neq 0$, then there exists $\varepsilon \in \mathbb{R}_{++}$ such that, with $A := \mathcal{U}_\varepsilon(x^0)$, there exists a function $f : A \mapsto \mathbb{R}$ such that*

(i) $f(x^0) = y^0$,
(ii) $F(x, f(x)) = 0 \quad \forall x \in A$,
(iii) f is partially differentiable with respect to x_i on A, and

$$f_{x_i}(x) = -\frac{F_{x_i}(x, f(x))}{F_y(x, f(x))} \quad \forall x \in A. \tag{4.10}$$

Note that (4.10) follows from the *chain rule.* By (ii), we have

$$F(x, f(x)) = 0 \quad \forall x \in A,$$

and differentiating with respect to x_i, we obtain

$$F_{x_i}(x, f(x)) + F_y(x, f(x)) f_{x_i}(x) = 0 \quad \forall x \in A,$$

which is equivalent to (4.10).

For example, consider (4.8). We have

$$F(x_1, x_2, y) = -y + 1 + (x_1)^2 + (x_2)^2 - e^y,$$

and, for $x^0 = (0, 0)$ and $y^0 = 0$, we obtain $F(x^0, y^0) = 0$. Furthermore, $F_{x_1}(x, y) = 2x_1$, $F_{x_2}(x, y) = 2x_2$, and $F_y(x, y) = -1 - e^y$ for all $(x, y) \in \mathbb{R}^3$. Because $F_y(x^0, y^0) \neq 0$, (4.8) defines an implicit function f in a neighborhood of x^0, and we obtain

$$f_{x_1}(x^0) = -\frac{2x_1^0}{-1 - e^{y^0}} = 0, \quad f_{x_2}(x^0) = -\frac{2x_2^0}{-1 - e^{y^0}} = 0.$$

The implicit function theorem can be used to determine the slope of the tangent to a level set of a function $f : A \mapsto \mathbb{R}$ where $A \subseteq \mathbb{R}^2$ is convex. Letting $c \in f(A)$, the level set of f for c is the set of all points $x \in A$ satisfying

$$f(x) - c = 0.$$

Under the assumptions of the implicit function theorem, this equation defines x_2 as an implicit function of x_1 in a neighborhood of a point x^0 in this level set, and the derivative of this implicit function at x^0 is given by $-f_{x_1}(x^0)/f_{x_2}(x^0)$.

4.4 Unconstrained Optimization

The definition of unconstrained maximization and minimization of functions of several variables is analogous to the corresponding definition for functions of one variable.

Definition 4.4.1 *Let $n \in \mathbb{N}$, let $A \subseteq \mathbb{R}^n$ be convex, and let $f : A \mapsto \mathbb{R}$. Furthermore, let $x^0 \in A$.*

(i) f has a global maximum *at $x^0 \Leftrightarrow f(x^0) \geq f(x) \;\; \forall x \in A$.*
(ii) f has a local maximum *at $x^0 \Leftrightarrow \exists \varepsilon \in \mathbb{R}_{++}$ such that $f(x^0) \geq f(x) \;\; \forall x \in \mathcal{U}_\varepsilon(x^0) \cap A$.*
(iii) f has a global minimum *at $x^0 \Leftrightarrow f(x^0) \leq f(x) \;\; \forall x \in A$.*
(iv) f has a local minimum *at $x^0 \Leftrightarrow \exists \varepsilon \in \mathbb{R}_{++}$ such that $f(x^0) \leq f(x) \;\; \forall x \in \mathcal{U}_\varepsilon(x^0) \cap A$.*

We will restrict attention to *interior* maxima and minima in this section. For functions that are partially differentiable with respect to all variables, we can again formulate *necessary* first-order conditions for local maxima and minima.

Theorem 4.4.2 *Let $n \in \mathbb{N}$, let $A \subseteq \mathbb{R}^n$ be convex, and let $f : A \mapsto \mathbb{R}$. Furthermore, let x^0 be an interior point of A, and suppose f is partially differentiable with respect to all variables at x^0.*

(i) f has a local maximum at $x^0 \Rightarrow f_{x_i}(x^0) = 0 \;\; \forall i = 1, \ldots, n$.
(ii) f has a local minimum at $x^0 \Rightarrow f_{x_i}(x^0) = 0 \;\; \forall i = 1, \ldots, n$.

Proof. (i) Suppose f has a local maximum at an interior point $x^0 \in A$. For $i \in \{1, \ldots, n\}$, let f^i be defined as in (4.3). Then the function f^i must have a local maximum at x_i^0 (otherwise, f could not have a local maximum at x^0). By Theorem 3.3.6, we must have $(f^i)'(x^0) = 0$. By definition, $(f^i)'(x^0) = f_{x_i}(x^0)$, and therefore, $f_{x_i}(x^0) = 0$ for all $i = 1, \ldots, n$.

The proof of (ii) is analogous. ‖

Note that, as is the case for functions of one variable, Theorem 4.4.2 provides necessary, but not sufficient conditions for local maxima and minima. Points that satisfy these conditions are often called *critical points* or *stationary points*.

To obtain sufficient conditions for maxima and minima, we again examine *second-order* derivatives of f at an interior critical point x^0. If all first-order and second-order partial derivatives of f exist in a neighborhood of x^0 and these partial derivatives are continuous at x^0, the second-order changes in f can, in a neighborhood of x^0, be approximated by a *second-order total differential*

$$d^2 f(x^0, h) := d(df(x^0, h)) = \sum_{j=1}^{n} \sum_{i=1}^{n} f_{x_i x_j}(x^0) h_i h_j.$$

Note that the above double sum can be written as

$$h' H(f(x^0)) h \tag{4.11}$$

where h' is the transpose vector of h and $H(f(x^0))$ is the Hessian matrix of f at x^0. Analogously to the case of functions of one variable, a *sufficient* condition for a local maximum is that this second-order change in f is *negative* for small h. Because of (4.11), we therefore obtain as a sufficient condition for a local maximum that the Hessian matrix of f at x^0 is *negative definite*. Similar considerations apply to local minima, and we obtain

Theorem 4.4.3 *Let $n \in \mathbb{N}$, let $A \subseteq \mathbb{R}^n$ be convex, and let $f : A \mapsto \mathbb{R}$. Furthermore, let x^0 be an interior point of A. Suppose f is twice partially differentiable with respect to all variables in a neighborhood of x^0, and these partial derivatives are continuous at x^0.*

(i) If $f_{x_i}(x^0) = 0$ for all $i = 1, \ldots, n$ and $H(f(x^0))$ is negative definite, then f has a local maximum at x^0.
(ii) If $f_{x_i}(x^0) = 0$ for all $i = 1, \ldots, n$ and $H(f(x^0))$ is positive definite, then f has a local minimum at x^0.
(iii) If $f_{x_i}(x^0) = 0$ for all $i = 1, \ldots, n$ and $H(f(x^0))$ is indefinite, then f has neither a local maximum nor a local minimum at x^0.

The conditions formulated in Theorem 4.4.3 (i) and (ii) are called sufficient *second-order* conditions for unconstrained maxima and minima.

Here is an example for an optimization problem with a function of two variables. Let

$$f : \mathbb{R}^2 \mapsto \mathbb{R},\ x \mapsto (x_1)^2 + (x_2)^2 + (x_1)^3.$$

The necessary first-order conditions for a maximum or minimum at $x^0 \in \mathbb{R}^2$ are

$$f_{x_1}(x^0) = 2x_1^0 + 3(x_1^0)^2 = 0$$

and

$$f_{x_2}(x^0) = 2x_2^0 = 0.$$

Therefore, we have two stationary points, namely, $(0, 0)$ and $(-2/3, 0)$. For $x \in \mathbb{R}^2$, the Hessian matrix of f at x is

$$H(f(x)) = \begin{pmatrix} 2 + 6x_1 & 0 \\ 0 & 2 \end{pmatrix}.$$

Therefore,

$$H(f(0,0)) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

which is a positive definite matrix. Therefore, f has a local minimum at $(0, 0)$. Furthermore,

$$H(f(-2/3,0)) = \begin{pmatrix} -2 & 0 \\ 0 & 2 \end{pmatrix}$$

which is an *indefinite* matrix. This implies that f has neither a maximum nor a minimum at $(-2/3, 0)$.

To formulate sufficient conditions for *global* maxima and minima, we define *concavity* and *convexity* of functions of several variables.

Definition 4.4.4 *Let $n \in \mathbb{N}$, let $A \subseteq \mathbb{R}^n$ be convex, and let $f : A \mapsto \mathbb{R}$.*

(i) f is concave *if and only if*

$$f(\lambda x + (1-\lambda)y) \geq \lambda f(x) + (1-\lambda)f(y) \quad \forall x, y \in A \text{ such that } x \neq y,\ \forall \lambda \in (0, 1).$$

(ii) f is strictly concave *if and only if*

$$f(\lambda x + (1-\lambda)y) > \lambda f(x) + (1-\lambda)f(y) \quad \forall x, y \in A \text{ such that } x \neq y,\ \forall \lambda \in (0, 1).$$

(iii) f is convex *if and only if*

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y) \quad \forall x, y \in A \text{ such that } x \neq y,\ \forall \lambda \in (0, 1).$$

(iv) f is strictly convex *if and only if*

$$f(\lambda x + (1-\lambda)y) < \lambda f(x) + (1-\lambda)f(y) \quad \forall x, y \in A \text{ such that } x \neq y,\ \forall \lambda \in (0, 1).$$

For *open* domains A, we obtain the following conditions for concavity and convexity if f has certain differentiability properties (we only consider open sets, because we restrict attention to interior points in this chapter).

Theorem 4.4.5 *Let $n \in \mathbb{N}$, let $A \subseteq \mathbb{R}^n$ be open and convex, and let $f : A \mapsto \mathbb{R}$. Suppose all first-order and second-order partial derivatives of f exist and are continuous.*

(i) $H(f(x))$ is negative semidefinite $\forall x \in A \Leftrightarrow f$ is concave.
(ii) $H(f(x))$ is negative definite $\forall x \in A \Rightarrow f$ is strictly concave.
(iii) $H(f(x))$ is positive semidefinite $\forall x \in A \Leftrightarrow f$ is convex.
(iv) $H(f(x))$ is positive definite $\forall x \in A \Rightarrow f$ is strictly convex.

Again, note that the reverse implications in parts (ii) and (iv) of Theorem 4.4.5 are *not* true.

Now we can state results that parallel Theorems 3.4.5 to 3.4.7 concerning global maxima and minima. The proofs of the following three theorems are analogous to the proofs of the above mentioned theorems for functions of one variable and are left as exercises.

Theorem 4.4.6 *Let $n \in I\!N$, let $A \subseteq I\!R^n$ be convex, and let $f : A \mapsto I\!R$. Furthermore, let $x^0 \in A$.*

(i) If f is concave and f has a local maximum at x^0, then f has a global maximum at x^0.
(ii) If f is convex and f has a local minimum at x^0, then f has a global minimum at x^0.

Theorem 4.4.7 *Let $n \in I\!N$, let $A \subseteq I\!R^n$ be convex, and let $f : A \mapsto I\!R$.*

(i) If f is strictly concave, then f has at most one local (and global) maximum.
(ii) If f is strictly convex, then f has at most one local (and global) minimum.

Theorem 4.4.8 *Let $n \in I\!N$, let $A \subseteq I\!R^n$ be convex, and let $f : A \mapsto I\!R$. Furthermore, let x^0 be an interior point of A, and suppose f is partially differentiable with respect to all variables at x^0.*

(i) $f_{x_i}(x_0) = 0 \;\; \forall i = 1, \ldots, n \;\wedge\; f$ is concave $\Rightarrow f$ has a global maximum at x^0.
(ii) $f_{x_i}(x_0) = 0 \;\; \forall i = 1, \ldots, n \;\wedge\; f$ is convex $\Rightarrow f$ has a global minimum at x^0.

We conclude this section with an economic example. Suppose a firm produces a good which is sold in a competitive market. The firm uses $n \in \mathbb{N}$ factors of production that can be bought in competitive factor markets. A *production function* $f : \mathbb{R}^n_{++} \mapsto \mathbb{R}$ is used to describe the *production technology* of the firm. If we interpret $x \in \mathbb{R}^n_{++}$ as a *vector of inputs* (or a *factor combination*), $f(x)$ is the maximal amount of output that can be produced with this factor combination according to the firm's technology.

Suppose $w \in \mathbb{R}^n_{++}$ is the vector of *factor prices*, where w_i is the price of factor i, $i = 1, \ldots, n$. Furthermore, suppose $p \in \mathbb{R}_{++}$ is the price of the good produced by the firm.

If the firm uses the factor combination $x \in \mathbb{R}^n_{++}$, the profit of the firm is $pf(x) - wx$, and therefore, the profit maximization problem of the firm is

$$\max_x \{pf(x) - wx\}.$$

Suppose f is concave and partially differentiable with respect to all variables. Therefore, the objective function is concave, and by Theorem 4.4.8, the conditions

$$pf_{x_i}(x^0) - w_i = 0 \quad \forall i = 1, \ldots, n$$

are sufficient for a global maximum at $x^0 \in \mathbb{R}^n_{++}$.

For example, let

$$f : \mathbb{R}^2_{++} \mapsto \mathbb{R},\; x \mapsto (x_1 x_2)^{1/4}.$$

We obtain the profit maximization problem

$$\max_x \{p(x_1 x_2)^{1/4} - w_1 x_1 - w_2 x_2\}.$$

First, we show that the function f is concave. The Hessian matrix of f at $x \in \mathbb{R}^2_{++}$ is

$$H(f(x)) = \begin{pmatrix} -\frac{3}{16}(x_1)^{-7/4}(x_2)^{1/4} & \frac{1}{16}(x_1 x_2)^{-3/4} \\ \frac{1}{16}(x_1 x_2)^{-3/4} & -\frac{3}{16}(x_1)^{1/4}(x_2)^{-7/4} \end{pmatrix}.$$

The leading principal minor of order one is

$$-\frac{3}{16}(x_1)^{-7/4}(x_2)^{1/4},$$

which is negative for all $x \in \mathbb{R}^2_{++}$, and the determinant of the Hessian matrix is

$$|H(f(x))| = \frac{1}{32}(x_1x_2)^{-3/2} > 0$$

for all $x \in \mathbb{R}^2_{++}$, and therefore, $H(f(x))$ is negative definite for all $x \in \mathbb{R}^2_{++}$. This implies that f is strictly concave, and therefore, the objective function is strictly concave. Hence, the first-order conditions

$$\frac{p}{4}(x_1^0)^{-3/4}(x_2^0)^{1/4} - w_1 = 0 \quad (4.12)$$

$$\frac{p}{4}(x_1^0)^{1/4}(x_2^0)^{-3/4} - w_2 = 0 \quad (4.13)$$

are sufficient for a global maximum at $x^0 \in \mathbb{R}^2_{++}$. From (4.13), it follows that

$$x_2^0 = (w_2)^{-4/3}\left(\frac{p}{4}\right)^{4/3}(x_1^0)^{1/3}. \quad (4.14)$$

Using (4.14) in (4.12), we obtain

$$\frac{p}{4}(x_1^0)^{-3/4}\left((w_2)^{-4/3}\left(\frac{p}{4}\right)^{4/3}(x_1^0)^{1/3}\right)^{1/4} - w_1 = 0. \quad (4.15)$$

Solving (4.15) for x_1^0, we obtain

$$x_1^0 = \frac{(p)^2}{16(w_1)^{3/2}(w_2)^{1/2}}, \quad (4.16)$$

and using (4.16) in (4.14), it follows that

$$x_2^0 = \frac{(p)^2}{16(w_1)^{1/2}(w_2)^{3/2}}. \quad (4.17)$$

(4.16) and (4.17) give us the optimal choices of the amounts used of each factor as *functions* of the prices, namely,

$$\bar{x}_1 : \mathbb{R}^3_{++} \mapsto \mathbb{R},\ (p,w) \mapsto \frac{(p)^2}{16(w_1)^{3/2}(w_2)^{1/2}}$$

$$\bar{x}_2 : \mathbb{R}^3_{++} \mapsto \mathbb{R},\ (p,w) \mapsto \frac{(p)^2}{16(w_1)^{1/2}(w_2)^{3/2}}.$$

These functions are called the *factor demand functions* corresponding to the technology described by the production function f. By substituting x^0 into f, we obtain the optimal amount of output to be produced, y^0, as a function $\bar{y}$ of the prices, namely,

$$\bar{y} : \mathbb{R}^3_{++} \mapsto \mathbb{R},\ (p,w) \mapsto \frac{p}{4\sqrt{w_1w_2}}.$$

$\bar{y}$ is the *supply function* corresponding to f. Finally, substituting x^0 into the objective function yields the *maximal profit* of the firm π^0 as a function $\bar{\pi}$ of the prices. In this example, this *profit function* is given by

$$\bar{\pi} : \mathbb{R}^3_{++} \mapsto \mathbb{R},\ (p,w) \mapsto \frac{(p)^2}{8\sqrt{w_1w_2}}.$$

4.5 Optimization with Equality Constraints

In many applications, optimization problems involving *constraints* have to be solved. A typical economic example is the *cost minimzation* problem of a firm, where production costs have to be minimized subject to the constraint that a certain amount is produced. We will discuss this example in more detail at the end of this section. First, general methods for solving constrained optimization problems are developed.

We will restrict attention to maximization and minimization problems involving functions of $n \geq 2$ variables and *one* equality constraint (it is possible to extend the results discussed here to problems with *more* than one constraint, as long as the number of constraints is smaller than the number of variables).

Suppose we have an *objective function* $f : A \mapsto \mathbb{R}$ such that $A \subseteq \mathbb{R}^n$ is convex, where $n \in \mathbb{N}$ and $n \geq 2$. Furthermore, suppose we have to satisfy an equality constraint that requires $g(x) = 0$, where $g : A \mapsto \mathbb{R}$. Note that *any* equality constraint can be expressed in this form. We now define constrained maxima and minima.

Definition 4.5.1 *Let $n \in \mathbb{N}$, $n \geq 2$, let $A \subseteq \mathbb{R}^n$ be convex, and let $f : A \mapsto \mathbb{R}$ and $g : A \mapsto \mathbb{R}$. Furthermore, let $x^0 \in A$.*

(i) f has a global constrained maximum subject to the constraint $g(x) = 0$ *at x^0 if and only if $g(x^0) = 0$ and $f(x^0) \geq f(x)$ for all $x \in A$ such that $g(x) = 0$.*
(ii) f has a local constrained maximum subject to the constraint $g(x) = 0$ *at x^0 if and only if there exists $\varepsilon \in \mathbb{R}_{++}$ such that $g(x^0) = 0$ and $f(x^0) \geq f(x)$ for all $x \in \mathcal{U}_\varepsilon(x^0) \cap A$ such that $g(x) = 0$.*
(iii) f has a global constrained minimum subject to the constraint $g(x) = 0$ *at x^0 if and only if $g(x^0) = 0$ and $f(x^0) \leq f(x)$ for all $x \in A$ such that $g(x) = 0$.*
(iv) f has a local constrained minimum subject to the constraint $g(x) = 0$ *at x^0 if and only if there exists $\varepsilon \in \mathbb{R}_{++}$ such that $g(x^0) = 0$ and $f(x^0) \leq f(x)$ for all $x \in \mathcal{U}_\varepsilon(x^0) \cap A$ such that $g(x) = 0$.*

Again, we only discuss *interior* constrained maxima and minima in this section. To obtain necessary conditions for local constrained maxima and minima, we first give an illustration for the case of a function of two variables. Suppose a function $f : A \mapsto \mathbb{R}$ has a local constrained maximum (or minimum) subject to the constraint $g(x) = 0$ at an interior point x^0 of A, where $A \subseteq \mathbb{R}^2$ is convex, f and g are twice partially differentiable with respect to all variables, and all partial derivatives are continuous at x^0. Furthermore, suppose $g_{x_2}(x^0) \neq 0$ (assuming $g_{x_1}(x^0) \neq 0$ would work as well). We will see shortly why such a condition—sometimes called a *constraint qualification* condition—is needed. We now derive an *unconstrained* optimization problem from this problem and then use the results for unconstrained optimization to draw conclusions about the solution of the constrained problem.

Because $g_{x_2}(x^0) \neq 0$, the implicit function theorem implies that the equation $g(x) = 0$ defines, in a neighborhood $\mathcal{U}_\varepsilon(x_1^0)$, an implicit function $k : \mathcal{U}_\varepsilon(x_1^0) \mapsto \mathbb{R}$, where $x_2 = k(x_1)$ for all $x_1 \in \mathcal{U}_\varepsilon(x_1^0)$. Because f has a local constrained maximum at x^0, it follows that the function

$$\hat{f} : \mathcal{U}_\varepsilon(x_1^0) \mapsto \mathbb{R},\ x_1 \mapsto f(x_1, k(x_1)) \tag{4.18}$$

has a local *unconstrained* maximum at x_1^0. Note that the constraint $g(x_1, x_2) = 0$ is already taken into account by definition of the (implicit) function k. Because $\hat{f}$ has a local interior maximum at x_1^0, it follows that

$$\hat{f}'(x_1^0) = 0$$

which, by definition of $\hat{f}$ and application of the chain rule, is equivalent to

$$f_{x_1}(x_1^0, k(x_1^0)) + f_{x_2}(x_1^0, k(x_1^0))k'(x_1^0) = 0. \tag{4.19}$$

By the implicit function theorem,

$$k'(x_1^0) = -\frac{g_{x_1}(x_1^0, k(x_1^0))}{g_{x_2}(x_1^0, k(x_1^0))}. \tag{4.20}$$

Defining

$$\lambda^0 := \frac{f_{x_2}(x_1^0, k(x_1^0))}{g_{x_2}(x_1^0, k(x_1^0))}, \tag{4.21}$$

(4.19) and (4.20) imply

$$f_{x_1}(x_1^0, k(x_1^0)) - \lambda^0 g_{x_1}(x_1^0, k(x_1^0)) = 0,$$

and (4.21) is equivalent to

$$f_{x_2}(x_1^0, k(x_1^0)) - \lambda^0 g_{x_2}(x_1^0, k(x_1^0)) = 0.$$

Because $x_2^0 = k(x_1^0)$ and $g(x^0) = 0$ (which is equivalent to $-g(x^0) = 0$), we obtain the following *necessary* conditions for an interior local constrained maximum or minimum at x^0.

$$-g(x_1^0, x_2^0) = 0 \tag{4.22}$$
$$f_{x_1}(x_1^0, x_2^0) - \lambda^0 g_{x_1}(x_1^0, x_2^0) = 0 \tag{4.23}$$
$$f_{x_2}(x_1^0, x_2^0) - \lambda^0 g_{x_2}(x_1^0, x_2^0) = 0. \tag{4.24}$$

Therefore, if f has an interior local constrained maximum (minimum) at x^0, it follows that there exists a real number λ^0 such that $(\lambda^0, x_1^0, x_2^0)$ satisfies the above system of equations. We can express these conditions in terms of the derivatives of the *Lagrange function* of this optimization problem.

The Lagrange function is defined by

$$L : \mathbb{R} \times A \mapsto \mathbb{R},\ (\lambda, x) \mapsto f(x) - \lambda g(x),$$

and the conditions (4.22), (4.23), and (4.24) are obtained by setting the partial derivatives of L with respect to its arguments λ, x_1, x_2, respectively, equal to zero. The additional variable $\lambda \in \mathbb{R}$ is called the *Lagrange multiplier* for this problem. Therefore, at a local constrained maximum or minimum, the point (λ^0, x^0) must be a *stationary point* of the Lagrange function L.

This procedure generalizes easily to functions of $n \geq 2$ variables. We obtain

Theorem 4.5.2 *Let $n \in \mathbb{N}$, $n \geq 2$, let $A \subseteq \mathbb{R}^n$ be convex, and let $f : A \mapsto \mathbb{R}$, $g : A \mapsto \mathbb{R}$. Furthermore, let x^0 be an interior point of A. Suppose f and g are partially differentiable with respect to all variables in a neighborhood of x^0, and these partial derivatives are continuous at x^0. Suppose there exists $i \in \{1, \ldots, n\}$ such that $g_{x_i}(x^0) \neq 0$.*

(i) If f has a local constrained maximum subject to the constraint $g(x) = 0$ at x^0, then there exists $\lambda^0 \in \mathbb{R}$ such that

$$g(x^0) = 0 \ \wedge\ f_{x_i}(x^0) - \lambda^0 g_{x_i}(x^0) = 0 \ \ \forall i = 1, \ldots, n.$$

(ii) If f has a local constrained minimum subject to the constraint $g(x) = 0$ at x^0, then there exists $\lambda^0 \in \mathbb{R}$ such that

$$g(x^0) = 0 \ \wedge\ f_{x_i}(x^0) - \lambda^0 g_{x_i}(x^0) = 0 \ \ \forall i = 1, \ldots, n.$$

Figure 4.6 provides a graphical illustration of a solution to a constrained optimization problem. Suppose we want to maximize $f : \mathbb{R}^2_{++} \mapsto \mathbb{R}$ subject to the constraint $g(x) = 0$, where $g : \mathbb{R}^2_{++} \mapsto \mathbb{R}$. The unique solution to this problem is at the point x^0. Note that, at that point, the level set of the objective function f passing through x^0 is tangent to the level set of g passing through x^0. Therefore, the slope of the tangent to the level set of f has to be equal to the slope of the tangent to the level set of g at x^0. Applying the implicit function theorem, these slopes are given by $-f_{x_1}(x^0)/f_{x_2}(x^0)$ and $-g_{x_1}(x^0)/g_{x_2}(x^0)$, respectively. Therefore, we obtain the condition

$$\frac{f_{x_1}(x^0)}{f_{x_2}(x^0)} = \frac{g_{x_1}(x^0)}{g_{x_2}(x^0)},$$

in addition to the requirement that $g(x^0) = 0$. For solutions where the value of the Lagrange multiplier λ^0 is different from zero, the above conditions are an immediate consequence of the first-order conditions stated in Theorem 4.5.2.

To obtain *sufficient* conditions for local constrained maxima and minima at interior points, we again have to use *second-order* derivatives. Whereas the definiteness properties of the *Hessian* matrix are important for *unconstrained* optimization problems, the so-called *bordered Hessian* matrix is relevant for constrained maxima and minima. The bordered Hessian matrix is the Hessian matrix of second-order partial derivatives of the Lagrange function L. By definition of the Lagrange function, the bordered Hessian at (λ^0, x^0) is given by

$$H(L(\lambda^0, x^0)) := \begin{pmatrix} 0 & -g_{x_1}(x^0) & \ldots & -g_{x_n}(x^0) \\ -g_{x_1}(x^0) & f_{x_1x_1}(x^0) - \lambda^0 g_{x_1x_1}(x^0) & \ldots & f_{x_1x_n}(x^0) - \lambda^0 g_{x_1x_n}(x^0) \\ \vdots & \vdots & & \vdots \\ -g_{x_n}(x^0) & f_{x_nx_1}(x^0) - \lambda^0 g_{x_nx_1}(x^0) & \ldots & f_{x_nx_n}(x^0) - \lambda^0 g_{x_nx_n}(x^0) \end{pmatrix}.$$

The sufficient second-order conditions for local constrained maxima and minima can be expressed in terms of the signs of the principal minors of this matrix. Note that the leading principal minor of order one is always equal to zero, and the leading principal minor of order two is always nonpositive. Therefore, only the higher-order leading principal minors will be of relevance.

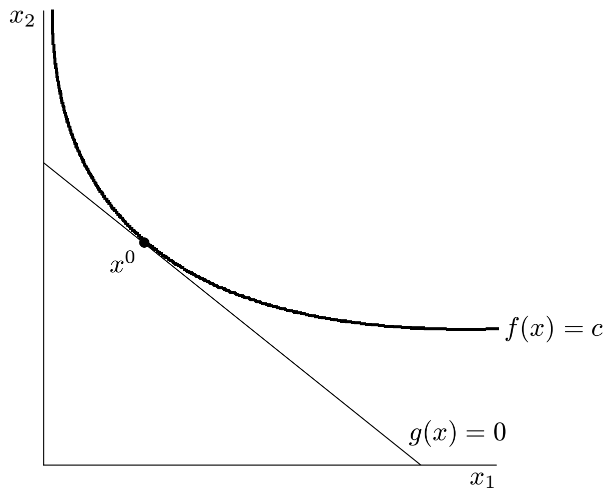


Figure 4.6: A constrained optimization problem.

To state these second-order conditions, we introduce the following notation. For $r \in \{2, \ldots, n\}$, let

$$H^r(L(\lambda^0, x^0)) := \begin{pmatrix} 0 & -g_{x_1}(x^0) & \ldots & -g_{x_r}(x^0) \\ -g_{x_1}(x^0) & f_{x_1x_1}(x^0) - \lambda^0 g_{x_1x_1}(x^0) & \ldots & f_{x_1x_r}(x^0) - \lambda^0 g_{x_1x_r}(x^0) \\ \vdots & \vdots & & \vdots \\ -g_{x_r}(x^0) & f_{x_rx_1}(x^0) - \lambda^0 g_{x_rx_1}(x^0) & \ldots & f_{x_rx_r}(x^0) - \lambda^0 g_{x_rx_r}(x^0) \end{pmatrix}.$$

The following theorem (which is proven for the case $n = 2$) gives sufficient conditions for local constrained optima at interior points.

Theorem 4.5.3 *Let $n \in \mathbb{N}$, $n \geq 2$, let $A \subseteq \mathbb{R}^n$ be convex, and let $f : A \mapsto \mathbb{R}$, $g : A \mapsto \mathbb{R}$. Let x^0 be an interior point of A. Suppose f and g are twice partially differentiable with respect to all variables in a neighborhood of x^0, and these partial derivatives are continuous at x^0. Suppose there exists $i \in \{1, \ldots, n\}$ such that $g_{x_i}(x^0) \neq 0$. Furthermore, suppose $g(x^0) = 0$ and $f_{x_i}(x^0) - \lambda^0 g_{x_i}(x^0) = 0$ for all $i = 1, \ldots, n$.*

(i) If $(-1)^r |H^r(L(\lambda^0, x^0))| > 0$ for all $r = 2, \ldots, n$, then f has a local constrained maximum subject to the constraint $g(x) = 0$ at x^0.
(ii) If $|H^r(L(\lambda^0, x^0))| < 0$ for all $r = 2, \ldots, n$, then f has a local constrained minimum subject to the constraint $g(x) = 0$ at x^0.

Proof (for $n = 2$). Theorem 4.5.3 is true for any $n \geq 2$, but for simplicity of exposition, we only prove the case $n = 2$ here.

(i) Consider the function $\hat{f}$ defined in (4.18). The sufficient second-order condition for a local maximum of $\hat{f}$ at x_1^0 is

$$\hat{f}''(x_1^0) < 0. \tag{4.25}$$

Recall that $\hat{f}'(x_1^0)$ is given by the left side of (4.19). Using (4.20), we obtain

$$\hat{f}'(x_1^0) = f_{x_1}(x_1^0, k(x_1^0)) - f_{x_2}(x_1^0, k(x_1^0)) \frac{g_{x_1}(x_1^0, k(x_1^0))}{g_{x_2}(x_1^0, k(x_1^0))}.$$

Differentiating again and using (4.20) and (4.21), we obtain

$$\begin{aligned} \hat{f}''(x_1^0) &= f_{x_1x_1}(x_1^0, k(x_1^0)) + f_{x_2x_2}(x_1^0, k(x_1^0)) \left(\frac{g_{x_1}(x_1^0, k(x_1^0))}{g_{x_2}(x_1^0, k(x_1^0))} \right)^2 \\ &- f_{x_1x_2}(x_1^0, k(x_1^0)) \frac{g_{x_1}(x_1^0, k(x_1^0))}{g_{x_2}(x_1^0, k(x_1^0))} - f_{x_2x_1}(x_1^0, k(x_1^0)) \frac{g_{x_1}(x_1^0, k(x_1^0))}{g_{x_2}(x_1^0, k(x_1^0))} \end{aligned}$$

$$
\begin{aligned}
&- \lambda^0 \left(g_{x_1x_1}(x_1^0, k(x_1^0)) - g_{x_1x_2}(x_1^0, k(x_1^0)) \frac{g_{x_1}(x_1^0, k(x_1^0))}{g_{x_2}(x_1^0, k(x_1^0))} \right) \\
&+ \lambda^0 \frac{g_{x_1}(x_1^0, k(x_1^0))}{g_{x_2}(x_1^0, k(x_1^0))} \left(g_{x_2x_1}(x_1^0, k(x_1^0)) - g_{x_2x_2}(x_1^0, k(x_1^0)) \frac{g_{x_1}(x_1^0, k(x_1^0))}{g_{x_2}(x_1^0, k(x_1^0))} \right),
\end{aligned}
$$

which is equivalent to

$$
\begin{aligned}
\hat{f}''(x_1^0) &= f_{x_1x_1}(x_1^0, k(x_1^0)) - \lambda^0 g_{x_1x_1}(x_1^0, k(x_1^0)) \\
&+ \left(f_{x_2x_2}(x_1^0, k(x_1^0)) - \lambda^0 g_{x_2x_2}(x_1^0, k(x_1^0)) \right) \left(\frac{g_{x_1}(x_1^0, k(x_1^0))}{g_{x_2}(x_1^0, k(x_1^0))} \right)^2 \\
&- \left(f_{x_1x_2}(x_1^0, k(x_1^0)) - \lambda^0 g_{x_1x_2}(x_1^0, k(x_1^0)) \right) \frac{g_{x_1}(x_1^0, k(x_1^0))}{g_{x_2}(x_1^0, k(x_1^0))} \\
&- \left(f_{x_2x_1}(x_1^0, k(x_1^0)) - \lambda^0 g_{x_2x_1}(x_1^0, k(x_1^0)) \right) \frac{g_{x_1}(x_1^0, k(x_1^0))}{g_{x_2}(x_1^0, k(x_1^0))}.
\end{aligned}
$$

Because $g_{x_2}(x_1^0, k(x_1^0)) \neq 0$, $\left(g_{x_2}(x_1^0, k(x_1^0))\right)^2$ is positive, and therefore, multiplying the inequality (4.25) by $\left(g_{x_2}(x_1^0, k(x_1^0))\right)^2$ yields

$$
\begin{aligned}
0 &> \left(f_{x_1x_1}(x_1^0, k(x_1^0)) - \lambda^0 g_{x_1x_1}(x_1^0, k(x_1^0)) \right) \left(g_{x_2}(x_1^0, k(x_1^0)) \right)^2 \\
&+ \left(f_{x_2x_2}(x_1^0, k(x_1^0)) - \lambda^0 g_{x_2x_2}(x_1^0, k(x_1^0)) \right) \left(g_{x_1}(x_1^0, k(x_1^0)) \right)^2 \\
&- \left(f_{x_1x_2}(x_1^0, k(x_1^0)) - \lambda^0 g_{x_1x_2}(x_1^0, k(x_1^0)) \right) g_{x_1}(x_1^0, k(x_1^0)) g_{x_2}(x_1^0, k(x_1^0)) \\
&- \left(f_{x_2x_1}(x_1^0, k(x_1^0)) - \lambda^0 g_{x_2x_1}(x_1^0, k(x_1^0)) \right) g_{x_1}(x_1^0, k(x_1^0)) g_{x_2}(x_1^0, k(x_1^0)). \qquad (4.26)
\end{aligned}
$$

The right side of (4.26) is equal to $-|H(L(\lambda^0, x^0))|$, and therefore, (4.25) is equivalent to

$$|H(L(\lambda^0, x^0))| > 0,$$

which, for $n = 2$, is the second-order sufficient condition stated in (i).

The proof of (ii) is analogous. ‖

For example, suppose we want to find all local constrained maxima and minima of the function

$$f : \mathbb{R}^2_{++} \mapsto \mathbb{R},\ x \mapsto x_1 x_2$$

subject to the constraint $x_1 + x_2 = 1$. The Lagrange function for this problem is

$$L : \mathbb{R} \times \mathbb{R}^2_{++} \mapsto \mathbb{R},\ (\lambda, x) \mapsto x_1 x_2 - \lambda(x_1 + x_2 - 1).$$

Differentiating the Lagrange function with respect to all variables and setting these partial derivatives equal to zero, we obtain the following necessary conditions for a constrained optimum.

$$
\begin{aligned}
-x_1 - x_2 + 1 &= 0 \\
x_2 - \lambda &= 0 \\
x_1 - \lambda &= 0.
\end{aligned}
$$

The unique solution to this system of equations is $\lambda^0 = 1/2$, $x^0 = (1/2, 1/2)$. To determine the nature of this stationary point, consider the bordered Hessian matrix

$$H(L(\lambda^0, x^0)) = \begin{pmatrix} 0 & -1 & -1 \\ -1 & 0 & 1 \\ -1 & 1 & 0 \end{pmatrix}.$$

The determinant of this matrix is $|H(L(\lambda^0, x^0))| = 2 > 0$, and therefore, according to Theorem 4.5.3, f has a local constrained maximum at x^0. The maximal value of the objective function subject to the constraint is $f(x^0) = 1/4$.

As an economic example, consider the following cost minimization problem of a competitive firm. Suppose $f : \mathbb{R}^n_{++} \mapsto \mathbb{R}$ is the production function of a firm. The firm wants to produce $y \in \mathbb{R}_{++}$ units of output in a way such that production costs are minimized. If all factors are variable and the factor prices are $w \in \mathbb{R}^n_{++}$, the costs to be minimized are given by

$$wx = \sum_{i=1}^{n} w_i x_i$$

where $x \in \mathbb{R}^2_{++}$ is a vector of factors of production. The constraint requires that x is chosen such that $f(x) = y$, or, equivalently, $f(x) - y = 0$.

Suppose the production function is given by

$$f : \mathbb{R}^2_{++} \mapsto \mathbb{R},\ x \mapsto (x_1 x_2)^{1/4}$$

(this is the same production function as in Section 4.4). Therefore, the firm wants to minimize the production costs $w_1 x_1 + w_2 x_2$ subject to the constraint $(x_1 x_2)^{1/4} - y = 0$. The Lagrange function for this problem is

$$L : \mathbb{R} \times \mathbb{R}^2_{++} \mapsto \mathbb{R},\ (\lambda, x) \mapsto w_1 x_1 + w_2 x_2 - \lambda \left((x_1 x_2)^{1/4} - y \right).$$

According to Theorem 4.5.2, the necessary first-order conditions for a minimum at $x^0 \in \mathbb{R}^2_{++}$ are

$$-(x_1^0 x_2^0)^{1/4} + y = 0 \tag{4.27}$$

$$w_1 - \frac{\lambda^0}{4}(x_1^0)^{-3/4}(x_2^0)^{1/4} = 0 \tag{4.28}$$

$$w_2 - \frac{\lambda^0}{4}(x_1^0)^{1/4}(x_2^0)^{-3/4} = 0. \tag{4.29}$$

(4.28) and (4.29) are quivalent to

$$w_1 = \frac{\lambda^0}{4}(x_1^0)^{-3/4}(x_2^0)^{1/4} \tag{4.30}$$

and

$$w_2 = \frac{\lambda^0}{4}(x_1^0)^{1/4}(x_2^0)^{-3/4}, \tag{4.31}$$

respectively. Dividing (4.30) by (4.31) yields

$$\frac{w_1}{w_2} = \frac{x_2^0}{x_1^0},$$

and therefore,

$$x_2^0 = \frac{w_1}{w_2} x_1^0. \tag{4.32}$$

Using (4.32) in (4.27), we obtain

$$\left(x_1^0 \frac{w_1}{w_2} x_1^0 \right)^{1/4} = y,$$

which implies

$$x_1^0 = (y)^2 \sqrt{\frac{w_2}{w_1}}. \tag{4.33}$$

Substituting (4.33) in (4.32), it follows that

$$x_2^0 = (y)^2 \sqrt{\frac{w_1}{w_2}}. \tag{4.34}$$

Substituting (4.33) and (4.34) into (4.28) (or (4.29)), we obtain $\lambda^0 = 4y\sqrt{w_1 w_2} > 0$. The bordered Hessian matrix at (λ^0, x^0) is given by

$$H(L(\lambda^0, x^0)) = \begin{pmatrix} 0 & -\frac{1}{4}(x_1^0)^{-3/4}(x_2^0)^{1/4} & -\frac{1}{4}(x_1^0)^{1/4}(x_2^0)^{-3/4} \\ -\frac{1}{4}(x_1^0)^{-3/4}(x_2^0)^{1/4} & \frac{3}{16}\lambda^0(x_1^0)^{-7/4}(x_2^0)^{1/4} & -\frac{1}{16}\lambda^0(x_1^0 x_2^0)^{-3/4} \\ -\frac{1}{4}(x_1^0)^{1/4}(x_2^0)^{-3/4} & -\frac{1}{16}\lambda^0(x_1^0 x_2^0)^{-3/4} & \frac{3}{16}\lambda^0(x_1^0)^{1/4}(x_2^0)^{-7/4} \end{pmatrix}.$$

The determinant of $H(L(\lambda^0, x^0))$ is $|H(L(\lambda^0, x^0))| = -\lambda^0 (x_1^0 x_2^0)^{-5/4}/32 < 0$, and therefore, the sufficient second-order conditions for a constrained minimum are satisfied.

The components of the solution vector (x_1^0, x_2^0) can be written as functions of w and y, namely,

$$\hat{x}_1 : \mathbb{R}^3_{++} \mapsto \mathbb{R},\ (w, y) \mapsto (y)^2 \sqrt{\frac{w_2}{w_1}} \quad \text{and} \quad \hat{x}_2 : \mathbb{R}^3_{++} \mapsto \mathbb{R},\ (w, y) \mapsto (y)^2 \sqrt{\frac{w_1}{w_2}}.$$

These functions are the *conditional factor demand functions* for the technology described by the production function f. Substituting x^0 into the objective function, we obtain the *cost function*

$$C : \mathbb{R}^3_{++} \mapsto \mathbb{R},\ (w, y) \mapsto 2(y)^2 \sqrt{w_1 w_2}$$

of the firm, which gives us the minimal cost of producing y units of output if the factor prices are given by the vector w. The profit of the firm is given by

$$py - C(w, y) = py - 2(y)^2 \sqrt{w_1 w_2},$$

which is a concave function of y (Exercise: prove this). If we want to maximize profit by choice of the optimal amount of output y^0, we obtain the first-order condition

$$p - 4y^0 \sqrt{w_1 w_2} = 0,$$

and it follows that

$$y^0 = \frac{p}{4\sqrt{w_1 w_2}}.$$

This defines the same supply function as the one we derived in Section 4.4 for this production function. Substituting y^0 into the objective function yields the same profit function as in Section 4.4.

4.6 Optimization with Inequality Constraints

The technique introduced in the previous section can be applied to many optimization problems which occur in economic models, but there are others for which a more general methodology has to be employed. In particular, it is often the case that constraints cannot be expressed as equalities, and *inequalities* have to be used instead. In particular, if there are two or more inequality constraints, it cannot be expected that all constraints are satisfied with an equality at a solution. Consider the example illustrated in Figure 4.7. Assume the objective function is $f : \mathbb{R}^2 \mapsto \mathbb{R}$, and there are two constraint functions g^1 and g^2 where $g^j : \mathbb{R}^2 \mapsto \mathbb{R}$ for all $i = 1, 2$. The straight lines in the diagram represent the points $x \in \mathbb{R}^2$ such that $g^1(x) = 0$ and $g^2(x) = 0$, respectively, and the area to the southwest of each line represents the points $x \in \mathbb{R}^2$ such that $g^1(x) < 0$ and $g^2(x) < 0$, respectively. If we want to maximize f subject to the constraints $g^1(x) \leq 0$ and $g^2(x) \leq 0$, we obtain the solution x^0. It is easy to see that $g^2(x^0) < 0$ and, therefore, the second constraint is *not binding* at this solution—it is not satisfied with an equality.

First, we give a formal definition of constrained maxima and minima subject to $m \in \mathbb{N}$ inequality constraints. Note that the constraints are formulated as '$\leq$' constraints for a maximization problem and as '$\geq$' constraints for a minimization problem. This is a matter of convention, but it is important for the formulation of optimality conditions. Clearly, any inequality constraint can be written in the required form.

Definition 4.6.1 *Let $n, m \in \mathbb{N}$, let $A \subseteq \mathbb{R}^n$ be convex, and let $f : A \mapsto \mathbb{R}$ and $g^j : A \mapsto \mathbb{R}$ for all $j = 1, \ldots, m$. Furthermore, let $x^0 \in A$.*

(i) f has a global constrained maximum subject to the constraints $g^j(x) \leq 0$ for all $j = 1, \ldots, m$ *at x^0 if and only if $g^j(x^0) \leq 0$ for all $j = 1, \ldots, m$ and $f(x^0) \geq f(x)$ for all $x \in A$ such that $g^j(x) \leq 0$ for all $j = 1, \ldots, m$.*
(ii) f has a local constrained maximum subject to the constraints $g^j(x) \leq 0$ for all $j = 1, \ldots, m$ *at x^0 if and only if there exists $\varepsilon \in \mathbb{R}_{++}$ such that $g^j(x^0) \leq 0$ for all $j = 1, \ldots, m$ and $f(x^0) \geq f(x)$ for all $x \in \mathcal{U}_\varepsilon(x^0) \cap A$ such that $g^j(x) \leq 0$ for all $j = 1, \ldots, m$.*
(iii) f has a global constrained minimum subject to the constraints $g^j(x) \geq 0$ for all $j = 1, \ldots, m$ *at x^0 if and only if $g^j(x^0) \geq 0$ for all $j = 1, \ldots, m$ and $f(x^0) \leq f(x)$ for all $x \in A$*

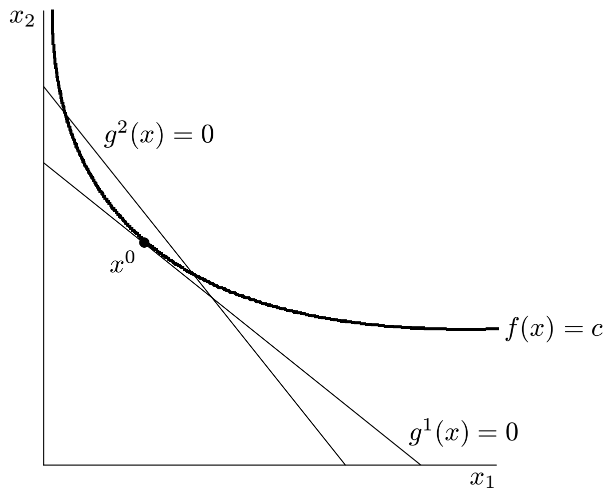


Figure 4.7: Two inequality constraints.

such that $g^j(x) \geq 0$ for all $j = 1, \ldots, m$.
(iv) f has a local constrained minimum subject to the constraints $g^j(x) \geq 0$ for all $j = 1, \ldots, m$ *at x^0 if and only if there exists $\varepsilon \in \mathbb{R}_{++}$ such that $g^j(x^0) \geq 0$ for all $j = 1, \ldots, m$ and $f(x^0) \leq f(x)$ for all $x \in \mathcal{U}_\varepsilon(x^0) \cap A$ such that $g^j(x) \geq 0$ for all $j = 1, \ldots, m$.*

To solve problems of that nature, we transform an optimization problem with inequality constraints into a problem with equality constraints and apply the results of the previous section. For simplicity, we provide a formal discussion only for problems involving two choice variables and one constraint and state the more general result for any number of choice variables and constraints without a proof. We focus on constrained *maximization* problems in this section—as usual, minimization problems can be viewed as maximization problems with reversed signs and thus can be dealt with analogously (see the exercises).

Let $A \subseteq \mathbb{R}^2$ be convex, and suppose we want to maximize $f : A \mapsto \mathbb{R}$ subject to the constraint $g(x) \leq 0$, where $g : A \mapsto \mathbb{R}$. As before, we assume that the constraint-qualification condition $g_{x_i}(x^0) \neq 0$ for at least one $i \in \{1, 2\}$ is satisfied at an interior point $x^0 \in A$ which we consider a candidate for a solution. Without loss of generality, suppose $g_{x_2}(x^0) \neq 0$. The inequality constraint $g(x) \leq 0$ can be transformed into an equality constraint by introducing a *slack variable* $s \in \mathbb{R}$. Given this additional variable, the constraint can equivalently be written as $g(x) + (s)^2 = 0$. Note that $(s)^2 = 0$ if $g(x) = 0$ and $(s)^2 > 0$ if $g(x) < 0$. We can now use the Lagrange method to solve the problem of maximizing f subject to the constraint $g(x) + (s)^2 = 0$. Note that the Lagrange function now has *four* arguments—the multiplier λ, the original choice variables x_1 and x_2, and the additional choice variable s. The necessary first-order conditions for a local constrained maximum at an interior point $x^0 \in A$ are

$$-g(x^0) - (s^0)^2 = 0 \tag{4.35}$$
$$f_{x_1}(x^0) - \lambda^0 g_{x_1}(x^0) = 0 \tag{4.36}$$
$$f_{x_2}(x^0) - \lambda^0 g_{x_2}(x^0) = 0 \tag{4.37}$$
$$-2\lambda^0 s^0 = 0. \tag{4.38}$$

Now we can eliminate the slack variable from these conditions in order to obtain first-order conditions in terms of the original variables. (4.35) is, of course, equivalent to the original constraint $g(x^0) \leq 0$. Furthermore, if $g(x) < 0$, it follows that $(s)^2 > 0$, and (4.38) requires that $\lambda^0 = 0$. Therefore, if the constraint is *not binding* at a solution, the corresponding value of the multiplier must be zero. This can be expressed by replacing (4.38) with the condition

$$\lambda^0 g(x^0) = 0.$$

Furthermore, the multiplier cannot be negative if we have a maximization problem with an inequality constraint (note that no sign restriction on the multiplier is implied in the case of an equality constraint).

To see why this is the case, suppose $\lambda^0 < 0$. Using our constraint qualification condition, (4.37) implies in this case that

$$\frac{f_{x_2}(x^0)}{g_{x_2}(x^0)} = \lambda^0 < 0.$$

This means that the two partial derivatives $f_{x_2}(x^0)$ and $g_{x_2}(x^0)$ have opposite signs. If $f_{x_2}(x^0) > 0$ and $g_{x_2}(x^0) < 0$, it follows that if we increase the value of x_2 from x_2^0 to $x_2^0 + h$ with $h \in \mathbb{R}_{++}$ sufficiently small, the value of f increases and the value of g decreases. But this means that $f(x_1^0, x_2^0 + h) > f(x^0)$ and $g(x_1^0, x_2^0 + h) < g(x^0) \leq 0$. Therefore, the point $(x_1^0, x_2^0 + h)$ satisfies the inequality constraint and leads to a higher value of the objective function f, contradicting the assumption that we have a local constrained maximum at x^0. Analogously, if $f_{x_2}(x^0) < 0$ and $g_{x_2}(x^0) > 0$, we can find a point satisfying the constraint and yielding a higher value of f than x^0 by decreasing the value of x_2. Therefore, the Lagrange multiplier must be nonnegative. To summarize these observations, if f has a local constrained maximum subject to the constraint $g(x) \leq 0$ at an interior point $x^0 \in A$, it follows that there exists $\lambda^0 \in \mathbb{R}_+$ such that

$$\begin{aligned}
g(x^0) &\leq 0 \\
\lambda^0 g(x^0) &= 0 \\
f_{x_1}(x^0) - \lambda^0 g_{x_1}(x^0) &= 0 \\
f_{x_2}(x^0) - \lambda^0 g_{x_2}(x^0) &= 0.
\end{aligned}$$

This technique can be used to deal with *nonnegativity constraints* imposed on the choice variables as well. Nonnegativity constraints appear naturally in economic models because the choice variables are often interpreted as nonnegative quantities. For example, suppose we want to maximize f subject to the constraint $x_2 \geq 0$. We can write this problem as a maximization problem with the equality constraint $-x_2 + (s)^2 = 0$, where, as before, $s \in \mathbb{R}$ is a slack variable. The Lagrange function for this problem is given by

$$L : \mathbb{R} \times A \times \mathbb{R} \mapsto \mathbb{R},\ (\lambda, x, s) \mapsto f(x) - \lambda(-x_2 + (s)^2).$$

We obtain the necessary first-order conditions

$$x_2^0 - (s^0)^2 = 0 \tag{4.39}$$
$$f_{x_1}(x^0) = 0 \tag{4.40}$$
$$f_{x_2}(x^0) + \lambda^0 = 0 \tag{4.41}$$
$$-2\lambda^0 s^0 = 0. \tag{4.42}$$

Again, it follows that $\lambda^0 \in \mathbb{R}_+$. If $\lambda^0 < 0$, (4.41) implies $f_{x_2}(x^0) = -\lambda^0 > 0$. Therefore, we can increase the value of the objective function by increasing the value of x_2, contradicting the assumption that we have a local constrained maximum at x^0. Therefore, (4.41) implies $f_{x_2}(x^0) \leq 0$.

If $x_2^0 > 0$, (4.39) and (4.42) imply $\lambda^0 = 0$, and by (4.41), it follows that $f_{x_2}(x^0) = 0$. Note that, in this case, we simply obtain the first-order conditions for an unconstrained maximum—the constraint is not binding. Therefore, in this special case of a single nonnegativity constraint, we can eliminate not only the slack variable, but also the multiplier λ^0. For that reason, nonnegativity constraints can be dealt with in a more simple manner than general inequality constraints. We obtain the conditions

$$\begin{aligned}
x_2^0 &\geq 0 \\
f_{x_1}(x^0) &= 0 \\
f_{x_2}(x^0) &\leq 0 \\
x_2 f_{x_2}(x^0) &= 0.
\end{aligned}$$

The above-described procedure can be generalized to problems involving any number of variables and inequality constraints, and nonnegativity constraints for all choice variables. In order to state a general result regarding necessary first-order conditions for these maximization problems, we first have to present a generalization of the constraint qualification condition.

Let $A \subseteq \mathbb{R}^n$ be convex, and suppose we want to maximize $f : A \mapsto \mathbb{R}$ subject to the nonnegativity constraints $x_i \geq 0$ for all $i = 1, \ldots, n$ and $m \in \mathbb{N}$ constraints $g^1(x) \leq 0, \ldots, g^m(x) \leq 0$, where $g^j : A \mapsto \mathbb{R}$ for all $j = 1, \ldots, m$. The *Jacobian matrix* of the functions $g^1, \ldots, g^m$ at the point $x^0 \in A$ is an $m \times n$

matrix $J(g^1(x^0), \ldots, g^m(x^0))$, where, for all $j = 1, \ldots, m$ and all $i = 1, \ldots, n$, the element in row j and column i is the partial derivative of g^j with respect to x_i. Therefore, this matrix can be written as

$$J(g^1(x^0), \ldots, g^m(x^0)) := \begin{pmatrix} g^1_{x_1}(x^0) & g^1_{x_2}(x^0) & \ldots & g^1 x_n(x^0) \\ g^2_{x_1}(x^0) & g^2_{x_2}(x^0) & \ldots & g^2 x_n(x^0) \\ \vdots & \vdots & & \vdots \\ g^m_{x_1}(x^0) & g^m_{x_2}(x^0) & \ldots & g^m x_n(x^0) \end{pmatrix}.$$

Let $\bar{J}(g^1(x^0), \ldots, g^m(x^0))$ be the submatrix of $J(g^1(x^0), \ldots, g^m(x^0))$ which is obtained by removing all rows j such that $g^j(x^0) < 0$ and all columns i such that $x^0_i = 0$. That is, $\bar{J}(g^1(x^0), \ldots, g^m(x^0))$ is the matrix of partial derivatives of all constraint functions such that the corresponding constraint is *binding* at x^0 with respect to all variables whose value is positive at x^0. The constraint qualification condition requires that $\bar{J}(g^1(x^0), \ldots, g^m(x^0))$ has its maximal possible rank. Notice that, in the case of a single constraint satisfied with an equality and no nonnegativity constraint, we get back the constraint qualification condition used earlier—at least one of the partial derivatives of the constraint function at x^0 must be nonzero.

The method to obtain necessary first-order conditions for a problem involving one inequality constraint can now be generalized. This is done by introducing a Lagrange multiplier for each constraint (except for the nonnegativity constraints which can be dealt with as described above), and then defining the Lagrange function by subtracting, for each constraint, the product of the multiplier and the constraint function from the objective function. Letting $\lambda = (\lambda_1, \ldots, \lambda_m) \in \mathbb{R}^m$ be the vector of multipliers, the Lagrange function is defined as

$$L : \mathbb{R}^m \times A \mapsto \mathbb{R}, \ (\lambda, x) \mapsto f(x) - \sum_{j=1}^{m} \lambda_j g^j(x).$$

The necessary first-order conditions for a local constrained maximum at an interior point $x^0 \in A$ require the existence of a vector of multipliers $\lambda^0 \in \mathbb{R}^m$ such that

$$\begin{aligned} L_{\lambda_j}(\lambda^0, x^0) &\geq 0 \quad \forall j = 1, \ldots, m \\ \lambda_j L_{\lambda_j}(\lambda^0, x^0) &= 0 \quad \forall j = 1, \ldots, m \\ L_{x_i}(\lambda^0, x^0) &\leq 0 \quad \forall i = 1, \ldots, n \\ x_i L_{x_i}(\lambda^0, x^0) &= 0 \quad \forall i = 1, \ldots, n \\ \lambda^0_j &\geq 0 \quad \forall j = 1, \ldots, m \\ x^0_i &\geq 0 \quad \forall i = 1, \ldots, n. \end{aligned}$$

Substituting the definition of the Lagrange function into these conditions, we obtain the following result.

Theorem 4.6.2 *Let $n, m \in \mathbb{N}$, let $A \subseteq \mathbb{R}^n$ be convex, and let $f : A \mapsto \mathbb{R}$, $g^j : A \mapsto \mathbb{R}$ for all $j = 1, \ldots, m$. Furthermore, let x^0 be an interior point of A. Suppose f and $g^1, \ldots, g^m$ are partially differentiable with respect to all variables in a neighborhood of x^0, and these partial derivatives are continuous at x^0. Suppose $\bar{J}(g^1(x^0), \ldots, g^m(x^0))$ has maximal rank. If f has a local constrained maximum subject to the constraints $g^j(x) \leq 0$ for all $j = 1, \ldots, m$ and $x_i \geq 0$ for all $i = 1, \ldots, n$ at x^0, then there exists $\lambda^0 \in \mathbb{R}^m$ such that*

$$\begin{aligned} g^j(x^0) &\leq 0 \quad \forall j = 1, \ldots, m \\ \lambda^0_j g^j(x^0) &= 0 \quad \forall j = 1, \ldots, m \\ f_{x_i}(x^0) - \sum_{j=1}^{m} \lambda^0_j g^j_{x_i}(x^0) &\leq 0 \quad \forall i = 1, \ldots, n \\ x^0_i \left(f_{x_i}(x^0) - \sum_{j=1}^{m} \lambda^0_j g^j_{x_i}(x^0) \right) &= 0 \quad \forall i = 1, \ldots, n \\ \lambda^0_j &\geq 0 \quad \forall j = 1, \ldots, m \\ x^0_i &\geq 0 \quad \forall i = 1, \ldots, n. \end{aligned}$$

The conditions stated in Theorem 4.6.2 are call the *Kuhn-Tucker conditions* for a maximization problem with inequality constraints.

Note that these first-order conditions are, in general, only *necessary* for a local constrained maximum. The following theorem states sufficient conditions for a *global* maximum. These conditions involve curvature properties of the objective function and the constraint functions. We obtain

Theorem 4.6.3 *Let $n, m \in \mathbb{N}$, let $A \subseteq \mathbb{R}^n$ be convex, and let $f : A \mapsto \mathbb{R}$, $g^j : A \mapsto \mathbb{R}$ for all $j = 1, \ldots, m$. Let x^0 be an interior point of A. Suppose f and $g^1, \ldots, g^m$ are twice partially differentiable with respect to all variables in a neighborhood of x^0, and these partial derivatives are continuous at x^0. Suppose $\bar{J}(g^1(x^0), \ldots, g^m(x^0))$ has maximal rank. If f is concave and g^j is convex for all $j = 1, \ldots, m$, then the Kuhn-Tucker conditions are necessary and sufficient for a global constrained maximum of f subject to the constraints $g^j(x) \leq 0$ for all $j = 1, \ldots, m$ and $x_i \geq 0$ for all $i = 1, \ldots, n$ at x^0.*

If the objective function f is *strictly* concave and the constraint functions are convex, it follows that f must have a *unique* global constrained maximum at a point satisfying the Kuhn-Tucker conditions. This is a consequence of the following theorem.

Theorem 4.6.4 *Let $n, m \in \mathbb{N}$, let $A \subseteq \mathbb{R}^n$ be convex, and let $f : A \mapsto \mathbb{R}$, $g^j : A \mapsto \mathbb{R}$ for all $j = 1, \ldots, m$. Let $x^0 \in A$. If f is strictly concave and g^j is convex for all $j = 1, \ldots, m$ and f has a global constrained maximum subject to the constraints $g^j(x) \leq 0$ for all $j = 1, \ldots, m$ and $x_i \geq 0$ for all $i = 1, \ldots, n$ at x^0, then f has a unique global constrained maximum at x^0.*

Proof. Suppose f has a global constrained maximum at x^0. By way of contradiction, suppose f has a global constrained maximum at $y \in A$ with $y \neq x^0$. By definition of a global constrained maximum, it follows that

$$\begin{aligned} f(x^0) &\geq f(x) \quad \forall x \in A \\ g^j(x^0) &\leq 0 \quad \forall j = 1, \ldots, m \\ x_i^0 &\geq 0 \quad \forall i = 1, \ldots, n \end{aligned}$$

and

$$\begin{aligned} f(y) &\geq f(x) \quad \forall x \in A \\ g^j(y) &\leq 0 \quad \forall j = 1, \ldots, m \\ y_i &\geq 0 \quad \forall i = 1, \ldots, n. \end{aligned}$$

Therefore, $f(x^0) \geq f(y)$ and $f(y) \geq f(x^0)$ and hence $f(x^0) = f(y)$. Let $z = x^0/2 + y/2$. By the convexity of A, $z \in A$. Because $x_i^0 \geq 0$ and $y_i \geq 0$ for all $i = 1, \ldots, n$, it follows that $z_i \geq 0$ for all $i = 1, \ldots, n$. Furthermore, because g^j is convex, we obtain $g^j(z) \leq g^j(x^0)/2 + g^j(y)/2$ and, because $g^j(x^0) \leq 0$ and $g^j(y) \leq 0$, $g^j(z) \leq 0$ for all $j = 1, \ldots, m$. Therefore, all required constraints are satisfied at z. Because f is strictly concave, it follows that $f(z) > f(x^0)/2 + f(y)/2 = f(x^0)$, contradicting the assumption that f has a global constrained maximum at x^0. ‖

Chapter 5

Difference Equations and Differential Equations

5.1 Complex Numbers

It is often desirable to have a richer set of numbers than $\mathbb{R}$ available. For example, the equation

$$x^2 + 1 = 0 \tag{5.1}$$

does not have a solution in $\mathbb{R}$: there exists no real number x such that the square of x is equal to -1. In order to obtain solutions to equations of that type, we introduce the *complex numbers*. The idea is to define a number such that the square of this number is equal to -1. We call this number i and it is characterized by the property

$$i^2 = (-i)^2 = -1.$$

Now we can define the set of complex numbers $\mathbb{C}$ as

$$\mathbb{C} := \{z \mid \exists a, b \in \mathbb{R} \text{ such that } z = a + bi\}.$$

The number a in this definition is the *real part* of the complex number z and b is the *imaginary part*. Clearly, $\mathbb{R} \subseteq \mathbb{C}$ because a real number is obtained whenever $b = 0$. Addition and multiplication of complex numbers is defined as follows.

Definition 5.1.1 *Let $z = (a + bi) \in \mathbb{C}$ and $z' = (a' + b'i) \in \mathbb{C}$.*
(i) The sum of z and z' is defined as $z + z' := (a + a') + (b + b')i$.
(ii) The product of z and z is defined as $zz' := (aa' - bb') + (ab' + a'b)i$.

For any complex number $z \in \mathbb{C}$, there exists a *conjugated complex number* $\overline{z}$, defined as follows.

Definition 5.1.2 *Let $z = (a + bi) \in \mathbb{C}$. The conjugated complex number of z is defined as $\overline{z} := a - bi$.*

The *absolute value* of a complex number is defined a the Euclidean norm of the two-dimensional vector composed of the real and the imaginary part of that number.

Definition 5.1.3 *Let $z = (a + bi) \in \mathbb{C}$. The absolute value of z is defined as $|z| := \sqrt{a^2 + b^2}$.*

As an immediate consequence of those definitions, we obain, for all $z = (a + bi) \in \mathbb{C}$,

$$z\overline{z} = (a + bi)(a - bi) = a^2 + b^2 = |z|^2.$$

In particular, this implies that $|\overline{z}| = |z|$ for all $z \in \mathbb{C}$ and

$$\frac{1}{z} = \frac{1}{|z|^2}\overline{z}$$

for all $z \in \mathbb{C}$ with $z \neq 0$.

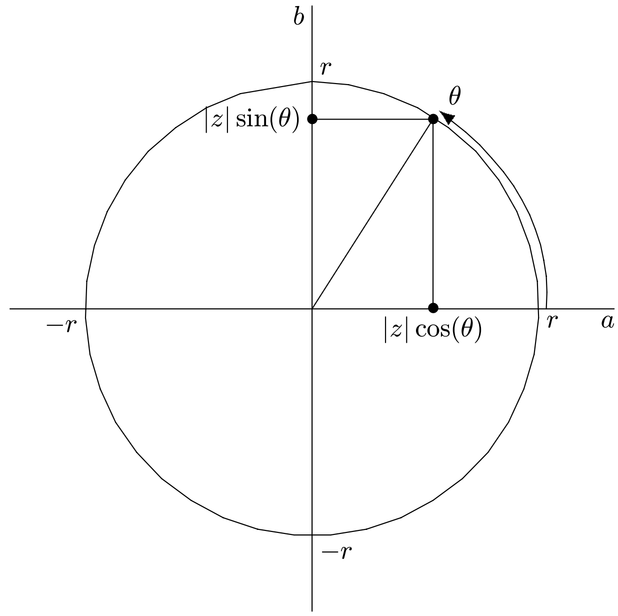


Figure 5.1: Complex numbers and polar coordinates.

Figure 5.1 contains a diagrammatic illustration of complex numbers. Using the trigonometric functions sin and cos, a representation of a complex number in terms of *polar coordinates* can be obtained. Given a complex number $z \neq 0$ represented by a vector $(a, b) \in \mathbb{R}^2$ (where a is the real part of z and b is the imaginary part of z), there exists a unique $\theta \in [0, 2\pi)$ such that $a = |z|\cos(\theta)$ and $b = |z|\sin(\theta)$. This θ is called the *argument* of $z \neq 0$. We define the argument of $z = 0$ as $\theta = 0$. With this definition of θ, we can write $z = (a + bi) \in \mathbb{C}$ as

$$z = |z|(\cos(\theta) + i\sin(\theta)). \tag{5.2}$$

The formulation of z in (5.2) is called the representation of z in terms of *polar coordinates*.

Using the properties of the trigonometric functions, we obtain

$$\begin{aligned} \overline{z} &= |z|(\cos(\theta) - i\sin(\theta)) \\ &= |z|(\cos(\theta) + i\sin(-\theta)) \\ &= |z|(\cos(-\theta) - i\sin(-\theta)) \end{aligned}$$

for all $z \in \mathbb{C}$, where θ is the argument of z. Thus, for any complex number z, the argument of its conjugated complex number $\overline{z}$ is obtained by multiplying the argument of z by -1.

We conclude this section with some important results regarding complex numbers, which we state without proofs. The first of those is *Moivre's theorem.* It is concerned with powers of complex numbers represented in terms of polar coordinates.

Theorem 5.1.4 *For all $\theta \in [0, 2\pi)$ and for all $n \in \mathbb{N}$,*

$$(\cos(\theta) + i\sin(\theta))^n = \cos(n\theta) + i\sin(n\theta).$$

Euler's theorem establishes a relationship between the exponential function and the trigonometric functions sin and cos.

Theorem 5.1.5 *For all $x \in \mathbb{R}$,*

$$e^{ix} = \cos(x) + i\sin(x).$$

The final result of this section returns to the question of the existence of solutions to specific equations. It turns out that any polynomial equation of degree $n \in \mathbb{N}$ has n (possibly multiple) solutions in $\mathbb{C}$. This result is called the *fundamental theorem of algebra*.

Theorem 5.1.6 *Let $n \in \mathbb{N}$, and let $a_0, \ldots, a_n \in \mathbb{C}$ be such that $a_n \neq 0$. The equation*

$$\sum_{j=0}^{n} a_j z_j = 0$$

has n solutions $z_1^*, \ldots, z_n^* \in \mathbb{C}$.

Multiple solutions are not ruled out by Theorem 5.1.6: it is possible to have $z_j^* = z_k^*$ for some $j, k \in \{1, \ldots, n\}$ with $j \neq k$.

5.2 Difference Equations

Many economic studies involve the behavior of economic variables over time. For example, macroeconomic models designed to describe changes in national income, in interest rates or in the rate of unemployment have to employ techniques that go beyond the static analysis carried out so far. If time is thought of as a discrete variable (that is, the values of the variables in question are measured at discrete times), *difference equations* can be used as important tools. If time is represented as a continuum, techniques for the solution of *differential equations* are essential. We discuss difference equations in this section and postpone our analysis of differential equations until the end of the chapter because the methods involved in finding their solutions make use of integration, to be discussed in the following section.

Suppose we observe the value of a variable in each period (a period could be a month or a year, for instance) starting at an initial period which we can, without loss of generality, call period zero. Formally, the development of this variable over time can then be expressed by means of a function $y : \mathbb{N}_0 \mapsto \mathbb{R}$ where, for each period $t \in \mathbb{N}_0$, $y(t)$ is the value of the variable in period t. Now we can define the notion of a difference equation.

Definition 5.2.1 *Let $n \in \mathbb{N}$. A* difference equation of order n *is an equation*

$$y(t+n) = F(t, y(t), y(t+1), \ldots, y(t+n-1)) \tag{5.3}$$

where $F : \mathbb{N}_0 \times \mathbb{R}^n \mapsto \mathbb{R}$ is a given function which is non-constant in its second argument. A function $y : \mathbb{N}_0 \mapsto \mathbb{R}$ is a solution *of this difference equation if and only if (5.3) is satisfied for all $t \in \mathbb{N}_0$.*

In some sources, more general differential equations are considered, where the equation cannot necessarily be solved explicitly for the value $y(t+n)$ and, instead, the relationship between the different values of y is expressed in implicit form only. However, the special case introduced in the above definition is sufficient for our purposes.

Note that the function F is assumed to be non-constant in its second argument $y(t)$. This is to ensure that the equation is indeed of order n. If F is constant in $y(t)$, the function is at most of order $n-1$ because, in this case, the value of y at $t+n$ depends (at most) on the values attained in the previous $n-1$ periods rather than in the previous n periods.

We begin with a simple observation regarding the solutions of difference equations. If we have a difference equation of order n and the n initial values $y(0), \ldots, y(n-1)$ are given, there exists a unique solution that can be obtained recursively.

Theorem 5.2.2 *If the values $y(0), \ldots, y(n-1)$ are determined, then (5.3) has a unique solution.*

Proof. Substituting $y(0), \ldots, y(n-1)$ into (5.3) for $t = 0$, we obtain

$$y(n) = F(0, y(0), \ldots, y(n-1)).$$

Now that $y(n)$ is determined, we have all the information necessary to calculate $y(n+1)$ and, substituting into (5.3) for $t = 1$, we obtain

$$y(n+1) = F(1, y(1), \ldots, y(n-1), y(n)).$$

This procedure can be repeated for $n+2, n+3, \ldots$ and the entire function y is uniquely determined. ‖

The recursive method illustrated in the above theorem is rather inconvenient when it comes to finding the function y explicitly—note that an infinite number of steps are involved. Moreover, it does not allow us to get a general idea about the relationship between t and the values of y described by the equation. For these reasons, it is desirable to obtain general methods to obtain solutions of difference equations in analytical form. Because this can be a rather difficult task for general functions F, we restrict attention to specific types of equations. First, we consider the case where F has a linear structure.

Definition 5.2.3 *Let $n \in \mathbb{N}$. A* linear difference equation of order n *is an equation*

$$y(t+n) = b(t) + a_0(t)y(t) + a_1(t)y(t+1) + \ldots + a_{n-1}(t)y(t+n-1) \tag{5.4}$$

where $b : \mathbb{N}_0 \mapsto \mathbb{R}$ and $a_t : \mathbb{N}_0 \mapsto \mathbb{R}$ for all $t \in \{0, \ldots, n-1\}$ are given functions, and there exists $t \in \mathbb{N}_0$ such that $a_0(t) \neq 0$. If $b(t) = 0$ for all $t \in \mathbb{N}_0$, the equation is a homogeneous linear difference equation of order n. *If there exists $t \in \mathbb{N}_0$ such that $b(t) \neq 0$, the equation (5.4) is an* inhomogeneous linear difference equation of order n.

The assumption that the function a_0 is not identically equal to zero ensures that the equation is indeed of order n.

The *homogeneous equation associated with (5.4)* is the equation

$$y(t+n) = a_0(t)y(t) + a_1(t)y(t+1) + \ldots + a_{n-1}(t)y(t+n-1). \tag{5.5}$$

The following result establishes that, given a particular solution of a linear difference equation, *any* solution of this equation can be obtained as the sum of the particular solution and a suitably chosen solution of the homogeneous equation associated with the original equation. Conversely, any sum of the particular solution and an arbitrary solution of the associated homogeneous equation is a solution of the original equation.

Theorem 5.2.4 *(i) Suppose $\hat{y}$ is a solution of (5.4). For each solution y of (5.4), there exists a solution z of the homogeneous equation (5.5) associated with (5.4) such that $y = z + \hat{y}$.*

(ii) If $\hat{y}$ is a solution of (5.4) and z is a solution of the homogeneous equation (5.5) associated with (5.4), then the function y defined by $y = z + \hat{y}$ is a solution of (5.4).

Proof. (i) Suppose $\hat{y}$ is a solution of (5.4). For an arbitrary solution y of (5.4), define the function $z : \mathbb{N}_0 \mapsto \mathbb{R}$ by

$$z(t) := y(t) - \hat{y}(t) \quad \forall t \in \mathbb{N}_0. \tag{5.6}$$

This implies

$$\begin{aligned} z(t+n) &= y(t+n) - \hat{y}(t+n) \\ &= b(t) + a_0(t)y(t) + a_1(t)y(t+1) + \ldots + a_{n-1}(t)y(t+n-1) \\ &- (b(t) + a_0(t)\hat{y}(t) + a_1(t)\hat{y}(t+1) + \ldots + a_{n-1}(t)\hat{y}(t+n-1)) \\ &= a_0(t)(y(t) - \hat{y}(t)) + a_1(t)(y(t+1) - \hat{y}(t+1)) + \ldots + a_{n-1}(t)(y(t+n-1) - \hat{y}(t+n-1)) \\ &= a_0(t)z(t) + a_1(t)z(t+1) + \ldots + a_{n-1}(t)z(t+n-1) \end{aligned}$$

for all $t \in \mathbb{N}$. Therefore, z is a solution of the homogeneous equation associated with (5.4). By (5.6), we have $y = z + \hat{y}$, which completes the proof of (i).

(ii) can be proven by substituting into (5.4). ‖

The methodology employed to identify solutions of linear difference equations makes use of Theorem 5.2.4. According to this theorem, it is sufficient to find *one* particular solution of the equation (5.4) and the general solution (that is, the set of all solutions) of the associated homogeneous equation (5.5)—all solutions of the (inhomogeneous) equation (5.4) can then be obtained according to the theorem. Our next result establishes that with any n solutions $z_1, \ldots, z_n$ of the homogeneous equation (5.5), all linear combinations of $z_1, \ldots, z_n$ are solutions of (5.5) as well.

Theorem 5.2.5 *If $z_1, \ldots, z_n$ are n solutions of (5.5) and $c = (c_1, \ldots, c_n) \in \mathbb{R}^n$ is a vector of arbitrary coefficients, then the function z defined by $z = \sum_{i=1}^n c_i z_i$ is a solution of (5.5).*

Proof. Suppose $z_1, \ldots, z_n$ are n solutions of (5.5) and $c = (c_1, \ldots, c_n) \in \mathbb{R}^n$ is a vector of arbitrary coefficients. Let $z = \sum_{i=1}^n c_i z_i$. Using this definition and (5.5), we obtain

$$\begin{aligned} z(t+n) &= \sum_{i=1}^n c_i z_i(t+n) \\ &= \sum_{i=1}^n c_i (a_0(t) z_i(t) + a_1(t) z_i(t+1) + \ldots + a_{n-1}(t) z_i(t+n-1)) \\ &= a_0(t) \sum_{i=1}^n c_i z_i(t) + a_1(t) \sum_{i=1}^n c_i z_i(t+1) + \ldots + a_{n-1}(t) \sum_{i=1}^n c_i z_i(t+n-1) \\ &= a_0(t) z(t) + a_1(t) z(t+1) + \ldots + a_{n-1}(t) z(t+n-1) \end{aligned}$$

for all $t \in \mathbb{N}_0$ and, therefore, z is a solution of (5.5). ||

Conversely, we can establish a necessary and sufficient condition under which the entire set of solutions of the homogeneous equation (5.5) can be obtained as a linear combination of n solutions. That is, given this condition, finding n solutions of a homogeneous linear difference equation of order n is sufficient to identify the general solution of this homogeneous equation. The condition requires that the n known solutions $z_1, \ldots, z_n$ are linearly independent in the sense specified in the following theorem.

Theorem 5.2.6 *Suppose $z_1, \ldots, z_n$ are n solutions of (5.5). The following two statements are equivalent.*
(i) For every solution z of (5.5), there exists a vector of coefficients $c \in \mathbb{R}^n$ such that $z = \sum_{i=1}^n c_i z_i$.
(ii)

$$\begin{vmatrix} z_1(0) & z_2(0) & \ldots & z_n(0) \\ z_1(1) & z_2(1) & \ldots & z_n(1) \\ \vdots & \vdots & & \vdots \\ z_1(n-1) & z_2(n-1) & \ldots & z_n(n-1) \end{vmatrix} \neq 0. \tag{5.7}$$

Proof. Suppose $z_1, \ldots, z_n$ are n solutions of (5.4). We first prove that (i) implies (ii). Suppose (ii) is violated, that is, the determinant in (5.7) is equal to zero. This implies that the row vectors of this matrix are linearly dependent and, thus, we can express one of them as a linear combination of the remaining row vectors. Without loss of generality, suppose the first row vector is such a linear combination of the others. Hence, there exist $n-1$ coefficients $\alpha_1, \ldots, \alpha_{n-1} \in \mathbb{R}$ such that

$$z_i(0) = \sum_{t=1}^{n-1} \alpha_t z_i(t) \quad \forall i \in \{1, \ldots, n\}. \tag{5.8}$$

We complete the proof of this part by showing that (i) must be violated as well, that is, there exists a solution z of (5.5) that cannot be expressed as a linear combination of the n solutions $z_1, \ldots, z_n$. Define $z(0) = 1$ and $z(t) = 0$ for all $t \in \{1, \ldots, n-1\}$, and let $z(t+n)$ be determined by (5.5) for all $t \in \mathbb{N}_0$. By definition, this function is a solution of (5.5). In order for z to be a linear combination of $z_1, \ldots, z_n$, there must exist n coefficients $c_1, \ldots, c_n \in \mathbb{R}$ such that $z(t) = \sum_{i=1}^n c_i z_i(t)$ for all $t \in \mathbb{N}_0$. This implies, in particular, that

$$\sum_{i=1}^n c_i z_i(0) = 1 \tag{5.9}$$

and

$$\sum_{i=1}^n c_i z_i(t) = 0 \quad \forall t \in \{1, \ldots, n-1\}. \tag{5.10}$$

Now successively add $-\alpha_1$ times (5.10) with $t = 1$, $-\alpha_2$ times (5.10) with $t = 2$,. . ., $-\alpha_{n-1}$ times (5.10) with $t = 2$ to (5.9). By (5.8) and the definition of z, (5.9) becomes

$$0 = 1,$$

which means that the system given by (5.9) and (5.10) cannot have a solution and, hence, z cannot be expressed as a linear combination of $z_1, \ldots, z_n$.

To prove that (ii) implies (i), suppose (5.7) is satisfied, and let z be an arbitrary solution of (5.5). By Theorem 5.2.2, the values $z(0), \ldots, z(n-1)$ are sufficient to determine all values of z. Thus, (i) is established if we can find a vector $c = (c_1, \ldots, c_n) \in \mathbb{R}^n$ such that

$$z(t) = \sum_{i=1}^{n} c_i z_i(t) \quad \forall t \in \{0, \ldots, n-1\}. \tag{5.11}$$

(5.11) is a system of n linear equations in the n variables $c_1, \ldots, c_n$, where the matrix of coefficients is given by the entries in (5.7). By assumption, the determinant of this matrix of coefficients is different from zero, which means that the matrix has full rank n. By Theorem 2.3.5, this implies that the system of equations (5.11) has a unique solution $(c_1, \ldots, c_n)$. By definition, these coefficients are such that $z = \sum_{i=1}^{n} c_i z_i$. ‖

We now turn to a special case of linear difference equations, where the functions $a_0, \ldots, a_{n-1}$ in (5.4) are constant.

Definition 5.2.7 *Let $n \in \mathbb{N}$. A* linear difference equation with constant coefficients of order n *is an equation*

$$y(t+n) = b(t) + a_0 y(t) + a_1 y(t+1) + \ldots + a_{n-1} y(t+n-1) \tag{5.12}$$

where $b : \mathbb{N}_0 \mapsto \mathbb{R}$, $a_0 \in \mathbb{R} \setminus \{0\}$ and $a_1, \ldots, a_{n-1} \in \mathbb{R}$. If $b(t) = 0$ for all $t \in \mathbb{N}_0$, the equation is a homogeneous linear difference equation with constant coefficients of order n. *If there exists $t \in \mathbb{N}_0$ such that $b(t) \neq 0$, the equation (5.12) is an* inhomogeneous linear difference equation with constant coefficients of order n.

We only consider linear difference equations with constant coefficients of order one and of order two. Moreover, in the case of equations of order two, we restrict attention to specific functional forms of the inhomogeneity b. This is the case because solving higher-order equations is a complex task and general methods are not easy to formulate. Using Theorem 5.2.4, we employ the following strategy. We determine the general solution of the associated homogeneous equation and then find a particular solution of the inhomogeneous equation. By Theorem 5.2.4, this yields the general solution of the inhomogeneous equation: any solution can be obtained as the sum of the particular solution and a suitably chosen solution of the homogeneous equation.

We begin with equations of order one. As mentioned above, we first discuss methods for determining all solutions of a homogeneous linear equation with constant coefficients of order one. Thus, the equations to be solved at this stage are of the form

$$y(t+1) = a_0 y(t) \quad \forall t \in \mathbb{N}_0 \tag{5.13}$$

where $a_0 \in \mathbb{R} \setminus \{0\}$.

An obvious solution is the function that assigns a value of zero to all $t \in \mathbb{N}_0$. According to Theorem 5.2.6, in the case of an equation of order $n = 1$, it is sufficient to obtain one solution z_1, *provided that the determinant in the theorem statement is non-zero* for this solution. Clearly, this is not the case for the solution that is identically equal to zero and, thus, it cannot be used to obtain the general solution of (5.13). Therefore, we have to search for a solution with the independence property stated in Theorem 5.2.6.

Due to the nature of the equation, a generalized-exponential function turns out to be suitable. That is, we assume that we have a solution z_1 with $z_1(t) = \lambda^t$ for all $t \in \mathbb{N}_0$ and determine whether we can find a value of the parameter $\lambda \in \mathbb{R} \setminus \{0\}$ such that (5.13) is satisfied. In that case, the equation requires that

$$z_1(t+1) = \lambda^{t+1} = a_0 z_1(t) = a_0 \lambda^t \quad \forall t \in \mathbb{N}_0.$$

Thus, we must have

$$\lambda^{t+1} - a_0 \lambda^t = 0 \quad \forall t \in \mathbb{N}_0. \tag{5.14}$$

Because $\lambda \neq 0$ by assumption, we can divide both sides of (5.14) by λ^t to obtain

$$\lambda - a_0 = 0. \tag{5.15}$$

The left side of (5.15) is the *characteristic polynomial* of the equation (5.13), and (5.15) is the corresponding *characteristic equation*. Solving, we obtain $\lambda = a_0$ and, substituting into (5.13), it is easy to verify that the function z_1 defined by $z_1(t) = a_0^t$ for all $t \in \mathbb{N}_0$ is indeed a solution of the homogeneous equation. Furthermore, the determinant of the 1×1 matrix $(z_1(0))$ is equal to one and, by Theorem 5.2.6, a function $z : \mathbb{N}_0 \mapsto \mathbb{R}$ is a solution of (5.13) if and only if there exists a constant $c_1 \in \mathbb{R}$ such that

$$z(t) = c_1 z_1(t) = c_1 a_0^t \quad \forall t \in \mathbb{N}_0. \tag{5.16}$$

Now we consider inhomogeneous equations with constant coefficients of order one. The general form of the equation is

$$y(t+1) = b(t) + a_0 y(t) \quad \forall t \in \mathbb{N}_0 \tag{5.17}$$

where $a_0 \in \mathbb{R} \setminus \{0\}$ and $b : \mathbb{N}_0 \mapsto \mathbb{R}$ is such that there exists $t \in \mathbb{N}_0$ with $b(t) \neq 0$. Because we already have the general solution of the associated homogeneous equation, it is sufficient to find one particular solution of (5.17). This is the case because, according to Theorem 5.2.4, any solution can be expressed as the sum of this particular solution and a solution of the corresponding homogeneous equation.

We now verify that the function $\hat{y} : \mathbb{N}_0 \mapsto \mathbb{R}$ defined by

$$\hat{y}(t) = \begin{cases} 0 & \text{if } t = 0, \\ \sum_{k=1}^{t} a_0^{t-k} b(k-1) & \text{if } t \in \mathbb{N} \end{cases}$$

satisfies (5.17). For $t = 0$, we obtain

$$\hat{y}(t+1) = \hat{y}(1) = b(0) = b(0) + a_0 \cdot 0 = b(0) + a_0 \hat{y}(0) = b(t) + a_0 \hat{y}(t)$$

and, thus, (5.17) is satisfied. For $t \in \mathbb{N}$, substitution into the definition of $\hat{y}$ yields

$$\begin{aligned} \hat{y}(t+1) &= \sum_{k=1}^{t+1} a_0^{t+1-k} b(k-1) = \sum_{k=1}^{t+1} a_0 a_0^{t-k} b(k-1) \\ &= a_0 \sum_{k=1}^{t} a_0^{t-k} b(k-1) + a_0 a_0^{-1} b(t) \\ &= b(t) + a_0 \hat{y}(t) \end{aligned}$$

and, again, (5.17) is satisfied. We have therefore found a particular solution of (5.17) and, according to Theorem 5.2.4, we obtain the general solution as the sum of $\hat{y}$ and the solution of the associated homogeneous equation (5.16). Thus, we have established

Theorem 5.2.8 *A function $y : \mathbb{N}_0 \mapsto \mathbb{R}$ is a solution of the linear difference equation with constant coefficients of order one (5.17) if and only if there exists $c_1 \in \mathbb{R}$ such that*

$$y(t) = \begin{cases} c_1 & \text{if } t = 0, \\ c_1 a_0^t + \sum_{k=1}^{t} a_0^{t-k} b(k-1) & \text{if } t \in \mathbb{N} \end{cases} \quad \forall t \in \mathbb{N}_0. \tag{5.18}$$

If an initial value y_0 for $y(0)$ is specified, we obtain a unique solution because this initial value determines the value of the parameter c_1; see also Theorem 5.2.2. In particular, if the initial value is given by $y(0) = y_0 \in \mathbb{R}$, substitution into (5.18) yields $c_1 = y_0$ and, thus, we obtain the unique solution

$$y(t) = \begin{cases} y_0 & \text{if } t = 0, \\ y_0 a_0^t + \sum_{k=1}^{t} a_0^{t-k} b(k-1) & \text{if } t \in \mathbb{N} \end{cases} \quad \forall t \in \mathbb{N}_0.$$

This covers all possibilities that can arise in the case of linear difference equations with constant coefficients of order one.

To solve linear difference equations with constant coefficients of order two, we again begin with a procedure for finding the general solution of the associated homogeneous equation. Thus, we consider equations of the form

$$y(t+2) = a_0 y(t) + a_1 y(t+1) \quad \forall t \in \mathbb{N}_0 \tag{5.19}$$

where $a_0 \in \mathbb{R} \setminus \{0\}$ and $a_1 \in \mathbb{R}$. Because $n = 2$, we now have to search for two solutions of (5.19) such that the independence condition of Theorem 5.2.6 is satisfied. Again, a generalized exponential function can be employed in the first stage. That is, we consider functions z of the form $z(t) = \lambda^t$ for all $t \in \mathbb{N}_0$

and determine whether we can find values of the parameter $\lambda \in \mathbb{R} \setminus \{0\}$ such that (5.19) is satisfied. We obtain

$$z(t+2) = \lambda^{t+2} = a_0\lambda^t + a_1\lambda^{t+1} = a_0 z(t) + a_1 z(t+1) \quad \forall t \in \mathbb{N}_0.$$

Thus, we must have

$$\lambda^{t+2} - a_0\lambda^t - a_1\lambda^{t+1} = 0 \quad \forall t \in \mathbb{N}_0.$$

Dividing both sides by λ^t, we obtain

$$\lambda^2 - a_0 - a_1\lambda = 0. \tag{5.20}$$

Again, (5.20) is called the characteristic equation and the left side of (5.20) is the characteristic polynomial for the homogeneous differential equation under consideration. Because we have an equation of order two, this polynomial is of order two as well.

Using the standard technique for finding the solution of a quadratic equation, we can distinguish three cases.

Case I: $a_1^2/4 + a_0 > 0$. In this case, the characteristic polynomial has two distinct real roots given by $\lambda_1 = a_1/2 + \sqrt{a_1^2/4 + a_0}$ and $\lambda_2 = a_1/2 - \sqrt{a_1^2/4 + a_0}$. Substituting back into the expression for a possible solution, we obtain the two functions z_1 and z_2 defined by

$$z_1(t) = \left(a_1/2 + \sqrt{a_1^2/4 + a_0}\right)^t \quad \forall t \in \mathbb{N}_0$$

and

$$z_2(t) = \left(a_1/2 - \sqrt{a_1^2/4 + a_0}\right)^t \quad \forall t \in \mathbb{N}_0.$$

We obtain $z_1(0) = z_2(0) = 1$, $z_1(1) = a_1/2 + \sqrt{a_1^2/4 + a_0}$ and $z_2(1) = a_1/2 - \sqrt{a_1^2/4 + a_0}$. Therefore,

$$\begin{vmatrix} z_1(0) & z_2(0) \\ z_1(1) & z_2(1) \end{vmatrix} = \begin{vmatrix} 1 & 1 \\ a_1/2 + \sqrt{a_1^2/4 + a_0} & a_1/2 - \sqrt{a_1^2/4 + a_0} \end{vmatrix} = -2\sqrt{a_1^2/4 + a_0}$$

which, by assumption, is negative and thus different from zero. Thus, according to Theorem 5.2.6, a function $z : \mathbb{N}_0 \mapsto \mathbb{R}$ is a solution of the homogeneous equation (5.19) if and only if there exist constants $c_1, c_2 \in \mathbb{R}$ such that

$$z(t) = c_1 z_1(t) + c_2 z_2(t) = c_1\left(a_1/2 + \sqrt{a_1^2/4 + a_0}\right)^t + c_2\left(a_1/2 - \sqrt{a_1^2/4 + a_0}\right)^t \quad \forall t \in \mathbb{N}_0. \tag{5.21}$$

Case II: $a_1^2/4 + a_0 = 0$. Now the characteristic polynomial has a double root at $\lambda = a_1/2$, and the corresponding solution is given by $z_1(t) = (a_1/2)^t$ for all $t \in \mathbb{N}_0$. Because $n = 2$, we need a second solution (which, together with z_1, satisfies the required independence condition) in order to obtain the set of all solutions. Clearly, the approach involving a generalized exponential function cannot be used for that purpose—only the solution z_1 can be obtained in this fashion. Therefore, we need to try another functional form. In the case under discussion, it turns out that the product of t and z_1 will work. By substitution into (5.19), it can easily be verified that the function z_2 defined by

$$z_2(t) = t(a_1/2)^t \quad \forall t \in \mathbb{N}_0$$

is another solution. To verify that the two solutions z_1 and z_2 satisfy the required independence condition, note that we have $z_1(0) = 1$, $z_2(0) = 0$ and $z_1(1) = z_2(1) = a_1/2$. Therefore,

$$\begin{vmatrix} z_1(0) & z_2(0) \\ z_1(1) & z_2(1) \end{vmatrix} = \begin{vmatrix} 1 & 0 \\ a_1/2 & a_1/2 \end{vmatrix} = a_1/2.$$

By assumption, $a_1^2/4 = -a_0 \neq 0$ and, therefore, $a_1/2 \neq 0$. Thus, according to Theorem 5.2.6, a function $z : \mathbb{N}_0 \mapsto \mathbb{R}$ is a solution of the homogeneous equation (5.19) if and only if there exist constants $c_1, c_2 \in \mathbb{R}$ such that

$$z(t) = c_1 z_1(t) + c_2 z_2(t) = c_1(a_1/2)^t + c_2 t(a_1/2)^t \quad \forall t \in \mathbb{N}_0. \tag{5.22}$$

Case III: $a_1^2/4 + a_0 < 0$. The characteristic polynomial has two complex roots, namely, $\lambda_1 = a_1/2 + i\sqrt{-a_1^2/4 - a_0}$ and $\lambda_2 = a_1/2 - i\sqrt{-a_1^2/4 - a_0}$. Note that λ_2 is the conjugated complex number associated with λ_1. Substituting back into the expression for a possible solution, we obtain the two functions $\hat{z}_1$ and $\hat{z}_2$ defined by

$$\hat{z}_1(t) = \left(a_1/2 + i\sqrt{-a_1^2/4 - a_0}\right)^t \quad \forall t \in \mathbb{N}_0$$

and

$$\hat{z}_2(t) = \left(a_1/2 - i\sqrt{-a_1^2/4 - a_0}\right)^t \quad \forall t \in \mathbb{N}_0.$$

It would be desirable to obtain solutions that can be expressed without having to resort to complex numbers. Especially if a solution represents the values of an economic variable, it is difficult to come up with a reasonable interpretation of complex numbers. Fortunately, these solutions can be reformulated in terms of real numbers by employing polar coordinates. Letting $\theta \in [0, 2\pi)$ be such that $a_1/2 = \sqrt{-a_0}\cos(\theta)$ and $\sqrt{-a_1^2/4 - a_0} = \sqrt{-a_0}\sin(\theta)$, the first root of the characteristic polynomial can be written as $\lambda_1 = \sqrt{-a_0}(\cos(\theta) + i\sin(\theta))$. Analogously, we obtain $\lambda_2 = \overline{\lambda_1} = \sqrt{-a_0}(\cos(\theta) - i\sin(\theta))$. Thus, using Moivre's theorem (Theorem 5.1.4), we obtain

$$\hat{z}_1(t) = (\sqrt{-a_0})^t\left((\cos(\theta) + i\sin(\theta)\right)^t = (\sqrt{-a_0})^t\left((\cos(t\theta) + i\sin(t\theta)\right)$$

and

$$\hat{z}_2(t) = (\sqrt{-a_0})^t\left((\cos(\theta) - i\sin(\theta)\right)^t = (\sqrt{-a_0})^t\left((\cos(t\theta) - i\sin(t\theta)\right)$$

for all $t \in \mathbb{N}_0$. As can be verified easily, the result of Theorem 5.2.5 remains true if the possible values of the solutions and the coefficients are extended to cover complex numbers as well. Thus, all linear combinations (with complex coefficients) of the two solutions $\hat{z}_1$ and $\hat{z}_2$ are solutions of (5.19) as well. In particular, choosing the vector of coefficients $(1/2, 1/2)$, we obtain that the function z_1 defined by

$$z_1(t) = \hat{z}_1(t)/2 + \hat{z}_2(t)/2 = \sqrt{-a_0}\cos(t\theta) \quad \forall t \in \mathbb{N}_0$$

is a solution. Analogously, choosing the vector of coefficients $(1/(2i), -1/(2i))$, it follows that

$$z_2(t) = \hat{z}_1(t)/(2i) - \hat{z}_2(t)/(2i) = \sqrt{-a_0}\sin(t\theta) \quad \forall t \in \mathbb{N}_0$$

is another solution. We have

$$\begin{vmatrix} z_1(0) & z_2(0) \\ z_1(1) & z_2(1) \end{vmatrix} = \begin{vmatrix} 1 & 0 \\ \sqrt{-a_0}\cos(\theta) & \sqrt{-a_0}\sin(\theta) \end{vmatrix} = \sqrt{-a_0}\sin(\theta)$$

which, by assumption, is non-zero. Therefore, z is a solution of (5.19) if and only if there exist $c_1, c_2 \in \mathbb{R}$ such that

$$z(t) = c_1 z_1(t) + c_2 z_2(t) = c_1(\sqrt{-a_0})^t\cos(t\theta) + c_2(\sqrt{-a_0})^t\sin(t\theta) \quad \forall t \in \mathbb{N}_0. \tag{5.23}$$

We summarize our observations regarding the solution of homogeneous linear difference equations with constant coefficients of order two in the following theorem.

Theorem 5.2.9 *(i) Suppose $a_0 \in \mathbb{R} \setminus \{0\}$ and $a_1 \in \mathbb{R}$ are such that $a_1^2/4 + a_0 > 0$. A function $z : \mathbb{N}_0 \mapsto \mathbb{R}$ is a solution of the homogeneous linear difference equation with constant coefficients of order two (5.19) if and only if there exist $c_1, c_2 \in \mathbb{R}$ such that*

$$z(t) = c_1\left(a_1/2 + \sqrt{a_1^2/4 + a_0}\right)^t + c_2\left(a_1/2 - \sqrt{a_1^2/4 + a_0}\right)^t \quad \forall t \in \mathbb{N}_0.$$

(ii) Suppose $a_0 \in \mathbb{R} \setminus \{0\}$ and $a_1 \in \mathbb{R}$ are such that $a_1^2/4 + a_0 = 0$. A function $z : \mathbb{N}_0 \mapsto \mathbb{R}$ is a solution of the homogeneous linear difference equation with constant coefficients of order two (5.19) if and only if there exist $c_1, c_2 \in \mathbb{R}$ such that

$$z(t) = c_1(a_1/2)^t + c_2 t(a_1/2)^t \quad \forall t \in \mathbb{N}_0.$$

(iii) Suppose $a_0 \in \mathbb{R} \setminus \{0\}$ and $a_1 \in \mathbb{R}$ are such that $a_1^2/4 + a_0 < 0$. A function $z : \mathbb{N}_0 \mapsto \mathbb{R}$ is a solution of the homogeneous linear difference equation with constant coefficients of order two (5.19) if and only if there exist $c_1, c_2 \in \mathbb{R}$ such that

$$z(t) = c_1(\sqrt{-a_0})^t\cos(t\theta) + c_2(\sqrt{-a_0})^t\sin(t\theta) \quad \forall t \in \mathbb{N}_0$$

where $\theta \in [0, 2\pi)$ is such that $a_1/2 = \sqrt{-a_0}\cos(\theta)$.

To illustrate the technique used to obtain the general solution of (5.19), we consider a few examples.

Let $y(t+2) = -2y(t) + 4y(t+1)$ for all $t \in \mathbb{N}_0$. We begin by trying to find two independent solutions using a generalized-exponential functional form, that is, $z(t) = \lambda^t$ for all $t \in \mathbb{N}_0$ with $\lambda \in \mathbb{R} \setminus \{0\}$ to be determined. Substituting into our equation, we obtain $\lambda^{t+2} = -2\lambda^t + 4\lambda^{t+1}$. Dividing by λ^t and rearranging results in the characteristic equation $\lambda^2 - 4\lambda + 2 = 0$. This quadratic equation has the two real solutions $\lambda_1 = 2 + \sqrt{2}$ and $\lambda_2 = 2 - \sqrt{2}$. Therefore, we obtain the solutions

$$z_1(t) = (2 + \sqrt{2})^t \quad \forall t \in \mathbb{N}_0$$

and

$$z_2(t) = (2 - \sqrt{2})^t \quad \forall t \in \mathbb{N}_0.$$

Because

$$\begin{vmatrix} z_1(0) & z_2(0) \\ z_1(1) & z_2(1) \end{vmatrix} = \begin{vmatrix} 1 & 1 \\ 2+\sqrt{2} & 2-\sqrt{2} \end{vmatrix} = -2\sqrt{2} \neq 0,$$

z is a solution if and only if there exist $c_1, c_2 \in \mathbb{R}$ such that

$$z(t) = c_1 z_1(t) + c_2 z_2(t) = c_1(2+\sqrt{2})^t + c_2(2-\sqrt{2})^t \quad \forall t \in \mathbb{N}_0.$$

In our next example, let $y(t+2) = -y(t)/16 + y(t+1)/2$ for all $t \in \mathbb{N}_0$. The approach employing a generalized exponential function now yields the characteristic polynomial $\lambda^2 - \lambda/2 + 1/16$ with the unique root $\lambda = 1/4$. Thus, two independent solutions of our homogeneous difference equation are given by

$$z_1(t) = (1/4)^t \quad \forall t \in \mathbb{N}_0$$

and

$$z_2(t) = t(1/4)^t \quad \forall t \in \mathbb{N}_0.$$

To prove that these solutions indeed satisfy the required independence property, note that

$$\begin{vmatrix} z_1(0) & z_2(0) \\ z_1(1) & z_2(1) \end{vmatrix} = \begin{vmatrix} 1 & 0 \\ 1/4 & 1/4 \end{vmatrix} = 1/4 \neq 0.$$

Thus, z is a solution if and only if there exist $c_1, c_2 \in \mathbb{R}$ such that

$$z(t) = c_1 z_1(t) + c_2 z_2(t) = c_1(1/4)^t + c_2 t(1/4)^t \quad \forall t \in \mathbb{N}_0.$$

As a final example, let $y(t+2) = -4y(t)$ for all $t \in \mathbb{N}_0$. We obtain the characteristic polynomial $\lambda^2 + 4$ with the two complex roots $\lambda_1 = 2i$ and $\lambda_2 = -2i$. Using polar coordinates, we obtain $\theta = \pi/2$ and these solutions can be written as $\lambda_1 = 2(\cos(\pi/2) + i\sin(\pi/2))$ and $\lambda_2 = 2(\cos(\pi/2) - i\sin(\pi/2))$. The corresponding solutions of our difference equation are given by

$$z_1(t) = 2^t \cos(t\pi/2) \quad \forall t \in \mathbb{N}_0$$

and

$$z_2(t) = 2^t \sin(t\pi/2) \quad \forall t \in \mathbb{N}_0.$$

We obtain

$$\begin{vmatrix} z_1(0) & z_2(0) \\ z_1(1) & z_2(1) \end{vmatrix} = \begin{vmatrix} 1 & 0 \\ 0 & 2 \end{vmatrix} = 2 \neq 0.$$

Thus, z is a solution if and only if there exist $c_1, c_2 \in \mathbb{R}$ such that

$$z(t) = c_1 z_1(t) + c_2 z_2(t) = c_1 2^t \cos(t\pi/2) + c_2 2^t \sin(t\pi/2) \quad \forall t \in \mathbb{N}_0.$$

The general form of a linear difference equation with constant coefficients of order two is given by

$$y(t+2) = b(t) + a_0 y(t) + a_1 y(t+1) \quad \forall t \in \mathbb{N}_0 \tag{5.24}$$

where $a_0 \in \mathbb{R} \setminus \{0\}$, $a_1 \in \mathbb{R}$ and $b : \mathbb{N}_0 \mapsto \mathbb{R}$ is such that there exists $t \in \mathbb{N}_0$ with $b(t) \neq 0$. Depending on the form of the inhomogeneity b, finding the solutions of inhomogeneous equations with constant coefficients of order two can be a quite difficult task; in fact, general solution methods for arbitrary inhomogeneities are not known. For that reason, we only consider specific functional forms for the inhomogeneity $b : \mathbb{N}_0 \mapsto \mathbb{R}$.

First, we consider the case of a constant function b. That is, there exists $b_0 \in \mathbb{R} \setminus \{0\}$ such that $b(t) = b_0$ for all $t \in \mathbb{N}_0$. The case $b_0 = 0$ is ruled out because this reduces to a homogeneous equation. Substituting this constant function b in (5.24), we obtain

$$y(t+2) = b_0 + a_0 y(t) + a_1 y(t+1) \quad \forall t \in \mathbb{N}_0. \tag{5.25}$$

Because we already have the general solution of the associated homogeneous equation, it is sufficient to find one particular solution of (5.25). We try to obtain a particular solution by using a functional structure that is analogous to that of the inhomogeneity. That is, in this case, we begin by attempting to find a constant particular solution. Suppose $c_0 \in \mathbb{R}$ is a constant, and consider the function $\hat{y}$ defined by $\hat{y}(t) = c_0$ for all $t \in \mathbb{N}_0$. Substituting into (5.25), we obtain

$$c_0 = b_0 + a_0 c_0 + a_1 c_0. \tag{5.26}$$

We can distinguish two cases.

Case I: $a_0 + a_1 \neq 1$. In this case, (5.26) can be solved for c_0 to obtain $c_0 = b_0/(1 - a_0 - a_1)$ and, thus, $\hat{y}(t) = b_0/(1 - a_0 - a_1)$ for all $t \in \mathbb{N}_0$. The general solution of (5.25) is thus

$$y(t) = z(t) + b_0/(1 - a_0 - a_1) \quad \forall t \in \mathbb{N}_0$$

where z is a solution of the corresponding homogeneous equation.

Case II: $a_0 + a_1 = 1$. In this case, there exists no $c_0 \in \mathbb{R}$ such that (5.26) is satisfied because $b_0 \neq 0$. Thus, we have to try another functional form for the desired particular solution $\hat{y}$. Given that $a_0 + a_1 = 1$, (5.25) becomes

$$y(t+2) = b_0 + a_0 y(t) + (1 - a_0) y(t+1) \quad \forall t \in \mathbb{N}_0. \tag{5.27}$$

We now attempt to find a particular solution using the functional form $\hat{y}(t) = c_0 t$ for all $t \in \mathbb{N}_0$ with the value of the constant $c_0 \in \mathbb{R}$ to be determined. Substituting into (5.27), we obtain

$$c_0(t+2) = b_0 + a_0 c_0 t + (1 - a_0) c_0 (t+1) \quad \forall t \in \mathbb{N}_0. \tag{5.28}$$

There are two subcases.

Subcase II.A: $a_0 \neq -1$. In this case, we can solve (5.28) for c_0 to obtain $c_0 = b_0/(1 + a_0)$ and, thus, $\hat{y}(t) = b_0 t/(1 + a_0)$ for all $t \in \mathbb{N}_0$. The general solution of (5.25) is

$$y(t) = z(t) + b_0 t/(1 + a_0) \quad \forall t \in \mathbb{N}_0$$

where, again, z is a solution of the corresponding homogeneous equation.

Subcase II.B: $a_0 = -1$. Because we are in case II, this implies $a_1 = 2$ and our equation becomes

$$y(t+2) = b_0 - y(t) + 2y(t+1) \quad \forall t \in \mathbb{N}_0$$

and the approach using a linear function given by $\hat{y} = c_0 t$ for all $t \in \mathbb{N}_0$ does not work—no c_0 such that (5.28) is satisfied exists. We therefore make an alternative attempt by setting $\hat{y}(t) = c_0 t^2$ for all $t \in \mathbb{N}_0$, where $c_0 \in \mathbb{R}$ is a constant to be determined. Substituting, we obtain

$$c_0(t+2)^2 = b_0 - c_0 t^2 + 2c_0(t+1)^2 \quad \forall t \in \mathbb{N}_0$$

and, solving, $c_0 = b_0/2$. This gives us the particular solution $\hat{y}(t) = b_0 t^2/2$ for all $t \in \mathbb{N}_0$ and, finally, the general solution

$$y(t) = z(t) + b_0 t^2/2 \quad \forall t \in \mathbb{N}_0$$

where z is a solution of the corresponding homogeneous equation.

We summarize our observations in the following theorem.

Theorem 5.2.10 *(i) Let $b_0 \in \mathbb{R} \setminus \{0\}$, and suppose $a_0 \in \mathbb{R} \setminus \{0\}$ and $a_1 \in \mathbb{R}$ are such that $a_0 + a_1 \neq 1$. A function $y : \mathbb{N}_0 \mapsto \mathbb{R}$ is a solution of the linear difference equation with constant coefficients of order two (5.25) if and only if*

$$y(t) = z(t) + b_0/(1 - a_0 - a_1) \quad \forall t \in \mathbb{N}_0$$

where z is a solution of the corresponding homogeneous equation (5.19).

(ii) Let $b_0 \in \mathbb{R} \setminus \{0\}$, and suppose $a_0 \in \mathbb{R} \setminus \{0\}$ and $a_1 \in \mathbb{R}$ are such that $a_0 + a_1 = 1$ and $a_0 \neq -1$. A function $y : \mathbb{N}_0 \mapsto \mathbb{R}$ is a solution of the linear difference equation with constant coefficients of order two (5.25) if and only if

$$y(t) = z(t) + b_0 t/(1 + a_0) \quad \forall t \in \mathbb{N}_0$$

where z is a solution of the corresponding homogeneous equation (5.19).

(iii) Let $b_0 \in \mathbb{R} \setminus \{0\}$, $a_0 = -1$ and $a_1 = 2$. A function $y : \mathbb{N}_0 \mapsto \mathbb{R}$ is a solution of the linear difference equation with constant coefficients of order two (5.25) if and only if

$$y(t) = z(t) + b_0 t^2/2 \quad \forall t \in \mathbb{N}_0$$

where z is a solution of the corresponding homogeneous equation (5.19).

The second inhomogeneity considered here is such that the function b is a generalized exponential function. Let $d_0 \in \mathbb{R} \setminus \{0\}$ be a parameter, and suppose $b : \mathbb{N}_0 \mapsto \mathbb{R}$ is such that $b(t) = d_0^t$ for all $t \in \mathbb{N}_0$. Substituting this function b into (5.24), we obtain

$$y(t + 2) = d_0^t + a_0 y(t) + a_1 y(t + 1) \quad \forall t \in \mathbb{N}_0. \tag{5.29}$$

Our first attempt to obtain a particular solution proceeds, again, by using a functional structure that is similar to that of the inhomogeneity. Suppose $c_0 \in \mathbb{R}$ is a constant, and consider the function $\hat{y}$ defined by $\hat{y}(t) = c_0 d_0^t$ for all $t \in \mathbb{N}_0$. Substituting into (5.29), we obtain

$$c_0 d_0^{t+2} = d_0^t + a_0 c_0 d_0^t + a_1 c_0 d_0^{t+1}. \tag{5.30}$$

We can distinguish two cases.

Case I: $d_0^2 - a_0 - a_1 d_0 \neq 0$. In this case, (5.30) can be solved for c_0 to obtain $c_0 = 1/(d_0^2 - a_0 - a_1 d_0)$ and, thus, $\hat{y}(t) = d_0^t/(d_0^2 - a_0 - a_1 d_0)$ for all $t \in \mathbb{N}_0$. The general solution of (5.29) is thus

$$y(t) = z(t) + d_0^t/(d_0^2 - a_0 - a_1 d_0) \quad \forall t \in \mathbb{N}_0$$

where z is a solution of the corresponding homogeneous equation.

Case II: $d_0^2 - a_0 - a_1 d_0 = 0$. In this case, there exists no $c_0 \in \mathbb{R}$ such that (5.30) is satisfied because $d_0 \neq 0$. Another functional form for the desired particular solution $\hat{y}$ is thus required. Given that $d_0^2 - a_0 - a_1 d_0 = 0$, (5.29) becomes

$$y(t + 2) = d_0^t + (d_0^2 - a_1 d_0) y(t) + a_1 y(t + 1) \quad \forall t \in \mathbb{N}_0. \tag{5.31}$$

We now attempt to find a particular solution using the functional form $\hat{y}(t) = c_0 t d_0^t$ for all $t \in \mathbb{N}_0$ with the value of the constant $c_0 \in \mathbb{R}$ to be determined. Substituting into (5.31) and dividing both sides by $d_0^t \neq 0$, we obtain

$$c_0(t + 2) d_0^2 = 1 + (d_0^2 - a_1 d_0) c_0 t + a_1 c_0 (t + 1) d_0 \quad \forall t \in \mathbb{N}_0.$$

This can be simplified to

$$c_0(2d_0 - a_1) d_0 = 1 \quad \forall t \in \mathbb{N}_0, \tag{5.32}$$

and there are two subcases.

Subcase II.A: $2d_0 \neq a_1$. In this case, we can solve (5.32) for c_0 to obtain $c_0 = 1/[d_0(2d_0 - a_1)]$ and, thus, $\hat{y}(t) = t d_0^{t-1}/(2d_0 - a_1)$ for all $t \in \mathbb{N}_0$. The general solution of (5.29) is

$$y(t) = z(t) + t d_0^{t-1}/(2d_0 - a_1) \quad \forall t \in \mathbb{N}_0$$

where, again, z is a solution of the corresponding homogeneous equation.

Subcase II.B: $2d_0 = a_1$. Because we are in case II, this implies $a_0 = -d_0^2$ and our equation becomes

$$y(t + 2) = d_0^t - d_0 y(t) + 2d_0 y(t + 1) \quad \forall t \in \mathbb{N}_0$$

and the approach using $\hat{y}(t) = c_0 d_0^t$ for all $t \in \mathbb{N}_0$ does not work—no c_0 such that (5.32) is satisfied exists. We therefore make an alternative attempt by setting $\hat{y}(t) = c_0 t^2 d_0^t$ for all $t \in \mathbb{N}_0$, where $c_0 \in \mathbb{R}$ is a constant to be determined. Substituting, we obtain

$$c_0(t + 2)^2 d_0^{t+2} = d_0^t - d_0^2 c_0 t^2 d_0^t + 2d_0 c_0 (t + 1)^2 d_0^{t+1} \quad \forall t \in \mathbb{N}_0$$

and, solving, $c_0 = 1/(2d_0^2)$. This gives us the particular solution $\hat{y}(t) = t^2 d_0^{t-2}/2$ for all $t \in \mathbb{N}_0$ and the general solution

$$y(t) = z(t) + t^2 d_0^{t-2}/2 \quad \forall t \in \mathbb{N}_0$$

where z is a solution of the corresponding homogeneous equation.

Thus, we have established the following theorem.

Theorem 5.2.11 *(i) Let $d_0 \in \mathbb{R} \setminus \{0\}$, and suppose $a_0 \in \mathbb{R} \setminus \{0\}$ and $a_1 \in \mathbb{R}$ are such that $d_0^2 - a_0 - a_1 d_0 \neq 0$. A function $y : \mathbb{N}_0 \mapsto \mathbb{R}$ is a solution of the linear difference equation with constant coefficients of order two (5.29) if and only if*

$$y(t) = z(t) + d_0^t/(d_0^2 - a_0 - a_1 d_0) \quad \forall t \in \mathbb{N}_0$$

where z is a solution of the corresponding homogeneous equation (5.19).

(ii) Let $d_0 \in \mathbb{R} \setminus \{0\}$, and suppose $a_0 \in \mathbb{R} \setminus \{0\}$ and $a_1 \in \mathbb{R}$ are such that $d_0^2 - a_0 - a_1 d_0 = 0$ and $2d_0 \neq a_1$. A function $y : \mathbb{N}_0 \mapsto \mathbb{R}$ is a solution of the linear difference equation with constant coefficients of order two (5.29) if and only if

$$y(t) = z(t) + t d_0^{t-1}/(2d_0 - a_1) \quad \forall t \in \mathbb{N}_0$$

where z is a solution of the corresponding homogeneous equation (5.19).

(iii) Let $d_0 \in \mathbb{R} \setminus \{0\}$, $a_0 = -d_0^2$ and $a_1 = 2d_0$. A function $y : \mathbb{N}_0 \mapsto \mathbb{R}$ is a solution of the linear difference equation with constant coefficients of order two (5.29) if and only if

$$y(t) = z(t) + t^2 d_0^{t-2}/2 \quad \forall t \in \mathbb{N}_0$$

where z is a solution of the corresponding homogeneous equation (5.19).

We conclude this section with some examples. Suppose a linear difference equation is given by

$$y(t+2) = 2^t + 3y(t) - 2y(t+1) \quad \forall t \in \mathbb{N}_0.$$

The corresponding homogeneous equation is

$$y(t+2) = 3y(t) - 2y(t+1) \quad \forall t \in \mathbb{N}_0.$$

Using the generalized exponential approach, we attempt to find a solution $z : \mathbb{N}_0 \mapsto \mathbb{R}$ such that $z(t) = \lambda^t$ for all $t \in \mathbb{N}_0$, where $\lambda \in \mathbb{R} \setminus \{0\}$ is a parameter to be determined (if possible). Substituting, we obtain the characteristic polynomial $\lambda^2 - 3 + 2\lambda$ which has the two real roots $\lambda_1 = 1$ and $\lambda_2 = -3$. Therefore, we obtain the two solutions

$$z_1(t) = 1 \quad \forall t \in \mathbb{N}_0$$

and

$$z_2(t) = (-3)^t \quad \forall t \in \mathbb{N}_0.$$

Because

$$\begin{vmatrix} z_1(0) & z_2(0) \\ z_1(1) & z_2(1) \end{vmatrix} = \begin{vmatrix} 1 & 1 \\ 1 & -3 \end{vmatrix} = -4 \neq 0,$$

z is a solution if and only if there exist $c_1, c_2 \in \mathbb{R}$ such that

$$z(t) = c_1 z_1(t) + c_2 z_2(t) = c_1 + c_2(-3)^t \quad \forall t \in \mathbb{N}_0.$$

To solve the inhomogeneous equation with the generalized exponential inhomogeneity defined by $b(t) = 2^t$ for all $t \in \mathbb{N}_0$, we begin by trying to find a particular solution $\hat{y}$ of the form $\hat{y}(t) = c_0 2^t$ for all $t \in \mathbb{N}_0$, where $c_0 \in \mathbb{R}$ is a constant. Substituting into the original equation, we obtain

$$c_0 2^{t+2} = 2^t + 3c_0 2^t - 2c_0 2^{t+1} \quad \forall t \in \mathbb{N}_0$$

and, solving, we obtain $c_0 = 1/5$. Therefore, thew particular solution is given by $\hat{y}(t) = 2^t/5$ for all $t \in \mathbb{N}_0$ and, together with the solution of the associated homogeneous equation, we obtain the general solution

$$y(t) = c_1 + c_2(-3)^t + 2^t/5 \quad \forall t \in \mathbb{N}_0$$

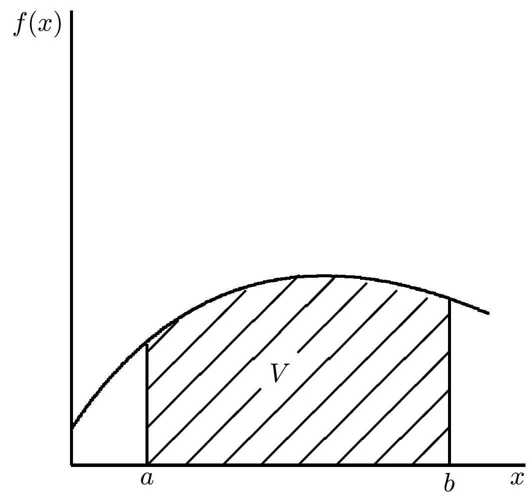


Figure 5.2: Geometric interpretation of integrals.

with parameters $c_1, c_2 \in \mathbb{R}$.

Now consider the following modification of this example, given by the equation

$$y(t+2) = (-3)^t + 3y(t) - 2y(t+1) \quad \forall t \in \mathbb{N}_0.$$

The associated homogeneous equation is the same as before, so we only need to find a particular solution to the inhomogeneous equation. Trying $\hat{y}(t) = c_0(-3)^t$ for all $t \in \mathbb{N}_0$ and substituting yields the equation

$$c_0(-3)^{t+2} = (-3)^t + 3c_0(-3)^t - 2c_0(-3)^{t+1} \quad \forall t \in \mathbb{N}_0$$

which does not have a solution. Therefore, we use the new functional form $\hat{y}(t) = c_0 t(-3)^t$ for all $t \in \mathbb{N}_0$ and, substituting again, we now obtain $c_0 = 1/12$. Thus, our particular solution is given by $\hat{y}(t) = t(-3)^{t-1}/4$ for all $t \in \mathbb{N}_0$. Combined with the solution of the corresponding homogeneous equation, we obtain the general solution

$$y(t) = c_1 + c_2(-3)^t + t(-3)^{t-1}/4 \quad \forall t \in \mathbb{N}_0.$$

5.3 Integration

Integration provides us with a method of finding the *areas* of certain subsets of $\mathbb{R}^2$. The subsets of $\mathbb{R}^2$ that we will consider here can be described by functions of one variable. As an illustration, consider Figure 5.2.

We will examine ways of finding the area of subsets of $\mathbb{R}^2$ such as

$$\{(x,y) \mid x \in [a,b] \wedge y \in [0, f(x)]\}$$

where $f : A \mapsto \mathbb{R}$ is a function such that $[a,b] \subseteq A$, $a < b$. In Figure 5.2, V denotes such an area.

There are two basic problems that have to be solved. We have to specify what types of subsets of $\mathbb{R}^2$ as discussed above can be assigned an area at all, and, for those subsets that do have a well-defined area, we have to find a general method which allows us to calculate this area.

As a first step, we can try to *approximate* the area of a subset of $\mathbb{R}^2$ (assuming this area is defined) by the areas of *other* sets which we are familiar with. Subsets of $\mathbb{R}^2$ that can be assigned an area very easily are *rectangles*. Consider again Figure 5.2. To obtain an approximation of the area V that *underestimates* the value of V, we could calculate the area of the rectangle formed by the points $(a,0)$, $(b,0)$, $(a, \inf\{f(x) \mid x \in [a,b]\})$, $(b, \inf\{f(x) \mid x \in [a,b]\})$. Analogously, to get an approximation that *overestimates* V, we could use the area of the rectangle given by the points $(a,0)$, $(b,0)$, $(a, \sup\{f(x) \mid x \in [a,b]\})$, $(b, \sup\{f(x) \mid x \in [a,b]\})$. Clearly, the areas of these rectangles are given by

$$(b-a)\inf\{f(x) \mid x \in [a,b]\} \text{ and } (b-a)\sup\{f(x) \mid x \in [a,b]\},$$

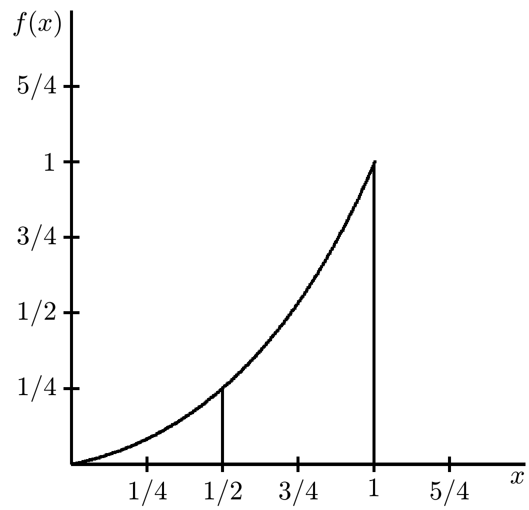


Figure 5.3: Approximation of an area.

respectively. This gives us the inequalities

$$(b-a)\inf\{f(x) \mid x \in [a,b]\} \leq V \leq (b-a)\sup\{f(x) \mid x \in [a,b]\}.$$

These approximations of V can, of course, be very inaccurate. One possibility to obtain a more precise approximation is to divide the interval $[a,b]$ into smaller *subintervals*, approximate the areas corresponding to these subintervals, and add up these approximations for all subintervals. The following example illustrates this idea. Consider the function

$$f : \mathbb{R} \mapsto \mathbb{R},\ x \mapsto x^2,$$

and let $a = 0$, $b = 1$. Suppose we want to find the area V of the set $\{(x,y) \mid x \in [0,1] \wedge y \in [0,x^2]\}$. We have $\inf\{x^2 \mid x \in [0,1]\} = 0$ and $\sup\{x^2 \mid x \in [0,1]\} = 1$. Therefore, $0 \leq V \leq 1$. To obtain a better approximation, we divide the interval $[0,1]$ into the two subintervals $[0,1/2]$ and $[1/2,1]$. We now obtain

$$\inf\{x^2 \mid x \in [0,1/2]\} = 0 \text{ and } \sup\{x^2 \mid x \in [0,1/2]\} = 1/4,$$

$$\inf\{x^2 \mid x \in [1/2,1]\} = 1/4 \text{ and } \sup\{x^2 \mid x \in [1/2,1]\} = 1.$$

This gives us new approximations of V, namely, $1/8 \leq V \leq 5/8$. See Figure 5.3 for an illustration.

In general, if we divide $[0,1]$ into $n \in \mathbb{N}$ subintervals $[0,1/n],[1/n,2/n],\ldots,[(n-1)/n,1]$, we can approximate V from below with the *lower sum*

$$L_n := \sum_{i=1}^{n}\left(\frac{i}{n} - \frac{i-1}{n}\right)\inf\{x^2 \mid x \in [(i-1)/n, i/n]\} = \sum_{i=1}^{n}\frac{1}{n}\left(\frac{i-1}{n}\right)^2,$$

and an approximation from above is given by the *upper sum*

$$U_n := \sum_{i=1}^{n}\left(\frac{i}{n} - \frac{i-1}{n}\right)\sup\{x^2 \mid x \in [(i-1)/n, i/n]\} = \sum_{i=1}^{n}\frac{1}{n}\left(\frac{i}{n}\right)^2.$$

For any $n \in \mathbb{N}$, we obtain $L_n \leq V \leq U_n$. We can rewrite L_n as

$$\begin{aligned} L_n &= \frac{1}{n^3}\sum_{i=1}^{n}(i-1)^2 = \frac{1}{n^3}\sum_{j=0}^{n-1}j^2 = \frac{1}{n^3}\frac{(n-1)n(2n-1)}{6} \\ &= \frac{(n-1)(2n-1)}{6n^2} = \frac{2n^2-3n+1}{6n^2} = \frac{1}{3} - \frac{1}{2n} + \frac{1}{6n^2}, \end{aligned}$$

and the upper sum is equal to

$$\begin{aligned} U_n &= \frac{1}{n^3}\sum_{i=1}^{n} i^2 = \frac{1}{n^3}\frac{n(n+1)(2n+1)}{6} = \frac{(n+1)(2n+1)}{6n^2} = \\ &= \frac{2n^2+3n+1}{6n^2} = \frac{1}{3} + \frac{1}{2n} + \frac{1}{6n^2}. \end{aligned}$$

Note that $\{L_n\}$ and $\{U_n\}$ are convergent sequences, and we obtain $\lim_{n\to\infty} L_n = \lim_{n\to\infty} U_n = 1/3$.

Because $\{L_n\}$ and $\{U_n\}$ converge to the same limit, it is natural to consider this limit the area V, which is therefore obtained as the limit of an approximation process.

This procedure can be generalized. First, we define *partitions* of intervals.

Definition 5.3.1 *Let $a, b \in I\!R$, $a < b$.*

(i) A partition of the interval $[a,b]$ *is a finite set of numbers $P = \{x_0, x_1, \ldots, x_n\}$ such that*

$$a = x_0 < x_1 < \ldots < x_n = b.$$

(ii) Suppose P and $\bar{P}$ are partitions of $[a,b]$. P is finer than $\bar{P}$ *if and only if $\bar{P} \subseteq P$.*

Using the notion of a partition, we can define lower and upper sums for functions defined on an interval.

Definition 5.3.2 *Let $A \subseteq I\!R$ be an interval, and let $a, b \in I\!R$ be such that $a < b$ and $[a,b] \subseteq A$. Furthermore, let $f : A \mapsto I\!R$ be bounded on $[a,b]$, and let $P = \{x_0, x_1, \ldots, x_n\}$ be a partition of $[a,b]$.*

(i) The lower sum of f with respect to P *is defined by*

$$L(P) := \sum_{i=1}^{n}(x_i - x_{i-1}) \inf\{f(x) \mid x \in [x_{i-1}, x_i]\}.$$

(ii) The upper sum of f with respect to P *is defined by*

$$U(P) := \sum_{i=1}^{n}(x_i - x_{i-1}) \sup\{f(x) \mid x \in [x_{i-1}, x_i]\}.$$

If we replace a partition by a *finer* partition, the lower sum cannot decrease, and the upper sum cannot increase.

Theorem 5.3.3 *Let $A \subseteq I\!R$ be an interval, and let $a, b \in I\!R$ be such that $a < b$ and $[a,b] \subseteq A$. Furthermore, let $f : A \mapsto I\!R$ be bounded on $[a,b]$, and let P and $\bar{P}$ be partitions of $[a,b]$.*

$$P \text{ is finer than } \bar{P} \;\Rightarrow\; L(P) \geq L(\bar{P}) \;\wedge\; U(P) \leq U(\bar{P}).$$

Proof. Suppose P is finer than $\bar{P}$, that is, $\bar{P} \subseteq P$. If $P = \bar{P}$, the result follows trivially. Now suppose $\bar{P} = \{x_0, \ldots, x_n\}$ and $P = \bar{P} \cup \{y\}$ with $y \notin \bar{P}$. Let $k \in \{1, \ldots, n\}$ be such that $x_{k-1} < y < x_k$ (such a k exists and is unique by the definition of a partition). Then we obtain

$$\begin{aligned} L(P) &= \sum_{i=1}^{k-1}(x_i - x_{i-1}) \inf\{f(x) \mid x \in [x_{i-1}, x_i]\} + (y - x_{k-1}) \inf\{f(x) \mid x \in [x_{k-1}, y]\} \\ &+ (x_k - y) \inf\{f(x) \mid x \in [y, x_k]\} + \sum_{i=k+1}^{n}(x_i - x_{i-1}) \inf\{f(x) \mid x \in [x_{i-1}, x_i]\}. \end{aligned}$$

By definition of an infimum,

$$(y - x_{k-1}) \inf\{f(x) \mid x \in [x_{k-1}, y]\} + (x_k - y) \inf\{f(x) \mid x \in [y, x_k]\} \geq$$
$$(x_k - x_{k-1}) \inf\{f(x) \mid x \in [x_{k-1}, x_k]\}.$$

Therefore,

$$L(P) \geq \sum_{i=1}^{n}(x_i - x_{i-1}) \inf\{f(x) \mid x \in [x_{i-1}, x_i]\} = L(\bar{P}).$$

If $P \setminus \bar{P}$ contains more than one element, the above argument can be applied repeatedly to conclude $L(P) \geq L(\bar{P})$.

The inequality $U(P) \leq U(\bar{P})$ is proven analogously. $\|$

For a given partition P, it is clear that

$$L(P) \leq U(P) \tag{5.33}$$

(by definition of the infimum and the supremum of a set). In addition, Theorem 5.3.3 implies that this inequality holds even for lower and upper sums that are calculated for *different* partitions.

Theorem 5.3.4 *Let $A \subseteq \mathbb{R}$ be an interval, and let $a, b \in \mathbb{R}$ be such that $a < b$ and $[a, b] \subseteq A$. Furthermore, let $f : A \mapsto \mathbb{R}$ be bounded on $[a, b]$. If P and $\bar{P}$ are partitions of $[a, b]$, then $L(P) \leq U(\bar{P})$.*

Proof. Let $\tilde{P} := P \cup \bar{P}$. Then $\tilde{P}$ is finer than P, and $\tilde{P}$ is finer than $\bar{P}$. Therefore, Theorem 5.3.3 and (5.33) imply

$$L(P) \leq L(\tilde{P}) \leq U(\tilde{P}) \leq U(\bar{P}). \quad \|$$

An interesting consequence of Theorem 5.3.4 is that the set

$$L := \{L(P) \mid P \text{ is a partition of } [a, b]\}$$

is bounded from above, because for any partition P of $[a, b]$, $U(P)$ is an upper bound for L. Similarly, the set

$$U := \{U(P) \mid P \text{ is a partition of } [a, b]\}$$

is bounded from below, because, for any partition P of $[a, b]$, $L(P)$ is a lower bound for U. Therefore, L has a supremum, and U has an infimum. By Theorem 5.3.4, $\sup(L) \leq \inf(U)$.

We now define

Definition 5.3.5 *Let $A \subseteq \mathbb{R}$ be an interval, and let $a, b \in \mathbb{R}$ be such that $a < b$ and $[a, b] \subseteq A$. Furthermore, let $f : A \mapsto \mathbb{R}$ be bounded on $[a, b]$.*

(i) The function f is Riemann integrable *on $[a, b]$ if and only if* $\sup(L) = \inf(U)$.

(ii) If f is Riemann integrable on $[a, b]$, the integral of f on $[a, b]$ *is defined by*

$$\int_a^b f(x)dx := \sup(L) = \inf(U).$$

(iii) We define

$$\int_a^a f(x)dx := 0 \quad \textit{and} \quad \int_b^a f(x)dx := -\int_a^b f(x)dx.$$

Because Riemann integrability is the only form of integrability discussed here, we will simply use the term "integrable" when referring to Riemann integrable functions.

We can now return to the problem of assigning an area to specific subsets of $\mathbb{R}^2$. Consider first the case of a function $f : A \mapsto \mathbb{R}$ such that $f([a, b]) \subseteq \mathbb{R}_+$, that is, on the interval $[a, b]$, f assumes nonnegative values only. If f is integrable on $[a, b]$, we define the area V (see Figure 5.2) to be given by the integral of f on $[a, b]$.

If f assumes only nonpositive values on $[a, b]$, the integral of f on $[a, b]$ is a nonpositive number (we will see why this is the case once we discuss methods to calculate integrals). Because we want areas to have nonnegative values, we will, in this case, use the *absolute value* of the integral of f on $[a, b]$ to represent this area.

Not all functions that are bounded on an interval are integrable on this interval. Consider, for example, the function

$$f : [0, 1] \mapsto \mathbb{R},\ x \mapsto \begin{cases} 0 & \text{if } x \in \mathcal{Q} \\ 1 & \text{if } x \notin \mathcal{Q}. \end{cases}$$

This function assigns the number zero to all rational numbers in the interval $[0, 1]$ and the number one to all irrational numbers in $[0, 1]$. To show that the function f is not integrable on $[0, 1]$, let $P = \{x_0, \ldots, x_n\}$ be an arbitrary partition of $[0, 1]$. Because *any* interval $[a, b]$ with $a < b$ contains rational and irrational numbers, we obtain

$$\inf\{f(x) \mid x \in [x_{i-1}, x_i]\} = 0 \ \wedge \ \sup\{f(x) \mid x \in [x_{i-1}, x_i]\} = 1 \quad \forall i = 1, \ldots, n,$$

and therefore, because this is true for *all* partitions of $[0, 1]$, $\sup(L) = 0$ and $\inf(U) = 1$, which shows that f is not integrable on $[0, 1]$.

Integrability is a condition that is *weaker* than continuity, that is, all functions that are continuous on an interval are integrable on this interval. Furthermore, all monotone functions are integrable. We summarize these observations in the following theorem, which we state without a proof.

Theorem 5.3.6 *Let $A \subseteq \mathbb{R}$ be an interval, and let $a, b \in \mathbb{R}$ be such that $a < b$ and $[a, b] \subseteq A$. Furthermore, let $f : A \mapsto \mathbb{R}$ be bounded on $[a, b]$.*

(i) f is continuous on $[a, b] \Rightarrow f$ is integrable on $[a, b]$.
(ii) f is nondecreasing on $[a, b] \Rightarrow f$ is integrable on $[a, b]$.
(iii) f is nonincreasing on $[a, b] \Rightarrow f$ is integrable on $[a, b]$.

The process of finding an integral can be expressed in terms of an operation that, in some sense, is the "reverse" operation of differentiation. We define

Definition 5.3.7 *Let $A \subseteq \mathbb{R}$ be an interval, and let $f : A \mapsto \mathbb{R}$. A differentiable function $F : A \mapsto \mathbb{R}$ such that $F'(x) = f(x)$ for all $x \in A$ is called an* integral function *of f.*

Clearly, if F is an integral function of f and $c \in \mathbb{R}$, the function

$$G : A \mapsto \mathbb{R}, \ x \mapsto F(x) + c$$

also is an integral function of f. Furthermore, if F and G are integral functions of f, then there must exist a constant $c \in \mathbb{R}$ such that $G(x) = F(x) + c$ for all $x \in A$. Therefore, integral functions—if they exist—are unique up to additive constants. If an integral function F of f exists, the *indefinite integral of f* is

$$\int f(x)dx := F(x) + c$$

where $c \in \mathbb{R}$ is a constant. Integrals of the form

$$\int_a^b f(x)dx$$

with $a, b \in \mathbb{R}$ are called *definite integrals* in order to distinguish them from indefinite integrals.

The following theorem is called the *fundamental theorem of calculus*. It describes how integral functions can be used to find definite integrals. We state this theorem without a proof.

Theorem 5.3.8 *Let $A \subseteq \mathbb{R}$ be an interval, and let $a, b \in \mathbb{R}$ be such that $a < b$ and $[a, b] \subseteq A$. Furthermore, let $f : A \mapsto \mathbb{R}$ be continuous on $[a, b]$.*

(i) The function

$$H : [a, b] \mapsto \mathbb{R}, \ y \mapsto \int_a^y f(x)dx$$

is differentiable on (a, b), and $H'(y) = f(y)$ for all $y \in (a, b)$.

(ii) If F is an integral function of f, then

$$\int_a^b f(x)dx = F(b) - F(a).$$

Theorem 5.3.8 (ii) shows how definite integrals can be obtained from indefinite integrals. Once an integral function F of f is known, we just have to calculate the *difference* of the values of F at the limits of integration. It is also common to use the notations

$$\int_a^b f(x)dx = F(x)|_a^b \quad \text{or} \quad \int_a^b f(x)dx = [F(x)]_a^b$$

for definite integrals, where F is an integral function of f.

We can now summarize some integration rules.

Theorem 5.3.9 *Let $c \in \mathbb{R}$ be a constant.*

(i) Let $A \subseteq \mathbb{R}$ be an interval, and let $n \in \mathbb{N}$. The integral function of

$$f : A \mapsto \mathbb{R},\ x \mapsto x^n$$

is given by

$$F : A \mapsto \mathbb{R},\ x \mapsto \frac{x^{n+1}}{n+1} + c.$$

(ii) Let $A \subseteq \mathbb{R}_{++}$ be an interval, and let $\alpha \in \mathbb{R} \setminus \{-1\}$. The integral function of

$$f : A \mapsto \mathbb{R},\ x \mapsto x^\alpha$$

is given by

$$F : A \mapsto \mathbb{R},\ x \mapsto \frac{x^{\alpha+1}}{\alpha+1} + c.$$

(iii) Let $A \subseteq \mathbb{R}_{++}$ be an interval. The integral function of

$$f : A \mapsto \mathbb{R},\ x \mapsto \frac{1}{x}$$

is given by

$$F : A \mapsto \mathbb{R},\ x \mapsto \ln(x) + c.$$

(iv) Let $A \subseteq \mathbb{R}$ be an interval. The integral function of

$$f : A \mapsto \mathbb{R},\ x \mapsto e^x$$

is given by

$$F : A \mapsto \mathbb{R},\ x \mapsto e^x + c.$$

(v) Let $A \subseteq \mathbb{R}$ be an interval, and let $\alpha \in \mathbb{R}_{++} \setminus \{1\}$. The integral function of

$$f : A \mapsto \mathbb{R},\ x \mapsto \alpha^x$$

is given by

$$F : A \mapsto \mathbb{R},\ x \mapsto \frac{\alpha^x}{\ln(\alpha)} + c.$$

The proof of Theorem 5.3.9 is obtained by differentiating the integral functions.

Furthermore, we obtain

Theorem 5.3.10 *Let $A \subseteq \mathbb{R}$ be an interval, and let $\alpha \in \mathbb{R}$. Furthermore, let $f : A \mapsto \mathbb{R}$ and $g : A \mapsto \mathbb{R}$.*

(i) If F is an integral function of f, then αF is an integral function of αf.
(ii) If F is an integral function of f and G is an integral function of g, then $F + G$ is an integral function of $f + g$.

The proof of Theorem 5.3.10 is left as an exercise.

Theorems 5.3.9 and 5.3.10 allow us to find integral functions for sums and multiples of given functions. For example, for a polynomial function such as

$$f : \mathbb{R} \mapsto \mathbb{R},\ x \mapsto 2 + x - 3x^2,$$

we obtain the integral function

$$F : \mathbb{R} \mapsto \mathbb{R},\ x \mapsto 2x + \frac{1}{2}x^2 - x^3 + c.$$

Now we calculate the definite integral of the function f on $[0, 1]$. We obtain

$$\int_0^1 f(x)dx = F(x)|_0^1 = 2x + \frac{1}{2}x^2 - x^3 + c\Big|_0^1 = 2 + \frac{1}{2} - 1 = \frac{3}{2}.$$

Theorem 5.3.8 only applies to *continuous* functions, but we can also find definite integrals involving noncontinuous functions, if it is possible to partition the interval under consideration into subintervals on which the function *is* continuous.

The following theorems provide some useful rules for calculating certain definite integrals.

Theorem 5.3.11 *Let $A \subseteq \mathbb{R}$ be an interval, and let $a, b, c \in \mathbb{R}$ be such that $a < c < b$ and $[a, b] \subseteq A$. Furthermore, let $f : A \mapsto \mathbb{R}$ be bounded on $[a, b]$. If f is integrable on $[a, b]$, then f is integrable on $[a, c]$ and on $[c, b]$, and*

$$\int_a^b f(x)dx = \int_a^c f(x)dx + \int_c^b f(x)dx.$$

Theorem 5.3.12 *Let $A \subseteq \mathbb{R}$ be an interval, and let $a, b \in \mathbb{R}$ be such that $a < b$ and $[a, b] \subseteq A$. Furthermore, let $\alpha, \beta \in \mathbb{R}$, and let $f : A \mapsto \mathbb{R}$ and $g : A \mapsto \mathbb{R}$ be bounded on $[a, b]$. If f and g are integrable on $[a, b]$, then $\alpha f + \beta g$ is integrable on $[a, b]$, and*

$$\int_a^b (\alpha f(x) + \beta g(x))dx = \alpha \int_a^b f(x)dx + \beta \int_a^b g(x)dx.$$

As an example, consider the following function

$$f : \mathbb{R}_+ \mapsto \mathbb{R},\ x \mapsto \begin{cases} x^2 & \text{if } x \in [0, 2] \\ x & \text{if } x \in (2, \infty). \end{cases}$$

This function is *not* continuous at $x_0 = 2$. The graph of f is illustrated in Figure 5.4.

Suppose we want to find $\int_0^4 f(x)dx$. This can be achieved by partitioning the interval $[0, 4]$ into $[0, 2]$ and $[2, 4]$. We obtain

$$\begin{aligned} \int_0^4 f(x)dx &= \int_0^2 f(x)dx + \int_2^4 f(x)dx \\ &= \frac{1}{3}x^3\Big|_0^2 + \frac{1}{2}x^2\Big|_2^4 = \frac{8}{3} + 8 - 2 = \frac{26}{3}. \end{aligned}$$

Some integral functions can be found using a method called *integration by parts*. This integration rule follows from the application of the *product rule* for differentiation. It is illustrated in the following theorem.

Theorem 5.3.13 *Let $A \subseteq \mathbb{R}$ be an interval. If $f : A \mapsto \mathbb{R}$ and $g : A \mapsto \mathbb{R}$ are differentiable, then*

$$\int f'(x)g(x)dx = f(x)g(x) - \int f(x)g'(x)dx + c. \tag{5.34}$$

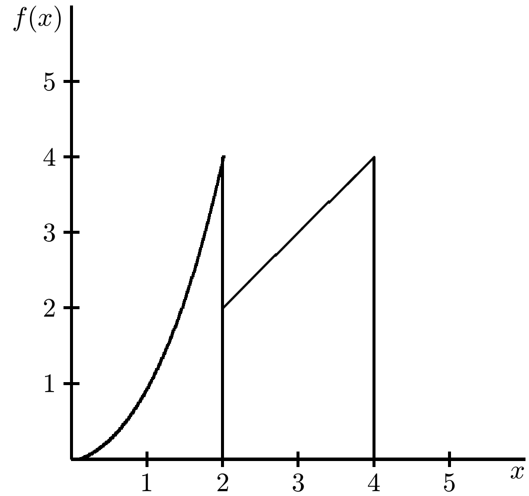


Figure 5.4: Piecewise integration.

Proof. By the product rule (see Theorem 3.2.5), $(fg)'(x) = f'(x)g(x) + f(x)g'(x)$ for all $x \in A$. Therefore, $\int (fg)'(x)dx = \int f'(x)g(x)dx + \int f(x)g'(x)dx + \bar{c}$. By definition of an integral function,

$$\int (fg)'(x)dx = (fg)(x) = f(x)g(x).$$

Setting $c := -\bar{c}$, (5.34) follows. ∥

As an example for the application of Theorem 5.3.13, we determine the indefinite integral $\int \ln(x)dx$ using integration by parts. Define $f : \mathbb{R}_{++} \mapsto \mathbb{R}$, $x \mapsto x$, and $g : \mathbb{R}_{++} \mapsto \mathbb{R}$, $x \mapsto \ln(x)$. Then $f'(x) = 1$ for all $x \in \mathbb{R}_{++}$, and $g'(x) = 1/x$ for all $x \in \mathbb{R}_{++}$. According to Theorem 5.3.13, we obtain

$$\int \ln(x)dx = \int f'(x)g(x)dx = x\ln(x) - \int dx = x\ln(x) - x + c = x(\ln(x) - 1) + c.$$

Another important rule of integration is the *substitution rule*, which is based on the *chain rule* for differentiation (see Theorem 3.2.6).

Theorem 5.3.14 *Let $A \subseteq \mathbb{R}$ be an interval, and let $f : A \mapsto \mathbb{R}$, $g : f(A) \mapsto \mathbb{R}$. If f is differentiable and G is an integral function of g, then*

$$\int g(f(x))f'(x)dx = G(f(x)) + c. \tag{5.35}$$

Proof. By the chain rule, $(G \circ f)'(x) = G'(f(x))f'(x) = g(f(x))f'(x)$ for all $x \in A$. Therefore,

$$\int g(f(x))f'(x)dx = \int (G \circ f)'(x)dx + c,$$

which is equivalent to (5.35). ∥

To illustrate the application of the substitution rule, we use this rule to find the indefinite integral $\int (x^2 + 1)2xdx$. Define $f : \mathbb{R} \mapsto \mathbb{R}$, $\mapsto x^2 + 1$, and $g : \mathbb{R} \mapsto \mathbb{R}$, $y \mapsto y$. We obtain $f'(x) = 2x$ for all $x \in \mathbb{R}$, and the integral function

$$G : \mathbb{R} \mapsto \mathbb{R},\ y \mapsto \frac{1}{2}y^2$$

for the function g. Using the substitution rule,

$$\int (x^2 + 1)2xdx = \int g(f(x))f'(x)dx = \frac{1}{2}(f(x))^2 + c = \frac{1}{2}(x^2 + 1)^2 + c.$$

So far, we restricted attention to functions that are bounded on a closed interval $[a, b]$. In some circumstances, we can derive integrals even if these conditions do not apply. Integrals of this kind are called *improper integrals*, and we can distinguish two types of them.

The first possibility to have an improper integral occurs if one of the *limits of integration* is not finite. We define

Definition 5.3.15 *(i) Let $A \subseteq \mathbb{R}$ be an interval, and let $a \in \mathbb{R}$ be such that $[a, \infty) \subseteq A$. Suppose $f : A \mapsto \mathbb{R}$ is integrable on $[a, b]$ for all $b \in (a, \infty)$. If*

$$\lim_{b\uparrow\infty} \int_a^b f(x)dx$$

exists and is finite, then f is integrable on $[a, \infty)$, *and the* improper integral

$$\int_a^\infty f(x)dx := \lim_{b\uparrow\infty} \int_a^b f(x)dx$$

exists (or converges). If $\int_a^\infty f(x)dx$ does not converge, we say that this improper integral diverges.

(ii) Let $A \subseteq \mathbb{R}$ be an interval, and let $b \in \mathbb{R}$ be such that $(-\infty, b] \subseteq A$. Suppose $f : A \mapsto \mathbb{R}$ is integrable on $[a, b]$ for all $a \in (-\infty, b)$. If

$$\lim_{a\downarrow-\infty} \int_a^b f(x)dx$$

exists and is finite, then f is integrable on $(-\infty, b]$, *and the* improper integral

$$\int_{-\infty}^b f(x)dx := \lim_{a\downarrow-\infty} \int_a^b f(x)dx$$

exists (or converges). If $\int_{-\infty}^b f(x)dx$ does not converge, we say that this improper integral diverges.

For example, consider

$$f : \mathbb{R}_{++} \mapsto \mathbb{R},\ x \mapsto \frac{1}{x^2}.$$

To determine whether the improper integral $\int_1^\infty f(x)dx$ exists, we have to find out whether $\int_1^b f(x)dx$ converges as b approaches infinity. We obtain

$$\int_1^b f(x)dx = \int_1^b \frac{dx}{x^2} = -\frac{1}{x}\Big|_1^b = -\frac{1}{b} + 1.$$

This implies

$$\lim_{b\uparrow\infty} \int_1^b \frac{dx}{x^2} = \lim_{b\uparrow\infty} \left(-\frac{1}{b} + 1\right) = 1,$$

and hence,

$$\int_1^\infty \frac{dx}{x^2} = 1.$$

As another example, define

$$f : \mathbb{R}_{++} \mapsto \mathbb{R},\ x \mapsto \frac{1}{x}.$$

Now we obtain

$$\int_1^b f(x)dx = \int_1^b \frac{dx}{x} = \ln(x)|_1^b = \ln(b) - \ln(1) = \ln(b).$$

Clearly, $\lim_{b\uparrow\infty} \ln(b) = \infty$, and therefore, $\int_1^\infty \frac{dx}{x}$ diverges.

Another type of improper integral is obtained if the function f is not bounded on an interval. This possibility is described in

Definition 5.3.16 *Let $A \subseteq I\!R$ be an interval, and let $a, b \in I\!R$ be such that $a < b$.*

(i) Suppose $[a, b) \subseteq A$, $b \notin A$, and $f : A \mapsto I\!R$ is integrable on $[a, c]$ for all $c \in (a, b)$. If

$$\lim_{c \uparrow b} \int_a^c f(x)dx$$

exists and is finite, then f is integrable on $[a, b)$, *and the* improper integral

$$\int_a^b f(x)dx := \lim_{c \uparrow b} \int_a^c f(x)dx$$

exists (or converges). If $\int_a^b f(x)dx$ does not converge, we say that this improper integral diverges.

(ii) Suppose $(a, b] \subseteq A$, $a \notin A$, and $f : A \mapsto I\!R$ is integrable on $[c, b]$ for all $c \in (a, b)$. If

$$\lim_{c \downarrow a} \int_c^b f(x)dx$$

exists and is finite, then f is integrable on $(a, b]$, *and the* improper integral

$$\int_a^b f(x)dx := \lim_{c \downarrow a} \int_c^b f(x)dx$$

exists (or converges). If $\int_a^b f(x)dx$ does not converge, we say that this improper integral diverges.

For example, consider the function

$$f : \mathbb{R}_{++} \mapsto \mathbb{R},\ x \mapsto \frac{1}{\sqrt{x}}.$$

For $a \in (0, 1)$, we obtain

$$\int_a^1 f(x)dx = \int_a^1 \frac{dx}{\sqrt{x}} = 2\sqrt{x}\Big|_a^1 = 2(1 - \sqrt{a}).$$

Therefore,

$$\lim_{a \downarrow 0} \int_a^1 f(x)dx = \lim_{a \downarrow 0} 2(1 - \sqrt{a}) = 2,$$

and the improper integral $\int_0^1 f(x)dx$ exists and is given by

$$\int_0^1 f(x)dx = \lim_{a \downarrow 0} \int_a^1 f(x)dx = 2.$$

5.4 Differential Equations

Differential equations are an alternative way to describe relationships between economic variables over time. Whereas difference equations work in a discrete setting, time is treated as a *continuous* variable in the analysis of differential equations. Therefore, the function y that describes the development of a variable over time now has as its domain a non-degenerate interval rather than the set $\mathbb{N}_0$. That is, we consider functions $y : A \mapsto \mathbb{R}$ where $A \subseteq \mathbb{R}$ is a non-degenerate interval. We assume that y is continuously differentiable as many times as required to ensure that the differential equations introduced below are well-defined.

Definition 5.4.1 *Let $A \subseteq I\!R$ be a non-degenerate interval, and let $n \in I\!N$. A* differential equation of order n *is an equation*

$$y^{(n)}(x) = G(x, y(x), y'(x), \ldots, y^{(n-1)}(x)) \tag{5.36}$$

where $G : I\!R \mapsto I\!R$ is a function and $y', y'', y^{(3)}, \ldots$ are the derivatives of y of order $1, 2, 3, \ldots$. A function $y : A \mapsto I\!R$ is a solution *of this differential equation if and only if (5.36) is satisfied for all $x \in A$.*

As in the case of difference equations, we restrict attention to certain types of differential equations.

Definition 5.4.2 *Let $A \subseteq I\!R$ be a non-degenerate interval, and let $n \in I\!N$. A* linear differential equation of order n *is an equation*

$$y^{(n)}(x) = b(x) + a_0(x)y(x) + a_1(x)y'(x) + \ldots + a_{n-1}(x)y^{(n-1)}(x) \tag{5.37}$$

where $b : A \mapsto I\!R$ and the $a_i : A \mapsto I\!R$ for all $i \in \{0, \ldots, n-1\}$ are continuous functions. If $b(x) = 0$ for all $x \in A$, the equation is a homogeneous linear differential equation of order n. *If there exists $x \in A$ such that $b(x) \neq 0$, the equation (5.37) is an* inhomogeneous linear differential equation of order n.

The continuity assumption regarding the functions b and $a_0, \ldots, a_{n-1}$ is not required for all of our results. For convenience, however, we impose it thoughout.

The homogeneous equation associated with the linear equation (5.37) is given by

$$y^{(n)}(x) = a_0(x)y(x) + a_1(x)y'(x) + \ldots + a_{n-1}(x)y^{(n-1)}(x). \tag{5.38}$$

Analogously to the results obtained for linear difference equations, there are some convenient properties of the set of solutions of a linear differential equations that will be useful in developing methods to solve these equations.

Theorem 5.4.3 *(i) Suppose $\hat{y}$ is a solution of (5.37). For each solution y of (5.37), there exists a solution z of the homogeneous equation (5.38) associated with (5.37) such that $y = z + \hat{y}$.*

(ii) If $\hat{y}$ is a solution of (5.37) and z is a solution of the homogeneous equation (5.38) associated with (5.37), then the function y defined by $y = z + \hat{y}$ is a solution of (5.37).

Proof. The proof of (i) is analogous to the proof of part (i) of Theorem 5.2.4.

To prove (ii), suppose $y = z + \hat{y}$ where $\hat{y}$ solves (5.37) and z solves (5.38). By definition,

$$y(x) = z(x) + \hat{y}(x) \quad \forall x \in A.$$

This is an identity and, differentiating both sides n times, we obtain

$$y^{(n)}(x) = z^{(n)}(x) = \hat{y}^{(n)}(x) \quad \forall x \in A. \tag{5.39}$$

Because $\hat{y}$ is a solution of (5.37), we have

$$\hat{y}^{(n)}(x) = b(x) + a_0(x)\hat{y}(x) + a_1(x)\hat{y}'(x) + \ldots + a_{n-1}(x)\hat{y}^{(n-1)}(x) \quad \forall x \in A$$

and because z is a solution of (5.38), it follows that

$$z^{(n)}(x) = a_0(x)z(x) + a_1(x)z'(x) + \ldots + a_{n-1}(x)z^{(n-1)}(x) \quad \forall x \in A.$$

Substituting into (5.39), we obtain

$$\begin{aligned} y^{(n)}(x) &= a_0(x)z(x) + a_1(x)z'(x) + \ldots + a_{n-1}(x)z^{(n-1)}(x) \\ &+ b(x) + a_0(x)\hat{y}(x) + a_1(x)\hat{y}'(x) + \ldots + a_{n-1}(x)\hat{y}^{(n-1)}(x) \\ &= b(x) + a_0(x)(z(x) + \hat{y}(x)) + a_1(x)(z'(x) + \hat{y}'(x)) + \ldots + a_{n-1}(x)(z^{(n-1)}(x) + \hat{y}^{(n-1)}(x)) \\ &= b(x) + a_0(x)y(x) + a_1(x)y'(x) + \ldots + a_{n-1}(x)y^{(n-1)}(x) \quad \forall x \in A \end{aligned}$$

and, thus, y satisfies (5.37). ‖

The next two results (the second of which we state without a proof) concern the structure of the set of solutions to the homogeneous equation (5.38). They are analogous to the corresponding theorems for linear difference equations.

Theorem 5.4.4 *If $z_1, \ldots, z_n$ are n solutions of (5.38) and $c = (c_1, \ldots, c_n) \in I\!R^n$ is a vector of arbitrary coefficients, then the function y defined by $y = \sum_{i=1}^n c_i z_i$ is a solution of (5.38).*

Proof. Suppose $z_1, \ldots, z_n$ are n solutions of (5.38) and $c = (c_1, \ldots, c_n) \in I\!R^n$ is a vector of arbitrary coefficients. Let $y = \sum_{i=1}^n c_i z_i$. Using this definition and differentiating n times, we obtain

$$y^{(n)}(x) = \sum_{i=1}^n c_i z_i^{(n)}(x)$$

$$\begin{aligned} &= \sum_{i=1}^{n} c_i(a_0(x)z_i(x) + a_1(t)z_i'(x) + \ldots + a_{n-1}(x)z_i^{(n-1)}(x)) \\ &= a_0(x)\sum_{i=1}^{n} c_i z_i(t) + a_1(t)\sum_{i=1}^{n} z_i'(x) + \ldots + a_{n-1}(x)\sum_{i=1}^{n} z_i^{(n-1)}(x) \\ &= a_0(x)y(x) + a_1(x)y'(x) + \ldots + a_{n-1}(x)y^{(n-1)}(x) \end{aligned}$$

for all $x \in A$ and, therefore, y is a solution of (5.38). ‖

Theorem 5.4.5 *Suppose $z_1, \ldots, z_n$ are n solutions of (5.38). The following two statements are equivalent.*

(i) For every solution y of (5.38), there exists a vector of coefficients $c \in I\!R^n$ such that $y = \sum_{i=1}^{n} c_i z_i$.

(ii) There exists $x_0 \in A$ such that

$$\begin{vmatrix} z_1(x_0) & z_2(x_0) & \ldots & z_n(x_0) \\ z_1'(x_0) & z_2'(x_0) & \ldots & z_n'(x_0) \\ \vdots & \vdots & & \vdots \\ z_1^{(n-1)}(x_0) & z_2^{(n-1)}(x_0) & \ldots & z_n^{(n-1)}(x_0) \end{vmatrix} \neq 0.$$

According to Theorem 5.4.5, it is sufficient to find a single point $x_0 \in A$ such that the determinant in statement (ii) of the theorem is non-zero in order to guarantee that the system of solutions $z_1, \ldots, z_n$ satisfies the required independence property.

We now consider solution methods for specific differential equations of order one. In general, a differential equation of order one can be expressed as

$$y'(x) = G(x, y(x)); \tag{5.40}$$

this is the special case obtained by setting $n = 1$ in (5.36). We discuss two types of differential equations of order one. The first type consists of separable equations, the second of linear equations.

Separable differential equations of order one are defined as follows.

Definition 5.4.6 *Let $A \subseteq I\!R$ be a non-degenerate interval. A* separable differential equation of order one *is an equation*

$$y'(x) = f(x)g(y(x)) \tag{5.41}$$

where $f : A \mapsto I\!R$ and $g : B \to I\!R$ are Riemann integrable functions, $B \subseteq y(A)$ is a nondegenerate interval and $g(y(x)) \neq 0$ for all $x \in A$ such that $y(x) \in B$.

Because the value of g is assumed to be different from zero for all points in the domain B of g, separable equations can be solved by a method that involves separating the equation in a way such that the function y to be obtained appears on one side of the equation only and the variable x appears on the other side only. This is a consequence of the substitution rule introduced in the previous section.

Dividing (5.41) by $g(y(x))$ (which is non-zero by the above hypothesis) yields

$$\frac{y'(x)}{g(y(x))} = f(x).$$

Integrating both sides with respect to x, we obtain

$$\int \frac{y'(x)dx}{g(y(x))} = \int f(x)dx. \tag{5.42}$$

Applying the substitution rule,

$$\int \frac{y'(x)dx}{g(y(x))} = \int \frac{dy(x)}{g(y(x))}. \tag{5.43}$$

Combining (5.42) and (5.43), we obtain

$$\int \frac{dy(x)}{g(y(x))} = \int f(x)dx.$$

Thus, we have proven

Theorem 5.4.7 *If y is a solution to the separable differential equation (5.41), then y satisfies*

$$\int \frac{dy(x)}{g(y(x))} = \int f(x)dx.$$

To illustrate the solution method introduced in this theorem, consider the example given by the following differential equation.

$$y'(x) = x^2(y(x))^2. \tag{5.44}$$

This is a separable equation with $f(x) = x^2$ for all $x \in A$ and $g(y(x)) = (y(x))^2$ for all $x \in A$ such that $y(x) \in B$. It follows that

$$\int \frac{dy(x)}{g(y(x))} = \int \frac{dy(x)}{(y(x))^2} = -\frac{1}{y(x)} + c_1$$

and

$$\int f(x)dx = \int x^2 dx = \frac{1}{3}x^3 + c_2$$

where $c_1, c_2 \in \mathbb{R}$ are constants of integration. According to Theorem 5.4.7, we must have

$$-\frac{1}{y(x)} + c_1 = \frac{1}{3}x^3 + c_2$$

or, defining $c := c_1 - c_2$ and simplifying,

$$y(x) = \frac{1}{c - x^3/3}$$

for all $x \in A$, where A and B must be chosen so that the denominator on the right side is non-zero.

If we impose an initial condition requiring the value of y at a specific point $x_0 \in A$ to be equal to a given value $y_0 \in B$, we can determine a unique solution of a separable equation. To obtain this solution, we use the definite integrals obtained by substituting the initial values into the equation of Theorem 5.4.7. We state the corresponding result without a proof.

Theorem 5.4.8 *Let $x_0 \in A$ and $y_0 \in B$. If y is a solution to the separable differential equation (5.41) and satisfies $y(x_0) = y_0$, then y satisfies*

$$\int_{y_0}^{y(x)} \frac{d\bar{y}(x)}{g(\bar{y}(x))} = \int_{x_0}^{x} f(\bar{x})d\bar{x}$$

for all $x \in A$ such that $y(x) \in B$.

Consider again the example given by (5.44), and suppose we require the initial condition $y(1) = 1$. We obtain

$$\int_{1}^{y(x)} \frac{d\bar{y}(x)}{g(\bar{y}(x))} = \int_{1}^{y(x)} \frac{d\bar{y}(x)}{(\bar{y}(x))^2} = -\frac{1}{(\bar{y}(x))}\Big|_{1}^{y(x)} = -\frac{1}{y(x)} + 1$$

and

$$\int_{1}^{x} f(\bar{x})d\bar{x} = \int_{1}^{x} \bar{x}^2 d\bar{x} = -\frac{1}{3}\bar{x}^3\Big|_{1}^{x} = -\frac{1}{3}x^3 - \frac{1}{3}.$$

According to Theorem 5.4.8, a solution y must satisfy

$$-\frac{1}{y(x)} + 1 = -\frac{1}{3}x^3 - \frac{1}{3}$$

and, solving, we obtain

$$y(x) = \frac{3}{4 - x^3}$$

for all $x \in A$ such that $y(x) \in B$. Again, the domains of f and g must be chosen so that the equation is required for values of x such that the denominator of the right side of this equation is non-zero.

Now we move on to linear differential equations of order one. We obtain these equations as the speical cases of (5.37) where $n = 1$, that is,

$$y'(x) = b(x) + a_0(x)y(x). \tag{5.45}$$

In contrast to linear difference equations of order one, there exists a solution method for linear differential equations of order one, including those equations where the coefficient a_0 is not constant.

Theorem 5.4.9 *A function $y : A \mapsto \mathbb{R}$ is a solution of the linear differential equation of order one (5.45) if and only if there exists $c \in \mathbb{R}$ such that*

$$y(x) = e^{\int a_0(x)dx} \left(\int b(x)e^{-\int a_0(x)dx}dx + c \right). \tag{5.46}$$

Proof. Clearly, (5.45) is equivalent to

$$y'(x) - a_0(x)y(x) = b(x).$$

Multiplying both sides by $e^{\int -a_0(x)dx}$, we obtain

$$e^{\int -a_0(x)dx} \left(y'(x) - a_0(x)y(x)\right) = b(x)e^{\int -a_0(x)dx}. \tag{5.47}$$

As can be verified easily by differentiating, the left side of (5.47) is equal to

$$\frac{d\left(y(x)e^{\int -a_0(x)dx}\right)}{dx}$$

and, thus, (5.47) is equivalent to

$$\frac{d\left(y(x)e^{\int -a_0(x)dx}\right)}{dx} = b(x)e^{\int -a_0(x)dx}.$$

Integrating both sides with respect to x, we obtain

$$y(x)e^{\int -a_0(x)dx} = \int b(x)e^{\int -a_0(x)dx}dx + c$$

where $c \in \mathbb{R}$ is the constant of integration. Solving for $y(x)$, we obtain (5.46). ‖

As an example, consider the equation

$$y'(x) = 3 - y(x).$$

This is a linear differential equation of order one with $b(x) = 3$ and $a_0(x) = -1$ for all $x \in A$. According to Theorem 5.4.9, we obtain

$$\begin{aligned} y(x) &= e^{-\int dx} \left(\int 3e^{\int dx}dx + c \right) \\ &= e^{-x} \left(\int 3e^{x}dx + c \right) \\ &= e^{-x}\left(3e^{x} + c\right) = 3 + ce^{-x} \end{aligned}$$

for all $x \in A$, where $c \in \mathbb{R}$ is a constant.

Again, an initial condition will determine the solution of the equation uniquely (by determining the value of the constant c). For instance, if we add the requirement $y(0) = 0$ in the above example, it follows that we must have

$$3 + ce^{-0} = 0$$

which implies $c = -3$. Thus, the unique solution satisfying the initial condition $y(0) = 0$ is given by

$$y(x) = 3 - 3e^{-x} \quad \forall x \in A.$$

In the case of linear differential equations of order two, solution methods are known for equations with constant coefficients but, unfortunately, these methods do not generalize to arbitrary linear equations. We therefore restrict attention to the class of equations with constant coefficients.

Definition 5.4.10 *Let $A \subseteq \mathbb{R}$ be a non-degenerate interval, and let $n \in \mathbb{N}$. A* linear differential equation with constant coefficients of order n *is an equation*

$$y^{(n)}(x) = b(x) + a_0 y(x) + a_1 y'(x) + \ldots + a_{n-1} y^{(n-1)}(x) \tag{5.48}$$

where $b : \mathbb{N}_0 \mapsto \mathbb{R}$ and $a_0, a_1, \ldots, a_{n-1} \in \mathbb{R}$. If $b(x) = 0$ for all $x \in A$, the equation is a homogeneous linear differential equation with constant coefficients of order n. *If there exists $x \in A$ such that $b(x) \neq 0$, the equation (5.48) is an* inhomogeneous linear differential equation with constant coefficients of order n.

The homogeneous equation associated with (5.48) is

$$y^{(n)}(x) = a_0 y(x) + a_1 y'(x) + \ldots + a_{n-1} y^{(n-1)}(x).$$

To solve linear differential equations with constant coefficients of order two, we proceed analogously to the way we did in the case of difference equations. First, we determine the general solution of the associated homogeneous equations, then we find a particular solution of the (inhomogeneous) equation to be solved and, finally, we use Theorem 5.4.3 to obtain the general solution.

For $n = 2$, the above definitions of linear differential equations with constant coefficients and their homogeneous counterparts become

$$y''(x) = b(x) + a_0 y(x) + a_1 y'(x) \tag{5.49}$$

and

$$y''(x) = a_0 y(x) + a_1 y'(x). \tag{5.50}$$

To solve (5.50), we set $z(x) = e^{\lambda x}$ for all $x \in A$. Substituting back, we obtain

$$\lambda^2 e^{\lambda x} = a_0 e^{\lambda x} + a_1 \lambda e^{\lambda x}$$

and the characteristic equation is

$$\lambda^2 - a_1 \lambda - a_0 = 0.$$

Analogously to linear difference equations with constant coefficients of order two, we have three possible cases.

Case I: $a_1^2/4 + a_0 > 0$. In this case, the characteristic polynomial has two distinct real roots given by $\lambda_1 = a_1/2 + \sqrt{a_1^2/4 + a_0}$ and $\lambda_2 = a_1/2 - \sqrt{a_1^2/4 + a_0}$, and we obtain the two solutions z_1 and z_2 defined by

$$z_1(x) = e^{\lambda_1 x} \quad \forall x \in A$$

and

$$z_2(x) = e^{\lambda_2 x} \quad \forall x \in A.$$

We obtain

$$\begin{vmatrix} z_1(0) & z_2(0) \\ z_1'(0) & z_2'(0) \end{vmatrix} = \begin{vmatrix} 1 & 1 \\ \lambda_1 & \lambda_2 \end{vmatrix} = \lambda_2 - \lambda_1 = -2\sqrt{a_1^2/4 + a_0} \neq 0.$$

Thus, the general solution of the homogeneous equation is given by

$$z(x) = c_1 z_1(t) + c_2 z_2(t) = c_1 e^{\lambda_1 x} + c_2 e^{\lambda_2 x} \quad \forall x \in A$$

where $c_1, c_2 \in \mathbb{R}$ are constants.

Case II: $a_1^2/4 + a_0 = 0$. Now the characteristic polynomial has a double root at $\lambda = a_1/2$, and the corresponding solutions are given by $z_1(x) = e^{a_1 x/2}$ and $z_2(x) = x e^{a_1 x/2}$ for all $x \in A$. We have

$$\begin{vmatrix} z_1(0) & z_2(0) \\ z_1'(0) & z_2'(0) \end{vmatrix} = \begin{vmatrix} 1 & 0 \\ a_1/2 & 1 \end{vmatrix} = 1 \neq 0.$$

Thus, a function $z : A \mapsto \mathbb{R}$ is a solution of the homogeneous equation if and only if there exist constants $c_1, c_2 \in \mathbb{R}$ such that

$$z(x) = c_1 z_1(t) + c_2 z_2(t) = c_1 e^{a_1 x/2} + c_2 x e^{a_1 x/2} \quad \forall x \in A.$$

Case III: $a_1^2/4 + a_0 < 0$. The characteristic polynomial has two complex roots, namely, $\lambda_1 = a + ib$ and $\lambda_2 = a - ib$, where $a = a_1/2$ and $b = \sqrt{-a_1^2/4 - a_0}$. Substituting back into the expression for a possible solution, we obtain the two functions $\hat{z}_1$ and $\hat{z}_2$ defined by

$$\hat{z}_1(x) = e^{(a+ib)x} = e^{ax}e^{ibx} \quad \forall x \in A$$

and

$$\hat{z}_2(x) = e^{(a-ib)x} = e^{ax}e^{-ibx} \quad \forall x \in A.$$

Using Euler's theorem (Theorem 5.1.5), the two solutions can be written as

$$\hat{z}_1(x) = e^{ax}(\cos(bx) + i\sin(bx)) \quad \forall x \in A$$

and

$$\hat{z}_2(x) = e^{ax}(\cos(bx) - i\sin(bx)) \quad \forall x \in A.$$

Because any linear combination of any two solutions is itself a solution, we obtain the solutions

$$\begin{aligned} z_1(x) &= \hat{z}_1(x)/2 + \hat{z}_2(x)/2 \\ &= e^{ax}(\cos(bx) + i\sin(bx))/2 + e^{ax}(\cos(bx) - i\sin(bx))/2 \\ &= e^{ax}\cos(bx) \end{aligned}$$

and

$$\begin{aligned} z_2(x) &= \hat{z}_1(x)/(2i) + \hat{z}_2(x)/(-2i) \\ &= e^{ax}(\cos(bx) + i\sin(bx))/(2i) - e^{ax}(\cos(bx) - i\sin(bx))/(2i) \\ &= e^{ax}\sin(bx) \end{aligned}$$

for all $x \in A$. We obtain

$$\begin{vmatrix} z_1(0) & z_2(0) \\ z_1'(0) & z_2'(0) \end{vmatrix} = \begin{vmatrix} 1 & 0 \\ a & b \end{vmatrix} = b = \sqrt{-a_1^2/4 - a_0} \neq 0$$

and, therefore, z is a solution of (5.50) if and only if there exist $c_1, c_2 \in \mathbb{R}$ such that

$$z(x) = c_1 z_1(x) + c_2 z_2(x) = c_1 e^{a_1 x/2}\cos\left(\sqrt{-a_1^2/4 - a_0}x\right) + c_2 e^{a_1 x/2}\sin\left(\sqrt{-a_1^2/4 - a_0}x\right) \quad \forall x \in A.$$

We summarize our observations regarding the solution of homogeneous linear differential equations with constant coefficients of order two in the following theorem.

Theorem 5.4.11 *(i) Suppose $a_0 \in \mathbb{R} \setminus \{0\}$ and $a_1 \in \mathbb{R}$ are such that $a_1^2/4 + a_0 > 0$. A function $z : A \mapsto \mathbb{R}$ is a solution of the homogeneous linear differential equation with constant coefficients of order two (5.50) if and only if there exist $c_1, c_2 \in \mathbb{R}$ such that*

$$z(x) = c_1 e^{\lambda_1 x} + c_2 e^{\lambda_2 x} \quad \forall x \in A$$

where $\lambda_1 = a_1/2 + \sqrt{a_1^2/4 + a_0}$ and $\lambda_2 = a_1/2 - \sqrt{a_1^2/4 + a_0}$

(ii) Suppose $a_0 \in \mathbb{R} \setminus \{0\}$ and $a_1 \in \mathbb{R}$ are such that $a_1^2/4 + a_0 = 0$. A function $z : A \mapsto \mathbb{R}$ is a solution of the homogeneous linear differential equation with constant coefficients of order two (5.50) if and only if there exist $c_1, c_2 \in \mathbb{R}$ such that

$$z(x) = c_1 e^{a_1 x/2} + c_2 x e^{a_1 x/2} \quad \forall x \in A.$$

(iii) Suppose $a_0 \in \mathbb{R} \setminus \{0\}$ and $a_1 \in \mathbb{R}$ are such that $a_1^2/4 + a_0 < 0$. A function $z : A \mapsto \mathbb{R}$ is a solution of the homogeneous linear differential equation with constant coefficients of order two (5.50) if and only if there exist $c_1, c_2 \in \mathbb{R}$ such that

$$z(x) = c_1 e^{a_1 x/2}\cos\left(\sqrt{-a_1^2/4 - a_0}x\right) + c_2 e^{a_1 x/2}\sin\left(\sqrt{-a_1^2/4 - a_0}x\right) \quad \forall x \in A.$$

We consider two types of inhomogeneity. The first is the case where the function b is a polynomial of degree two. In this case, our equation (5.49) becomes

$$y''(x) = b_0 + b_1 x + b_2 x^2 + a_0 y(x) + a_1 y'(x).$$

where $b_0, b_1 \in \mathbb{R}$ and $b_2 \in \mathbb{R} \setminus \{0\}$ are parameters (the case where $b_2 = 0$ leads to a polynomial of degree one or less, and the solution method to be employed is analogous; working out the details is left as an exercise).

To obtain a particular solution, we begin by checking whether a quadratic function will work. We set

$$\hat{y}(x) = \gamma_0 + \gamma_1 x + \gamma_2 x^2 \quad \forall x \in A$$

and, substituting into (5.49) and rearranging, we obtain

$$2\gamma_2 - a_0\gamma_0 - a_1\gamma_1 - (a_0\gamma_1 + 2a_1\gamma_2)x - a_0\gamma_2 x^2 = b_0 + b_1 x + b_2 x^2.$$

This has to be true for all values of $x \in A$ and, therefore, the coefficients of x^0, x^1 and x^2 have to be the same on both sides. Comparing coefficients, we obtain the system of equations

$$\begin{pmatrix} -a_0 & -a_1 & 2 \\ 0 & -a_0 & -2a_1 \\ 0 & 0 & -a_0 \end{pmatrix} \begin{pmatrix} \gamma_0 \\ \gamma_1 \\ \gamma_2 \end{pmatrix} = \begin{pmatrix} b_0 \\ b_1 \\ b_2 \end{pmatrix}. \tag{5.51}$$

This is a system of linear equations in γ_0, γ_1 and γ_2. There are two possible cases.

Case I: $a_0 \neq 0$. In this case, the matrix of coefficients in (5.51) is nonsingular and, consequently, (5.51) has a unique solution $(\gamma_0, \gamma_1, \gamma_2)$, and we obtain the particular solution

$$\hat{y}(x) = \gamma_0 + \gamma_1 x + \gamma_2 x^2 \quad \forall x \in A.$$

Case II: $a_0 = 0$. Because $b_2 \neq 0$, the system (5.51) has no solution and we have to find an alternative approach. Setting $\hat{y}(x) = x(\gamma_0 + \gamma_1 x + \gamma_2 x^2)$ for all $x \in A$, substituting into (5.49) and comparing coefficients, we obtain the system of equations

$$\begin{pmatrix} -a_1 & 2 & 0 \\ 0 & -2a_1 & 6 \\ 0 & 0 & -3a_1 \end{pmatrix} \begin{pmatrix} \gamma_0 \\ \gamma_1 \\ \gamma_2 \end{pmatrix} = \begin{pmatrix} b_0 \\ b_1 \\ b_2 \end{pmatrix}. \tag{5.52}$$

We have two subcases.

Subcase II.A: $a_1 \neq 0$. In this case, the matrix of coefficients is nonsingular and, consequently, (5.52) has a unique solution $(\gamma_0, \gamma_1, \gamma_2)$, and we obtain the particular solution

$$\hat{y}(x) = x(\gamma_0 + \gamma_1 x + \gamma_2 x^2) \quad \forall x \in A.$$

Subcase II.B: $a_1 = 0$. Because $b_2 \neq 0$, the system (5.51) has no solution and we have to try yet another functional form for a particular solution. Setting $\hat{y}(x) = x^2(\gamma_0 + \gamma_1 x + \gamma_2 x^2)$ for all $x \in A$, substituting into (5.49) and comparing coefficients, we now obtain the system of equations

$$\begin{pmatrix} 2 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 12 \end{pmatrix} \begin{pmatrix} \gamma_0 \\ \gamma_1 \\ \gamma_2 \end{pmatrix} = \begin{pmatrix} b_0 \\ b_1 \\ b_2 \end{pmatrix}$$

which has the unique solution $\gamma_0 = b_0/2$, $\gamma_1 = b_1/6$ and $\gamma_2 = b_2/12$. Substituting back, we obtain the particular solution

$$\hat{y}(x) = x^2(b_0/2 + b_1 x/6 + b_2 x^2/12) \quad \forall x \in A.$$

In all cases, the general solution is obtained as the sum of the particular solution and the general solution of the associated homogeneous equation. Thus, we obtain

Theorem 5.4.12 *(i) Let $b_0, b_1 \in \mathbb{R}$, $b_2 \in \mathbb{R} \setminus \{0\}$, $a_0 \in \mathbb{R} \setminus \{0\}$ and $a_1 \in \mathbb{R}$. A function $y : A \mapsto \mathbb{R}$ is a solution of the linear differential equation with constant coefficients of order two (5.49) if and only if*

$$y(x) = z(x) + \gamma_0 + \gamma_1 x + \gamma_2 x^2 \quad \forall x \in A$$

where $(\gamma_0, \gamma_1, \gamma_2)$ is the unique solution of (5.51) and z is a solution of the corresponding homogeneous equation (5.50).

(ii) Let $b_0, b_1 \in \mathbb{R}$, $b_2 \in \mathbb{R} \setminus \{0\}$, $a_0 = 0$ and $a_1 \in \mathbb{R} \setminus \{0\}$. A function $y : A \mapsto \mathbb{R}$ is a solution of the linear differential equation with constant coefficients of order two (5.49) if and only if

$$y(x) = z(x) + x(\gamma_0 + \gamma_1 x + \gamma_2 x^2) \quad \forall x \in A$$

where $(\gamma_0, \gamma_1, \gamma_2)$ is the unique solution of (5.52) and z is a solution of the corresponding homogeneous equation (5.50).

(iii) Let $b_0, b_1 \in \mathbb{R}$, $b_2 \in \mathbb{R} \setminus \{0\}$ and $a_0 = a_1 = 0$. A function $y : A \mapsto \mathbb{R}$ is a solution of the linear differential equation with constant coefficients of order two (5.49) if and only if

$$y(x) = z(x) + x^2(b_0/2 + b_1 x/6 + b_2 x^2/12) \quad \forall x \in A$$

where z is a solution of the corresponding homogeneous equation (5.50).

As an example, consider the equation

$$y''(x) = 1 + 2x + x^2 + y'(x).$$

The associated homogeneous equation is

$$y''(x) = y'(x).$$

Setting $z(x) = e^{\lambda x}$ for all $x \in A$, we obtain the characteristic polynomial $\lambda^2 - \lambda$ with the two distinct real roots $\lambda_1 = 1$ and $\lambda_0 = 0$. The corresponding solutions of the homogeneous equation are given by $z_1(x) = e^x$ and $z_2(x) = 1$ for all $x \in A$. We obtain

$$\begin{vmatrix} z_1(0) & z_2(0) \\ z_1'(0) & z_2'(0) \end{vmatrix} = \begin{vmatrix} 1 & 1 \\ 1 & 0 \end{vmatrix} = -1 \neq 0$$

and, therefore, the general solution of the homogeneous equation is given by

$$z(x) = c_1 e^x + c_2 \quad \forall x \in A$$

with constants $c_1, c_2 \in \mathbb{R}$. To obtain a particular solution of the inhomogeneous equation, we first set

$$\hat{y}(x) = \gamma_0 + \gamma_1 x + \gamma_2 x^2 \quad \forall x \in A.$$

Substituting back and comparing coefficients, we see that there exists no $(\gamma_0, \gamma_1, \gamma_2) \in \mathbb{R}^3$ satisfying the required system of equations (provide the details as an exercise). Thus, we use the alternative functional form

$$\hat{y}(x) = x(\gamma_0 + \gamma_1 x + \gamma_2 x^2) \quad \forall x \in A,$$

and comparing coefficients yields $\gamma_0 = -5$, $\gamma_1 = -2$ and $\gamma_2 = -1/3$. Therefore, the general solution is

$$y(x) = c_1 e^x + c_2 - 5x - 2x^2 - x^3/3 \quad \forall x \in A$$

where $c_1, c_2 \in \mathbb{R}$ are constants.

The second type of inhomogeneity is the case where b is an exponential function. Letting $b_0, b_1 \in \mathbb{R}$ be constants, the resulting differential equation is

$$y''(x) = b_0 e^{b_1 x} + a_0 y(x) + a_1 y'(x). \tag{5.53}$$

Because we have already solved the associated homogeneous equation (5.50), all that is left to do is to find a particular solution $\hat{y}$ of (5.53). We begin with the functional form $\hat{y}(x) = c_0 e^{b_1 x}$ for all $x \in A$, where $c_0 \in \mathbb{R}$ is a constant to be determined. Substituting into (5.53) and simplifying, we obtain

$$c_0(b_1^2 - a_0 - a_1 b_1) = b_0.$$

There are two cases to be considered.

Case I: $b_1^2 - a_0 - a_1 b_1 \neq 0$. In this case, we can solve for c_0 to obtain $c_0 = b_0/(b_1^2 - a_0 - a_1 b_1)$ and, thus, $\hat{y}(x) = b_0 e^{b_1 x}/(b_1^2 - a_0 - a_1 b_1)$ for all $x \in A$. The general solution of (5.53) is thus

$$y(x) = z(x) + b_0 e^{b_1 x}/(b_1^2 - a_0 - a_1 b_1) \quad \forall x \in A$$

where z is a solution of the corresponding homogeneous equation.

Case II: $b_1^2 - a_0 - a_1 b_1 = 0$. In this case, we have to try another functional form for the desired particular solution $\hat{y}$. We now use $\hat{y}(x) = c_0 x e^{b_1 x}$ for all $x \in A$ with the value of the constant $c_0 \in \mathbb{R}$ to be determined. Substituting $\hat{y}$ and the equation defining this case into (5.53) and simplifying, we obtain

$$c_0(2b_1 - a_1) = b_0.$$

There are two subcases.

Subcase II.A: $2b_1 - a_1 \neq 0$. In this case, we can solve for c_0 to obtain $c_0 = b_0/(2b_1 - a_1)$ and, thus, $\hat{y}(x) = b_0 x e^{b_1 x}/(2b_1 - a_1)$ for all $x \in A$. The general solution of (5.53) is

$$y(x) = z(x) + b_0 x e^{b_1 x}/(2b_1 - a_1) \quad \forall x \in A$$

where, again, z is a solution of the corresponding homogeneous equation.

Subcase II.B: $2b_1 - a_1 = 0$. Now the approach of subcase II.A does not yield a solution, and we try the functional form $\hat{y}(x) = c_0 x^2 e^{b_1 x}$ for all $x \in A$, where $c_0 \in \mathbb{R}$ is a constant to be determined. Substituting and simplifying, we obtain $c_0 = b_0/2$. This gives us the particular solution $\hat{y}(x) = b_0 x^2 e^{b_1 x}/2$ for all $x \in A$ and, finally, the general solution

$$y(x) = z(x) + b_0 x^2 e^{b_1 x}/2 \quad \forall x \in A$$

where z is a solution of the corresponding homogeneous equation. In summary, we obtain

Theorem 5.4.13 *(i) Let $b_0, b_1, a_1 \in \mathbb{R}$ and $a_0 \in \mathbb{R} \setminus \{b_1^2 - a_1 b_1\}$. A function $y : A \mapsto \mathbb{R}$ is a solution of the linear differential equation with constant coefficients of order two (5.53) if and only if*

$$y(x) = z(x) + b_0 e^{b_1 x}/(b_1^2 - a_0 - a_1 b_1) \quad \forall x \in A$$

where z is a solution of the corresponding homogeneous equation (5.50).

(ii) Let $b_0, b_1 \in \mathbb{R}$, $a_1 \in \mathbb{R} \setminus \{2b_1\}$ and $a_0 = b_1^2 - a_1 b_1$. A function $y : A \mapsto \mathbb{R}$ is a solution of the linear differential equation with constant coefficients of order two (5.53) if and only if

$$y(x) = z(x) + b_0 x e^{b_1 x}/(2b_1 - a_1) \quad \forall x \in A$$

where z is a solution of the corresponding homogeneous equation (5.50).

(iii) Let $b_0, b_1 \in \mathbb{R}$, $a_0 = -b_1^2$ and $a_1 = 2b_1$. A function $y : A \mapsto \mathbb{R}$ is a solution of the linear differential equation with constant coefficients of order two (5.53) if and only if

$$y(x) = z(x) + b_0 x^2 e^{b_1 x}/2 \quad \forall x \in A$$

where z is a solution of the corresponding homogeneous equation (5.50).

As an example, consider the equation

$$y''(x) = 2e^{3x} - y(x) - 2y'(x). \tag{5.54}$$

The corresponding homogeneous equation is

$$y''(x) = y(x) - 2y'(x). \tag{5.55}$$

Using the functional form $z(x) = e^{\lambda x}$ for all $x \in A$, we obtain the characteristic polynomial $\lambda^2 + 2\lambda + 1$ which has the double real root $\lambda = -1$. Thus, the two functions z_1 and z_2 given by $z_1(x) = e^{-x}$ and $z_2(x) = xe^{-x}$ are solutions of (5.55). We obtain

$$\begin{vmatrix} z_1(0) & z_2(0) \\ z_1'(0) & z_2'(0) \end{vmatrix} = \begin{vmatrix} 1 & 0 \\ -1 & -2 \end{vmatrix} = -2 \neq 0$$

and, therefore, the general solution of (5.55) is given by

$$z(x) = c_1 e^{-x} + c_2 x e^{-x} \quad \forall x \in A$$

with constants $c_1, c_2 \in \mathbb{R}$.

To obtain a particular solution of (5.54), we begin with the functional form $\hat{y}(x) = c_0 e^{3x}$ for all $x \in A$. Substituting into (5.54) and solving, we obtain $c_0 = 1/8$ and, thus, the particular solution given by $\hat{y}(x) = e^{3x}/8$ for all $x \in A$. Thus, the general solution is

$$y(x) = c_1 e^{-x} + c_2 x e^{-x} + e^{3x}/8 \quad \forall x \in A$$

with constants $c_1, c_2 \in \mathbb{R}$.

Now consider the equation

$$y''(x) = 2e^{3x} + 3y(x) + 2y'(x). \tag{5.56}$$

The corresponding homogeneous equation is

$$y''(x) = 3y(x) + 2y'(x). \tag{5.57}$$

Using the functional form $z(x) = e^{\lambda x}$ for all $x \in A$, we obtain the characteristic polynomial $\lambda^2 - 2\lambda - 3$ which has the two distinct real roots $\lambda_1 = 3$ and $\lambda_2 = -1$. Thus, the two functions z_1 and z_2 given by $z_1(x) = e^{3x}$ and $z_2(x) = e^{-x}$ are solutions of (5.55). We obtain

$$\begin{vmatrix} z_1(0) & z_2(0) \\ z_1'(0) & z_2'(0) \end{vmatrix} = \begin{vmatrix} 1 & 1 \\ 3 & -1 \end{vmatrix} = -4 \neq 0$$

and, therefore, the general solution of (5.55) is given by

$$z(x) = c_1 e^{3xx} + c_2 e^{-x} \quad \forall x \in A$$

with constants $c_1, c_2 \in \mathbb{R}$.

To obtain a particular solution of (5.56), we begin with the functional form $\hat{y}(x) = c_0 e^{3x}$ for all $x \in A$. Substituting into (5.56) and simplifying, it follows that this approach does not yield a solution. Next, we try the functional form $\hat{y}(x) = c_0 x e^{3x}$ for all $x \in A$. Substituting into (5.56) and solving, we obtain $c_0 = 1/2$ and, thus, the particular solution given by $\hat{y}(x) = xe^{3x}/2$ for all $x \in A$. Therefore, the general solution is

$$y(x) = c_1 e^{3x} + c_2 e^{-x} + xe^{3x}/2 \quad \forall x \in A$$

with constants $c_1, c_2 \in \mathbb{R}$.

The final topic discussed in this chapter is the question whether the solutions to given linear differential equations with constant coefficients of order two possess a *stability* property. Stability deals with the question how the long-run behavior of a solution is affected by changes in the initial conditions. For reasons that will become apparent, we assume that the domain A is not bounded from above.

Definition 5.4.14 *Let $A \subseteq \mathbb{R}$ be an interval such that $(a, \infty) \subseteq A$ for some $a \in \mathbb{R}$. The equation (5.49) is* stable *if and only if*

$$\lim_{x \to \infty} z(x) = 0$$

for all solutions z of the associated homogeneous equation.

According to this definition, every solution of the homogeneous equation corresponding to a linear differential equation of order two must converge to zero as x approaches infinity in order for the equation to be stable. Recall that any solution y of (5.49) can be expressed as the sum $z + \hat{y}$ for a particular solution $\hat{y}$ of (5.49) and a suitably chosen solution z of the corresponding homogeneous equation. Thus, if the equation is stable and $\hat{y}$ has a limit as x approaches infinity, we obtain

$$\lim_{x\to\infty} y(x) = \lim_{x\to\infty} (z(x) + \hat{y}(x)) = \lim_{x\to\infty} z(x) + \lim_{x\to\infty} \hat{y}(x) = 0 + \lim_{x\to\infty} \hat{y}(x) = \lim_{x\to\infty} \hat{y}(x).$$

Recall that any solution z of the associated homogeneous equation can be written as $z(x) = c_1 z_1(x) + c_2 z_2(x)$ for all $x \in A$, where z_1 and z_2 are linearly independent solutions and $c_1, c_2 \in \mathbb{R}$ are parameters. Intital conditions determine the values of the parameters c_1 and c_2 but do not affect the particular solution $\hat{y}$. Therefore, if the equation is stable, its limiting behavior is independent of the intial conditions, which is what stability is about. As a preliminary result, we obtain

Theorem 5.4.15 *Let $A \subseteq I\!R$ be an interval such that $(a, \infty) \subseteq A$ for some $a \in I\!R$, and let $z_1 : A \mapsto I\!R$ and $z_2 : A \mapsto I\!R$. Then*

$$\left[\lim_{x\to\infty}(c_1 z_1(x) + c_2 z_2(x)) = 0 \quad \forall c_1, c_2 \in I\!R\right] \Leftrightarrow \left[\lim_{x\to\infty} z_1(x) = 0 \text{ and } \lim_{x\to\infty} z_2(x) = 0\right].$$

Proof. Suppose first that $\lim_{x\to\infty}(c_1 z_1(x) + c_2 z_2(x)) = 0$ for all $c_1, c_2 \in I\!R$. This implies, in particular, that this limit is equal to zero for $c_1 = 1$ and $c_2 = 0$. Substituting, we obtain $\lim_{x\to\infty} z_1(x) = 0$. Analogously, for $c_1 = 0$ and $c_2 = 1$, we obtain $\lim_{x\to\infty} z_2(x) = 0$.

To prove the converse implication, suppose that $\lim_{x\to\infty} z_1(x) = 0$ and $\lim_{x\to\infty} z_2(x) = 0$. This implies immediately that $\lim_{x\to\infty}(c_1 z_1(x) + c_2 z_2(x)) = 0$ for all $c_1, c_2 \in I\!R$. ‖

By Theorem 5.4.15, all that is required to check a linear differential equation of order two for stability is to examine the limits of the two linearly independent solutions of the corresponding homogeneous equation: if both limits exist and are equal to zero, the equation is stable, and if one of the limits does not exist or is different from zero, it is not.

The characteristic polynomial of the homogeneous equation corresponding to (5.49) is

$$\lambda^2 - a_1\lambda - a_0. \tag{5.58}$$

If this polynomial has two distinct real roots λ_1 and λ_2, we obtain the solutions $z_1(x) = e^{\lambda_1 x}$ and $z_2(x) = e^{\lambda_2 x}$ for all $x \in A$. If $\lambda_1 > 0$, we obtain $\lim_{x\to\infty} z_1(x) = \infty$ and if $\lambda_1 = 0$, it follows that $\lim_{x\to\infty} z_1(x) = 1$. For $\lambda_1 < 0$, we obtain $\lim_{x\to\infty} z_1(x) = 0$. Analogously, $\lim_{x\to\infty} z_2(x) = \infty$ for $\lambda_2 > 0$, $\lim_{x\to\infty} z_2(x) = 1$ for $\lambda_2 = 0$ and $\lim_{x\to\infty} z_2(x) = 0$ for $\lambda_2 < 0$. Therefore, in this case, the equation is stable if and only if both real roots are negative. Because λ_1 and λ_2 are roots of the characteristic polynomial, it follows that $\lambda_1 + \lambda_2 = a_1$ and $\lambda_1\lambda_2 = -a_0$ (verify this as an exercise). If λ_1 and λ_2 are both negative, it follows that $a_1 = \lambda_1 + \lambda_2 < 0$ and $a_0 = -\lambda_1\lambda_2 < 0$. Conversely, if a_0 and a_1 are both negative, we obtain $\lambda_1 + \lambda_2 = a_1 < 0$ which implies that at least one of the two numbers λ_1 and λ_2 must be negative. Thus, because $\lambda_1\lambda_2 = -a_0 > 0$, it follows that both λ_1 and λ_2 must be negative. Therefore, in the case of two distinct real roots of the characteristic polynomial, the equation (5.49) is stable if and only if $a_0 < 0$ and $a_1 < 0$.

Now suppose that (5.58) has one double root λ. In this case, substituting the corresponding solutions $z_1(x) = e^{\lambda x}$ and $z_2(x) = xe^{\lambda x}$ for all $x \in A$, it follows again that the equation is stable if and only if $\lambda < 0$. Because $\lambda = a_1/2$ and $a_0 = -a_1^2/4$ in this case, this is again equivalent to $a_0 < 0$ and $a_1 < 0$.

Finally, for two complex roots λ_1 and λ_2, we again obtain $\lambda_1 + \lambda_2 = a_1$ and $\lambda_1\lambda_2 = -a_0$ and the equation (5.49) is stable if and only if $a_0 < 0$ and $a_1 < 0$.

Summarizing, we obtain

Theorem 5.4.16 *Let $A \subseteq I\!R$ be an interval such that $(a, \infty) \subseteq A$ for some $a \in I\!R$. The equation (5.49) is stable if and only if $a_0 < 0$ and $a_1 < 0$.*

Chapter 6

Exercises

6.1 Chapter 1

1.1.1 Find the negations of the following statements.

(i) All students registered in this course are female;
(ii) $x > 3 \ \vee \ x < 2$;
(iii) For any real number x, there exists a real number y such that $x + y = 0$.

1.1.2 Which of the following statements is (are) true, which is (are) false?

(i) $3 < 2 \ \Rightarrow$ Ontario is a province in Canada;
(ii) $3 < 2 \ \Rightarrow$ Ontario is not a province in Canada;
(iii) $3 > 2 \ \Rightarrow$ Ontario is a province in Canada;
(iv) $3 > 2 \ \Rightarrow$ Ontario is not a province in Canada.

1.1.3 Find a statement which is equivalent to the statement $a \Leftrightarrow b$ using negation and conjunction only.

1.1.4 Suppose x and y are natural numbers. Prove:

xy is odd $\Leftrightarrow$ (x is odd) $\wedge$ (y is odd).

1.2.1 Let A and B be sets. Prove:

(i) $A \cap (A \cap B) = A \cap B$;
(ii) $A \cup (A \cap B) = A$.

1.2.2 Let $X = \mathbb{N}$ be the universal set. Find the complements of the following sets.

(i) $A = \{1, 2\}$;
(ii) $B = \mathbb{N}$;
(iii) $C = \{x \in \mathbb{N} \mid (x \text{ is odd}) \ \wedge \ x \geq 6\}$.

1.2.3 Let X be a universal set, and let $A, B \subseteq X$. Prove:

$X \setminus (A \cup B) = (X \setminus A) \cap (X \setminus B)$.

1.2.4 Let $A := \{x \in \mathbb{N} \mid x \text{ is even}\}$ and $B := \{x \in \mathbb{N} \mid x \text{ is odd}\}$. Which of the following pairs are elements of $A \times B$, which are not?

(i) $(2, 2)$;
(ii) $(2, 3)$;
(iii) $(3, 2)$;
(iv) $(3, 3)$;
(v) $(2, 1)$;
(vi) $(1, 2)$.

1.3.1 Which of the following subsets of $\mathbb{R}$ is (are) open in $\mathbb{R}$?

(i) $A = \{2\}$;

(ii) $B = [0, 1)$;
(iii) $C = (0, 1)$;
(iv) $D = (0, 1) \cup \{2\}$.

1.3.2 Which of the sets in Problem 1.3.1 is (are) closed in $\mathbb{R}$?

1.3.3 (i) Prove that the intersection of two convex sets is convex.
(ii) Give an example of two convex sets such that the union of these sets is not convex.

1.3.4 For each of the following subsets of $\mathbb{R}$, determine whether or not it has an infimum (a supremum, a minimum, a maximum).

(i) $A = \mathbb{N}$;
(ii) $B = \mathcal{Z}$;
(iii) $C = (0, 1)$;
(iv) $D = (-\infty, 0]$.

1.4.1 Consider the function $f : \mathbb{N} \mapsto \mathbb{N},\ x \mapsto 2x - 1$. Find $f(\mathbb{N})$ and $f(\{x \in \mathbb{N} \mid x \text{ is odd}\})$.

1.4.2 Use a diagram to illustrate the graph of the following function.

$$f : [0, 2] \mapsto \mathbb{R},\ x \mapsto \begin{cases} x & \text{if } x \in [0, 1] \\ 0 & \text{if } x \in (1, 2]. \end{cases}$$

1.4.3 Which of the following functions is (are) surjective (injective, bijective)?

(i) $f : \mathbb{R} \mapsto \mathbb{R},\ x \mapsto |x|$;
(ii) $f : \mathbb{R} \mapsto \mathbb{R}_+,\ x \mapsto |x|$;
(iii) $f : \mathbb{R}_+ \mapsto \mathbb{R},\ x \mapsto |x|$;
(iv) $f : \mathbb{R}_+ \mapsto \mathbb{R}_+,\ x \mapsto |x|$.

1.4.4 Find all permutations of $A = \{1, 2, 3\}$.

1.5.1 For each of the following sequences, determine whether or not it converges. In case of convergence, find the limit.

(i) $a_n = (-1)^n/n \ \ \forall n \in \mathbb{N}$;
(ii) $b_n = n(-1)^n \ \ \forall n \in \mathbb{N}$.

1.5.2 Show that the sequence defined by $a_n = n^2 - n$ for all $n \in \mathbb{N}$ diverges to ∞.

1.5.3 Which of the following sequences is (are) monotone nondecreasing (monotone nonincreasing, bounded, convergent)?

(i) $a_n = 1 + 1/n \ \ \forall n \in \mathbb{N}$;
(ii) $b_n = 1 + (-1)^n/n \ \ \forall n \in \mathbb{N}$.

1.5.4 Find the sequences $\{a_n + b_n\}$ and $\{a_n b_n\}$, where the sequences $\{a_n\}$ and $\{b_n\}$ are defined by

$a_n = 2n^2 \ \ \forall n \in \mathbb{N}, \ \ b_n = 1/n \ \ \forall n \in \mathbb{N}$.

6.2 Chapter 2

2.1.1 For $n \in \mathbb{N}$ and $x \in \mathbb{R}^n$, let $\|x\|$ denote the norm of x. Prove that, for all $x \in \mathbb{R}^n$,

(i) $\|x\| \geq 0$;
(ii) $\|x\| = 0 \ \Leftrightarrow \ x = \mathbf{0}$;
(iii) $\|x\| = \|-x\|$.

2.1.2 For $n \in \mathbb{N}$ and $x, y \in \mathbb{R}^n$, let $d(x, y)$ denote the Euclidean distance between x and y. Prove that, for all $x, y \in \mathbb{R}^n$,

(i) $d(x, y) \geq 0$;
(ii) $d(x, y) = 0 \ \Leftrightarrow \ x = y$;

(iii) $d(x, y) = d(y, x)$.

2.1.3 Let $x = (1, 3, -5)$ and $y = (2, 0, 1)$. Find (i) the sum of x and y; (ii) the inner product of x and y.

2.1.4 Let $x^1 = (1, 0, 2, -1)$, $x^2 = (0, 0, 1, 2)$, $x^3 = (2, -4, 1, 9)$. Are the vectors x^1, x^2, x^3 linearly independent?

2.2.1 For each of the following matrices, determine whether it is (i) a square matrix; (ii) a symmetric matrix.

$$A = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 0 \\ 1 & 2 & 1 \end{pmatrix}; \; B = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}; \; C = \begin{pmatrix} -1 & 1 & 0 \\ 1 & -2 & -1 \\ 0 & -1 & 2 \end{pmatrix}.$$

2.2.2 Find the matrix product AB, where

$$A = \begin{pmatrix} 1 & 2 & 1 & -3 \\ 0 & 1 & 0 & 8 \\ 1 & 2 & 0 & 1 \end{pmatrix}; \; B = \begin{pmatrix} -1 & 0 & 4 \\ 2 & 1 & -2 \\ 0 & 1 & 6 \\ 3 & 2 & -5 \end{pmatrix}.$$

2.2.3 For each of the following matrices, determine its rank.

$$A = (0); \; B = (1); \; C = \begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix}; \; D = \begin{pmatrix} 2 & 1 & 1 \\ 0 & 1 & -1 \end{pmatrix}.$$

2.2.4 Find two 2×2 matrices A and B such that $R(A) = R(B) = 2$ and $R(A + B) = 0$.

2.3.1 Use the elimination method to solve the system of equations $Ax = b$, where

$$A = \begin{pmatrix} 1 & 2 & 0 & 4 & 2 \\ 0 & 2 & 2 & -1 & 3 \\ 1 & -1 & 0 & 1 & 1 \\ 1 & 1 & 2 & 0 & 4 \end{pmatrix}; \; b = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 2 \end{pmatrix}.$$

2.3.2 Use the elimination method to solve the system of equations $Ax = b$, where

$$A = \begin{pmatrix} 2 & 1 & 4 & 0 \\ 0 & 0 & 1 & 1 \\ 4 & -1 & 1 & 0 \\ -2 & 2 & 5 & 2 \end{pmatrix}; \; b = \begin{pmatrix} 2 \\ -1 \\ -2 \\ 2 \end{pmatrix}.$$

2.3.3 Use the elimination method to solve the system of equations $Ax = b$, where

$$A = \begin{pmatrix} 2 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & -2 & 2 \end{pmatrix}; \; b = \begin{pmatrix} 1 \\ 0 \\ -2 \\ 0 \end{pmatrix}.$$

2.3.4 Prove: If $\alpha \in \mathbb{R}$ and x^* solves $Ax = b$, then αx^* solves $Ax = \alpha b$.

2.4.1 Which of the following matrices is (are) nonsingular?

$$A = (0); \; B = (1); \; C = \begin{pmatrix} 2 & 1 \\ -2 & 1 \end{pmatrix}; \; D = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}.$$

2.4.2 Find a 3×3 matrix A such that $A \neq E$ and $AA' = E$.

2.4.3 Calculate the product $AA'A^{-1}$, where

$$A = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}.$$

2.4.4 Suppose A is a nonsingular square matrix, and A^{-1} is the inverse of A. Let $\alpha \in \mathbb{R}$, $\alpha \neq 0$. Show that αA is nonsingular, and find $(\alpha A)^{-1}$.

2.5.1 Use the definition of a determinant to find $|A|$, where

$$A = \begin{pmatrix} 2 & 1 & 0 & 3 \\ 3 & 1 & 1 & 1 \\ 0 & 2 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}.$$

2.5.2 Find the determinant of the matrix in 2.5.1 by expansion along (i) the first column; (ii) the last row.

2.5.3 Show that A is nonsingular, and use Cramer's rule to solve the system $Ax = b$, where

$$A = \begin{pmatrix} 3 & 1 & 0 \\ 1 & 0 & 1 \\ 2 & -1 & 0 \end{pmatrix}; \; b = \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix}.$$

2.5.4 Find the adjoint and the inverse of the matrix in 2.5.3.

2.6.1 Determine the definiteness properties of the matrix

$$A = \begin{pmatrix} 3 & 1 & 1 \\ 1 & 2 & 0 \\ 1 & 0 & 1 \end{pmatrix}.$$

2.6.2 Find all principal minors of the following matrix A. Is A (i) positive semidefinite? (ii) negative semidefinite?

$$A = \begin{pmatrix} 2 & 3 & -1 \\ 3 & 1 & 2 \\ -1 & 2 & 0 \end{pmatrix}.$$

2.6.3 Let A be a symmetric square matrix. Prove:
A is positive definite $\Leftrightarrow$ $(-1)A$ is negative definite.

2.6.4 Give an example of a symmetric 3×3 matrix which is positive semidefinite, but not positive definite and not negative semidefinite.

6.3 Chapter 3

3.1.1 Consider the function f defined by

$$f : \mathbb{R} \mapsto \mathbb{R}, \; x \mapsto \begin{cases} 0 & \text{if } x \leq 0 \\ 1 & \text{if } x > 0. \end{cases}$$

Use Theorem 3.1.2 to show that this function is not continuous at $x_0 = 0$.

3.1.2 Let

$$f : \mathbb{R} \mapsto \mathbb{R}, \; x \mapsto \begin{cases} x & \text{if } x \leq 1 \\ 2x - 1 & \text{if } x > 1. \end{cases}$$

Is f continuous at $x_0 = 1$? Justify your answer rigorously.

3.1.3 Let

$$f : \mathbb{R} \mapsto \mathbb{R}, \; x \mapsto \begin{cases} x & \text{if } x \leq 1 \\ x - 1 & \text{if } x > 1. \end{cases}$$

Is f continuous at $x_0 = 1$? Justify your answer rigorously.

3.1.4 Let

$$f : [0, 3] \mapsto \mathbb{R}, \; x \mapsto \begin{cases} 0 & \text{if } x \in [0, 1) \\ x - 1 & \text{if } x \in [1, 2) \\ 2 & \text{if } x \in [2, 3]. \end{cases}$$

Is f (i) monotone nondecreasing? (ii) monotone increasing? (iii) monotone nonincreasing? (iv) monotone decreasing? Justify your answers rigorously.

3.2.1 Let

$$f : \mathbb{R} \mapsto \mathbb{R},\ x \mapsto \begin{cases} x & \text{if } x \leq 1 \\ 2x - 1 & \text{if } x > 1. \end{cases}$$

Is f differentiable at $x_0 = 1$? Justify your answer rigorously. If yes, find the derivative of f at $x_0 = 1$.

3.2.2 Find the derivative of the function $f : \mathbb{R}_{++} \mapsto \mathbb{R},\ x \mapsto \left(\sqrt{x+1} + x^2\right)^4$.

3.2.3 Find the derivative of the function $f : \mathbb{R} \mapsto \mathbb{R},\ x \mapsto e^{2x^2} \ln(x^2 + 1)$.

3.2.4 Find the derivative of the function $f : \mathbb{R} \mapsto \mathbb{R},\ x \mapsto e^{2(\cos(x)+1)} + 4(\sin(x))^3$.

3.3.1 (i) Find all local maxima and minima of $f : \mathbb{R} \mapsto \mathbb{R},\ x \mapsto 2 - x - 4x^2$.

(ii) Find all local and global maxima and minima of $f : [0, 4] \mapsto \mathbb{R},\ x \mapsto 2 - x - 4x^2$.

(iii) Find all local and global maxima and minima of $f : (0, 4] \mapsto \mathbb{R},\ x \mapsto x^2$.

3.3.2 Consider the function $f : (0, 1) \mapsto \mathbb{R},\ x \mapsto \sqrt{x} - x/2$. Is this function (i) monotone nondecreasing? (ii) monotone increasing? (iii) monotone nonincreasing? (iv) monotone decreasing? Use Theorem 3.3.11 to answer this question.

3.3.3 Let $f : (0, 1) \mapsto \mathbb{R},\ x \mapsto 2\ln(x)$ and $g : (0, 1) \mapsto \mathbb{R},\ x \mapsto x - 1$. Find

$$\lim_{x \uparrow 1} \frac{f(x)}{g(x)}.$$

3.3.4 Let $f : \mathbb{R}_{++} \mapsto \mathbb{R},\ x \mapsto \ln(x)$. Find the second-order Taylor polynomial of f around $x_0 = 1$, and use this polynomial to approximate the value of f at $x = 2$.

3.4.1 Consider the function $f : \mathbb{R} \mapsto \mathbb{R},\ x \mapsto ax + b$, where $a, b \in \mathbb{R}$ are constants and $a \neq 0$. Use Definition 3.4.1 to show that f is concave and convex. Is f (i) strictly concave? (ii) strictly convex?

3.4.2 Let $f : \mathbb{R}_{++} \mapsto \mathbb{R},\ x \mapsto x^\alpha$ with $\alpha \in (0, 1)$. Use Theorem 3.4.4 to show that f is strictly concave.

3.4.3 Let $f : [0, 4] \mapsto \mathbb{R},\ x \mapsto 1 - x - x^2$. Show that f must have a unique global maximum, and find this maximum.

3.4.4 The cost function of a perfectly competitive firm is given by $C : \mathbb{R}_+ \mapsto \mathbb{R},\ y \mapsto y + 2y^2$. Find the supply function and the profit function of this firm.

6.4 Chapter 4

4.1.1 Consider the distance function

$$d(x, y) = \max(\{|x_i - y_i| \mid i \in \{1, \ldots, n\}\})\ \ \forall x, y \in \mathbb{R}^n.$$

Prove that, for all $x, y \in \mathbb{R}^n$,

(i) $d(x, y) \geq 0$;
(ii) $d(x, y) = 0 \ \Leftrightarrow\ x = y$;
(iii) $d(x, y) = d(y, x)$.

4.1.2 For $x^0 \in \mathbb{R}^n$ and $\varepsilon \in \mathbb{R}_{++}$, let $\mathcal{U}_\varepsilon(x^0)$ be the ε-neighborhood of x^0 as defined in (4.1), and let $\mathcal{U}^E_\varepsilon(x^0)$ be the ε-neighborhood of x^0 as defined in (4.2). Let $x^0 = (0, 0) \in \mathbb{R}^2$ and $\varepsilon = 1$.

(i) Find a $\delta \in \mathbb{R}_{++}$ such that $\mathcal{U}_\delta(x^0) \subseteq \mathcal{U}^E_\varepsilon(x^0)$.

(ii) Find a $\delta \in \mathbb{R}_{++}$ such that $\mathcal{U}^E_\delta(x^0) \subseteq \mathcal{U}_\varepsilon(x^0)$.

4.1.3 Which of the following sets is (are) open in $\mathbb{R}^2$?

(i) $A = (0,1) \times (1,2)$;
(ii) $B = \{x \in \mathbb{R}^2 \mid 0 < x_1 < 1 \ \wedge \ x_1 = x_2\}$;
(iii) $C = \{x \in \mathbb{R}^2 \mid (x_1)^2 + (x_2)^2 < 1\}$.

4.1.4 Consider the sequence $\{a^m\}$ in $\mathbb{R}^2$ defined by

$$a^m = \left(1 + \left(\frac{1}{m}\right)(-1)^m, 2 - \frac{m}{m+1}\right) \quad \forall m \in \mathbb{N}.$$

Show that this sequence converges, and find its limit.

4.2.1 Which of the following subsets of $\mathbb{R}^2$ is (are) convex?

(i) $A = (0,1) \times (1,2)$;
(ii) $B = \{x \in \mathbb{R}^2 \mid 0 < x_1 < 1 \ \wedge \ x_1 = x_2\}$;
(iii) $C = \{x \in \mathbb{R}^2 \mid x_1 = 0 \ \vee \ x_2 = 0\}$.

4.2.2 Let $f : \mathbb{R}^2 \mapsto \mathbb{R}, \ x \mapsto \min(\{x_1, x_2\})$. Find the level set of f for $y = 1$, and illustrate this level set in a diagram.

4.2.3 Let

$$f : \mathbb{R}^2 \mapsto \mathbb{R}, \ x \mapsto \begin{cases} x_1 + x_2 & \text{if } x \neq (0,0) \\ 1 & \text{if } x = (0,0). \end{cases}$$

Show that f is not continuous at $x^0 = (0,0)$.

4.2.4 Let $f : \mathbb{R}^2 \mapsto \mathbb{R}, \ x \mapsto x_1 x_2$. Illustrate the graph of the function $f^1 : \mathbb{R} \mapsto \mathbb{R}, \ x \mapsto f(x,2)$ in a diagram.

4.3.1 Find all partial derivatives of the function

$$f : \mathbb{R}^3_{++} \mapsto \mathbb{R}, \ x \mapsto \sqrt{x_1 x_2} + e^{x_1 x_3}.$$

4.3.2 Find the total differential of the function defined in 4.3.1 at the point $x^0 = (1,1,1)$.

4.3.3 Find the Hessian matrix of the function defined in 4.3.1 at a point $x \in \mathbb{R}^3_{++}$.

4.3.4 Use the implicit function theorem to show that the equation

$$e^{yx_1} + yx_1 x_2 - e^y = 0$$

defines an implicit function in a neighborhood of $x^0 = (1,1)$, and find the partial derivatives of this implicit function at $(1,1)$.

4.4.1 Find all local maxima and minima of $f : \mathbb{R}^2 \mapsto \mathbb{R}, \ x \mapsto (x_1)^2 + (x_2)^2 - x_1 x_2$.

4.4.2 Let $f : \mathbb{R}^2_{++} \mapsto \mathbb{R}, \ x \mapsto x_1 + \sqrt{x_2} + \ln(x_1)$. Use Theorem 4.4.5 to show that f is strictly concave.

4.4.3 Show that $f : \mathbb{R}^2_{++} \mapsto \mathbb{R}, \ x \mapsto (x_1 x_2)^{1/4} - x_1 - x_2$ has a unique global maximum, and find this maximum.

4.4.4 The production function of a firm is given by

$$f : \mathbb{R}^2_{++} \mapsto \mathbb{R}, \ x \mapsto \sqrt{x_1} + \sqrt{x_2}.$$

Find the factor demand functions, the supply function, and the profit function of this firm.

4.5.1 Let $f : \mathbb{R}^2_{++} \mapsto \mathbb{R}, \ x \mapsto \sqrt{x_1} + \sqrt{x_2}$, and $g : \mathbb{R}^2_{++} \mapsto \mathbb{R}, \ x \mapsto x_1 + x_2 - 4$. Consider the problem of maximizing f subject to the constraint $g(x) = 0$. Find the Lagrange function for this problem.

4.5.2 Find the stationary point of the Lagrange function defined in 4.5.1.

4.5.3 Use Theorem 4.5.3 to show that f has a constrained maximum at the stationary point found in 4.5.2.

4.5.4 The production function of a firm is given by

$$f : \mathbb{R}^2_{++} \mapsto \mathbb{R},\ x \mapsto x_1 x_2.$$

Find the conditional factor demand functions and the cost function of this firm.

4.6.1 Consider the functions

$$f : \mathbb{R}^2 \mapsto \mathbb{R},\ x \mapsto 2x_1x_2 + 4x_1 - (x_1)^2 - 2(x_2)^2$$
$$g^1 : \mathbb{R}^2 \mapsto \mathbb{R},\ x \mapsto x_1 + 4x_2 - 4$$
$$g^2 : \mathbb{R}^2 \mapsto \mathbb{R},\ x \mapsto x_1 + x_2 - 2.$$

Show that f is strictly concave and g^1 and g^2 are convex.

4.6.2 Consider the functions f, g^1, and g^2 defined in Exercise 4.6.1. Show that the matrix $\bar{J}(g^1(x), g^2(x))$ has maximal rank for all $x \in \mathbb{R}^2$.

4.6.3 Consider the functions f, g^1, and g^2 defined in Exercise 4.6.1. Use the Kuhn-Tucker conditions to maximize f subject to the constraints $g^1(x) \leq 0$, $g^2(x) \leq 0$, $x_1 \geq 0$, $x_2 \geq 0$.

4.6.4 Let $n, m \in \mathbb{N}$, let $A \subseteq \mathbb{R}^n$ be convex, and let $f : A \mapsto \mathbb{R}$, $g^j : A \mapsto \mathbb{R}$ for all $j = 1, \ldots, m$. Furthermore, let x^0 be an interior point of A. Suppose f and $g^1, \ldots, g^m$ are partially differentiable with respect to all variables in a neighborhood of x^0, and these partial derivatives are continuous at x^0. Suppose $\bar{J}(g^1(x^0), \ldots, g^m(x^0))$ has maximal rank. Formulate the Kuhn-Tucker conditions for a local constrained minimum of f subject to the constraints $g^j(x) \geq 0$ for all $j = 1, \ldots, m$ and $x_i \geq 0$ for all $i = 1, \ldots, n$ at x^0.

6.5 Chapter 5

5.1.1 Consider the complex numbers $z = 1 - 2i$ and $z' = 3 + i$. Calculate
(a) $|2z + iz'|$;
(b) $\overline{zz'}$.

5.1.2 Prove that the following statements are true for all $z, z' \in \mathbb{C}$.
(a) $\overline{z + z'} = \overline{z} + \overline{z'}$.
(b) $\overline{zz'} = \overline{z}\,\overline{z'}$.

5.1.3 Find the representations in terms of polar coordinates for the following complex numbers.
(a) $z = 2$.
(b) $z' = -2i$.

5.1.4 Find all solutions to the equation $z^4 = 1$.

5.2.1 A difference equation is given by

$$y(t+2) = (y(t))^2 \quad \forall t \in \mathbb{N}_0.$$

(a) Of what order is this difference equation? Is it linear?
(b) Suppose we have the initial conditions $y(0) = 1$ and $y(1) = 0$. Find all solutions of this equation satisfying these initial conditions.

5.2.2 Find all solutions of the difference equation

$$y(t+2) = 3 + y(t) - 2y(t+1) \quad \forall t \in \mathbb{N}_0.$$

5.2.3 Find all solutions of the difference equation

$$y(t+2) = 3 + 3y(t) - 2y(t+1) \quad \forall t \in \mathbb{N}_0.$$

5.2.4 Consider the following macroeconomic model. For each period $t \in \mathbb{N}_0$, national income in t is given by $Y(t)$, private consumption in t is given by $C(t)$, investment in t is $I(t)$ and government expenses are

$G(t)$, where Y, C, I and G are functions with domain $\mathbb{N}_0$ and range $\mathbb{R}$. Suppose these functions are given by

$$\begin{aligned} C(t) &= \frac{1}{4}Y(t) \quad \forall t \in \mathbb{N}_0, \\ I(t+2) &= Y(t+1) - Y(t) \quad \forall t \in \mathbb{N}_0, \\ G(t) &= 3 \quad \forall t \in \mathbb{N}_0. \end{aligned}$$

An equilibrium in the economy is described by the condition

$$Y(t) = C(t) + I(t) + G(t) \quad \forall t \in \mathbb{N}_0.$$

(a) Define a difference equation describing this model.
(b) Find all solutions of this equation.
(c) Suppose the initial conditions are $Y(0) = 6$ and $Y(1) = 16/3$. Find the unique solution satisfying these conditions.

5.3.1 Find the definite integral

$$\int_3^{e+2} \frac{dx}{x-2}.$$

5.3.2 Find the indefinite integral

$$\int xe^x dx.$$

5.3.3 Find the definite integral

$$\int_0^2 \frac{3x^2}{(x^3+1)^2} dx.$$

5.3.4 Determine whether the following improper integral converges. If it converges, find its value.

$$\int_{-\infty}^0 \frac{dx}{(4-x)^2}$$

5.4.1 Let $y : \mathbb{R}_{++} \mapsto \mathbb{R}$ be a function. A differential equation is given by

$$y'(x) = 1 + \frac{2}{x^2} - \frac{y(x)}{x}.$$

(a) Find all solutions of this equation.
(b) Find the unique solution satisfying the initial condition $y(1) = 3$.

5.4.2 Let $y : \mathbb{R} \mapsto \mathbb{R}$ be a function. A differential equation is given by

$$y''(x) = 4 - x^2 - y(x).$$

(a) Find all solutions of this equation.
(b) Find the solution satisfying the conditions $y(0) = 0$ and $y'(0) = 1$.

5.4.3 Let $y : \mathbb{R}_{++} \mapsto \mathbb{R}$ be a function. A differential equation is given by

$$y''(x) = -\frac{2y'(x)}{x}.$$

Find all solutions of this equation.

5.4.4 Let $y : \mathbb{R}_{++} \mapsto \mathbb{R}$ be a function. A differential equation is given by

$$y''(x) = -\frac{y'(x)}{x}.$$

(a) Find all solutions of this equation.
(b) Find the solution such that $y(1) = 2$ and $y(e) = 6$.

6.6 Answers

1.1.1 (i) Not all students registered in this course are female;
(ii) $2 \leq x \leq 3$;
(iii) There exists a real number x such that $x + y \neq 0$ for all $y \in \mathbb{R}$.

1.1.2 (i) True; (ii) True; (iii) True; (iv) False.

1.1.3 $[\neg(a \wedge \neg b)] \wedge [\neg(b \wedge \neg a)]$.

1.1.4 "$\Rightarrow$": By way of contradiction, suppose

$$(xy \text{ is odd }) \ \wedge \ ((x \text{ is even}) \vee (y \text{ is even})).$$

Without loss of generality, suppose x is even. Then there exists $n \in \mathbb{N}$ such that $x = 2n$. Then $xy = 2ny = 2r$ with $r := ny$, which shows that xy is even. This is a contradiction.

"$\Leftarrow$": Suppose x is odd and y is odd. Then there exist $n, m \in \mathbb{N}$ such that $x = 2n-1$ and $y = 2m-1$. Therefore,

$$\begin{aligned} xy &= (2n-1)(2m-1) = 4nm - 2m - 2n + 1 \\ &= 2(2nm - m - n) + 1 = 2(2nm - m - n + 1) - 1 \\ &= 2r - 1 \end{aligned}$$

where $r := 2nm - m - n + 1$. Therefore, xy is odd. $\|$

1.2.1

$$\begin{aligned} \text{(i) } A \cap (A \cap B) &= \{x \mid x \in A \wedge x \in (A \cap B)\} \\ &= \{x \mid x \in A \wedge x \in A \wedge x \in B\} \\ &= \{x \mid x \in A \wedge x \in B\} \\ &= A \cap B. \quad \| \end{aligned}$$

$$\begin{aligned} \text{(ii) } A \cup (A \cap B) &= \{x \mid x \in A \vee x \in (A \cap B)\} \\ &= \{x \mid x \in A \vee (x \in A \wedge x \in B)\} \\ &= \{x \mid x \in A\} \\ &= A. \quad \| \end{aligned}$$

1.2.2 (i) $\overline{A} = \{x \in \mathbb{N} \mid x > 2\}$;
(ii) $\overline{B} = \emptyset$;
(iii) $\overline{C} = \{x \in \mathbb{N} \mid x \text{ is even}\} \cup \{1, 3, 5\}$.

1.2.3

$$\begin{aligned} X \setminus (A \cup B) &= \{x \in X \mid x \notin (A \cup B)\} \\ &= \{x \in X \mid x \notin A \wedge x \notin B\} \\ &= \{x \in X \mid x \notin A\} \cap \{x \in X \mid x \notin B\} \\ &= (X \setminus A) \cap (X \setminus B). \quad \| \end{aligned}$$

1.2.4 (i) $(2, 2) \notin A \times B$, because $2 \notin B$;
(ii) $(2, 3) \in A \times B$, because $2 \in A \ \wedge 3 \ \in B$;
(iii) $(3, 2) \notin A \times B$, because $3 \notin A$, $2 \notin B$;
(iv) $(3, 3) \notin A \times B$, because $3 \notin A$;
(v) $(2, 1) \in A \times B$, because $2 \in A \ \wedge 1 \ \in B$;
(vi) $(1, 2) \notin A \times B$, because $1 \notin A$, $2 \notin B$.

1.3.1 (i) A is not open in $\mathbb{R}$, because $2 \in A$ is not an interior point of A;
(ii) B is not open in $\mathbb{R}$, because $0 \in B$ is not an interior point of B;

(iii) C is open—all points in C are interior poins of C (see Section 1.3);
(iv) D is not open in $\mathbb{R}$, because $2 \in D$ is not an interior point of D.

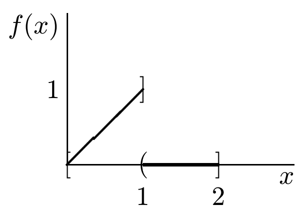
1.3.2 (i) A is closed in $\mathbb{R}$, because $\overline{A}$ is open in $\mathbb{R}$;
(ii) B is not closed in $\mathbb{R}$, because $\overline{B}$ is not open in $\mathbb{R}$;
(iii) C is not closed in $\mathbb{R}$, because $\overline{C}$ is not open in $\mathbb{R}$;
(iv) D is not closed in $\mathbb{R}$, because $\overline{D}$ is not open in $\mathbb{R}$.

1.3.3 (i) Let A, B be convex. Let $x, y \in A \cap B$ and $\lambda \in [0, 1]$. We have to show that $[\lambda x + (1-\lambda)y] \in A \cap B$. Because A and B are convex, we have $[\lambda x + (1-\lambda)y] \in A$ and $[\lambda x + (1-\lambda)y] \in B$. Therefore, $[\lambda x + (1-\lambda)y] \in A \cap B$. $\|$
(ii) $A = \{0\}$, $B = \{1\}$.

1.3.4 (i) $\inf(A) = \min(A) = 1$, no supremum, no maximum;
(ii) no infimum, no minimum, no supremum, no maximum;
(iii) $\inf(C) = 0$, $\sup(C) = 1$, no minimum, no maximum;
(iv) $\sup(D) = \max(D) = 0$, no infimum, no minimum.

1.4.1 $f(\mathbb{N}) = \{x \in \mathbb{N} \mid x \text{ is odd }\}$; $f(\{x \in \mathbb{N} \mid x \text{ is odd}\}) = \{1, 5, 9, 13, \ldots\}$.

1.4.2



1.4.3 (i) f is not surjective, because $-1 \in \mathbb{R}$ but $\not\exists x \in \mathbb{R}$ such that $f(x) = -1$. f is not injective, because $1 \neq -1$, but $f(1) = f(-1) = 1$. f is not bijective, because it is not surjective, not injective.

(ii) f is surjective. For any $y \in \mathbb{R}_+$, let $x = y$ to obtain $f(x) = y$. f is not injective, because $1 \neq -1$, but $f(1) = f(-1) = 1$. f is not bijective, because it is not injective.

(iii) f is not surjective, because $-1 \in \mathbb{R}$ but $\not\exists x \in \mathbb{R}$ such that $f(x) = -1$. f is injective, because $x \neq y$ implies $f(x) = x \neq y = f(y)$ for all $x, y \in \mathbb{R}_+$. f is not bijective, because it is not surjective.

(iv) f is surjective. For any $y \in \mathbb{R}_+$, let $x = y$ to obtain $f(x) = y$. f is injective, because $x \neq y$ implies $f(x) = x \neq y = f(y)$ for all $x, y \in \mathbb{R}_+$. f is bijective, because it is surjective and injective.

1.4.4

$$\pi : \{1,2,3\} \mapsto \{1,2,3\},\ x \mapsto \begin{cases} 1 & \text{if } x = 1 \\ 2 & \text{if } x = 2 \\ 3 & \text{if } x = 3; \end{cases}$$

$$\pi : \{1,2,3\} \mapsto \{1,2,3\},\ x \mapsto \begin{cases} 1 & \text{if } x = 1 \\ 2 & \text{if } x = 3 \\ 3 & \text{if } x = 2; \end{cases}$$

$$\pi : \{1,2,3\} \mapsto \{1,2,3\},\ x \mapsto \begin{cases} 1 & \text{if } x = 2 \\ 2 & \text{if } x = 1 \\ 3 & \text{if } x = 3; \end{cases}$$

$$\pi : \{1,2,3\} \mapsto \{1,2,3\},\ x \mapsto \begin{cases} 1 & \text{if } x = 2 \\ 2 & \text{if } x = 3 \\ 3 & \text{if } x = 1; \end{cases}$$

$$\pi : \{1,2,3\} \mapsto \{1,2,3\},\ x \mapsto \begin{cases} 1 & \text{if } x = 3 \\ 2 & \text{if } x = 1 \\ 3 & \text{if } x = 2; \end{cases}$$

$$\pi : \{1,2,3\} \mapsto \{1,2,3\},\ x \mapsto \begin{cases} 1 & \text{if } x = 3 \\ 2 & \text{if } x = 2 \\ 3 & \text{if } x = 1. \end{cases}$$

1.5.1 (i) The sequence $\{a_n\}$ converges to 0. Proof: For any $\varepsilon \in \mathbb{R}_{++}$, choose $n_0 \in \mathbb{N}$ such that $n_0 > 1/\varepsilon$. For $n \geq n_0$, we obtain $n > 1/\varepsilon$, and therefore,

$$\varepsilon > \frac{1}{n} = |(-1)^n/n| = |(-1)^n/n - 0| = |a_n - 0|,$$

and therefore, $\lim_{n\to\infty} a_n = 0$. ‖

(ii) The sequence $\{b_n\}$ does not converge. Proof: Suppose, by way of contradiction, $\alpha \in \mathbb{R}$ is the limit of $\{b_n\}$. Let $\varepsilon = 1$. Then we have

$$|n(-1)^n - \alpha| > 1$$

for all $n \in \mathbb{N}$ such that n is even and $n > |\alpha| + 1$. Therefore, there cannot exist $n_0 \in \mathbb{N}$ such that $|b_n - \alpha| < 1$ for all $n \geq n_0$, which contradicts the assumption that $\{b_n\}$ converges to α. ‖

1.5.2 Let $c \in \mathbb{R}$. Choose $n_0 \in \mathbb{N}$ such that $n_0 \geq 1 + c$. Then, for all $n \geq n_0$,

$$a_n = n^2 - n = n(n-1) \geq (1+c)c = c + c^2 \geq c,$$

and therefore, $\{a_n\}$ diverges to ∞. ‖

1.5.3 (i) $\{a_n\}$ is monotone nonincreasing, because

$$a_{n+1} = 1 + 1/(n+1) < 1 + 1/n = a_n$$

for all $n \in \mathbb{N}$. $\{a_n\}$ is bounded, because $1 < a_n \leq 2$ for all $n \in \mathbb{N}$. Because $\{a_n\}$ is monotone and bounded, $\{a_n\}$ must be convergent.

(ii) $\{b_n\}$ is neither monotone nondecreasing nor monotone nonincreasing, because

$$b_1 = 0 < 3/2 = b_2 > 2/3 = b_3.$$

$\{b_n\}$ is bounded, because $0 \leq b_n \leq 3/2$ for all $n \in \mathbb{N}$. $\{a_n\}$ converges to 1—see Problem 1.5.1 (i) and apply Theorem 1.5.12.

1.5.4 $a_n + b_n = 2n^2 + 1/n \;\; \forall n \in \mathbb{N}$; $a_n b_n = 2n \;\; \forall n \in \mathbb{N}$.

2.1.1 (i) $\|x\| = \sqrt{\sum_{i=1}^n x_i^2}$. Because $x_i^2 \geq 0$ for all $x_i \in \mathbb{R}$, $\sum_{i=1}^n x_i^2 \geq 0$, and therefore, $\|x\| \geq 0$. ‖

$$\begin{aligned} \text{(ii) } \|x\| = 0 \quad &\Leftrightarrow \quad \sqrt{\sum_{i=1}^n x_i^2} = 0 \\ &\Leftrightarrow \quad x_i^2 = 0 \;\; \forall i = 1, \ldots, n \\ &\Leftrightarrow \quad x_i = 0 \;\; \forall i = 1, \ldots, n \\ &\Leftrightarrow \quad x = \mathbf{0}. \quad \| \end{aligned}$$

$$\begin{aligned} \text{(iii) } \|-x\| \quad &= \quad \sqrt{\sum_{i=1}^n (-x_i)^2} \\ &= \quad \sqrt{\sum_{i=1}^n (-1)^2 x_i^2} \\ &= \quad \sqrt{\sum_{i=1}^n x_i^2} \\ &= \quad \|x\|. \quad \| \end{aligned}$$

2.1.2 (i) $d(x,y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$. Because $(x_i - y_i)^2 \geq 0$ for all $x_i, y_i \in \mathbb{R}$, $\sum_{i=1}^n (x_i - y_i)^2 \geq 0$, and therefore, $d(x,y) \geq 0$. ‖

$$\begin{aligned}
\text{(ii) } d(x,y)=0 \quad &\Leftrightarrow \quad \sqrt{\sum_{i=1}^{n}(x_i-y_i)^2}=0\\
&\Leftrightarrow \quad (x_i-y_i)^2=0 \ \ \forall i=1,\ldots,n\\
&\Leftrightarrow \quad x_i-y_i=0 \ \ \forall i=1,\ldots,n\\
&\Leftrightarrow \quad x=y. \quad \|
\end{aligned}$$

$$\begin{aligned}
\text{(iii) } d(y,x) \quad &= \quad \sqrt{\sum_{i=1}^{n}(y_i-x_i)^2}\\
&= \quad \sqrt{\sum_{i=1}^{n}(-1)^2(x_i-y_i)^2}\\
&= \quad \sqrt{\sum_{i=1}^{n}(x_i-y_i)^2}\\
&= \quad d(x,y). \quad \|
\end{aligned}$$

2.1.3 (i) $x+y=(1+2,3+0,-5+1)=(3,3,-4)$;
(ii) $xy=1\cdot 2+3\cdot 0+(-5)\cdot 1=-3$.

2.1.4 We have

$$\alpha_1 x^1+\alpha_2 x^2+\alpha_3 x^3=\mathbf{0} \Leftrightarrow \alpha_1\begin{pmatrix}1\\0\\2\\-1\end{pmatrix}+\alpha_2\begin{pmatrix}0\\0\\1\\2\end{pmatrix}+\alpha_3\begin{pmatrix}2\\-4\\1\\9\end{pmatrix}=\begin{pmatrix}0\\0\\0\\0\end{pmatrix}.$$

This implies $\alpha_3=0$ (second equation), $\alpha_1=0$ (first equation), and $\alpha_2=0$ (third equation). Therefore, the vectors x^1, x^2, x^3 are linearly independent.

2.2.1 A is a square matrix, A is not symmetric;
B is not a square matrix, and therefore not symmetric;
C is a symmetric square matrix.

2.2.2

$$AB=\begin{pmatrix}1&2&1&-3\\0&1&0&8\\1&2&0&1\end{pmatrix}\begin{pmatrix}-1&0&4\\2&1&-2\\0&1&6\\3&2&-5\end{pmatrix}=\begin{pmatrix}-6&-3&21\\26&17&-42\\6&4&-5\end{pmatrix}.$$

2.2.3 $R(A)=0$, $R(B)=1$, $R(C)=2$, $R(D)=2$.

2.2.4

$$A=\begin{pmatrix}1&0\\0&1\end{pmatrix},\ B=\begin{pmatrix}-1&0\\0&-1\end{pmatrix}.$$

2.3.1

$$\left(\begin{array}{ccccc|c}x_1&x_2&x_3&x_4&x_5&\\ \hline 1&2&0&4&2&0\\0&2&2&-1&3&1\\1&-1&0&1&1&0\\1&1&2&0&4&2\end{array}\right)\sim\left(\begin{array}{ccccc|c}x_1&x_2&x_3&x_4&x_5&\\ \hline 1&2&0&4&2&0\\0&1&1&-1/2&3/2&1/2\\0&-3&0&-3&-1&0\\0&-1&2&-4&2&2\end{array}\right)$$

(multiply Equation 2 by 1/2, add −1 times Equation 1 to Equation 3, add −1 times Equation 1 to Equation 4)

$$\sim \left(\begin{array}{ccccc|c} x_1 & x_2 & x_3 & x_4 & x_5 & \\ \hline 1 & 2 & 0 & 4 & 2 & 0 \\ 0 & 1 & 1 & -1/2 & 3/2 & 1/2 \\ 0 & 0 & 3 & -9/2 & 7/2 & 3/2 \\ 0 & 0 & 3 & -9/2 & 7/2 & 5/2 \end{array}\right) \sim \left(\begin{array}{ccccc|c} x_1 & x_2 & x_3 & x_4 & x_5 & \\ \hline 1 & 2 & 0 & 4 & 2 & 0 \\ 0 & 1 & 1 & -1/2 & 3/2 & 1/2 \\ 0 & 0 & 3 & -9/2 & 7/2 & 3/2 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right)$$

(add 3 times Equation 2 to Equation 3, add Equation 2 to Equation 4; add −1 times Equation 3 to Equation 4). Equation 4 in the last system requires $0 = 1$, which is impossible. Therefore, the system $Ax = b$ has no solution.

2.3.2

$$\left(\begin{array}{cccc|c} x_1 & x_2 & x_3 & x_4 & \\ \hline 2 & 1 & 4 & 0 & 2 \\ 0 & 0 & 1 & 1 & -1 \\ 4 & -1 & 1 & 0 & -2 \\ -2 & 2 & 5 & 2 & 2 \end{array}\right) \sim \left(\begin{array}{cccc|c} x_2 & x_4 & x_1 & x_3 & \\ \hline 1 & 0 & 2 & 4 & 2 \\ 0 & 1 & 0 & 1 & -1 \\ -1 & 0 & 4 & 1 & -2 \\ 2 & 2 & -2 & 5 & 2 \end{array}\right)$$

(interchange columns)

$$\sim \left(\begin{array}{cccc|c} x_2 & x_4 & x_1 & x_3 & \\ \hline 1 & 0 & 2 & 4 & 2 \\ 0 & 1 & 0 & 1 & -1 \\ 0 & 0 & 6 & 5 & 0 \\ 0 & 2 & -6 & -3 & -2 \end{array}\right) \sim \left(\begin{array}{cccc|c} x_2 & x_4 & x_1 & x_3 & \\ \hline 1 & 0 & 2 & 4 & 2 \\ 0 & 1 & 0 & 1 & -1 \\ 0 & 0 & 6 & 5 & 0 \\ 0 & 0 & -6 & -5 & 0 \end{array}\right)$$

(add Equation 1 to Equation 3, add −2 times Equation 1 to Equation 4; add −2 times Equation 2 to Equation 4)

$$\sim \left(\begin{array}{cccc|c} x_2 & x_4 & x_1 & x_3 & \\ \hline 1 & 0 & 2 & 4 & 2 \\ 0 & 1 & 0 & 1 & -1 \\ 0 & 0 & 1 & 5/6 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right) \sim \left(\begin{array}{cccc|c} x_2 & x_4 & x_1 & x_3 & \\ \hline 1 & 0 & 0 & 7/3 & 2 \\ 0 & 1 & 0 & 1 & -1 \\ 0 & 0 & 1 & 5/6 & 0 \end{array}\right)$$

(multiply Equation 3 by 1/6; add 1/6 times Equation 3 to Equation 4; add −2 times Equation 3 to Equation 1). Let $\alpha := x_3$, which implies $x_2 = 2 - 7\alpha/3$, $x_4 = -1 - \alpha$, and $x_1 = -5\alpha/6$. Therefore, the solution is

$$x^* = \begin{pmatrix} 0 \\ 2 \\ 0 \\ -1 \end{pmatrix} + \alpha \begin{pmatrix} -5/6 \\ -7/3 \\ 1 \\ -1 \end{pmatrix}$$

with $\alpha \in \mathbb{R}$.

2.3.3

$$\left(\begin{array}{ccc|c} x_1 & x_2 & x_3 & \\ \hline 2 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 2 & -2 \\ 0 & -2 & 2 & 0 \end{array}\right) \sim \left(\begin{array}{ccc|c} x_1 & x_2 & x_3 & \\ \hline 1 & 0 & 1/2 & 1/2 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & -2 & 2 & 0 \end{array}\right)$$

(multiply Equation 1 by 1/2, multiply Equation 3 by 1/2)

$$\sim \left(\begin{array}{ccc|c} x_1 & x_2 & x_3 & \\ \hline 1 & 0 & 1/2 & 1/2 \\ 0 & 1 & -1/2 & -1/2 \\ 0 & 0 & 1 & -1 \\ 0 & -2 & 2 & 0 \end{array}\right) \sim \left(\begin{array}{ccc|c} x_1 & x_2 & x_3 & \\ \hline 1 & 0 & 1/2 & 1/2 \\ 0 & 1 & -1/2 & -1/2 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & -1 \end{array}\right)$$

(add −1 times Equation 1 to Equation 2; add 2 times Equation 2 to Equation 4)

$$\sim \left(\begin{array}{ccc|c} x_1 & x_2 & x_3 & \\ \hline 1 & 0 & 1/2 & 1/2 \\ 0 & 1 & -1/2 & -1/2 \\ 0 & 0 & 1 & -1 \end{array}\right) \sim \left(\begin{array}{ccc|c} x_1 & x_2 & x_3 & \\ \hline 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -1 \end{array}\right)$$

(add −1 times Equation 3 to Equation 4; add 1/2 times Equation 3 to Equation 2, add −1/2 times Equation 3 to Equation 1). Therefore, the unique solution vector is $x^* = (1, -1, -1)$.

2.3.4 Suppose x^* solves $Ax = b$. Then we have $Ax^* = b$. Multiplying both sides by $\alpha \in \mathbb{R}$, we obtain $\alpha Ax^* = \alpha b$, which is equivalent to $A(\alpha x^*) = \alpha b$. Therefore, αx^* solves $Ax = \alpha b$. ‖

2.4.1 A is singular;
B is nonsingular with inverse $B^{-1} = (1)$;
C is nonsingular with inverse

$$C^{-1} = \begin{pmatrix} 1/4 & -1/4 \\ 1/2 & 1/2 \end{pmatrix};$$

D is nonsingular with inverse

$$D^{-1} = \begin{pmatrix} 1 & 1 & -1 \\ -1 & 0 & 1 \\ 0 & -1 & 1 \end{pmatrix}.$$

2.4.2

$$A = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

2.4.3

$$A' = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}, \quad A^{-1} = \begin{pmatrix} 1/2 & -1/2 \\ 1/2 & 1/2 \end{pmatrix}.$$

Therefore,

$$AA'A^{-1} = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1/2 & -1/2 \\ 1/2 & 1/2 \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}.$$

2.4.4

$$\begin{aligned} AA^{-1} = E \quad &\Rightarrow \quad \alpha AA^{-1} = \alpha E \\ &\Rightarrow \quad (\alpha A)\left(\frac{1}{\alpha}A^{-1}\right) = E \\ &\Rightarrow \quad (\alpha A)^{-1} = \frac{1}{\alpha}A^{-1}, \end{aligned}$$

which shows that αA is nonsingular. ‖

2.5.1

$$\begin{aligned} |A| &= (-1)^0 a_{11}a_{22}a_{33}a_{44} + (-1)^1 a_{11}a_{22}a_{34}a_{43} + (-1)^1 a_{11}a_{23}a_{32}a_{44} \\ &+ (-1)^2 a_{11}a_{23}a_{34}a_{42} + (-1)^2 a_{11}a_{24}a_{32}a_{43} + (-1)^3 a_{11}a_{24}a_{33}a_{42} \\ &+ (-1)^1 a_{12}a_{21}a_{33}a_{44} + (-1)^2 a_{12}a_{21}a_{34}a_{43} + (-1)^2 a_{12}a_{23}a_{31}a_{44} \\ &+ (-1)^3 a_{12}a_{23}a_{34}a_{41} + (-1)^3 a_{12}a_{24}a_{31}a_{43} + (-1)^4 a_{12}a_{24}a_{33}a_{41} \\ &+ (-1)^2 a_{13}a_{21}a_{32}a_{44} + (-1)^3 a_{13}a_{21}a_{34}a_{42} + (-1)^3 a_{13}a_{22}a_{31}a_{44} \\ &+ (-1)^4 a_{13}a_{22}a_{34}a_{41} + (-1)^4 a_{13}a_{24}a_{31}a_{42} + (-1)^5 a_{13}a_{24}a_{32}a_{41} \\ &+ (-1)^3 a_{14}a_{21}a_{32}a_{43} + (-1)^4 a_{14}a_{21}a_{33}a_{42} + (-1)^4 a_{14}a_{22}a_{31}a_{43} \\ &+ (-1)^5 a_{14}a_{22}a_{33}a_{41} + (-1)^5 a_{14}a_{23}a_{31}a_{42} + (-1)^6 a_{14}a_{23}a_{32}a_{41} \\ &= 2 \cdot 1 \cdot 1 \cdot 0 - 2 \cdot 1 \cdot 0 \cdot 0 - 2 \cdot 1 \cdot 2 \cdot 0 + 2 \cdot 1 \cdot 0 \cdot 0 + 2 \cdot 1 \cdot 2 \cdot 0 - 2 \cdot 1 \cdot 1 \cdot 0 \\ &- 1 \cdot 3 \cdot 1 \cdot 0 + 1 \cdot 3 \cdot 0 \cdot 0 + 1 \cdot 1 \cdot 0 \cdot 0 - 1 \cdot 1 \cdot 0 \cdot 1 - 1 \cdot 1 \cdot 0 \cdot 0 + 1 \cdot 1 \cdot 1 \cdot 1 \\ &+ 0 \cdot 3 \cdot 2 \cdot 0 - 0 \cdot 3 \cdot 0 \cdot 0 - 0 \cdot 1 \cdot 0 \cdot 0 + 0 \cdot 1 \cdot 0 \cdot 1 + 0 \cdot 1 \cdot 0 \cdot 0 - 0 \cdot 1 \cdot 2 \cdot 1 \\ &- 3 \cdot 3 \cdot 2 \cdot 0 + 3 \cdot 3 \cdot 1 \cdot 0 + 3 \cdot 1 \cdot 0 \cdot 0 - 3 \cdot 1 \cdot 1 \cdot 1 - 3 \cdot 1 \cdot 0 \cdot 0 + 3 \cdot 1 \cdot 2 \cdot 1 \\ &= 1 - 3 + 6 = 4. \end{aligned}$$

2.5.2

$$\begin{aligned}
\text{(i) } |A| &= 2\cdot\begin{vmatrix} 1 & 1 & 1 \\ 2 & 1 & 0 \\ 0 & 0 & 0 \end{vmatrix} - 3\cdot\begin{vmatrix} 1 & 0 & 3 \\ 2 & 1 & 0 \\ 0 & 0 & 0 \end{vmatrix} + 0\cdot\begin{vmatrix} 1 & 0 & 3 \\ 1 & 1 & 1 \\ 0 & 0 & 0 \end{vmatrix} - 1\cdot\begin{vmatrix} 1 & 0 & 3 \\ 1 & 1 & 1 \\ 2 & 1 & 0 \end{vmatrix} \\
&= (-1)\cdot\left(1\cdot\begin{vmatrix} 1 & 3 \\ 2 & 0 \end{vmatrix} - 1\cdot\begin{vmatrix} 1 & 3 \\ 1 & 1 \end{vmatrix}\right) \\
&= (-1)\cdot(-6+2) = 4.
\end{aligned}$$

$$\text{(ii) } |A| = (-1)\cdot\begin{vmatrix} 1 & 0 & 3 \\ 1 & 1 & 1 \\ 2 & 1 & 0 \end{vmatrix} = 4.$$

2.5.3

$$|A| = (-1)\cdot\begin{vmatrix} 3 & 1 \\ 2 & -1 \end{vmatrix} = (-1)\cdot(-5) = 5 \neq 0.$$

Therefore, A is nonsingular. By Cramer's rule,

$$x_1^* = \frac{\begin{vmatrix} 2 & 1 & 0 \\ 2 & 0 & 1 \\ 1 & -1 & 0 \end{vmatrix}}{5} = \frac{(-1)\cdot\begin{vmatrix} 2 & 1 \\ 1 & -1 \end{vmatrix}}{5} = \frac{3}{5};$$

$$x_2^* = \frac{\begin{vmatrix} 3 & 2 & 0 \\ 1 & 2 & 1 \\ 2 & 1 & 0 \end{vmatrix}}{5} = \frac{(-1)\cdot\begin{vmatrix} 3 & 2 \\ 2 & 1 \end{vmatrix}}{5} = \frac{1}{5};$$

$$x_3^* = \frac{\begin{vmatrix} 3 & 1 & 2 \\ 1 & 0 & 2 \\ 2 & -1 & 1 \end{vmatrix}}{5} = \frac{(-1)\cdot\begin{vmatrix} 1 & 2 \\ 2 & 1 \end{vmatrix} + 1\cdot\begin{vmatrix} 3 & 2 \\ 1 & 2 \end{vmatrix}}{5} = \frac{7}{5}.$$

2.5.4

$$|C_{11}| = \begin{vmatrix} 0 & 1 \\ -1 & 0 \end{vmatrix} = 1,\ |C_{12}| = \begin{vmatrix} 1 & 1 \\ 2 & 0 \end{vmatrix}\cdot(-1) = 2,\ |C_{13}| = \begin{vmatrix} 1 & 0 \\ 2 & -1 \end{vmatrix} = -1,$$

$$|C_{21}| = \begin{vmatrix} 1 & 0 \\ -1 & 0 \end{vmatrix}\cdot(-1) = 0,\ |C_{22}| = \begin{vmatrix} 3 & 0 \\ 2 & 0 \end{vmatrix} = 0,\ |C_{23}| = \begin{vmatrix} 3 & 1 \\ 2 & -1 \end{vmatrix}\cdot(-1) = 5,$$

$$|C_{31}| = \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = 1,\ |C_{32}| = \begin{vmatrix} 3 & 0 \\ 1 & 1 \end{vmatrix}\cdot(-1) = -3,\ |C_{33}| = \begin{vmatrix} 3 & 1 \\ 1 & 0 \end{vmatrix} = -1.$$

Therefore,

$$adj(A) = \begin{pmatrix} 1 & 0 & 1 \\ 2 & 0 & -3 \\ -1 & 5 & -1 \end{pmatrix},\quad A^{-1} = \frac{1}{|A|}adj(A) = \begin{pmatrix} 1/5 & 0 & 1/5 \\ 2/5 & 0 & -3/5 \\ -1/5 & 1 & -1/5 \end{pmatrix}.$$

2.6.1 We obtain

$$|3| = 3 > 0,\quad \begin{vmatrix} 3 & 1 \\ 1 & 2 \end{vmatrix} = 5 > 0,\quad |A| = 3 > 0,$$

and therefore, all leading principal minors of A are positive. This implies that A is positive definite.

2.6.2 Principal minors of order one:

$$|2| = 2 > 0,\ |1| = 1 > 0,\ |0| = 0;$$

principal minors of order two:

$$\begin{vmatrix} 2 & 3 \\ 3 & 1 \end{vmatrix} = -7 < 0,\ \begin{vmatrix} 2 & -1 \\ -1 & 0 \end{vmatrix} = -1 < 0,\ \begin{vmatrix} 1 & 2 \\ 2 & 0 \end{vmatrix} = -4 < 0;$$

principal minor of order three:

$$|A| = -21 < 0.$$

Therefore, A is indefinite, which implies that A is neither positive semidefinite nor negative semidefinite.

2.6.3

$$\begin{aligned} A \text{ is positive definite} &\Leftrightarrow x'Ax > 0 \;\; \forall x \neq \mathbf{0} \\ &\Leftrightarrow (-1)x'Ax < 0 \;\; \forall x \neq \mathbf{0} \\ &\Leftrightarrow x'(-1)Ax < 0 \;\; \forall x \neq \mathbf{0} \\ &\Leftrightarrow (-1)A \text{ is negative definite.} \quad \| \end{aligned}$$

2.6.4

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

3.1.1 Define the sequence $\{x_n\}$ by $x_n = 1/n$ for all $n \in \mathbb{N}$. Then we have $\lim_{n\to\infty} x_n = 0$, but

$$\lim_{n\to\infty} f(x_n) = \lim_{n\to\infty} 1 = 1 \neq 0 = f(0),$$

and therefore, f is not continuous at $x_0 = 0$. $\quad \|$

3.1.2 We have

$$\lim_{x\downarrow 1} f(x) = \lim_{x\downarrow 1}(2x-1) = 1, \quad \lim_{x\uparrow 1} f(x) = \lim_{x\uparrow 1} x = 1,$$

and therefore,

$$\lim_{x\to 1} f(x) = 1 = f(1),$$

which implies that f is continuous at $x_0 = 1$.

3.1.3 We obtain

$$\lim_{x\downarrow 1} f(x) = \lim_{x\downarrow 1}(x-1) = 0, \quad \lim_{x\uparrow 1} f(x) = \lim_{x\uparrow 1} x = 1.$$

Therefore, $\lim_{x\to 1} f(x)$ does not exist, which implies that f is not continuous at $x_0 = 1$.

3.1.4 (i) f is monotone nondecreasing. Suppose $x, y \in [0,3]$ and $x > y$. We have six possible cases:
(a) $x \in [0,1) \wedge y \in [0,1)$. In this case, we have $f(x) = 0 \geq 0 = f(y)$.
(b) $x \in [1,2) \wedge y \in [0,1)$. Now we obtain $f(x) = x - 1 \geq 0 = f(y)$.
(c) $x \in [2,3] \wedge y \in [0,1)$. In this case, $f(x) = 2 \geq 0 = f(y)$.
(d) $x \in [1,2) \wedge y \in [1,2)$. We obtain $f(x) = x - 1 \geq y - 1 = f(y)$.
(e) $x \in [2,3] \wedge y \in [1,2)$. We have $f(x) = 2 \geq y - 1 = f(y)$.
(f) $x \in [2,3] \wedge y \in [2,3]$. In this case, $f(x) = 2 \geq 2 = f(y)$.
(ii) f is not monotone increasing, because $f(0) = f(1/2) = 0$.
(iii) f is not monotone nonincreasing, because $f(2) = 2 > 0 = f(0)$.
(iv) f is not monotone decreasing (see (iii)).

3.2.1 We have

$$\lim_{h\downarrow 0} \frac{f(1+h) - f(1)}{h} = \lim_{h\downarrow 0} \frac{2(1+h) - 1 - 1}{h} = 2$$

and

$$\lim_{h\uparrow 0} \frac{f(1+h) - f(1)}{h} = \lim_{h\uparrow 0} \frac{(1+h) - 1}{h} = 1,$$

which implies that f is not differentiable at $x_0 = 1$.

3.2.2

$$f'(x) = 4\left(\sqrt{x+1} + x^2\right)^3 \left(\frac{1}{2\sqrt{x+1}} + 2x\right) \quad \forall x \in \mathbb{R}_{++}.$$

3.2.3
$$f'(x) = e^{2x^2} 2x \left(2\ln(x^2+1) + \frac{1}{x^2+1} \right) \quad \forall x \in \mathbb{R}.$$

3.2.4
$$f'(x) = 2\sin(x)[6\sin(x)\cos(x) - e^{2(\cos(x)+1)}] \quad \forall x \in \mathbb{R}.$$

3.3.1 (i) The first-order condition for an interior solution is

$$f'(x) = -1 - 8x = 0,$$

which implies that we have a critical point at $x_0 = -1/8$. The second derivative of f at x_0 is $f''(-1/8) = -8 < 0$, and therefore, we have a local maximum at $x_0 = -1/8$. The domain of f is an open set, and therefore, there are no boundary points to be checked. f has no local minimum.

(ii) The first-order condition for an interior maximum or minimum (see part (i)) is not satisfied for any $x \in (0,4)$. We have $f'(0) = -1 < 0$ and $f'(4) = -33 < 0$. Therefore, we have a local maximum at $x_0 = 0$ and a local minimum at $y_0 = 4$. There are no other critical points, and furthermore, the domain of f is closed and bounded, and f is continuous. Therefore, f has a global maximum at x_0 and a global minimum at y_0. The maximal and minimal values of f are $f(x_0) = 2$ and $f(y_0) = -66$.

(iii) We have $f'(x) = 2x$ and $f''(x) = 2 > 0$ for all $x \in (0,4]$. The first-order condition for an interior maximum or minimum is $2x = 0$, which cannot be satisfied for any $x \in (0,4)$. The only possibility for a boundary solution is at $x = 4$. We have $f'(4) = 8 > 0$, and therefore, f has a local maximum at $x_0 = 4$. Because f is increasing, this is a global maximum, and the maximal value of f is $f(x_0) = 16$. f has no local and no global minimum.

3.3.2 We have
$$f'(x) = \frac{1}{2\sqrt{x}} - \frac{1}{2} > 0 \quad \forall x \in (0,1).$$
Therefore, f is increasing, which implies that f is nondecreasing. f is not nonincreasing and not decreasing.

3.3.3 We have $\lim_{x\uparrow 1} f(x) = \lim_{x\uparrow 1} 2\ln(x) = 0$ and $\lim_{x\uparrow 1} g(x) = \lim_{x\uparrow 1}(x-1) = 0$. Therefore, we can apply l'Hôpital's rule. We obtain $f'(x) = 2/x$ and $g'(x) = 1$ for all $x \in (0,1)$. Therefore,

$$\lim_{x\uparrow 1} \frac{f(x)}{g(x)} = \lim_{x\uparrow 1} \frac{f'(x)}{g'(x)} = \lim_{x\uparrow 1} \frac{2}{x} = 2.$$

3.3.4 We obtain $f'(x) = 1/x$ and $f''(x) = -1/x^2$ for all $x \in \mathbb{R}_{++}$. Therefore, $f(1) = 0$, $f'(1) = 1$, and $f''(1) = -1$. The second-order Taylor polynomial of f around $x_0 = 1$ is

$$f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)(x-x_0)^2}{2} = (x-1) - \frac{(x-1)^2}{2}.$$

Setting $x = 2$, we obtain
$$(2-1) - \frac{(2-1)^2}{2} = \frac{1}{2}$$
as a second-order Taylor approximation of $f(2)$.

3.4.1 We have

$$\begin{aligned} f(\lambda x + (1-\lambda)y) &= a(\lambda x + (1-\lambda)y) + b \\ &= \lambda a x + \lambda b + (1-\lambda)ay + (1-\lambda)b \\ &= \lambda(ax+b) + (1-\lambda)(ay+b) \\ &= \lambda f(x) + (1-\lambda)f(y) \end{aligned}$$

for all $x, y \in \mathbb{R}$, for all $\lambda \in (0,1)$. Therefore, f is concave and convex, but not strictly concave and not strictly convex.

3.4.2 Differentiating, we obtain $f'(x) = \alpha x^{\alpha-1}$ and $f''(x) = \alpha(\alpha-1)x^{\alpha-2}$ for all $x \in \mathbb{R}_{++}$. Because $\alpha \in (0,1)$, $f''(x) < 0$ for all $x \in \mathbb{R}_{++}$. Therefore, f is strictly concave. ‖

3.4.3 We have $f'(x) = -1 - 2x$ and $f''(x) = -2 < 0$ for all $x \in [0, 4]$. Therefore, f is strictly concave. Because f is continuous and its domain is closed and bounded, f must have a global maximum. Because f is strictly concave, this maximum is unique.

The first-order condition for an interior maximum is

$$-1 - 2x = 0,$$

which cannot be satisfied for any $x \in (0, 4)$. At the boundary point 0, we have $f'(0) = -1 < 0$, and therefore, f has a local and global maximum at $x^0 = 0$. The maximal value of f is $f(x^0) = 1$.

3.4.4 We have to solve

$$\max_y \{\pi(y)\}$$

where $\pi : \mathbb{R}_+ \mapsto \mathbb{R},\ y \mapsto py - y - 2y^2$. Differentiating, we obtain

$$\pi'(y) = p - 1 - 4y, \quad \pi''(y) = -4 < 0 \quad \forall y \in \mathbb{R}_+.$$

Therefore, π is strictly concave, and the first-order conditions are sufficient for a unique global maximum. The first-order condition for an interior maximum is

$$p - 1 - 4y = 0,$$

which implies $y_0 = (p - 1)/4$. y_0 is positive if and only if $p > 1$. For a boundary solution $y_0 = 0$, we obtain the first-order condition

$$p - 1 \leq 0.$$

Therefore, we obtain the supply function

$$\bar{y} : \mathbb{R}_{++} \mapsto \mathbb{R},\ p \mapsto \begin{cases} (p-1)/4 & \text{if } p > 1 \\ 0 & \text{if } p \leq 1 \end{cases}$$

and the profit function

$$\bar{\pi} : \mathbb{R}_{++} \mapsto \mathbb{R},\ p \mapsto \begin{cases} (p-1)^2/8 & \text{if } p > 1 \\ 0 & \text{if } p \leq 1. \end{cases}$$

4.1.1 (i) $d(x, y) = \max(\{|x_i - y_i| \mid i \in \{1, \ldots, n\}\})$. Because $|x_i - y_i| \geq 0$ for all $x_i, y_i \in \mathbb{R}$, $\max(\{|x_i - y_i| \mid i \in \{1, \ldots, n\}\}) \geq 0$, and therefore, $d(x, y) \geq 0$. $\quad\|$

$$\begin{aligned} \text{(ii) } d(x, y) = 0 \quad &\Leftrightarrow \quad \max(\{|x_i - y_i| \mid i \in \{1, \ldots, n\}\}) = 0 \\ &\Leftrightarrow \quad |x_i - y_i| = 0 \ \ \forall i = 1, \ldots, n \\ &\Leftrightarrow \quad x_i - y_i = 0 \ \ \forall i = 1, \ldots, n \\ &\Leftrightarrow \quad x = y. \quad \| \end{aligned}$$

$$\begin{aligned} \text{(iii) } d(y, x) \quad &= \quad \max(\{|y_i - x_i| \mid i \in \{1, \ldots, n\}\}) \\ &= \quad \max(\{|(-1)(x_i - y_i)| \mid i \in \{1, \ldots, n\}\}) \\ &= \quad \max(\{|x_i - y_i| \mid i \in \{1, \ldots, n\}\}) \\ &= \quad d(x, y). \quad \| \end{aligned}$$

4.1.2 (i) We have to find a $\delta \in \mathbb{R}_{++}$ such that

$$\sqrt{(x_1 - 0)^2 + (x_2 - 0)^2} < 1$$

for all $x \in \mathbb{R}^2$ such that $\max(\{|x_1 - 0|, |x_2 - 0|\}) < \delta$. Let $\delta = 1/2$. Then $\max(\{|x_1|, |x_2|\}) < \delta$ implies $|x_1| < 1/2$ and $|x_2| < 1/2$. Therefore, $(x_1)^2 < 1/4$ and $(x_2)^2 < 1/4$, which implies

$$\sqrt{(x_1)^2 + (x_2)^2} < \sqrt{1/4 + 1/4} = \sqrt{1/2} < 1.$$

(ii) Now we have to find a $\delta \in \mathbb{R}_{++}$ such that

$$\max(\{|x_1 - 0|, |x_2 - 0|\}) < 1$$

for all $x \in \mathbb{R}^2$ such that $\sqrt{(x_1 - 0)^2 + (x_2 - 0)^2} < \delta$. Again, let $\delta = 1/2$. Then $\sqrt{(x_1)^2 + (x_2)^2} < \delta$ implies $(x_1)^2 + (x_2)^2 < 1/4$. Therefore, $(x_1)^2 < 1/4$ and $(x_2)^2 < 1/4$. This implies $|x_1| < 1/2$ and $|x_2| < 1/2$, and therefore, $\max(\{|x_1|, |x_2|\}) < 1/2 < 1$.

4.1.3 (i) A is open;

(ii) B is not open—the point $(1/2, 1/2) \in B$ is not an interior point of B;

(iii) C is open.

4.1.4

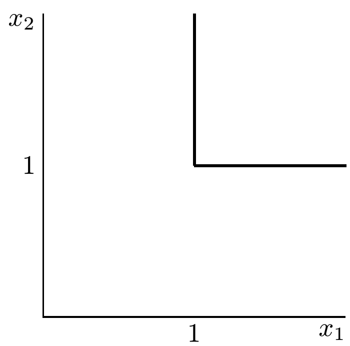
$$\lim_{m\to\infty} a^m = \left(\lim_{m\to\infty}\left[1 + \left(\frac{1}{m}\right)(-1)^m\right], \lim_{m\to\infty}\left[2 - \frac{m}{m+1}\right]\right) = (1, 1).$$

4.2.1 (i) Let $x, y \in A$ and $\lambda \in [0, 1]$. Then $\lambda x_1 + (1 - \lambda)y_1 \in (0, 1)$ and $\lambda x_2 + (1 - \lambda)y_2 \in (1, 2)$, and therefore, $\lambda x + (1 - \lambda)y \in (0, 1) \times (1, 2) = A$. Therefore, A is convex.

(ii) Let $x, y \in B$ and $\lambda \in [0, 1]$. Then $x_1, y_1 \in (0, 1)$ and $x_1 = x_2$ and $y_1 = y_2$. Therefore, $\lambda x_1 + (1 - \lambda)y_1 \in (0, 1)$ and $\lambda x_2 + (1 - \lambda)y_2 = \lambda x_1 + (1 - \lambda)y_1$, and hence, $\lambda x + (1 - \lambda)y \in B$. Therefore, B is convex.

(iii) C is not convex. Let $x = (0, 1)$, $y = (1, 0)$, and $\lambda = 1/2$. Then we have $x, y \in C$, but $\lambda x + (1 - \lambda)y = (1/2, 1/2) \notin C$.

4.2.2 The level set of f for $y = 1$ is $\{x \in \mathbb{R}^2 \mid \min(\{x_1, x_2\}) = 1\}$. Illustration:



4.2.3 Let $x^m = (1/m, 1/m)$ for all $m \in \mathbb{N}$. Then we have

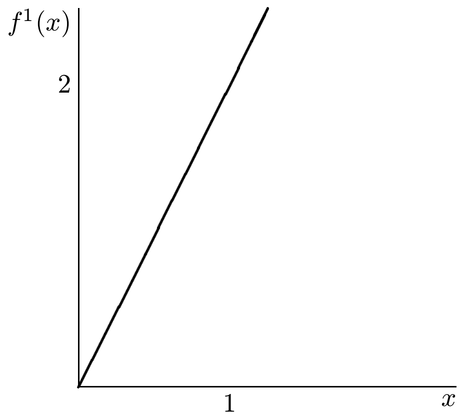
$$\lim_{m\to\infty} x^m = (0, 0) = x^0$$

and $f(x^m) = x_1^m + x_2^m$ for all $m \in \mathbb{N}$, but

$$\lim_{m\to\infty} f(x^m) = 0 \neq 1 = f(x^0).$$

Therefore, f is not continuous at $x^0 = (0, 0)$. $\|$

4.2.4 We have $f^1(x) = f(x, 2) = 2x$ for all $x \in \mathbb{R}$. Illustration:



4.3.1

$$\frac{\partial f(x)}{\partial x_1} = \frac{1}{2}\sqrt{\frac{x_2}{x_1}} + x_3 e^{x_1 x_3}, \;\; \frac{\partial f(x)}{\partial x_2} = \frac{1}{2}\sqrt{\frac{x_1}{x_2}}, \;\; \frac{\partial f(x)}{\partial x_3} = x_1 e^{x_1 x_3} \quad \forall x \in \mathbb{R}^3_{++}.$$

4.3.2

$$df(x^0, h) = \frac{\partial f(x^0)}{\partial x_1} h_1 + \frac{\partial f(x^0)}{\partial x_2} h_2 + \frac{\partial f(x^0)}{\partial x_3} h_3 = (1/2 + e)h_1 + h_2/2 + e h_3.$$

4.3.3 For all $x \in \mathbb{R}^3_{++}$,

$$H(f(x)) = \begin{pmatrix} -\frac{\sqrt{x_2}}{4x_1\sqrt{x_1}} + (x_3)^2 e^{x_1 x_3} & \frac{1}{4\sqrt{x_1 x_2}} & (1 + x_1 x_3) e^{x_1 x_3} \\ \frac{1}{4\sqrt{x_1 x_2}} & -\frac{\sqrt{x_1}}{4x_2\sqrt{x_2}} & 0 \\ (1 + x_1 x_3) e^{x_1 x_3} & 0 & (x_1)^2 e^{x_1 x_3} \end{pmatrix}.$$

4.3.4 Let $y^0 = 0$. We obtain

$$e^{y^0 x_1^0} + y^0 x_1^0 x_2^0 - e^{y^0} = 1 + 0 - 1 = 0$$

and

$$\frac{\partial F(x^0, y^0)}{\partial y} = x_1^0 e^{y^0 x_1^0} + x_1^0 x_2^0 - e^{y^0} = 1 + 1 - 1 = 1 \neq 0.$$

Therefore, there exists an implicit function in a neighborhood of $x^0 = (1, 1)$. Furthermore, we have

$$\frac{\partial F(x^0, y^0)}{\partial x_1} = y^0 e^{y^0 x_1^0} + y^0 x_2^0 = 0$$

and

$$\frac{\partial F(x^0, y^0)}{\partial x_2} = y^0 x_1^0 = 0.$$

Therefore, the partial derivatives of this implicit function at x^0 are

$$\frac{\partial f(x^0)}{\partial x_1} = \frac{\partial f(x^0)}{\partial x_2} = 0.$$

4.4.1 The partial derivatives of f are

$$\frac{\partial f(x)}{\partial x_1} = 2x_1 - x_2, \;\; \frac{\partial f(x)}{\partial x_2} = 2x_2 - x_1$$

for all $x \in \mathbb{R}^2$, and the Hessian matrix at $x \in \mathbb{R}^2$ is

$$H(f(x)) = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}.$$

$H(f(x))$ is positive definite for all $x \in \mathbb{R}^2$, and therefore, f is strictly convex. This implies that the first-order conditions are sufficient for a unique minimum. The only stationary point is $x^0 = (0, 0)$, and therefore, f has a global (and therefore, local) minimum at x^0 with $f(x^0) = 0$. f has no local and no global maximum.

4.4.2 The Hessian matrix of f at $x \in \mathbb{R}^2_{++}$ is

$$H(f(x)) = \begin{pmatrix} -1/(x_1)^2 & 0 \\ 0 & -1/(4x_2\sqrt{x_2}) \end{pmatrix}$$

which is negative definite for all $x \in \mathbb{R}^2_{++}$, and therefore, f is strictly concave. $\|$

4.4.3 The partial derivatives of f are

$$\frac{\partial f(x)}{\partial x_1} = \frac{1}{4}(x_1)^{-3/4}(x_2)^{1/4} - 1, \quad \frac{\partial f(x)}{\partial x_2} = \frac{1}{4}(x_1)^{1/4}(x_2)^{-3/4} - 1$$

for all $x \in \mathbb{R}^2_{++}$, and the Hessian matrix of f at $x \in \mathbb{R}^2_{++}$ is

$$H(f(x)) = \begin{pmatrix} -\frac{3}{16}(x_1)^{-7/4}(x_2)^{1/4} & \frac{1}{16}(x_1x_2)^{-3/4} \\ \frac{1}{16}(x_1x_2)^{-3/4} & -\frac{3}{16}(x_1)^{1/4}(x_2)^{-7/4} \end{pmatrix}.$$

The principal minors of order one are negative for all $x \in \mathbb{R}^2_{++}$, and the determinant of $H(f(x))$ is

$$|H(f(x))| = \frac{1}{32}(x_1x_2)^{-3/2} > 0$$

for all $x \in \mathbb{R}^2_{++}$, which implies that f is strictly concave. Therefore, f has at most one global maximum. Using the first-order conditions, we obtain the unique stationary point $x^0 = (1/16, 1/16)$, and therefore, f has a unique global maximum at x^0. The maximal value of f is $f(x^0) = 1/8$.

4.4.4 We have to solve

$$\max_x \{p(\sqrt{x_1} + \sqrt{x_2}) - w_1x_1 - w_2x_2\}.$$

The first-order partial derivatives of the objective function are

$$\frac{\partial \pi(x)}{\partial x_1} = \frac{p}{2\sqrt{x_1}} - w_1, \quad \frac{\partial \pi(x)}{\partial x_2} = \frac{p}{2\sqrt{x_2}} - w_2$$

for all $x \in \mathbb{R}^2_{++}$, and the Hessian matrix of π at $x \in \mathbb{R}^2_{++}$ is

$$H(\pi(x)) = \begin{pmatrix} -\frac{p}{4x_1\sqrt{x_1}} & 0 \\ 0 & -\frac{p}{4x_2\sqrt{x_2}} \end{pmatrix}.$$

This matrix is negative definite for all $x \in \mathbb{R}^2_{++}$, and therefore, the first-order conditions are sufficient for a unique global maximum. The first-order conditions for an interior solution are

$$\frac{p}{2\sqrt{x_1}} - w_1 = 0 \quad \text{and} \quad \frac{p}{2\sqrt{x_2}} - w_2 = 0.$$

Solving, we obtain $x_1^0 = (p/(2w_1))^2$ and $x_2^0 = (p/(2w_2))^2$. Therefore, the factor demand functions are given by

$$\bar{x}_1 : \mathbb{R}^3_{++} \mapsto \mathbb{R}, \ (p, w) \mapsto (p/(2w_1))^2,$$
$$\bar{x}_2 : \mathbb{R}^3_{++} \mapsto \mathbb{R}, \ (p, w) \mapsto (p/(2w_2))^2.$$

The supply function is

$$\bar{y} : \mathbb{R}^3_{++} \mapsto \mathbb{R}, \ (p, w) \mapsto p/(2w_1) + p/(2w_2),$$

and the profit function is

$$\bar{\pi} : \mathbb{R}^3_{++} \mapsto \mathbb{R}, \ (p, w) \mapsto (p)^2/(4w_1) + (p)^2/(4w_2).$$

4.5.1 $L : \mathbb{R} \times \mathbb{R}^2_{++} \mapsto \mathbb{R}, \ (\lambda, x) \mapsto \sqrt{x_1} + \sqrt{x_2} - \lambda(x_1 + x_2 - 4)$.

4.5.2 The stationary points of the Lagrange function must satisfy

$$\begin{aligned} -x_1 - x_2 + 4 &= 0 \\ 1/(2\sqrt{x_1}) - \lambda &= 0 \\ 1/(2\sqrt{x_2}) - \lambda &= 0. \end{aligned}$$

Solving, we obtain $x^0 = (2, 2)$ and $\lambda^0 = 1/(2\sqrt{2})$.

4.5.3 The bordered Hessian at (λ^0, x^0) is

$$H(L(\lambda^0, x^0)) = \begin{pmatrix} 0 & -1 & -1 \\ -1 & -1/(8\sqrt{2}) & 0 \\ -1 & 0 & -1/(8\sqrt{2}) \end{pmatrix}.$$

Therefore, $|H(L(\lambda^0, x^0))| = 1/(4\sqrt{2}) > 0$, which implies that f has a constrained maximum at x^0.

4.5.4 We have to minimize $w_1x_1 + w_2x_2$ by choice of x subject to the constraint $y = x_1x_2$. The Lagrange function for this problem is

$$L : \mathbb{R} \times \mathbb{R}^2_{++} \mapsto \mathbb{R}, \ (\lambda, x) \mapsto w_1x_1 + w_2x_2 - \lambda(x_1x_2 - y).$$

The necessary first-order conditions for a constrained minimum are

$$\begin{aligned} -x_1x_2 + y &= 0 \\ w_1 - \lambda x_2 &= 0 \\ w_2 - \lambda x_1 &= 0. \end{aligned}$$

Solving, we obtain $x^0 = (\sqrt{yw_2/w_1}, \sqrt{yw_1/w_2})$ and $\lambda^0 = \sqrt{w_1w_2/y}$. The bordered Hessian at (λ^0, x^0) is

$$H(L(\lambda^0, x^0)) = \begin{pmatrix} 0 & -x_2^0 & -x_1^0 \\ -x_2^0 & 0 & -\lambda^0 \\ -x_1^0 & -\lambda^0 & 0 \end{pmatrix}.$$

Therefore, $|H(L(\lambda^0, x^0))| = -2\sqrt{w_1w_2y} < 0$, which implies that the objective function has a constrained minimum at x^0. The conditional factor demand functions are

$$\begin{aligned} \hat{x}_1 : \mathbb{R}^3_{++} \mapsto \mathbb{R}, \ (w, y) \mapsto \sqrt{yw_2/w_1} \\ \hat{x}_2 : \mathbb{R}^3_{++} \mapsto \mathbb{R}, \ (w, y) \mapsto \sqrt{yw_1/w_2} \end{aligned}$$

and the cost function is

$$C : \mathbb{R}^3_{++} \mapsto \mathbb{R}, \ (w, y) \mapsto 2\sqrt{w_1w_2y}.$$

4.6.1 The Hessian matrix of f at $x \in \mathbb{R}^2$ is given by

$$H(f(x)) = \begin{pmatrix} -2 & 2 \\ 2 & -4 \end{pmatrix}.$$

Because this matrix is negative definite for all $x \in \mathbb{R}^2$, f is strictly concave. The Hessian matrix of g^1 and g^2 at $x \in \mathbb{R}^2$ is given by

$$H(g^1(x)) = H(g^2(x)) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

This matrix is positive semidefinite for all $x \in \mathbb{R}^2$ and, therefore, g^1 and g^2 are convex.

4.6.2 The Jacobian matrix of g^1 and g^2 at $x \in \mathbb{R}^2$ is given by

$$J(g^1(x), g^2(x)) = \begin{pmatrix} 1 & 4 \\ 1 & 1 \end{pmatrix}.$$

This matrix has rank 2, which is its maximal possible rank. If a row or a column is removed from this matrix, the resulting matrix has its maximal possible rank 1. If two rows or two columns are removed, the rank condition is trivially satisfied. Therefore, in all possible cases, $\bar{J}(g^1(x), g^2(x))$ has maximal rank for all $x \in \mathbb{R}^2$.

4.6.3 The Lagrange function for this problem is given by

$$L : \mathbb{R}^2 \times \mathbb{R}^2 \mapsto \mathbb{R},\ (\lambda, x) \mapsto 2x_1x_2 + 4x_1 - (x_1)^2 - 2(x_2)^2 - \lambda_1(x_1 + 4x_2 - 4) - \lambda_2(x_1 + x_2 - 2).$$

The Kuhn-Tucker conditions are

$$\begin{aligned}
x_1^0 + 4x_2^0 - 4 &\leq 0 &(6.1)\\
\lambda_1^0(x_1^0 + 4x_2^0 - 4) &= 0 &(6.2)\\
x_1^0 + x_2^0 - 2 &\leq 0 &(6.3)\\
\lambda_2^0(x_1^0 + x_2^0 - 2) &= 0 &(6.4)\\
2x_2^0 + 4 - 2x_1^0 - \lambda_1^0 - \lambda_2^0 &\leq 0 &(6.5)\\
x_1^0(2x_2^0 + 4 - 2x_1^0 - \lambda_1^0 - \lambda_2^0) &= 0 &(6.6)\\
2x_1^0 - 4x_2^0 - 4\lambda_1^0 - \lambda_2^0 &\leq 0 &(6.7)\\
x_2^0(2x_1^0 - 4x_2^0 - 4\lambda_1^0 - \lambda_2^0) &= 0 &(6.8)\\
\lambda_1^0 &\geq 0 &(6.9)\\
\lambda_2^0 &\geq 0 &(6.10)\\
x_1^0 &\geq 0 &(6.11)\\
x_2^0 &\geq 0. &(6.12)
\end{aligned}$$

We have to consider all possible cases regarding whether or not the values of the choice variables and multipliers are equal to zero. Because the objective function f is strictly concave and the constraint functions are convex, we can stop as soon as we find a point (λ^0, x^0) satisfying the Kuhn-Tucker conditions. By Theorems 4.6.3 and 4.6.4, we know that, in this case, f has a unique global constrained maximum at x^0.

(a) $x_1^0 = 0 \wedge x_2^0 = 0$. In this case, (6.2) and (6.4) imply $\lambda_1^0 = \lambda_2^0 = 0$. Substituting, we obtain a contradiction to (6.5).

(b) $x_1^0 = 0 \wedge x_2^0 > 0$. In this case, (6.8) requires

$$-4x_2^0 - 4\lambda_1^0 - \lambda_2^0 = 0$$

which, because of (6.9), (6.10), and (6.12), implies $\lambda_1^0 = \lambda_2^0 = x_2^0 = 0$, contradicting our assumption that $x_2^0 > 0$.

(c) $x_1^0 > 0 \wedge x_2^0 = 0$. (6.3) implies $x_1^0 \leq 2$ and, by (6.2), $\lambda_1^0 = 0$. Therefore, (6.6) implies $4 - 2x_1^0 - \lambda_2^0 = 0$, and we obtain

$$x_1^0 = 2 - \lambda_2^0/2 \tag{6.13}$$

or, equivalently, $\lambda_2^0 = 4 - 2x_1^0$. Substituting into (6.7), we find $x_1^0 \leq 1$ and, by (6.4), $\lambda_2^0 = 0$. Therefore, by (6.13), $x_1^0 = 2$, contradicting $x_1^0 \leq 1$.

Therefore, the only remaining possibility is

(d) $x_1^0 > 0 \wedge x_2^0 > 0$. We now go through the possible cases regarding whether or not the multipliers are equal to zero.

(i) $\lambda_1^0 = 0 \wedge \lambda_2^0 = 0$. By (6.6) and (6.8), we obtain the system of equations

$$\begin{aligned}
2x_2^0 + 4 - 2x_1^0 &= 0\\
2x_1^0 - 4x_2^0 &= 0.
\end{aligned}$$

The unique solution to this system of equations is $x_1^0 = 4$, $x_2^0 = 2$. But this leads to a contradiction of (6.1).

(ii) $\lambda_1^0 = 0 \wedge \lambda_2^0 > 0$. By (6.4), (6.6), and (6.8), we obtain the system of equations

$$\begin{aligned}
x_1^0 + x_2^0 - 2 &= 0\\
2x_2^0 + 4 - 2x_1^0 - \lambda_2^0 &= 0\\
2x_1^0 - 4x_2^0 - \lambda_2^0 &= 0.
\end{aligned}$$

The unique solution is given by $x_1^0 = 8/5$, $x_2^0 = 2/5$, $\lambda_2^0 = 8/5$. Substituting $x^0 = (8/5, 2/5)$ and $\lambda^0 = (0, 8/5)$, we find that all Kuhn-Tucker conditions are satisfied. Therefore, f has a unique global constrained maximum at x^0. The maximal value of the objective function is $f(x^0) = 24/5$.

4.6.4 There exists $\lambda^0 = (\lambda_1^0, \ldots, \lambda_m^0) \in \mathbb{R}^m$ such that

$$\begin{aligned}
g^j(x^0) &\geq 0 \quad \forall j = 1, \ldots, m \\
\lambda_j^0 g^j(x^0) &= 0 \quad \forall j = 1, \ldots, m \\
f_{x_i}(x^0) - \sum_{j=1}^m \lambda_j^0 g_{x_i}^j(x^0) &\geq 0 \quad \forall i = 1, \ldots, n \\
x_i^0 \left(f_{x_i}(x^0) - \sum_{j=1}^m \lambda_j^0 g_{x_i}^j(x^0) \right) &= 0 \quad \forall i = 1, \ldots, n \\
\lambda_j^0 &\geq 0 \quad \forall j = 1, \ldots, m \\
x_i^0 &\geq 0 \quad \forall i = 1, \ldots, n.
\end{aligned}$$

5.1.1 (a) $\sqrt{2}$.
(b) $5 + 5i$.

5.1.2 (a) $\overline{z + z'} = \overline{(a + a') - (b + b')i} = (a - bi) + (a' - b'i) = \overline{z} + \overline{z'}$. ‖
(b) $\overline{zz'} = \overline{(aa' - bb') + (ab + a'b)i} = (aa' - bb') - (ab' + a'b)i = (a - bi)(a' - b'i) = \overline{z}\overline{z'}$. ‖

5.1.3 (a) $|z| = 2$ and $\theta = 0$. Therefore, $z = 2(\cos(0) + i\sin(0))$.
(b) $|z'| = 2$ and $\theta = 3\pi/2$. Therefore, $z = 2(\cos(3\pi/2) + i\sin(3\pi/2))$.

5.1.4 $z_1^* = 1$, $z_2^* = -1$, $z_3^* = i$, $z_4^* = -i$.

5.2.1 (a) Order two, non-linear.
(b) $y(t) = 1$ for all $t \in \{0, 2, 4, \ldots\}$ and $y(t) = 0$ for all $t \in \{1, 3, 5, \ldots\}$.

5.2.2 The associated homogeneous equation is

$$y(t + 2) = y(t) - 2y(t + 1).$$

We obtain the characteristic equation $\lambda^2 + 2\lambda - 1 = 0$ with the two real roots $\lambda_1 = -1 + \sqrt{2}$ and $\lambda_2 = -1 - \sqrt{2}$. Thus, we have the two solutions $z_1(t) = (-1 + \sqrt{2})^t$ and $z_2(t) = (-1 - \sqrt{2})^t$ for all $t \in \mathbb{N}_0$. The two solutions are linearly independent because

$$\begin{vmatrix} z_1(0) & z_2(0) \\ z_1(1) & z_2(1) \end{vmatrix} = \begin{vmatrix} 1 & 1 \\ -1 + \sqrt{2} & -1 - \sqrt{2} \end{vmatrix} = -2\sqrt{2} \neq 0.$$

Thus, the general solution of the homogeneous equation is

$$z(t) = c_1(-1 + \sqrt{2})^2 + c_2(-1 - \sqrt{2})^t \quad \forall t \in \mathbb{N}_0$$

where $c_1, c_2 \in \mathbb{R}$ are constants. To obtain a particular solution of the inhomogeneous equation, we set $\hat{y}(t) = c_0$ for all $t \in \mathbb{N}_0$ where $c_0 \in \mathbb{R}$ is a constant. Substituting into the equation yields $c_0 = 3/2$ and, thus, the general solution is

$$y(t) = c_1(-1 + \sqrt{2})^t + c_2(-1 - \sqrt{2})^t + 3/2 \quad \forall t \in \mathbb{N}_0.$$

5.2.3 The associated homogeneous equation is

$$y(t + 2) = 3y(t) - 2y(t + 1).$$

We obtain the characteristic equation $\lambda^2 + 2\lambda - 3 = 0$ with the two real roots $\lambda_1 = 1$ and $\lambda_2 = -3$. Thus, we have the two solutions $z_1(t) = 1$ and $z_2(t) = (-3)^t$ for all $t \in \mathbb{N}_0$. The two solutions are linearly independent because

$$\begin{vmatrix} z_1(0) & z_2(0) \\ z_1(1) & z_2(1) \end{vmatrix} = \begin{vmatrix} 1 & 1 \\ 1 & -3 \end{vmatrix} = -4 \neq 0.$$

Thus, the general solution of the homogeneous equation is

$$z(t) = c_1 + c_2(-3)^t \quad \forall t \in \mathbb{N}_0$$

where $c_1, c_2 \in \mathbb{R}$ are constants. To obtain a particular solution of the inhomogeneous equation, we set $\hat{y}(t) = c_0$ for all $t \in \mathbb{N}_0$ where $c_0 \in \mathbb{R}$ is a constant. Substituting into the equation yields $c_0 = 3+3c_0-2c_0$ which cannot be satisfied for any $c_0 \in \mathbb{R}$. Therefore, we try $\hat{y}(t) = c_0 t$ for all $t \in \mathbb{N}_0$. Now we obtain $c_0 = 3/4$ and, thus, the general solution is

$$y(t) = c_1 + c_2(-3)^t + 3/4t \quad \forall t \in \mathbb{N}_0.$$

5.2.4 (a) Substituting into the equilibrium condition yields

$$Y(t+2) = -4Y(t)/3 + 4Y(t+1)/3 + 4.$$

(b) The associated homogeneous equation is

$$Y(t+2) = -4Y(t)/3 + 4Y(t+1)/3.$$

The characteristic equation is $\lambda^2 - 4\lambda/3 + 4/3 = 0$ with the complex solutions $\lambda_1 = 2/3 + i2\sqrt{2}/3$ and $\lambda_2 = \overline{\lambda_1} = 2/3 - i2\sqrt{2}/3$. Using polar coordinates, the general solution of the homogeneous equation is

$$z(t) = c_1(2/\sqrt{3})^t \cos(t\theta) + c_2(2/\sqrt{3})^t \sin(t\theta) \quad \forall t \in \mathbb{N}_0$$

where $\theta \in [0, 2\pi)$ is such that $\cos(\theta) = 1/\sqrt{3}$ (and $\sin(\theta) = \sqrt{2/3}$) and $c_1, c_2 \in \mathbb{R}$ are constants. To obtain a particular solution of the inhomogeneous equation, we set $\hat{y}(t) = c_0$ with $c_0 \in \mathbb{R}$ constant. Substituting into the equation yields $c_0 = 4$, and the general solution is given by

$$Y(t) = c_1(2/\sqrt{3})^t \cos(t\theta) + c_2(2/\sqrt{3})^t \sin(t\theta) + 4 \quad \forall t \in \mathbb{N}_0.$$

(c) Substituting $Y(0) = 6$ and $Y(1) = 16/3$ into the solution obtained in part (b) and solving for the parameter values, we obtain $c_1 = 2$ and $c_2 = 0$. Therefore, the unique solution satisfying the initial conditions is

$$Y(t) = 2(2/\sqrt{3})^t \cos(t\theta) + 4 \quad \forall t \in \mathbb{N}_0.$$

5.3.1 Let $f(x) = x-2$ and $g(y) = 1/y$, $y > 0$. Then we obtain $f'(x) = 1$ and $G(y) = \ln(y)+c$. Therefore,

$$\begin{aligned}\int_3^{e+2} \frac{dx}{x-2} &= \int_3^{e+2} g(f(x))f'(x)dx \\ &= G(f(x))|_3^{e+2} = \ln(x-2)|_3^{e+2} \\ &= \ln(e) - \ln(1) = 1 - 0 = 1.\end{aligned}$$

5.3.2 Let $f(x) = e^x$ and $g(x) = x$. Then we have $f'(x) = e^x$ and $g'(x) = 1$. Therefore,

$$\begin{aligned}\int xe^x dx &= \int f'(x)g(x)dx \\ &= f(x)g(x) - \int f(x)g'(x)dx \\ &= xe^x - \int e^x dx = xe^x - e^x + c \\ &= (x-1)e^x + c.\end{aligned}$$

5.3.3 Define $f(x) = x^3 + 1$ and $g(y) = y^{-2}$, $y > 0$. We obtain $f'(x) = 3x^2$ and $G(y) = -y^{-1} + c$. Hence,

$$\begin{aligned}\int_0^2 \frac{3x^2}{(x^3+1)^2} dx &= \int_0^2 g(f(x))f'(x)dx = G(f(x))|_0^2 \\ &= \left.\frac{-1}{x^3+1}\right|_0^2 = -1/9 + 1 = 8/9.\end{aligned}$$

5.3.4 Let $f(x) = 4 - x$ and $g(y) = y^{-2}$, $y > 0$. Then it follows that $f'(x) = -1$ and $G(y) = -y^{-1} + c$. Therefore, for $a < 0$, we obtain

$$\begin{aligned} \int_a^0 \frac{dx}{(4-x)^2} &= -\int_a^0 g(f(x))f'(x)dx \\ &= -G(f(x))|_a^0 = \left.\frac{1}{4-x}\right|_a^0 = \frac{1}{4} - \frac{1}{4-a}, \end{aligned}$$

and

$$\int_{-\infty}^0 \frac{dx}{(4-x)^2} = \lim_{a \downarrow -\infty}\left(\frac{1}{4} - \frac{1}{4-a}\right) = \frac{1}{4}.$$

5.4.1 (a) We obtain

$$\begin{aligned} y(x) &= e^{-\int dx/x}\left(\int (1+2/x^2)e^{\int dx/x}dx + c\right) \\ &= \frac{1}{x}\left(\int (1+2/x^2)xdx + c\right) = \frac{1}{x}\left(\int (1+2/x)dx + c\right) \\ &= \frac{1}{x}\left(\frac{1}{2}x^2 + 2\ln(x) + c\right) = x/2 + 2\ln(x)/x + c/x \end{aligned}$$

for all $x \in \mathbb{R}_{++}$, where $c \in \mathbb{R}$ is a constant.

(b) Substituting $y(1) = 3$ into the solution found in part (a), we obtain $c = 5/2$ and, thus, the solution

$$y(x) = x/2 + 2\ln(x)/x + 5/(2x) \quad \forall x \in \mathbb{R}_{++}.$$

5.4.2 (a) We obtain the associated homogeneous equation

$$y''(x) = -y(x).$$

Setting $z(x) = e^{\lambda x}$ for all $x \in \mathbb{R}$, we obtain the characteristic equation $\lambda^2 + 1 = 0$ with the complex solutions $\lambda_1 = i$ and $\lambda_2 = \overline{\lambda_1} = -i$. Therefore, the general solution of the homogeneous equation is

$$z(x) = c_1\cos(x) + c_2\sin(x) \quad \forall x \in \mathbb{R}$$

where $c_1, c_2 \in \mathbb{R}$ are constants. To find a particular solution of the inhomogenenous equation, we set $\hat{y}(x) = \gamma_0 + \gamma_1 x + \gamma_2 x^2$ for all $x \in \mathbb{R}$ and, comparing coefficients, we obtain $\gamma_0 = 6$, $\gamma_1 = 0$ and $\gamma_2 = -1$. Therefore, the general solution is

$$y(x) = c_1\cos(x) + c_2\sin(x) + 6 - x^2 \quad \forall x \in \mathbb{R}$$

where $c_1, c_2 \in \mathbb{R}$ are constants.

(b) Substituting $y(0) = 0$ and $y'(0) = 1$ into the solution found in part (a), we obtain $c_1 = -6$ and $c_2 = 1$ and, thus, the solution

$$y(x) = -6\cos(x) + \sin(x) + 6 - x^2 \quad \forall x \in \mathbb{R}_{++}.$$

5.4.3 Defining $w(x) = y'(x)$ for all $x \in \mathbb{R}_{++}$, the differential equation becomes

$$w'(x) = -2w(x)/x$$

which is a separable equation because $w'(x) = f(x)g(w(x))$ with $f(x) = -2/x$ and $g(w(x)) = w(x)$ for all $x \in \mathbb{R}_{++}$. Therefore, we must have

$$\int \frac{dw(x)}{g(w(x))} = \int f(x)dx$$

which is equivalent to

$$\ln(w(x)) = -2\ln(x) + \bar{c}$$

where $\bar{c}$ is a constant of integration. Solving, we obtain $w(x) = c/x^2$ for all $x \in \mathbb{R}_{++}$, where $c := e^{\bar{c}} \in \mathbb{R}_{++}$. By definition of w, we obtain

$$y(x) = c \int dx/x^2$$

and, thus,

$$y(x) = -c/x + k \quad \forall x \in \mathbb{R}_{++}$$

where $c \in \mathbb{R}_{++}$ and $k \in \mathbb{R}$ are constants.

5.4.4 (a) Analogously to **5.4.4**, we obtain

$$y(x) = c \int dx/x$$

and, thus,

$$y(x) = c \ln(x) + k \quad \forall x \in \mathbb{R}_{++}$$

where $c \in \mathbb{R}_{++}$ and $k \in \mathbb{R}$ are constants.

(b) Substituting $y(1) = 2$ and $y(e) = 6$ into the solution found in part (a), we obtain $c = 4$ and $k = 2$ and, thus,

$$y(x) = 4 \ln(x) + 2 \quad \forall x \in \mathbb{R}_{++}.$$