

CHAPTER 5

Infinite Sequences and Series

The study of the theory of infinite sequences and series is an integral part of advanced calculus. All limiting processes, such as differentiation and integration, can be investigated on the basis of this theory.

The first example of an infinite series is attributed to Archimedes, who showed that the sum

$$1 + \frac{1}{4} + \cdots + \frac{1}{4^n}$$

was less than $\frac{4}{3}$ for any value of n . However, it was not until the nineteenth century that the theory of infinite series was firmly established by Augustin-Louis Cauchy (1789–1857).

In this chapter we shall study the theory of infinite sequences and series, and investigate their convergence. Unless otherwise stated, the terms of all sequences and series considered in this chapter are real-valued.

5.1. INFINITE SEQUENCES

In Chapter 1 we introduced the general concept of a function. An infinite sequence is a particular function $f: J^+ \rightarrow R$ defined on the set of all positive integers. For a given $n \in J^+$, the value of this function, namely $f(n)$, is called the n th term of the infinite sequence and is denoted by a_n . The sequence itself is denoted by the symbol $\{a_n\}_{n=1}^\infty$. In some cases, the integer with which the infinite sequence begins is different from one. For example, it may be equal to zero or to some other integer. For the sake of simplicity, an infinite sequence will be referred to as just a sequence.

Since a sequence is a function, then, in particular, the sequence $\{a_n\}_{n=1}^\infty$ can have the following properties:

1. It is bounded if there exists a constant $K > 0$ such that $|a_n| \leq K$ for all n .

2. It is monotone increasing if $a_n \leq a_{n+1}$ for all n , and is monotone decreasing if $a_n \geq a_{n+1}$ for all n .
3. It converges to a finite number c if $\lim_{n \rightarrow \infty} a_n = c$, that is, for a given $\epsilon > 0$ there exists an integer N such that

$$|a_n - c| < \epsilon \quad \text{if } n > N.$$

In this case, c is called the limit of the sequence and this fact is denoted by writing $a_n \rightarrow c$ as $n \rightarrow \infty$. If the sequence does not converge to a finite limit, then it is said to be divergent.

4. It is said to oscillate if it does not converge to a finite limit, nor to $+\infty$ or $-\infty$ as $n \rightarrow \infty$.

EXAMPLE 5.1.1. Let $a_n = (n^2 + 2n)/(2n^2 + 3)$. Then $a_n \rightarrow \frac{1}{2}$ as $n \rightarrow \infty$, since

$$\begin{aligned} \lim_{n \rightarrow \infty} a_n &= \lim_{n \rightarrow \infty} \frac{1 + 2/n}{2 + 3/n^2} \\ &= \frac{1}{2}. \end{aligned}$$

EXAMPLE 5.1.2. Consider $a_n = \sqrt{n+1} - \sqrt{n}$. This sequence converges to zero, since

$$\begin{aligned} a_n &= \frac{(\sqrt{n+1} - \sqrt{n})(\sqrt{n+1} + \sqrt{n})}{\sqrt{n+1} + \sqrt{n}} \\ &= \frac{1}{\sqrt{n+1} + \sqrt{n}}. \end{aligned}$$

Hence, $a_n \rightarrow 0$ as $n \rightarrow \infty$.

EXAMPLE 5.1.3. Suppose that $a_n = 2^n/n^3$. Here, the sequence is divergent, since by Example 4.2.3,

$$\lim_{n \rightarrow \infty} \frac{2^n}{n^3} = \infty.$$

EXAMPLE 5.1.4. Let $a_n = (-1)^n$. This sequence oscillates, since it is equal to 1 when n is even and to -1 when n is odd.

Theorem 5.1.1. Every convergent sequence is bounded.

Proof. Suppose that $\{a_n\}_{n=1}^{\infty}$ converges to c . Then, there exists an integer N such that

$$|a_n - c| < 1 \quad \text{if } n > N.$$

For such values of n , we have

$$|a_n| < \max(|c - 1|, |c + 1|).$$

It follows that

$$|a_n| < K$$

for all n , where

$$K = \max(|a_1| + 1, |a_2| + 1, \dots, |a_N| + 1, |c - 1|, |c + 1|). \quad \square$$

The converse of Theorem 5.1.1 is not necessarily true. That is, if a sequence is bounded, then it does not have to be convergent. As a counterexample, consider the sequence given in Example 5.1.4. This sequence is bounded, but is not convergent. To guarantee convergence of a bounded sequence we obviously need an additional condition.

Theorem 5.1.2. Every bounded monotone (increasing or decreasing) sequence converges.

Proof. Suppose that $\{a_n\}_{n=1}^{\infty}$ is a bounded and monotone increasing sequence (the proof is similar if the sequence is monotone decreasing). Since the sequence is bounded, it must be bounded from above and hence has a least upper bound c (see Theorem 1.5.1). Thus $a_n \leq c$ for all n . Furthermore, for any given $\epsilon > 0$ there exists an integer N such that

$$c - \epsilon < a_N \leq c;$$

otherwise $c - \epsilon$ would be an upper bound of $\{a_n\}_{n=1}^{\infty}$. Now, because the sequence is monotone increasing,

$$c - \epsilon < a_N \leq a_{N+1} \leq a_{N+2} \leq \dots \leq c,$$

that is,

$$c - \epsilon < a_n \leq c \quad \text{for } n \geq N.$$

We can write

$$c - \epsilon < a_n < c + \epsilon,$$

or equivalently,

$$|a_n - c| < \epsilon \quad \text{if } n \geq N.$$

This indicates that $\{a_n\}_{n=1}^{\infty}$ converges to c . $\square$

Using Theorem 5.1.2 it is easy to prove the following corollary.

Corollary 5.1.1.

1. If $\{a_n\}_{n=1}^{\infty}$ is bounded from above and is monotone increasing, then $\{a_n\}_{n=1}^{\infty}$ converges to $c = \sup_{n \geq 1} a_n$.
2. If $\{a_n\}_{n=1}^{\infty}$ is bounded from below and is monotone decreasing, then $\{a_n\}_{n=1}^{\infty}$ converges to $d = \inf_{n \geq 1} a_n$.

EXAMPLE 5.1.5. Consider the sequence $\{a_n\}_{n=1}^{\infty}$, where $a_1 = \sqrt{2}$ and $a_{n+1} = \sqrt{2 + \sqrt{a_n}}$ for $n \geq 1$. This sequence is bounded, since $a_n < 2$ for all n , as can be easily shown using mathematical induction: We have $a_1 = \sqrt{2} < 2$. If $a_n < 2$, then $a_{n+1} < \sqrt{2 + \sqrt{2}} < 2$. Furthermore, the sequence is monotone increasing, since $a_n \leq a_{n+1}$ for $n = 1, 2, \dots$, which can also be shown by mathematical induction. Hence, by Theorem 5.1.2 $\{a_n\}_{n=1}^{\infty}$ must converge. To find its limit, we note that

$$\begin{aligned} \lim_{n \rightarrow \infty} a_{n+1} &= \lim_{n \rightarrow \infty} \sqrt{2 + \sqrt{a_n}} \\ &= \sqrt{2 + \sqrt{\lim_{n \rightarrow \infty} a_n}}. \end{aligned}$$

If c denotes the limit of a_n as $n \rightarrow \infty$, then

$$c = \sqrt{2 + \sqrt{c}}.$$

By solving this equation under the condition $c \geq \sqrt{2}$ we find that the only solution is $c = 1.831$.

Definition 5.1.1. Consider the sequence $\{a_n\}_{n=1}^{\infty}$. An infinite collection of its terms, picked out in a manner that preserves the original order of the terms of the sequence, is called a subsequence of $\{a_n\}_{n=1}^{\infty}$. More formally, any sequence of the form $\{b_n\}_{n=1}^{\infty}$, where $b_n = a_{k_n}$ such that $k_1 < k_2 < \dots < k_n < \dots$ is a subsequence of $\{a_n\}_{n=1}^{\infty}$. Note that $k_n \geq n$ for $n \geq 1$. $\square$

Theorem 5.1.3. A sequence $\{a_n\}_{n=1}^{\infty}$ converges to c if and only if every subsequence of $\{a_n\}_{n=1}^{\infty}$ converges to c .

Proof. The proof is left to the reader. $\square$

It should be noted that if a sequence diverges, then it does not necessarily follow that every one of its subsequences must diverge. A sequence may fail to converge, yet several of its subsequences converge. For example, the

sequence whose n th term is $a_n = (-1)^n$ is divergent, as was seen earlier. However, the two subsequences $\{b_n\}_{n=1}^{\infty}$ and $\{c_n\}_{n=1}^{\infty}$, where $b_n = a_{2n} = 1$ and $c_n = a_{2n-1} = -1$ ($n = 1, 2, \dots$), are both convergent.

We have noted earlier that a bounded sequence may not converge. It is possible, however, that one of its subsequences is convergent. This is shown in the next theorem.

Theorem 5.1.4. Every bounded sequence has a convergent subsequence.

Proof. Suppose that $\{a_n\}_{n=1}^{\infty}$ is a bounded sequence. Without loss of generality we can consider that the number of distinct terms of the sequence is infinite. (If this is not the case, then there exists an infinite subsequence of $\{a_n\}_{n=1}^{\infty}$ that consists of terms that are equal. Obviously, such a subsequence converges.) Let G denote the set consisting of all terms of the sequence. Then G is a bounded infinite set. By Theorem 1.6.2, G must have a limit point, say c . Also, by Theorem 1.6.1, every neighborhood of c must contain infinitely many points of G . It follows that we can find integers $k_1 < k_2 < k_3 < \dots$ such that

$$|a_{k_n} - c| < \frac{1}{n} \quad \text{for } n = 1, 2, \dots$$

Thus for a given $\epsilon > 0$ there exists an integer $N > 1/\epsilon$ such that $|a_{k_n} - c| < \epsilon$ if $n > N$. This indicates that the subsequence $\{a_{k_n}\}_{n=1}^{\infty}$ converges to c . $\square$

We conclude from Theorem 5.1.4 that a bounded sequence can have several convergent subsequences. The limit of each of these subsequences is called a subsequential limit. Let E denote the set of all subsequential limits of $\{a_n\}_{n=1}^{\infty}$. This set is bounded, since the sequence is bounded (why?).

Definition 5.1.2. Let $\{a_n\}_{n=1}^{\infty}$ be a bounded sequence, and let E be the set of all its subsequential limits. Then the least upper bound of E is called the upper limit of $\{a_n\}_{n=1}^{\infty}$ and is denoted by $\limsup_{n \rightarrow \infty} a_n$. Similarly, the greatest lower bound of E is called the lower limit of $\{a_n\}_{n=1}^{\infty}$ and is denoted by $\liminf_{n \rightarrow \infty} a_n$. For example, the sequence $\{a_n\}_{n=1}^{\infty}$, where $a_n = (-1)^n[1 + (1/n)]$, has two subsequential limits, namely -1 and 1 . Thus $E = \{-1, 1\}$, and $\limsup_{n \rightarrow \infty} a_n = 1$, $\liminf_{n \rightarrow \infty} a_n = -1$. $\square$

Theorem 5.1.5. The sequence $\{a_n\}_{n=1}^{\infty}$ converges to c if and only if

$$\liminf_{n \rightarrow \infty} a_n = \limsup_{n \rightarrow \infty} a_n = c.$$

Proof. The proof is left to the reader. $\square$

Theorem 5.1.5 implies that when a sequence converges, the set of all its subsequential limits consists of a single element, namely the limit of the sequence.

5.1.1. The Cauchy Criterion

We have seen earlier that the definition of convergence of a sequence $\{a_n\}_{n=1}^{\infty}$ requires finding the limit of a_n as $n \rightarrow \infty$. In some cases, such a limit may be difficult to figure out. For example, consider the sequence whose n th term is

$$a_n = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \cdots + \frac{(-1)^{n-1}}{2n-1}, \quad n = 1, 2, \dots \quad (5.1)$$

It is not easy to calculate the limit of a_n in order to find out if the sequence converges. Fortunately, however, there is another convergence criterion for sequences, known as the *Cauchy criterion* after Augustin-Louis Cauchy (it was known earlier to Bernhard Bolzano, 1781–1848, a Czechoslovakian priest whose mathematical work was undeservedly overlooked by his lay and clerical contemporaries; see Boyer, 1968, page 566).

Theorem 5.1.6 (The Cauchy Criterion). The sequence $\{a_n\}_{n=1}^{\infty}$ converges if and only if it satisfies the following condition, known as the ϵ -condition: For each $\epsilon > 0$ there is an integer N such that

$$|a_m - a_n| < \epsilon \quad \text{for all } m > N, n > N.$$

Proof. Necessity: If the sequence converges, then it must satisfy the ϵ -condition. Let $\epsilon > 0$ be given. Since the sequence $\{a_n\}_{n=1}^{\infty}$ converges, then there exists a number c and an integer N such that

$$|a_n - c| < \frac{\epsilon}{2} \quad \text{if } n > N.$$

Hence, for $m > N, n > N$ we must have

$$\begin{aligned} |a_m - a_n| &= |a_m - c + c - a_n| \\ &\leq |a_m - c| + |a_n - c| < \epsilon. \end{aligned}$$

Sufficiency: If the sequence satisfies the ϵ -condition, then it must converge. If the ϵ -condition is satisfied, then there is an integer N such that for any given $\epsilon > 0$,

$$|a_n - a_{N+1}| < \epsilon$$

for all values of $n \geq N + 1$. Thus for such values of n ,

$$a_{N+1} - \epsilon < a_n < a_{N+1} + \epsilon. \quad (5.2)$$

The sequence $\{a_n\}_{n=1}^{\infty}$ is therefore bounded, since from the double inequality (5.2) we can assert that

$$|a_n| < \max(|a_1| + 1, |a_2| + 1, \dots, |a_N| + 1, |a_{N+1} - \epsilon|, |a_{N+1} + \epsilon|)$$

for all n . By Theorem 5.1.4, $\{a_n\}_{n=1}^{\infty}$ has a convergent subsequence $\{a_{k_n}\}_{n=1}^{\infty}$. Let c be the limit of this subsequence. If we invoke again the ϵ -condition, we can find an integer N' such that

$$|a_m - a_{k_n}| < \epsilon' \quad \text{if } m > N', k_n \geq n \geq N',$$

where $\epsilon' < \epsilon$. By fixing m and letting $k_n \rightarrow \infty$ we get

$$|a_m - c| \leq \epsilon' < \epsilon \quad \text{if } m > N'.$$

This indicates that the sequence $\{a_n\}_{n=1}^{\infty}$ is convergent and has c as its limit. $\square$

Definition 5.1.3. A sequence $\{a_n\}_{n=1}^{\infty}$ that satisfies the ϵ -condition of the Cauchy criterion is said to be a Cauchy sequence. $\square$

EXAMPLE 5.1.6. With the help of the Cauchy criterion it is now possible to show that the sequence $\{a_n\}_{n=1}^{\infty}$ whose n th term is defined by formula (5.1) is a Cauchy sequence and is therefore convergent. To do so, let $m > n$. Then,

$$a_m - a_n = (-1)^n \left[\frac{1}{2n+1} - \frac{1}{2n+3} + \cdots + \frac{(-1)^{p-1}}{2n+2p-1} \right], \quad (5.3)$$

where $p = m - n$. We claim that the quantity inside brackets in formula (5.3) is positive. This can be shown by grouping successive terms in pairs. Thus if p is even, the quantity is equal to

$$\begin{aligned} & \left(\frac{1}{2n+1} - \frac{1}{2n+3} \right) + \left(\frac{1}{2n+5} - \frac{1}{2n+7} \right) + \cdots \\ & + \left(\frac{1}{2n+2p-3} - \frac{1}{2n+2p-1} \right), \end{aligned}$$

which is positive, since the difference inside each parenthesis is positive. If $p = 1$, the quantity is obviously positive, since it is then equal to $1/(2n+1)$. If $p \geq 3$ is an odd integer, the quantity can be written as

$$\begin{aligned} & \left(\frac{1}{2n+1} - \frac{1}{2n+3} \right) + \left(\frac{1}{2n+5} - \frac{1}{2n+7} \right) + \cdots \\ & + \left(\frac{1}{2n+2p-5} - \frac{1}{2n+2p-3} \right) + \frac{1}{2n+2p-1}, \end{aligned}$$

which is also positive. Hence, for any p ,

$$|a_m - a_n| = \frac{1}{2n+1} - \frac{1}{2n+3} + \cdots + \frac{(-1)^{p-1}}{2n+2p-1}. \quad (5.4)$$

We now claim that

$$|a_m - a_n| < \frac{1}{2n+1}.$$

To prove this claim, let us again consider two cases. If p is even, then

$$\begin{aligned} |a_m - a_n| &= \frac{1}{2n+1} - \left(\frac{1}{2n+3} - \frac{1}{2n+5} \right) - \cdots \\ &\quad - \left(\frac{1}{2n+2p-5} - \frac{1}{2n+2p-3} \right) - \frac{1}{2n+2p-1} \\ &< \frac{1}{2n+1}, \end{aligned} \quad (5.5)$$

since all the quantities inside parentheses in (5.5) are positive. If p is odd, then

$$\begin{aligned} |a_m - a_n| &= \frac{1}{2n+1} - \left(\frac{1}{2n+3} - \frac{1}{2n+5} \right) - \cdots \\ &\quad - \left(\frac{1}{2n+2p-3} - \frac{1}{2n+2p-1} \right) < \frac{1}{2n+1}, \end{aligned}$$

which proves our claim. On the basis of this result we can assert that for a given $\epsilon > 0$,

$$|a_m - a_n| < \epsilon \quad \text{if } m > n > N,$$

where N is such that

$$\frac{1}{2N+1} < \epsilon,$$

or equivalently,

$$N > \frac{1}{2\epsilon} - \frac{1}{2}.$$

This shows that $\{a_n\}_{n=1}^{\infty}$ is a Cauchy sequence.

EXAMPLE 5.1.7. Consider the sequence $\{a_n\}_{n=1}^{\infty}$, where

$$a_n = (-1)^n \left(1 + \frac{1}{n}\right).$$

We have seen earlier that $\liminf_{n \rightarrow \infty} a_n = -1$ and $\limsup_{n \rightarrow \infty} a_n = 1$. Thus by Theorem 5.1.5 this sequence is not convergent. We can arrive at the same conclusion using the Cauchy criterion by showing that the ϵ -condition is not satisfied. This occurs whenever we can find an $\epsilon > 0$ such that for however N may be chosen,

$$|a_m - a_n| \geq \epsilon$$

for some $m > N$, $n > N$. In our example, if N is any positive integer, then the inequality

$$|a_m - a_n| \geq 2 \tag{5.6}$$

can be satisfied by choosing $m = \nu$ and $n = \nu + 1$, where ν is an odd integer greater than N .

5.2. INFINITE SERIES

Let $\{a_n\}_{n=1}^{\infty}$ be a given sequence. Consider the symbolic expression

$$\sum_{n=1}^{\infty} a_n = a_1 + a_2 + \cdots + a_n + \cdots. \tag{5.7}$$

By definition, this expression is called an infinite series, or just a series for simplicity, and a_n is referred to as the n th term of the series. The finite sum

$$s_n = \sum_{i=1}^n a_i, \quad n = 1, 2, \dots,$$

is called the n th partial sum of the series.

Definition 5.2.1. Consider the series $\sum_{n=1}^{\infty} a_n$. Let s_n be its n th partial sum ($n = 1, 2, \dots$).

1. The series is said to be convergent if the sequence $\{s_n\}_{n=1}^{\infty}$ converges. In this case, if $\lim_{n \rightarrow \infty} s_n = s$, where s is finite, then we say that the series converges to s , or that s is the sum of the series. Symbolically, this is

expressed by writing

$$s = \sum_{n=1}^{\infty} a_n.$$

2. If s_n does not tend to a finite limit, then the series is said to be divergent. $\square$

Definition 5.2.1 formulates convergence of a series in terms of convergence of the associated sequence of its partial sums. By applying the Cauchy criterion (Theorem 5.1.6) to the latter sequence, we arrive at the following condition of convergence for a series:

Theorem 5.2.1. The series $\sum_{n=1}^{\infty} a_n$, converges if and only if for a given $\epsilon > 0$ there is an integer N such that

$$\left| \sum_{i=m+1}^n a_i \right| < \epsilon \quad \text{for all } n > m > N. \quad (5.8)$$

Inequality (5.8) follows from applying Theorem 5.1.6 to the sequence $\{s_n\}_{n=1}^{\infty}$ of partial sums of the series and noting that

$$|s_n - s_m| = \left| \sum_{i=m+1}^n a_i \right| \quad \text{for } n > m.$$

In particular, if $n = m + 1$, then inequality (5.8) becomes

$$|a_{m+1}| < \epsilon \quad (5.9)$$

for all $m > N$. This implies that $\lim_{m \rightarrow \infty} a_{m+1} = 0$, and hence $\lim_{n \rightarrow \infty} a_n = 0$. We therefore conclude the following result:

RESULT 5.2.1. If $\sum_{n=1}^{\infty} a_n$ is a convergent series, then $\lim_{n \rightarrow \infty} a_n = 0$.

It is important here to note that the convergence of the n th term of a series to zero as $n \rightarrow \infty$ is a necessary condition for the convergence of the series. It is not, however, a sufficient condition, that is, if $\lim_{n \rightarrow \infty} a_n = 0$, then it does not follow that $\sum_{n=1}^{\infty} a_n$ converges. For example, as we shall see later, the series $\sum_{n=1}^{\infty} (1/n)$ is divergent, and its n th term goes to zero as $n \rightarrow \infty$. It is true, however, that if $\lim_{n \rightarrow \infty} a_n \neq 0$, then $\sum_{n=1}^{\infty} a_n$ is divergent. This follows from applying the law of contraposition to the necessary condition of convergence. We conclude the following:

1. If $a_n \rightarrow 0$ as $n \rightarrow \infty$, then no conclusion can be reached regarding convergence or divergence of $\sum_{n=1}^{\infty} a_n$.

2. If $a_n \not\rightarrow 0$ as $n \rightarrow \infty$, then $\sum_{n=1}^{\infty} a_n$ is divergent. For example, the series $\sum_{n=1}^{\infty} [n/(n+1)]$ is divergent, since

$$\lim_{n \rightarrow \infty} \frac{n}{n+1} = 1 \neq 0.$$

EXAMPLE 5.2.1. One of the simplest series is the geometric series, $\sum_{n=1}^{\infty} a^n$. This series is divergent if $|a| \geq 1$, since $\lim_{n \rightarrow \infty} a^n \neq 0$. It is convergent if $|a| < 1$ by the Cauchy criterion: Let $n > m$. Then

$$s_n - s_m = a^{m+1} + a^{m+2} + \cdots + a^n. \quad (5.10)$$

By multiplying the two sides of (5.10) by a , we get

$$a(s_n - s_m) = a^{m+2} + a^{m+3} + \cdots + a^{n+1}. \quad (5.11)$$

If we now subtract (5.11) from (5.10), we obtain

$$s_n - s_m = \frac{a^{m+1} - a^{n+1}}{1 - a}. \quad (5.12)$$

Since $|a| < 1$, we can find an integer N such that for $m > N$, $n > N$,

$$|a|^{m+1} < \frac{\epsilon(1-a)}{2},$$

$$|a|^{n+1} < \frac{\epsilon(1-a)}{2}.$$

Hence, for a given $\epsilon > 0$,

$$|s_n - s_m| < \epsilon \quad \text{if } n > m > N.$$

Formula (5.12) can actually be used to find the sum of the geometric series when $|a| < 1$. Let $m = 1$. By taking the limits of both sides of (5.12) as $n \rightarrow \infty$ we get

$$\begin{aligned} \lim_{n \rightarrow \infty} s_n &= s_1 + \frac{a^2}{1-a}, & \text{since } \lim_{n \rightarrow \infty} a^{n+1} &= 0, \\ &= a + \frac{a^2}{1-a} \\ &= \frac{a}{1-a}. \end{aligned}$$

EXAMPLE 5.2.2. Consider the series $\sum_{n=1}^{\infty} (1/n!)$. This series converges by the Cauchy criterion. To show this, we first note that

$$\begin{aligned} n! &= n(n-1)(n-2) \times \cdots \times 3 \times 2 \times 1 \\ &\geq 2^{n-1} \quad \text{for } n = 1, 2, \dots \end{aligned}$$

Hence, for $n > m$,

$$\begin{aligned} |s_n - s_m| &= \frac{1}{(m+1)!} + \frac{1}{(m+2)!} + \cdots + \frac{1}{n!} \\ &\leq \frac{1}{2^m} + \frac{1}{2^{m+1}} + \cdots + \frac{1}{2^{n-1}} \\ &= 2 \sum_{i=m+1}^n \frac{1}{2^i}. \end{aligned}$$

This is a partial sum of a convergent geometric series with $a = \frac{1}{2} < 1$ [see formula (5.10)]. Consequently, $|s_n - s_m|$ can be made smaller than any given $\epsilon > 0$ by choosing m and n large enough.

Theorem 5.2.2. If $\sum_{n=1}^{\infty} a_n$ and $\sum_{n=1}^{\infty} b_n$ are two convergent series, and if c is a constant, then the following series are also convergent:

1. $\sum_{n=1}^{\infty} (ca_n) = c \sum_{n=1}^{\infty} a_n$.
2. $\sum_{n=1}^{\infty} (a_n + b_n) = \sum_{n=1}^{\infty} a_n + \sum_{n=1}^{\infty} b_n$.

Proof. The proof is left to the reader. □

Definition 5.2.2. The series $\sum_{n=1}^{\infty} a_n$ is absolutely convergent if $\sum_{n=1}^{\infty} |a_n|$ is convergent. □

For example, the series $\sum_{n=1}^{\infty} [(-1)^n/n!]$ is absolutely convergent, since $\sum_{n=1}^{\infty} (1/n!)$ is convergent, as was seen in Example 5.2.2.

Theorem 5.2.3. Every absolutely convergent series is convergent.

Proof. Consider the series $\sum_{n=1}^{\infty} a_n$, and suppose that $\sum_{n=1}^{\infty} |a_n|$ is convergent. We have that

$$\left| \sum_{i=m+1}^n a_i \right| \leq \sum_{i=m+1}^n |a_i|. \quad (5.13)$$

By applying the Cauchy criterion to $\sum_{n=1}^{\infty} |a_n|$ we can find an integer N such that for a given $\epsilon > 0$,

$$\sum_{i=m+1}^n |a_i| < \epsilon \quad \text{if } n > m > N. \quad (5.14)$$

From (5.13) and (5.14) we conclude that $\sum_{n=1}^{\infty} a_n$ satisfies the Cauchy criterion and is therefore convergent by Theorem 5.2.1. $\square$

Note that it is possible that $\sum_{n=1}^{\infty} a_n$ is convergent while $\sum_{n=1}^{\infty} |a_n|$ is divergent. In this case, the series $\sum_{n=1}^{\infty} a_n$ is said to be conditionally convergent. Examples of this kind of series will be seen later.

In the next section we shall discuss convergence of series whose terms are positive.

5.2.1. Tests of Convergence for Series of Positive Terms

Suppose that the terms of the series $\sum_{n=1}^{\infty} a_n$ are such that $a_n > 0$ for $n > K$, where K is a constant. Without loss of generality we shall consider that $K = 1$. Such a series is called a series of positive terms.

Series of positive terms are interesting because the study of their convergence is comparatively simple and can be used in the determination of convergence of more general series whose terms are not necessarily positive. It is easy to see that a series of positive terms diverges if and only if its sum is $+\infty$.

In what follows we shall introduce techniques that simplify the process of determining whether or not a given series of positive terms is convergent. We refer to these techniques as tests of convergence. The advantage of these tests is that they are in general easier to apply than the Cauchy criterion. This is because evaluating or obtaining inequalities involving the expression $\sum_{i=m+1}^n a_i$ in Theorem 5.2.1 can be somewhat difficult. The tests of convergence, however, have the disadvantage that they can sometime fail to determine convergence or divergence, as we shall soon find out. It should be remembered that these tests apply only to series of positive terms.

The Comparison Test

This test is based on the following theorem:

Theorem 5.2.4. Let $\sum_{n=1}^{\infty} a_n$ and $\sum_{n=1}^{\infty} b_n$ be two series of positive terms such that $a_n \leq b_n$ for $n > N_0$, where N_0 is a fixed integer.

- i. If $\sum_{n=1}^{\infty} b_n$ converges, then so does $\sum_{n=1}^{\infty} a_n$.
- ii. If $\sum_{n=1}^{\infty} a_n$ is divergent, then $\sum_{n=1}^{\infty} b_n$ is divergent too.

Proof. We have that

$$\sum_{i=m+1}^n a_i \leq \sum_{i=m+1}^n b_i \quad \text{for } n > m > N_0. \quad (5.15)$$

If $\sum_{n=1}^{\infty} b_n$ is convergent, then for a given $\epsilon > 0$ there exists an integer N_1 such that

$$\sum_{i=m+1}^n b_i < \epsilon \quad \text{for } n > m > N_1. \quad (5.16)$$

From (5.15) and (5.16) it follows that if $n > m > N$, where $N = \max(N_0, N_1)$, then

$$\sum_{i=m+1}^n a_i < \epsilon,$$

which proves (i).

The proof of (ii) follows from applying the law of contraposition to (i). $\square$

To determine convergence or divergence of $\sum_{n=1}^{\infty} a_n$ we thus need to have in our repertoire a collection of series of positive terms whose behavior (with regard to convergence or divergence) is known. These series can then be compared against $\sum_{n=1}^{\infty} a_n$. For this purpose, the following series can be useful:

- a. $\sum_{n=1}^{\infty} 1/n$. This is a divergent series called the harmonic series.
- b. $\sum_{n=1}^{\infty} 1/n^k$. This is divergent if $k < 1$ and is convergent if $k > 1$.

To prove that the harmonic series is divergent, let us consider its n th partial sum, namely,

$$s_n = \sum_{i=1}^n \frac{1}{i}.$$

Let $A > 0$ be an arbitrary positive number. Choose n large enough so that $n > 2^m$, where $m > 2A$. Then for such values of n ,

$$\begin{aligned} s_n &> \left(1 + \frac{1}{2}\right) + \left(\frac{1}{3} + \frac{1}{4}\right) + \left(\frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8}\right) + \cdots \\ &\quad + \left(\frac{1}{2^{m-1}+1} + \cdots + \frac{1}{2^m}\right) \\ &> \frac{1}{2} + \frac{2}{4} + \frac{4}{8} + \cdots + \frac{2^{m-1}}{2^m} = \frac{m}{2} > A. \end{aligned} \quad (5.17)$$

Since $\mathcal{A}$ is arbitrary and s_n is a monotone increasing function of n , inequality (5.17) implies that $s_n \rightarrow \infty$ as $n \rightarrow \infty$. This proves divergence of the harmonic series.

Let us next consider the series in (b). If $k < 1$, then $1/n^k > 1/n$ and $\sum_{n=1}^{\infty} (1/n^k)$ must be divergent by Theorem 5.2.4(ii). Suppose now that $k > 1$. Consider the n th partial sum of the series, namely,

$$s'_n = \sum_{i=1}^n \frac{1}{i^k}.$$

Then, by choosing m large enough so that $2^m > n$ we get

$$\begin{aligned} s'_n &\leq \sum_{i=1}^{2^m-1} \frac{1}{i^k} \\ &= 1 + \left(\frac{1}{2^k} + \frac{1}{3^k} \right) + \left(\frac{1}{4^k} + \frac{1}{5^k} + \frac{1}{6^k} + \frac{1}{7^k} \right) + \cdots \\ &\quad + \left[\frac{1}{(2^{m-1})^k} + \cdots + \frac{1}{(2^m-1)^k} \right] \\ &\leq 1 + \left(\frac{1}{2^k} + \frac{1}{2^k} \right) + \left(\frac{1}{4^k} + \frac{1}{4^k} + \frac{1}{4^k} + \frac{1}{4^k} \right) + \cdots \\ &\quad + \left[\frac{1}{(2^{m-1})^k} + \cdots + \frac{1}{(2^{m-1})^k} \right] \\ &= 1 + \frac{2}{2^k} + \frac{4}{4^k} + \cdots + \frac{2^{m-1}}{(2^{m-1})^k} \\ &= \sum_{i=1}^m a^{i-1}, \end{aligned} \tag{5.18}$$

where $a = 1/2^{k-1}$. But the right-hand side of (5.18) represents the m th partial sum of a convergent geometric series (since $a < 1$). Hence, as $m \rightarrow \infty$, the right-hand side of (5.18) converges to

$$\sum_{i=1}^{\infty} a^{i-1} = \frac{1}{1-a} \quad (\text{see Example 5.2.1}).$$

Thus the sequence $\{s'_n\}_{n=1}^{\infty}$ is bounded. Since it is also monotone increasing, it must be convergent (see Theorem 5.1.2). This proves convergence of the series $\sum_{n=1}^{\infty} (1/n^k)$ for $k > 1$.

Another version of the comparison test in Theorem 5.2.4 that is easier to implement is given by the following theorem:

Theorem 5.2.5. Let $\sum_{n=1}^{\infty} a_n$ and $\sum_{n=1}^{\infty} b_n$ be two series of positive terms. If there exists a positive constant l such that a_n and lb_n are asymptotically equal, $a_n \sim lb_n$ as $n \rightarrow \infty$ (see Section 3.3), that is,

$$\lim_{n \rightarrow \infty} \frac{a_n}{b_n} = l,$$

then the two series are either both convergent or both divergent.

Proof. There exists an integer N such that

$$\left| \frac{a_n}{b_n} - l \right| < \frac{l}{2} \quad \text{if } n > N,$$

or equivalently,

$$\frac{l}{2} < \frac{a_n}{b_n} < \frac{3l}{2} \quad \text{whenever } n > N.$$

If $\sum_{n=1}^{\infty} a_n$ is convergent, then $\sum_{n=1}^{\infty} b_n$ is convergent by a combination of Theorem 5.2.2(1) and 5.2.4(i), since $b_n < (2/l)a_n$. Similarly, if $\sum_{n=1}^{\infty} b_n$ converges, then so does $\sum_{n=1}^{\infty} a_n$, since $a_n < (3l/2)b_n$. If $\sum_{n=1}^{\infty} a_n$ is divergent, then $\sum_{n=1}^{\infty} b_n$ is divergent too by a combination of Theorems 5.2.2(1) and 5.2.4(ii), since $b_n > (2/3l)a_n$. Finally, $\sum_{n=1}^{\infty} a_n$ diverges if the same is true of $\sum_{n=1}^{\infty} b_n$, since $a_n > (l/2)b_n$. $\square$

EXAMPLE 5.2.3. The series $\sum_{n=1}^{\infty} (n+2)/(n^3+2n+1)$ is convergent, since

$$\frac{n+2}{n^3+2n+1} \sim \frac{1}{n^2} \quad \text{as } n \rightarrow \infty,$$

which is the n th term of a convergent series [recall that $\sum_{n=1}^{\infty} (1/n^k)$ is convergent if $k > 1$].

EXAMPLE 5.2.4. $\sum_{n=1}^{\infty} 1/\sqrt{n(n+1)}$ is divergent, because

$$\frac{1}{\sqrt{n(n+1)}} \sim \frac{1}{n} \quad \text{as } n \rightarrow \infty,$$

which is the n th term of the divergent harmonic series.

The Ratio or d'Alembert's Test

This test is usually attributed to the French mathematician Jean Baptiste d'Alembert (1717–1783), but is also known as Cauchy's ratio test after Augustin-Louis Cauchy (1789–1857).

Theorem 5.2.6. Let $\sum_{n=1}^{\infty} a_n$ be a series of positive terms. Then the following hold:

1. The series converges if $\limsup_{n \rightarrow \infty} (a_{n+1}/a_n) < 1$ (see Definition 5.1.2).
2. The series diverges if $\liminf_{n \rightarrow \infty} (a_{n+1}/a_n) > 1$ (see Definition 5.1.2).
3. If $\liminf_{n \rightarrow \infty} (a_{n+1}/a_n) \leq 1 \leq \limsup_{n \rightarrow \infty} (a_{n+1}/a_n)$, no conclusion can be made regarding convergence or divergence of the series (that is, the ratio test fails).

In particular, if $\lim_{n \rightarrow \infty} (a_{n+1}/a_n) = r$ exists, then the following hold:

1. The series converges if $r < 1$.
2. The series diverges if $r > 1$.
3. The test fails if $r = 1$.

Proof. Let $p = \liminf_{n \rightarrow \infty} (a_{n+1}/a_n)$, $q = \limsup_{n \rightarrow \infty} (a_{n+1}/a_n)$.

1. If $q < 1$, then by the definition of the upper limit (Definition 5.1.2), there exists an integer N such that

$$\frac{a_{n+1}}{a_n} < q' \quad \text{for } n \geq N, \quad (5.19)$$

where q' is chosen such that $q < q' < 1$. (If $a_{n+1}/a_n \geq q'$ for infinitely many values of n , then the sequence $\{a_{n+1}/a_n\}_{n=1}^{\infty}$ has a subsequential limit greater than or equal to q' , which exceeds q . This contradicts the definition of q .) From (5.19) we then get

$$\begin{aligned} a_{N+1} &< a_N q' \\ a_{N+2} &< a_{N+1} q' < a_N q'^2, \\ &\vdots \\ a_{N+m} &< a_{N+m-1} q' < a_N q'^m, \end{aligned}$$

where $m \geq 1$. Thus for $n > N$,

$$a_n < a_N q'^{(n-N)} = a_N (q')^{-N} q'^n.$$

Hence, the series converges by comparison with the convergent geometric series $\sum_{n=1}^{\infty} q'^n$, since $q' < 1$.

2. If $p > 1$, then in an analogous manner we can find an integer N such that

$$\frac{a_{n+1}}{a_n} > p' \quad \text{for } n \geq N, \quad (5.20)$$

where p' is chosen such that $p > p' > 1$. But this implies that a_n cannot tend to zero as $n \rightarrow \infty$, and the series is therefore divergent by Result 5.2.1.

3. If $p \leq 1 \leq q$, then we can demonstrate by using an example that the ratio test is inconclusive: Consider the two series $\sum_{n=1}^{\infty} (1/n)$, $\sum_{n=1}^{\infty} (1/n^2)$. For both series, $p = q = 1$ and hence $p \leq 1 \leq q$, since $\lim_{n \rightarrow \infty} (a_{n+1}/a_n) = 1$. But the first series is divergent while the second is convergent, as was seen earlier. $\square$

EXAMPLE 5.2.5. Consider the same series as in Example 5.2.2. This series was shown to be convergent by the Cauchy criterion. Let us now apply the ratio test. In this case,

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{a_{n+1}}{a_n} &= \lim_{n \rightarrow \infty} \frac{1}{(n+1)!} \bigg/ \frac{1}{n!} \\ &= \lim_{n \rightarrow \infty} \frac{1}{n+1} = 0 < 1, \end{aligned}$$

which indicates convergence by Theorem 5.2.6(1).

Nurcombe (1979) stated and proved the following extension of the ratio test:

Theorem 5.2.7. Let $\sum_{n=1}^{\infty} a_n$ be a series of positive terms, and k be a fixed positive integer.

1. If $\lim_{n \rightarrow \infty} (a_{n+k}/a_n) < 1$, then the series converges.
2. If $\lim_{n \rightarrow \infty} (a_{n+k}/a_n) > 1$, then the series diverges.

This test reduces to the ratio test when $k = 1$.

The Root or Cauchy's Test

This is a more powerful test than the ratio test. It is based on the following theorem:

Theorem 5.2.8. Let $\sum_{n=1}^{\infty} a_n$ be a series of positive terms. Let $\limsup_{n \rightarrow \infty} a_n^{1/n} = \rho$. Then we have the following:

1. The series converges if $\rho < 1$.
2. The series diverges if $\rho > 1$.
3. The test is inconclusive if $\rho = 1$.

In particular, if $\lim_{n \rightarrow \infty} a_n^{1/n} = \tau$ exists, then we have the following:

1. The series converges if $\tau < 1$.
2. The series diverges if $\tau > 1$.
3. The test is inconclusive if $\tau = 1$.

Proof.

1. As in Theorem 5.2.6(1), if $\rho < 1$, then there is an integer N such that

$$a_n^{1/n} < \rho' \quad \text{for } n \geq N,$$

where ρ' is chosen such that $\rho < \rho' < 1$. Thus

$$a_n < \rho'^n \quad \text{for } n \geq N.$$

The series is therefore convergent by comparison with the convergent geometric series $\sum_{n=1}^{\infty} \rho'^n$, since $\rho' < 1$.

2. Suppose that $\rho > 1$. Let $\epsilon > 0$ be such that $\epsilon < \rho - 1$. Then

$$a_n^{1/n} > \rho - \epsilon > 1$$

for infinitely many values of n (why?). Thus for such values of n ,

$$a_n > (\rho - \epsilon)^n,$$

which implies that a_n cannot tend to zero as $n \rightarrow \infty$ and the series is therefore divergent by Result 5.2.1.

3. Consider again the two series $\sum_{n=1}^{\infty} (1/n)$, $\sum_{n=1}^{\infty} (1/n^2)$. In both cases $\rho = 1$ (see Exercise 5.18). The test therefore fails, since the first series is divergent and the second is convergent. $\square$

NOTE 5.2.1. We have mentioned earlier that the root test is more powerful than the ratio test. By this we mean that whenever the ratio test shows convergence or divergence, then so does the root test; whenever the root test is inconclusive, the ratio test is inconclusive too. However, there are situations where the ratio test fails, but the root test does not (see Example 5.2.6). This fact is based on the following theorem:

Theorem 5.2.9. If $a_n > 0$, then

$$\liminf_{n \rightarrow \infty} \frac{a_{n+1}}{a_n} \leq \liminf_{n \rightarrow \infty} a_n^{1/n} \leq \limsup_{n \rightarrow \infty} a_n^{1/n} \leq \limsup_{n \rightarrow \infty} \frac{a_{n+1}}{a_n}.$$

Proof. It is sufficient to prove the two inequalities

$$\limsup_{n \rightarrow \infty} a_n^{1/n} \leq \limsup_{n \rightarrow \infty} \frac{a_{n+1}}{a_n}, \quad (5.21)$$

$$\liminf_{n \rightarrow \infty} \frac{a_{n+1}}{a_n} \leq \liminf_{n \rightarrow \infty} a_n^{1/n}. \quad (5.22)$$

Inequality (5.21): Let $q = \limsup_{n \rightarrow \infty} (a_{n+1}/a_n)$. If $q = \infty$, then there is nothing to prove. Let us therefore consider that q is finite. If we choose q' such that $q < q'$, then as in the proof of Theorem 5.2.6(1), we can find an integer N such that

$$a_n < a_N (q')^{-N} q'^n \quad \text{for } n > N.$$

Hence,

$$a_n^{1/n} < \left[a_N (q')^{-N} \right]^{1/n} q'. \quad (5.23)$$

As $n \rightarrow \infty$, the limit of the right-hand side of inequality (5.23) is q' . It follows that

$$\limsup_{n \rightarrow \infty} a_n^{1/n} \leq q'. \quad (5.24)$$

Since (5.24) is true for any $q' > q$, then we must also have

$$\limsup_{n \rightarrow \infty} a_n^{1/n} \leq q.$$

Inequality (5.22): Let $p = \liminf_{n \rightarrow \infty} (a_{n+1}/a_n)$. We can consider p to be finite (if $p = \infty$, then $q = \infty$ and the proof of the theorem will be complete; if $p = -\infty$, then there is nothing to prove). Let p' be chosen such that $p' < p$. As in the proof of Theorem 5.2.6(2), we can find an integer N such that

$$\frac{a_{n+1}}{a_n} > p' \quad \text{for } n \geq N. \quad (5.25)$$

From (5.25) it is easy to show that

$$a_n > a_N (p')^{-N} p'^n \quad \text{for } n \geq N.$$

Hence, for such values of n ,

$$a_n^{1/n} > \left[a_N (p')^{-N} \right]^{1/n} p'.$$

Consequently,

$$\liminf_{n \rightarrow \infty} a_n^{1/n} \geq p'. \quad (5.26)$$

Since (5.26) is true for any $p' < p$, then

$$\liminf_{n \rightarrow \infty} a_n^{1/n} \geq p.$$

From Theorem 5.2.9 we can easily see that whenever $q < 1$, then $\limsup_{n \rightarrow \infty} a_n^{1/n} < 1$; whenever $p > 1$, then $\limsup_{n \rightarrow \infty} a_n^{1/n} > 1$. In both cases, if convergence or divergence of the series is resolved by the ratio test, then it can also be resolved by the root test. If, however, the root test fails (when $\limsup_{n \rightarrow \infty} a_n^{1/n} = 1$), then the ratio test fails too by Theorem 5.2.6(3). On the other hand, it is possible for the ratio test to be inconclusive whereas the root test is not. This occurs when

$$\liminf_{n \rightarrow \infty} \frac{a_{n+1}}{a_n} \leq \liminf_{n \rightarrow \infty} a_n^{1/n} \leq \limsup_{n \rightarrow \infty} a_n^{1/n} < 1 \leq \limsup_{n \rightarrow \infty} \frac{a_{n+1}}{a_n}. \quad \square$$

EXAMPLE 5.2.6. Consider the series $\sum_{n=1}^{\infty} (a^n + b^n)$, where $0 < a < b < 1$. This can be written as $\sum_{n=1}^{\infty} c_n$, where for $n \geq 1$,

$$c_n = \begin{cases} a^{(n+1)/2} & \text{if } n \text{ is odd,} \\ b^{n/2} & \text{if } n \text{ is even.} \end{cases}$$

Now,

$$\frac{c_{n+1}}{c_n} = \begin{cases} (b/a)^{(n+1)/2} & \text{if } n \text{ is odd,} \\ a(a/b)^{n/2} & \text{if } n \text{ is even,} \end{cases}$$

$$c_n^{1/n} = \begin{cases} a^{(n+1)/(2n)} & \text{if } n \text{ is odd,} \\ b^{1/2} & \text{if } n \text{ is even.} \end{cases}$$

As $n \rightarrow \infty$, c_{n+1}/c_n has two limits, namely 0 and ∞ ; $c_n^{1/n}$ has two limits, $a^{1/2}$ and $b^{1/2}$. Thus

$$\begin{aligned}\liminf_{n \rightarrow \infty} \frac{c_{n+1}}{c_n} &= 0, \\ \limsup_{n \rightarrow \infty} \frac{c_{n+1}}{c_n} &= \infty, \\ \limsup_{n \rightarrow \infty} c_n^{1/n} &= b^{1/2} < 1.\end{aligned}$$

Since $0 \leq 1 \leq \infty$, we can clearly see that the ratio test is inconclusive, whereas the root test indicates that the series is convergent.

Maclaurin's (or Cauchy's) Integral Test

This test was introduced by Colin Maclaurin (1698–1746) and then rediscovered by Cauchy. The description and proof of this test will be given in Chapter 6.

Cauchy's Condensation Test

Let us consider the following theorem:

Theorem 5.2.10. Let $\sum_{n=1}^{\infty} a_n$ be a series of positive terms, where a_n is a monotone decreasing function of n ($= 1, 2, \dots$). Then $\sum_{n=1}^{\infty} a_n$ converges or diverges if and only if the same is true of the series $\sum_{n=1}^{\infty} 2^n a_{2^n}$.

Proof. Let s_n and t_m be the n th and m th partial sums, respectively, of $\sum_{n=1}^{\infty} a_n$ and $\sum_{n=1}^{\infty} 2^n a_{2^n}$. If m is such that $n < 2^m$, then

$$\begin{aligned}s_n &\leq a_1 + (a_2 + a_3) + (a_4 + a_5 + a_6 + a_7) \\ &\quad + \cdots + (a_{2^m} + a_{2^m+1} + \cdots + a_{2^{m+2^m-1}}) \\ &\leq a_1 + 2a_2 + 4a_4 + \cdots + 2^m a_{2^m} = t_m.\end{aligned}\tag{5.27}$$

Furthermore, if $n > 2^m$, then

$$\begin{aligned}s_n &\geq a_1 + a_2 + (a_3 + a_4) + \cdots + (a_{2^{m-1}+1} + \cdots + a_{2^m}) \\ &\geq \frac{a_1}{2} + a_2 + 2a_4 + \cdots + 2^{m-1} a_{2^m} = \frac{t_m}{2}.\end{aligned}\tag{5.28}$$

If $\sum_{n=1}^{\infty} 2^n a_{2^n}$ diverges, then $t_m \rightarrow \infty$ as $m \rightarrow \infty$. Hence, from (5.28), $s_n \rightarrow \infty$ as $n \rightarrow \infty$, and the series $\sum_{n=1}^{\infty} a_n$ is also divergent.

Now, if $\sum_{n=1}^{\infty} 2^n a_{2^n}$ converges, then the sequence $\{t_m\}_{m=1}^{\infty}$ is bounded. From (5.27), the sequence $\{s_n\}_{n=1}^{\infty}$ is also bounded. It follows that $\sum_{n=1}^{\infty} a_n$ is a convergent series (see Exercise 5.13). $\square$

EXAMPLE 5.2.7. Consider again the series $\sum_{n=1}^{\infty} (1/n^k)$. We have already seen that this series converges if $k > 1$ and diverges if $k \leq 1$. Let us now apply Cauchy's condensation test. In this case,

$$\sum_{n=1}^{\infty} 2^n a_{2^n} = \sum_{n=1}^{\infty} 2^n \frac{1}{2^{nk}} = \sum_{n=1}^{\infty} \frac{1}{2^{n(k-1)}}$$

is a geometric series $\sum_{n=1}^{\infty} b^n$, where $b = 1/2^{k-1}$. If $k \leq 1$, then $b \geq 1$ and the series diverges. If $k > 1$, then $b < 1$ and the series converges. It is interesting to note that in this example, both the ratio and the root tests fail.

The following tests enable us to handle situations where the ratio test fails. These tests are particular cases on a general test called Kummer's test.

Kummer's Test

This test is named after the German mathematician Ernst Eduard Kummer (1810–1893).

Theorem 5.2.11. Let $\sum_{n=1}^{\infty} a_n$ and $\sum_{n=1}^{\infty} b_n$ be two series of positive terms. Suppose that the series $\sum_{n=1}^{\infty} b_n$ is divergent. Let

$$\lim_{n \rightarrow \infty} \left(\frac{1}{b_n} \frac{a_n}{a_{n+1}} - \frac{1}{b_{n+1}} \right) = \lambda,$$

Then $\sum_{n=1}^{\infty} a_n$ converges if $\lambda > 0$ and diverges if $\lambda < 0$.

Proof. Suppose that $\lambda > 0$. We can find an integer N such that for $n > N$,

$$\frac{1}{b_n} \frac{a_n}{a_{n+1}} - \frac{1}{b_{n+1}} > \frac{\lambda}{2}. \quad (5.29)$$

Inequality (5.29) can also be written as

$$a_{n+1} < \frac{2}{\lambda} \left(\frac{a_n}{b_n} - \frac{a_{n+1}}{b_{n+1}} \right). \quad (5.30)$$

If s_n is the n th partial sum of $\sum_{n=1}^{\infty} a_n$, then from (5.30) and for $n > N$,

$$s_{n+1} < s_{N+1} + \frac{2}{\lambda} \sum_{i=N+2}^{n+1} \left(\frac{a_{i-1}}{b_{i-1}} - \frac{a_i}{b_i} \right),$$

that is,

$$\begin{aligned} s_{n+1} &< s_{N+1} + \frac{2}{\lambda} \left(\frac{a_{N+1}}{b_{N+1}} - \frac{a_{n+1}}{b_{n+1}} \right), \\ s_{n+1} &< s_{N+1} + \frac{2}{\lambda} \frac{a_{N+1}}{b_{N+1}} \quad \text{for } n > N. \end{aligned} \tag{5.31}$$

Inequality (5.31) indicates that the sequence $\{s_n\}_{n=1}^{\infty}$ is bounded. Hence, the series $\sum_{n=1}^{\infty} a_n$ is convergent (see Exercise 5.13).

Now, let us suppose that $\lambda < 0$. We can find an integer N such that

$$\frac{1}{b_n} \frac{a_n}{a_{n+1}} - \frac{1}{b_{n+1}} < 0 \quad \text{for } n > N.$$

Thus for such values of n ,

$$a_{n+1} > \frac{a_n}{b_n} b_{n+1}. \tag{5.32}$$

It is easy to verify that because of (5.32),

$$a_n > \frac{a_{N+1}}{b_{N+1}} b_n \tag{5.33}$$

for $n \geq N+2$. Since $\sum_{n=1}^{\infty} b_n$ is divergent, then from (5.33) and the use of the comparison test we conclude that $\sum_{n=1}^{\infty} a_n$ is divergent too. $\square$

Two particular cases of Kummer's test are Raabe's test and Gauss's test.

Raabe's Test

This test was established in 1832 by J. L. Raabe.

Theorem 5.2.12. Suppose that $\sum_{n=1}^{\infty} a_n$ is a series of positive terms and that

$$\frac{a_n}{a_{n+1}} = 1 + \frac{\tau}{n} + o\left(\frac{1}{n}\right) \quad \text{as } n \rightarrow \infty.$$

Then $\sum_{n=1}^{\infty} a_n$ converges if $\tau > 1$ and diverges if $\tau < 1$.

Proof. We have that

$$\frac{a_n}{a_{n+1}} = 1 + \frac{\tau}{n} + o\left(\frac{1}{n}\right).$$

This means that

$$n\left(\frac{a_n}{a_{n+1}} - 1 - \frac{\tau}{n}\right) \rightarrow 0 \quad (5.34)$$

as $n \rightarrow \infty$. Equivalently, (5.34) can be expressed as

$$\lim_{n \rightarrow \infty} \left(\frac{na_n}{a_{n+1}} - n - 1 \right) = \tau - 1. \quad (5.35)$$

Let $b_n = 1/n$ in (5.35). This is the n th term of a divergent series. If we now apply Kummer's test, we conclude that the series $\sum_{n=1}^{\infty} a_n$ converges if $\tau - 1 > 0$ and diverges if $\tau - 1 < 0$. $\square$

Gauss's Test

This test is named after Carl Friedrich Gauss (1777–1855). It provides a slight improvement over Raabe's test in that it usually enables us to handle the case $\tau = 1$. For such a value of τ , Raabe's test is inconclusive.

Theorem 5.2.13. Let $\sum_{n=1}^{\infty} a_n$ be a series of positive terms. Suppose that

$$\frac{a_n}{a_{n+1}} = 1 + \frac{\theta}{n} + O\left(\frac{1}{n^{\delta+1}}\right), \quad \delta > 0.$$

Then $\sum_{n=1}^{\infty} a_n$ converges if $\theta > 1$ and diverges if $\theta \leq 1$.

Proof. Since

$$O\left(\frac{1}{n^{\delta+1}}\right) = o\left(\frac{1}{n}\right),$$

then by Raabe's test, $\sum_{n=1}^{\infty} a_n$ converges if $\theta > 1$ and diverges if $\theta < 1$. Let us therefore consider $\theta = 1$. We have

$$\frac{a_n}{a_{n+1}} = 1 + \frac{1}{n} + O\left(\frac{1}{n^{\delta+1}}\right).$$

Put $b_n = 1/(n \log n)$, and consider

$$\begin{aligned} \lim_{n \rightarrow \infty} \left(\frac{1}{b_n} \frac{a_n}{a_{n+1}} - \frac{1}{b_{n+1}} \right) \\ = \lim_{n \rightarrow \infty} \left\{ n \log n \left[1 + \frac{1}{n} + O\left(\frac{1}{n^{\delta+1}} \right) \right] - (n+1) \log(n+1) \right\} \\ = \lim_{n \rightarrow \infty} \left[(n+1) \log \frac{n}{n+1} + (n \log n) O\left(\frac{1}{n^{\delta+1}} \right) \right] = -1. \end{aligned}$$

This is true because

$$\lim_{n \rightarrow \infty} \left[(n+1) \log \frac{n}{n+1} \right] = -1 \quad (\text{by l'Hospital's rule})$$

and

$$\lim_{n \rightarrow \infty} (n \log n) O\left(\frac{1}{n^{\delta+1}} \right) = 0 \quad [\text{see Example 4.2.3(2)}].$$

Since $\sum_{n=1}^{\infty} [1/(n \log n)]$ is a divergent series (this can be shown by using Cauchy's condensation test), then by Kummer's test, the series $\sum_{n=1}^{\infty} a_n$ is divergent. $\square$

EXAMPLE 5.2.8. Gauss established his test in order to determine the convergence of the so-called hypergeometric series. He managed to do so in an article published in 1812. This series is of the form $1 + \sum_{n=1}^{\infty} a_n$, where

$$a_n = \frac{\alpha(\alpha+1)(\alpha+2) \cdots (\alpha+n-1) \beta(\beta+1)(\beta+2) \cdots (\beta+n-1)}{n! \gamma(\gamma+1)(\gamma+2) \cdots (\gamma+n-1)},$$

$$n = 1, 2, \dots,$$

where α, β, γ are real numbers, and none of them is zero or a negative integer. We have

$$\begin{aligned} \frac{a_n}{a_{n+1}} &= \frac{(n+1)(n+\gamma)}{(n+\alpha)(n+\beta)} = \frac{n^2 + (\gamma+1)n + \gamma}{n^2 + (\alpha+\beta)n + \alpha\beta} \\ &= 1 + \frac{\gamma+1-\alpha-\beta}{n} + O\left(\frac{1}{n^2} \right). \end{aligned}$$

In this case, $\theta = \gamma+1-\alpha-\beta$ and $\delta=1$. By Gauss's test, this series is convergent if $\theta > 1$, or $\gamma > \alpha + \beta$, and is divergent if $\theta \leq 1$, or $\gamma \leq \alpha + \beta$.

5.2.2. Series of Positive and Negative Terms

Consider the series $\sum_{n=1}^{\infty} a_n$, where a_n may be positive or negative for $n \geq 1$. The convergence of this general series can be determined by the Cauchy criterion (Theorem 5.1.6). However, it is more convenient to consider the series $\sum_{n=1}^{\infty} |a_n|$ of absolute values, to which the tests of convergence in Section 5.2.1 can be applied. We recall from Definition 5.2.2 that if the latter series converges, then the series $\sum_{n=1}^{\infty} a_n$ is absolutely convergent. This is a stronger type of convergence than the one given in Definition 5.2.1, since by Theorem 5.2.3 convergence of $\sum_{n=1}^{\infty} |a_n|$ implies convergence of $\sum_{n=1}^{\infty} a_n$. The converse, however, is not necessarily true, that is, convergence of $\sum_{n=1}^{\infty} a_n$ does not necessarily imply convergence of $\sum_{n=1}^{\infty} |a_n|$. For example, consider the series

$$\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{2n-1} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \cdots. \quad (5.36)$$

This series is convergent by the result of Example 5.1.6. It is not, however, absolutely convergent, since $\sum_{n=1}^{\infty} [1/(2n-1)]$ is divergent by comparison with the harmonic series $\sum_{n=1}^{\infty} (1/n)$, which is divergent. We recall that a series such as (5.36) that converges, but not absolutely, is called a conditionally convergent series.

The series in (5.36) belongs to a special class of series known as alternating series.

Definition 5.2.3. The series $\sum_{n=1}^{\infty} (-1)^{n-1} a_n$, where $a_n > 0$ for $n \geq 1$, is called an alternating series. $\square$

The following theorem, which was established by Gottfried Wilhelm Leibniz (1646–1716), can be used to determine convergence of alternating series:

Theorem 5.2.14. Let $\sum_{n=1}^{\infty} (-1)^{n-1} a_n$ be an alternating series such that the sequence $\{a_n\}_{n=1}^{\infty}$ is monotone decreasing and converges to zero as $n \rightarrow \infty$. Then the series is convergent.

Proof. Let s_n be the n th partial sum of the series, and let m be an integer such that $m < n$. Then

$$\begin{aligned} s_n - s_m &= \sum_{i=m+1}^n (-1)^{i-1} a_i \\ &= (-1)^m [a_{m+1} - a_{m+2} + \cdots + (-1)^{n-m-1} a_n]. \end{aligned} \quad (5.37)$$

Since $\{a_n\}_{n=1}^{\infty}$ is monotone decreasing, it is easy to show that the quantity inside brackets in (5.37) is nonnegative. Hence,

$$|s_n - s_m| = a_{m+1} - a_{m+2} + \cdots + (-1)^{n-m-1} a_n.$$

Now, if $n - m$ is odd, then

$$\begin{aligned} |s_n - s_m| &= a_{m+1} - (a_{m+2} - a_{m+3}) - \cdots - (a_{n-1} - a_n) \\ &\leq a_{m+1}. \end{aligned}$$

If $n - m$ is even, then

$$\begin{aligned} |s_n - s_m| &= a_{m+1} - (a_{m+2} - a_{m+3}) - \cdots - (a_{n-2} - a_{n-1}) - a_n \\ &\leq a_{m+1}. \end{aligned}$$

Thus in both cases

$$|s_n - s_m| \leq a_{m+1}.$$

Since the sequence $\{a_n\}_{n=1}^{\infty}$ converges to zero, then for a given $\epsilon > 0$ there exists an integer N such that for $m \geq N$, $a_{m+1} < \epsilon$. Consequently,

$$|s_n - s_m| < \epsilon \quad \text{if } n > m \geq N.$$

By Theorem 5.2.1, the alternating series is convergent. $\square$

EXAMPLE 5.2.9. The series given by formula (5.36) was shown earlier to be convergent. This result can now be easily verified with the help of Theorem 5.2.14.

EXAMPLE 5.2.10. The series $\sum_{n=1}^{\infty} (-1)^n / n^k$ is absolutely convergent if $k > 1$, is conditionally convergent if $0 < k \leq 1$, and is divergent if $k \leq 0$ (since the n th term does not go to zero).

EXAMPLE 5.2.11. The series $\sum_{n=2}^{\infty} (-1)^n / (\sqrt{n} \log n)$ is conditionally convergent, since it converges by Theorem 5.2.14, but the series of absolute values diverges by Cauchy's condensation test (Theorem 5.2.10).

5.2.3. Rearrangement of Series

One of the main differences between infinite series and finite series is that whereas the latter are amenable to the laws of algebra, the former are not necessarily so. In particular, if the order of terms of an infinite series is altered, its sum (assuming it converges) may, in general, change; or worse, the

altered series may even diverge. Before discussing this rather disturbing phenomenon, let us consider the following definition:

Definition 5.2.4. Let J^+ denote the set of positive integers and $\sum_{n=1}^{\infty} a_n$ be a given series. Then a second series such as $\sum_{n=1}^{\infty} b_n$ is said to be a rearrangement of $\sum_{n=1}^{\infty} a_n$ if there exists a one-to-one and onto function $f: J^+ \rightarrow J^+$ such that $b_n = a_{f(n)}$ for $n \geq 1$.

For example, the series

$$1 + \frac{1}{3} - \frac{1}{2} + \frac{1}{5} + \frac{1}{7} - \frac{1}{4} + \cdots, \quad (5.38)$$

where two positive terms are followed by one negative term, is a rearrangement of the alternating harmonic series

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \cdots. \quad (5.39)$$

The series in (5.39) is conditionally convergent, as is the series in (5.38). However, the two series have different sums (see Exercise 5.21). $\square$

Fortunately, for absolutely convergent series we have the following theorem:

Theorem 5.2.15. If the series $\sum_{n=1}^{\infty} a_n$ is absolutely convergent, then any rearrangement of it remains absolutely convergent and has the same sum.

Proof. Suppose that $\sum_{n=1}^{\infty} a_n$ is absolutely convergent and that $\sum_{n=1}^{\infty} b_n$ is a rearrangement of it. By Theorem 5.2.1, for a given $\epsilon > 0$, there exists an integer N such that for all $n > m > N$,

$$\sum_{i=m+1}^n |a_i| < \frac{\epsilon}{2}.$$

We then have

$$\sum_{k=1}^{\infty} |a_{m+k}| \leq \frac{\epsilon}{2} \quad \text{if } m > N.$$

Now, let us choose an integer M large enough so that

$$\{1, 2, \dots, N+1\} \subset \{f(1), f(2), \dots, f(M)\}.$$

It follows that if $n > M$, then $f(n) \geq N+2$. Consequently, for $n > m > M$,

$$\begin{aligned} \sum_{i=m+1}^n |b_i| &= \sum_{i=m+1}^n |a_{f(i)}| \\ &\leq \sum_{k=1}^{\infty} |a_{N+k+1}| \leq \frac{\epsilon}{2}. \end{aligned}$$

This implies that the series $\sum_{n=1}^{\infty} |b_n|$ satisfies the Cauchy criterion of Theorem 5.2.1. Therefore, $\sum_{n=1}^{\infty} b_n$ is absolutely convergent.

We now show that the two series have the same sum. Let $s = \sum_{n=1}^{\infty} a_n$, and s_n be its n th partial sum. Then, for a given $\epsilon > 0$ there exists an integer N large enough so that

$$|s_{N+1} - s| < \frac{\epsilon}{2}.$$

If t_n is the n th partial sum of $\sum_{n=1}^{\infty} b_n$, then

$$|t_n - s| \leq |t_n - s_{N+1}| + |s_{N+1} - s|.$$

By choosing M large enough as was done earlier, and by taking $n > M$, we get

$$\begin{aligned} |t_n - s_{N+1}| &= \left| \sum_{i=1}^n b_i - \sum_{i=1}^{N+1} a_i \right| \\ &= \left| \sum_{i=1}^n a_{f(i)} - \sum_{i=1}^{N+1} a_i \right| \\ &\leq \sum_{k=1}^{\infty} |a_{N+k+1}| \leq \frac{\epsilon}{2}, \end{aligned}$$

since if $n > M$,

$$\{a_1, a_2, \dots, a_{N+1}\} \subset \{a_{f(1)}, a_{f(2)}, \dots, a_{f(n)}\}.$$

Hence, for $n > M$,

$$|t_n - s| < \epsilon,$$

which shows that the sum of the series $\sum_{n=1}^{\infty} b_n$ is s . $\square$

Unlike absolutely convergent series, those that are conditionally convergent are susceptible to rearrangements of their terms. To demonstrate this, let us consider the following alternating series:

$$\sum_{n=1}^{\infty} a_n = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{\sqrt{n}}.$$

This series is conditionally convergent, since it is convergent by Theorem 5.2.14 while $\sum_{n=1}^{\infty} (1/\sqrt{n})$ is divergent. Let us consider the following rearrangement:

$$\sum_{n=1}^{\infty} b_n = 1 + \frac{1}{\sqrt{3}} - \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{5}} + \frac{1}{\sqrt{7}} - \frac{1}{\sqrt{4}} + \dots \quad (5.40)$$

in which two positive terms are followed by one that is negative. Let s'_{3n} denote the sum of the first $3n$ terms of (5.40). Then

$$\begin{aligned}
 s'_{3n} &= \left(1 + \frac{1}{\sqrt{3}} - \frac{1}{\sqrt{2}}\right) + \left(\frac{1}{\sqrt{5}} + \frac{1}{\sqrt{7}} - \frac{1}{\sqrt{4}}\right) + \cdots \\
 &\quad + \left(\frac{1}{\sqrt{4n-3}} + \frac{1}{\sqrt{4n-1}} - \frac{1}{\sqrt{2n}}\right) \\
 &= \left(1 - \frac{1}{\sqrt{2}}\right) + \left(\frac{1}{\sqrt{3}} - \frac{1}{\sqrt{4}}\right) + \cdots + \left(\frac{1}{\sqrt{2n-1}} - \frac{1}{\sqrt{2n}}\right) \\
 &\quad + \frac{1}{\sqrt{2n+1}} + \frac{1}{\sqrt{2n+3}} + \cdots + \frac{1}{\sqrt{4n-3}} + \frac{1}{\sqrt{4n-1}} \\
 &= s_{2n} + \frac{1}{\sqrt{2n+1}} + \frac{1}{\sqrt{2n+3}} + \cdots + \frac{1}{\sqrt{4n-1}},
 \end{aligned}$$

where s_{2n} is the sum of the first $2n$ terms of the original series. We note that

$$s'_{3n} > s_{2n} + \frac{n}{\sqrt{4n-1}}. \quad (5.41)$$

If s is the sum of the original series, then $\lim_{n \rightarrow \infty} s_{2n} = s$ in (5.41). But, since

$$\lim_{n \rightarrow \infty} \frac{n}{\sqrt{4n-1}} = \infty,$$

the sequence $\{s'_{3n}\}_{n=1}^{\infty}$ is not convergent, which implies that the series in (5.40) is divergent. This clearly shows that a rearrangement of a conditionally convergent series can change its character. This rather unsettling characteristic of conditionally convergent series is depicted in the following theorem due to Georg Riemann (1826–1866):

Theorem 5.2.16. A conditionally convergent series can always be rearranged so as to converge to any given number s , or to diverge to $+\infty$ or to $-\infty$.

Proof. The proof can be found in several books, for example, Apostol (1964, page 368), Fuls (1978, page 489), Knopp (1951, page 318), and Rudin (1964, page 67). $\square$

5.2.4. Multiplication of Series

Suppose that $\sum_{n=1}^{\infty} a_n$ and $\sum_{n=1}^{\infty} b_n$ are two series. We recall from Theorem 5.2.2 that if these series are convergent, then their sum is a convergent series

obtained by adding the two series term by term. The product of these two series, however, requires a more delicate operation. There are several ways to define this product. We shall consider the so-called Cauchy's product.

Definition 5.2.5. Let $\sum_{n=0}^{\infty} a_n$ and $\sum_{n=0}^{\infty} b_n$ be two series in which the summation index starts at zero instead of one. Cauchy's product of these two series is the series $\sum_{n=0}^{\infty} c_n$, where

$$c_n = \sum_{k=0}^n a_k b_{n-k}, \quad n = 0, 1, 2, \dots,$$

that is,

$$\sum_{n=0}^{\infty} c_n = a_0 b_0 + (a_0 b_1 + a_1 b_0) + (a_0 b_2 + a_1 b_1 + a_2 b_0) + \dots$$

Other products could have been defined by simply adopting different arrangements of the terms that make up the series $\sum_{n=0}^{\infty} c_n$. $\square$

The question now is: under what condition will Cauchy's product of two series converge? The answer to this question is given in the next theorem.

Theorem 5.2.17. Let $\sum_{n=0}^{\infty} c_n$ be Cauchy's product of $\sum_{n=0}^{\infty} a_n$ and $\sum_{n=0}^{\infty} b_n$. Suppose that these two series are convergent and have sums equal to s and t , respectively.

1. If at least one of $\sum_{n=0}^{\infty} a_n$ and $\sum_{n=0}^{\infty} b_n$ converges absolutely, then $\sum_{n=0}^{\infty} c_n$ converges and its sum is equal to st (this result is known as Mertens's theorem).
2. If both series are absolutely convergent, then $\sum_{n=0}^{\infty} c_n$ converges absolutely to the product st (this result is due to Cauchy).

Proof.

1. Suppose that $\sum_{n=0}^{\infty} a_n$ is the series that converges absolutely. Let s_n , t_n , and u_n denote the partial sums $\sum_{i=0}^n a_i$, $\sum_{i=0}^n b_i$, and $\sum_{i=0}^n c_i$, respectively. We need to show that $u_n \rightarrow st$ as $n \rightarrow \infty$. We have that

$$\begin{aligned} u_n &= a_0 b_0 + (a_0 b_1 + a_1 b_0) + \dots + (a_0 b_n + a_1 b_{n-1} + \dots + a_n b_0) \\ &= a_0 t_n + a_1 t_{n-1} + \dots + a_n t_0. \end{aligned} \tag{5.42}$$

Let β_n denote the remainder of the series $\sum_{n=0}^{\infty} b_n$ with respect to t_n , that is, $\beta_n = t - t_n$ ($n = 0, 1, 2, \dots$). By making the proper substitution in

(5.42) we get

$$\begin{aligned} u_n &= a_0(t - \beta_n) + a_1(t - \beta_{n-1}) + \cdots + a_n(t - \beta_0) \\ &= ts_n - (a_0 \beta_n + a_1 \beta_{n-1} + \cdots + a_n \beta_0). \end{aligned} \quad (5.43)$$

Since $s_n \rightarrow s$ as $n \rightarrow \infty$, the proof of (1) will be complete if we can show that the sum inside parentheses in (5.43) goes to zero as $n \rightarrow \infty$. We now proceed to show that this is the case.

Let $\epsilon > 0$ be given. Since the sequence $\{\beta_n\}_{n=0}^\infty$ converges to zero, there exists an integer N such that

$$|\beta_n| < \epsilon \quad \text{if } n > N.$$

Hence,

$$\begin{aligned} &|a_0 \beta_n + a_1 \beta_{n-1} + \cdots + a_n \beta_0| \\ &\leq |a_n \beta_0 + a_{n-1} \beta_1 + \cdots + a_{n-N} \beta_N| \\ &\quad + |a_{n-N-1} \beta_{N+1} + a_{n-N-2} \beta_{N+2} + \cdots + a_0 \beta_n| \\ &< |a_n \beta_0 + a_{n-1} \beta_1 + \cdots + a_{n-N} \beta_N| + \epsilon \sum_{i=0}^{n-N-1} |a_i| \\ &< B \sum_{i=n-N}^n |a_i| + \epsilon s^*, \end{aligned} \quad (5.44)$$

where $B = \max(|\beta_0|, |\beta_1|, \dots, |\beta_N|)$ and s^* is the sum of the series $\sum_{n=0}^\infty |a_n|$ ($\sum_{n=0}^\infty a_n$ is absolutely convergent). Furthermore, because of this and by the Cauchy criterion we can find an integer M such that

$$\sum_{i=n-N}^n |a_i| < \epsilon \quad \text{if } n - N > M + 1.$$

Thus when $n > N + M + 1$ we get from inequality (5.44)

$$|a_0 \beta_n + a_1 \beta_{n-1} + \cdots + a_n \beta_0| \leq \epsilon(B + s^*).$$

Since ϵ can be arbitrarily small, we conclude that

$$\lim_{n \rightarrow \infty} (a_0 \beta_n + a_1 \beta_{n-1} + \cdots + a_n \beta_0) = 0.$$

2. Let v_n denote the n th partial sum of $\sum_{i=0}^{\infty} |c_i|$. Then

$$\begin{aligned} v_n &= |a_0 b_0| + |a_0 b_1 + a_1 b_0| + \cdots + |a_0 b_n + a_1 b_{n-1} + \cdots + a_n b_0| \\ &\leq |a_0| |b_0| + |a_0| |b_1| + |a_1| |b_0| \\ &\quad + \cdots + |a_0| |b_n| + |a_1| |b_{n-1}| + \cdots + |a_n| |b_0| \\ &= |a_0| t_n^* + |a_1| t_{n-1}^* + \cdots + |a_n| t_0^*, \end{aligned}$$

where $t_k^* = \sum_{i=0}^k |b_i|$, $k = 0, 1, 2, \dots, n$. Thus,

$$\begin{aligned} v_n &\leq (|a_0| + |a_1| + \cdots + |a_n|) t_n^* \\ &\leq s^* t^* \quad \text{for all } n, \end{aligned}$$

where t^* is the sum of the series $\sum_{n=0}^{\infty} |b_n|$, which is convergent by assumption. We conclude that the sequence $\{v_n\}_{n=0}^{\infty}$ is bounded. Since $v_n \geq 0$, then by Exercise 5.12 this sequence is convergent, and therefore $\sum_{n=0}^{\infty} c_n$ converges absolutely. By part (1), the sum of this series is st . $\square$

It should be noted that absolute convergence of at least one of $\sum_{n=0}^{\infty} a_n$ and $\sum_{n=0}^{\infty} b_n$ is an essential condition for the validity of part (1) of Theorem 5.2.17. If this condition is not satisfied, then $\sum_{n=0}^{\infty} c_n$ may not converge. For example, consider the series $\sum_{n=0}^{\infty} a_n$, $\sum_{n=0}^{\infty} b_n$, where

$$a_n = b_n = \frac{(-1)^n}{\sqrt{n+1}}, \quad n = 0, 1, \dots$$

These two series are convergent by Theorem 5.2.14. They are not, however, absolutely convergent, and their Cauchy's product is divergent (see Exercise 5.22).

5.3. SEQUENCES AND SERIES OF FUNCTIONS

All the sequences and series considered thus far in this chapter had constant terms. We now extend our study to sequences and series whose terms are functions of x .

Definition 5.3.1. Let $\{f_n(x)\}_{n=1}^{\infty}$ be a sequence of functions defined on a set $D \subset \mathbb{R}$.

1. If there exists a function $f(x)$ defined on D such that for every x in D ,

$$\lim_{n \rightarrow \infty} f_n(x) = f(x),$$

then the sequence $\{f_n(x)\}_{n=1}^{\infty}$ is said to converge to $f(x)$ on D . Thus for a given $\epsilon > 0$ there exists an integer N such that $|f_n(x) - f(x)| < \epsilon$ if $n > N$. In general, N depends on ϵ as well as on x .

2. If $\sum_{n=1}^{\infty} f_n(x)$ converges for every x in D to $s(x)$, then $s(x)$ is said to be the sum of the series. In this case, for a given $\epsilon > 0$ there exists an integer N such that

$$|s_n(x) - s(x)| < \epsilon \quad \text{if } n > N,$$

where $s_n(x)$ is the n th partial sum of the series $\sum_{n=1}^{\infty} f_n(x)$. The integer N depends on ϵ and, in general, on x also.

3. In particular, if N in (1) depends on ϵ but not on $x \in D$, then the sequence $\{f_n(x)\}_{n=1}^{\infty}$ is said to converge uniformly to $f(x)$ on D . Similarly, if N in (2) depends on ϵ , but not on $x \in D$, then the series $\sum_{n=1}^{\infty} f_n(x)$ converges uniformly to $s(x)$ on D . $\square$

The Cauchy criterion for sequences (Theorem 5.1.6) and its application to series (Theorem 5.2.1) apply to sequences and series of functions. In case of uniform convergence, the integer N described in this criterion depends only on ϵ .

Theorem 5.3.1. Let $\{f_n(x)\}_{n=1}^{\infty}$ be a sequence of functions defined on $D \subset \mathbb{R}$ and converging to $f(x)$. Define the number λ_n as

$$\lambda_n = \sup_{x \in D} |f_n(x) - f(x)|.$$

Then the sequence converges uniformly to $f(x)$ on D if and only if $\lambda_n \rightarrow 0$ as $n \rightarrow \infty$.

Proof. Sufficiency: Suppose that $\lambda_n \rightarrow 0$ as $n \rightarrow \infty$. To show that $f_n(x) \rightarrow f(x)$ uniformly on D . Let $\epsilon > 0$ be given. Then there exists an integer N such that for $n > N$, $\lambda_n < \epsilon$. Hence, for such values of n ,

$$|f_n(x) - f(x)| \leq \lambda_n < \epsilon$$

for all $x \in D$. Since N depends only on ϵ , the sequence $\{f_n(x)\}_{n=1}^{\infty}$ converges uniformly to $f(x)$ on D .

Necessity: Suppose that $f_n(x) \rightarrow f(x)$ uniformly on D . To show that $\lambda_n \rightarrow 0$. Let $\epsilon > 0$ be given. There exists an integer N that depends only on ϵ such that for $n > N$,

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

for all $x \in D$. It follows that

$$\lambda_n = \sup_{x \in D} |f_n(x) - f(x)| \leq \frac{\epsilon}{2}.$$

Thus $\lambda_n \rightarrow 0$ as $n \rightarrow \infty$. $\square$

Theorem 5.3.1 can be applied to convergent series of functions by replacing $f_n(x)$ and $f(x)$ with $s_n(x)$ and $s(x)$, respectively, where $s_n(x)$ is the n th partial sum of the series and $s(x)$ is its sum.

EXAMPLE 5.3.1. Let $f_n(x) = \sin(2\pi x/n)$, $0 \leq x \leq 1$. Then $f_n(x) \rightarrow 0$ as $n \rightarrow \infty$. Furthermore,

$$\left| \sin\left(\frac{2\pi x}{n}\right) \right| \leq \frac{2\pi x}{n} \leq \frac{2\pi}{n}.$$

In this case,

$$\lambda_n = \sup_{0 \leq x \leq 1} \left| \sin\left(\frac{2\pi x}{n}\right) \right| \leq \frac{2\pi}{n}.$$

Thus $\lambda_n \rightarrow 0$ as $n \rightarrow \infty$, and the sequence $\{f_n(x)\}_{n=1}^{\infty}$ converges uniformly to $f(x) = 0$ on $[0, 1]$.

The next theorem provides a simple test for uniform convergence of series of functions. It is due to Karl Weierstrass (1815–1897).

Theorem 5.3.2 (Weierstrass's M -test). Let $\sum_{n=1}^{\infty} f_n(x)$ be a series of functions defined on $D \subset \mathbb{R}$. If there exists a sequence $\{M_n\}_{n=1}^{\infty}$ of constants such that

$$|f_n(x)| \leq M_n, \quad n = 1, 2, \dots,$$

for all $x \in D$, and if $\sum_{n=1}^{\infty} M_n$ converges, then $\sum_{n=1}^{\infty} f_n(x)$ converges uniformly on D .

Proof. Let $\epsilon > 0$ be given. By the Cauchy criterion (Theorem 5.2.1), there exists an integer N such that

$$\sum_{i=m+1}^n M_i < \epsilon$$

for all $n > m > N$. Hence, for all such values of m , n , and for all $x \in D$,

$$\begin{aligned} \left| \sum_{i=m+1}^n f_i(x) \right| &\leq \sum_{i=m+1}^n |f_i(x)| \\ &\leq \sum_{i=m+1}^n M_i < \epsilon. \end{aligned}$$

This implies that $\sum_{n=1}^{\infty} f_n(x)$ converges uniformly on D by the Cauchy criterion. $\square$

We note that Weierstrass's M -test is easier to apply than Theorem 5.3.1, since it does not require specifying the sum $s(x)$ of the series.

EXAMPLE 5.3.2. Let us investigate convergence of the sequence $\{f_n(x)\}_{n=1}^{\infty}$, where $f_n(x)$ is defined as

$$f_n(x) = \begin{cases} 2x + \frac{1}{n}, & 0 \leq x < 1, \\ \exp\left(\frac{x}{n}\right), & 1 \leq x < 2, \\ 1 - \frac{1}{n}, & x \geq 2. \end{cases}$$

This sequence converges to

$$f(x) = \begin{cases} 2x, & 0 \leq x < 1, \\ 1, & x \geq 1. \end{cases}$$

Now,

$$|f_n(x) - f(x)| = \begin{cases} 1/n, & 0 \leq x < 1, \\ \exp(x/n) - 1, & 1 \leq x < 2, \\ 1/n, & x \geq 2. \end{cases}$$

However, for $1 \leq x < 2$,

$$\exp(x/n) - 1 < \exp(2/n) - 1.$$

Furthermore, by Maclaurin's series expansion,

$$\exp\left(\frac{2}{n}\right) - 1 = \sum_{k=1}^{\infty} \frac{(2/n)^k}{k!} > \frac{2}{n} > \frac{1}{n}.$$

Thus,

$$\sup_{0 \leq x < \infty} |f_n(x) - f(x)| = \exp(2/n) - 1,$$

which tends to zero as $n \rightarrow \infty$. Therefore, the sequence $\{f_n(x)\}_{n=1}^{\infty}$ converges uniformly to $f(x)$ on $[0, \infty)$.

EXAMPLE 5.3.3. Consider the series $\sum_{n=1}^{\infty} f_n(x)$, where

$$f_n(x) = \frac{x^n}{n^3 + nx^n}, \quad 0 \leq x \leq 1.$$

The function $f_n(x)$ is monotone increasing with respect to x . It follows that for $0 \leq x \leq 1$,

$$|f_n(x)| = f_n(x) \leq \frac{1}{n^3 + n}, \quad n = 1, 2, \dots$$

But the series $\sum_{n=1}^{\infty} [1/(n^3 + n)]$ is convergent. Hence, $\sum_{n=1}^{\infty} f_n(x)$ is uniformly convergent on $[0, 1]$ by Weierstrass's M -test.

5.3.1. Properties of Uniformly Convergent Sequences and Series

Sequences and series of functions that are uniformly convergent have several interesting properties. We shall study some of these properties in this section.

Theorem 5.3.3. Let $\{f_n(x)\}_{n=1}^{\infty}$ be uniformly convergent to $f(x)$ on a set D . If for each n , $f_n(x)$ has a limit τ_n as $x \rightarrow x_0$, where x_0 is a limit point of D , then the sequence $\{\tau_n\}_{n=1}^{\infty}$ converges to $\tau_0 = \lim_{x \rightarrow x_0} f(x)$. This is equivalent to stating that

$$\lim_{n \rightarrow \infty} \left[\lim_{x \rightarrow x_0} f_n(x) \right] = \lim_{x \rightarrow x_0} \left[\lim_{n \rightarrow \infty} f_n(x) \right].$$

Proof. Let us first show that $\{\tau_n\}_{n=1}^{\infty}$ is a convergent sequence. By the Cauchy criterion (Theorem 5.1.6), there exists an integer N such that for a given $\epsilon > 0$,

$$|f_m(x) - f_n(x)| < \frac{\epsilon}{2} \quad \text{for all } m > N, n > N. \quad (5.45)$$

The integer N depends only on ϵ , and inequality (5.45) is true for all $x \in D$, since the sequence is uniformly convergent. By taking the limit as $x \rightarrow x_0$ in (5.45) we get

$$|\tau_m - \tau_n| \leq \frac{\epsilon}{2} \quad \text{if } m > N, n > N,$$

which indicates that $\{\tau_n\}_{n=1}^\infty$ is a Cauchy sequence and is therefore convergent. Let $\tau_0 = \lim_{n \rightarrow \infty} \tau_n$. We now need to show that $f(x)$ has a limit and that this limit is equal to τ_0 . Let $\epsilon > 0$ be given. There exists an integer N_1 such that for $n > N_1$,

$$|f(x) - f_n(x)| < \frac{\epsilon}{4}$$

for all $x \in D$, by the uniform convergence of the sequence. Furthermore, there exists an integer N_2 such that

$$|\tau_n - \tau_0| < \frac{\epsilon}{4} \quad \text{if } n > N_2.$$

Thus for $n > \max(N_1, N_2)$,

$$|f(x) - f_n(x)| + |\tau_n - \tau_0| < \frac{\epsilon}{2}$$

for all $x \in D$. Then

$$\begin{aligned} |f(x) - \tau_0| &\leq |f(x) - f_n(x)| + |f_n(x) - \tau_n| + |\tau_n - \tau_0| \\ &< |f_n(x) - \tau_n| + \frac{\epsilon}{2} \end{aligned} \quad (5.46)$$

if $n > \max(N_1, N_2)$ for all $x \in D$. By taking the limit as $x \rightarrow x_0$ in (5.46) we get

$$\left| \lim_{x \rightarrow x_0} f(x) - \tau_0 \right| \leq \frac{\epsilon}{2} \quad (5.47)$$

by the fact that

$$\lim_{x \rightarrow x_0} |f_n(x) - \tau_n| = 0 \quad \text{for } n = 1, 2, \dots$$

Since ϵ is arbitrarily small, inequality (5.47) implies that

$$\lim_{x \rightarrow x_0} f(x) = \tau_0. \quad \square$$

Corollary 5.3.1. Let $\{f_n(x)\}_{n=1}^\infty$ be a sequence of continuous functions that converges uniformly to $f(x)$ on a set D . Then $f(x)$ is continuous on D .

Proof. The proof follows directly from Theorem 5.3.3, since $\tau_n = f_n(x_0)$ for $n \geq 1$ and $\tau_0 = \lim_{n \rightarrow \infty} \tau_n = \lim_{n \rightarrow \infty} f_n(x_0) = f(x_0)$. $\square$

Corollary 5.3.2. Let $\sum_{n=1}^{\infty} f_n(x)$ be a series of functions that converges uniformly to $s(x)$ on a set D . If for each n , $f_n(x)$ has a limit τ_n as $x \rightarrow x_0$, then the series $\sum_{n=1}^{\infty} \tau_n$ converges and has a sum equal to $s_0 = \lim_{x \rightarrow x_0} s(x)$, that is,

$$\lim_{x \rightarrow x_0} \sum_{n=1}^{\infty} f_n(x) = \sum_{n=1}^{\infty} \lim_{x \rightarrow x_0} f_n(x).$$

Proof. The proof is left to the reader. $\square$

By combining Corollaries 5.3.1 and 5.3.2 we conclude the following corollary:

Corollary 5.3.3. Let $\sum_{n=1}^{\infty} f_n(x)$ be a series of continuous functions that converges uniformly to $s(x)$ on a set D . Then $s(x)$ is continuous on D .

EXAMPLE 5.3.4. Let $f_n(x) = x^2/(1+x^2)^{n-1}$ be defined on $[1, \infty)$ for $n \geq 1$. Let $s_n(x)$ be the n th partial sum of the series $\sum_{n=1}^{\infty} f_n(x)$. Then,

$$s_n(x) = \sum_{k=1}^n \frac{x^2}{(1+x^2)^{k-1}} = x^2 \left[\frac{1 - \frac{1}{(1+x^2)^n}}{1 - \frac{1}{1+x^2}} \right],$$

by using the fact that the sum of the finite geometric series $\sum_{k=1}^n a^{k-1}$ is

$$\sum_{k=1}^n a^{k-1} = \frac{1-a^n}{1-a}. \quad (5.48)$$

Since $1/(1+x^2) < 1$ for $x \geq 1$, then as $n \rightarrow \infty$,

$$s_n(x) \rightarrow \frac{x^2}{1 - \frac{1}{1+x^2}} = 1+x^2.$$

Thus,

$$\sum_{n=1}^{\infty} \frac{x^2}{(1+x^2)^{n-1}} = 1+x^2.$$

Now, let $x_0 = 1$, then

$$\begin{aligned} \sum_{n=1}^{\infty} \lim_{x \rightarrow 1} \frac{x^2}{(1+x^2)^{n-1}} &= \sum_{n=1}^{\infty} \frac{1}{2^{n-1}} = \frac{1}{1 - \frac{1}{2}} = 2 \\ &= \lim_{x \rightarrow 1} (1+x^2), \end{aligned}$$

which results from applying formula (5.48) with $a = \frac{1}{2}$ and then letting $n \rightarrow \infty$. This provides a verification to Corollary 5.3.2. Note that the series $\sum_{n=1}^{\infty} f_n(x)$ is uniformly convergent by Weierstrass's M -test (why?).

Corollaries 5.3.2 and 5.3.3 clearly show that the properties of the function $f_n(x)$ carry over to the sum $s(x)$ of the series $\sum_{n=1}^{\infty} f_n(x)$ when the series is uniformly convergent.

Another property that $s(x)$ shares with the $f_n(x)$'s is given by the following theorem:

Theorem 5.3.4. Let $\sum_{n=1}^{\infty} f_n(x)$ be a series of functions, where $f_n(x)$ is differentiable on $[a, b]$ for $n \geq 1$. Suppose that $\sum_{n=1}^{\infty} f_n(x)$ converges at least at one point $x_0 \in [a, b]$ and that $\sum_{n=1}^{\infty} f'_n(x)$ converges uniformly on $[a, b]$. Then we have the following:

1. $\sum_{n=1}^{\infty} f_n(x)$ converges uniformly to $s(x)$ on $[a, b]$.
2. $s'(x) = \sum_{n=1}^{\infty} f'_n(x)$, that is, the derivative of $s(x)$ is obtained by a term-by-term differentiation of the series $\sum_{n=1}^{\infty} f_n(x)$.

Proof.

1. Let $x \neq x_0$ be a point in $[a, b]$. By the mean value theorem (Theorem 4.2.2), there exists a point ξ_n between x and x_0 such that for $n \geq 1$,

$$f_n(x) - f_n(x_0) = (x - x_0)f'_n(\xi_n). \quad (5.49)$$

Since $\sum_{n=1}^{\infty} f'_n(x)$ is uniformly convergent on $[a, b]$, then by the Cauchy criterion, there exists an integer N such that

$$\left| \sum_{i=m+1}^n f'_i(x) \right| < \frac{\epsilon}{b-a}$$

for all $n > m > N$ and for any $x \in [a, b]$. From (5.49) we get

$$\begin{aligned} \left| \sum_{i=m+1}^n [f_i(x) - f_i(x_0)] \right| &= |x - x_0| \left| \sum_{i=m+1}^n f'_i(\xi_i) \right| \\ &< \frac{\epsilon}{b-a} |x - x_0| \\ &< \epsilon \end{aligned}$$

for all $n > m > N$ and for any $x \in [a, b]$. This shows that

$$\sum_{n=1}^{\infty} [f_n(x) - f_n(x_0)]$$

is uniformly convergent on D . Consequently,

$$\sum_{n=1}^{\infty} f_n(x) = \sum_{n=1}^{\infty} [f_n(x) - f_n(x_0)] + s(x_0)$$

is uniformly convergent to $s(x)$ on D , where $s(x_0)$ is the sum of the series $\sum_{n=1}^{\infty} f_n(x_0)$, which was assumed to be convergent.

2. Let $\phi_n(h)$ denote the ratio

$$\phi_n(h) = \frac{f_n(x+h) - f_n(x)}{h}, \quad n = 1, 2, \dots,$$

where both x and $x+h$ belong to $[a, b]$. By invoking the mean value theorem again, $\phi_n(h)$ can be written as

$$\phi_n(h) = f'_n(x + \theta_n h), \quad n = 1, 2, \dots,$$

where $0 < \theta_n < 1$. Furthermore, by the uniform convergence of $\sum_{n=1}^{\infty} f'_n(x)$ we can deduce that $\sum_{n=1}^{\infty} \phi_n(h)$ is also uniformly convergent on $[-r, r]$ for some $r > 0$. But

$$\begin{aligned} \sum_{n=1}^{\infty} \phi_n(h) &= \sum_{n=1}^{\infty} \frac{f_n(x+h) - f_n(x)}{h} \\ &= \frac{s(x+h) - s(x)}{h}, \end{aligned} \quad (5.50)$$

where $s(x)$ is the sum of the series $\sum_{n=1}^{\infty} f_n(x)$. Let us now apply Corollary 5.3.2 to $\sum_{n=1}^{\infty} \phi_n(h)$. We get

$$\lim_{h \rightarrow 0} \sum_{n=1}^{\infty} \phi_n(h) = \sum_{n=1}^{\infty} \lim_{h \rightarrow 0} \phi_n(h). \quad (5.51)$$

From (5.50) and (5.51) we then have

$$\lim_{h \rightarrow 0} \frac{s(x+h) - s(x)}{h} = \sum_{n=1}^{\infty} f'_n(x).$$

Thus,

$$s'(x) = \sum_{n=1}^{\infty} f'_n(x). \quad \square$$

5.4. POWER SERIES

A power series is a special case of the series of functions discussed in Section 5.3. It is of the form $\sum_{n=0}^{\infty} a_n x^n$, where the a_n 's are constants. We have already encountered such series in connection with Taylor's and Maclaurin's series in Section 4.3.

Obviously, just as with any series of functions, the convergence of a power series depends on the values of x . By definition, if there exists a number $\rho > 0$ such that $\sum_{n=0}^{\infty} a_n x^n$ is convergent if $|x| < \rho$ and is divergent if $|x| > \rho$, then ρ is said to be the radius of convergence of the series, and the interval $(-\rho, \rho)$ is called the interval of convergence. The set of all values of x for which the power series converges is called its region of convergence.

The definition of the radius of convergence implies that $\sum_{n=0}^{\infty} a_n x^n$ is absolutely convergent within its interval of convergence. This is shown in the next theorem.

Theorem 5.4.1. Let ρ be the radius of convergence of $\sum_{n=0}^{\infty} a_n x^n$. Suppose that $\rho > 0$. Then $\sum_{n=0}^{\infty} a_n x^n$ converges absolutely for all x inside the interval $(-\rho, \rho)$.

Proof. Let x be such that $|x| < \rho$. There exists a point $x_0 \in (-\rho, \rho)$ such that $|x| < |x_0|$. Then, $\sum_{n=0}^{\infty} a_n x_0^n$ is a convergent series. By Result 5.2.1, $a_n x_0^n \rightarrow 0$ as $n \rightarrow \infty$, and hence $\{a_n x_0^n\}_{n=0}^{\infty}$ is a bounded sequence by Theorem 5.1.1. Thus

$$|a_n x_0^n| < K \quad \text{for all } n.$$

Now,

$$\begin{aligned} |a_n x^n| &= \left| a_n \left(\frac{x}{x_0} \right)^n x_0^n \right| \\ &< K \eta^n, \end{aligned}$$

where

$$\eta = \left| \frac{x}{x_0} \right| < 1.$$

Since the geometric series $\sum_{n=0}^{\infty} \eta^n$ is convergent, then by the comparison test (see Theorem 5.2.4), the series $\sum_{n=0}^{\infty} |a_n x^n|$ is convergent. $\square$

To determine the radius of convergence we shall rely on some of the tests of convergence given in Section 5.2.1.

Theorem 5.4.2. Let $\sum_{n=0}^{\infty} a_n x^n$ be a power series. Suppose that

$$\lim_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| = p.$$

Then the radius of convergence of the power series is

$$\rho = \begin{cases} 1/p, & 0 < p < \infty, \\ 0, & p = \infty, \\ \infty, & p = 0. \end{cases}$$

Proof. The proof follows from applying the ratio test given in Theorem 5.2.6 to the series $\sum_{n=0}^{\infty} |a_n x^n|$: We have that if

$$\lim_{n \rightarrow \infty} \left| \frac{a_{n+1} x^{n+1}}{a_n x^n} \right| < 1,$$

then $\sum_{n=0}^{\infty} a_n x^n$ is absolutely convergent. This inequality can be written as

$$p|x| < 1. \quad (5.52)$$

If $0 < p < \infty$, then absolute convergence occurs if $|x| < 1/p$ and the series diverges when $|x| > 1/p$. Thus $\rho = 1/p$. If $p = \infty$, the series diverges whenever $x \neq 0$. In this case, $\rho = 0$. If $p = 0$, then (5.52) holds for any value of x , that is, $\rho = \infty$. $\square$

Theorem 5.4.3. Let $\sum_{n=0}^{\infty} a_n x^n$ be a power series. Suppose that

$$\limsup_{n \rightarrow \infty} |a_n|^{1/n} = q.$$

Then,

$$\rho = \begin{cases} 1/q, & 0 < q < \infty, \\ 0, & q = \infty, \\ \infty, & q = 0. \end{cases}$$

Proof. This result follows from applying the root test in Theorem 5.2.8 to the series $\sum_{n=0}^{\infty} |a_n x^n|$. Details of the proof are similar to those given in Theorem 5.4.2. $\square$

The determination of the region of convergence of $\sum_{n=0}^{\infty} a_n x^n$ depends on the value of ρ . We know that the series converges if $|x| < \rho$ and diverges if $|x| > \rho$. The convergence of the series at $x = \rho$ and $x = -\rho$ has to be determined separately. Thus the region of convergence can be $(-\rho, \rho)$, $[-\rho, \rho)$, $(-\rho, \rho]$, or $[-\rho, \rho]$.

EXAMPLE 5.4.1. Consider the geometric series $\sum_{n=0}^{\infty} x^n$. By applying either Theorem 5.4.2 or Theorem 5.4.3, it is easy to show that $\rho = 1$. The series diverges if $x = 1$ or -1 . Thus the region of convergence is $(-1, 1)$. The sum

of this series can be obtained from formula (5.48) by letting n go to infinity. Thus

$$\sum_{n=0}^{\infty} x^n = \frac{1}{1-x}, \quad -1 < x < 1. \quad (5.53)$$

EXAMPLE 5.4.2. Consider the series $\sum_{n=0}^{\infty} (x^n/n!)$. Here,

$$\begin{aligned} \lim_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| &= \lim_{n \rightarrow \infty} \frac{n!}{(n+1)!} \\ &= \lim_{n \rightarrow \infty} \frac{1}{n+1} = 0. \end{aligned}$$

Thus $\rho = \infty$, and the series converges absolutely for any value of x . This particular series is Maclaurin's expansion of e^x , that is,

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}.$$

EXAMPLE 5.4.3. Suppose we have the series $\sum_{n=1}^{\infty} (x^n/n)$. Then

$$\lim_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| = \lim_{n \rightarrow \infty} \frac{n}{n+1} = 1,$$

and $\rho = 1$. When $x = 1$ we get the harmonic series, which is divergent. When $x = -1$ we get the alternating harmonic series, which is convergent by Theorem 5.2.14. Thus the region of convergence is $[-1, 1)$.

In addition to being absolutely convergent within its interval of convergence, a power series is also uniformly convergent there. This is shown in the next theorem.

Theorem 5.4.4. Let $\sum_{n=0}^{\infty} a_n x^n$ be a power series with a radius of convergence ρ (> 0). Then we have the following:

1. The series converges uniformly on the interval $[-r, r]$, where $r < \rho$.
2. If $s(x) = \sum_{n=0}^{\infty} a_n x^n$, then $s(x)$ (i) is continuous on $[-r, r]$; (ii) is differentiable on $[-r, r]$ and has derivative

$$s'(x) = \sum_{n=1}^{\infty} n a_n x^{n-1}, \quad -r \leq x \leq r;$$

and (iii) has derivatives of all orders on $[-r, r]$ and

$$\frac{d^k s(x)}{dx^k} = \sum_{n=k}^{\infty} \frac{a_n n!}{(n-k)!} x^{n-k}, \quad k = 1, 2, \dots, -r \leq x \leq r.$$

Proof.

1. If $|x| \leq r$, then $|a_n x^n| \leq |a_n| r^n$ for $n \geq 0$. Since $\sum_{n=0}^{\infty} |a_n| r^n$ is convergent by Theorem 5.4.1, then by the Weierstrass M -test (Theorem 5.3.2), $\sum_{n=0}^{\infty} a_n x^n$ is uniformly convergent on $[-r, r]$.
2. (i) Continuity of $s(x)$ follows directly from Corollary 5.3.3. (ii) To show this result, we first note that the two series $\sum_{n=0}^{\infty} a_n x^n$ and $\sum_{n=1}^{\infty} n a_n x^{n-1}$ have the same radius of convergence. This is true by Theorem 5.4.3 and the fact that

$$\limsup_{n \rightarrow \infty} |n a_n|^{1/n} = \limsup_{n \rightarrow \infty} |a_n|^{1/n},$$

since $\lim_{n \rightarrow \infty} n^{1/n} = 1$ as $n \rightarrow \infty$. We can then assert that $\sum_{n=1}^{\infty} n a_n x^{n-1}$ is uniformly convergent on $[-r, r]$. By Theorem 5.3.4, $s(x)$ is differentiable on $[-r, r]$, and its derivative is obtained by a term-by-term differentiation of $\sum_{n=0}^{\infty} a_n x^n$. (iii) This follows from part (ii) by repeated differentiation of $s(x)$. $\square$

Under a certain condition, the interval on which the power series converges uniformly can include the end points of the interval of convergence. This is discussed in the next theorem.

Theorem 5.4.5. Let $\sum_{n=0}^{\infty} a_n x^n$ be a power series with a finite nonzero radius of convergence ρ . If $\sum_{n=0}^{\infty} a_n \rho^n$ is absolutely convergent, then the power series is uniformly convergent on $[-\rho, \rho]$.

Proof. The proof is similar to that of part 1 of Theorem 5.4.4. In this case, for $|x| \leq \rho$, $|a_n x^n| \leq |a_n| \rho^n$. Since $\sum_{n=0}^{\infty} |a_n| \rho^n$ is convergent, then $\sum_{n=0}^{\infty} a_n x^n$ is uniformly convergent on $[-\rho, \rho]$ by the Weierstrass M -test. $\square$

EXAMPLE 5.4.4. Consider the geometric series of Example 5.4.1. This series is uniformly convergent on $[-r, r]$, where $r < 1$. Furthermore, by differentiating the two sides of (5.53) we get

$$\sum_{n=1}^{\infty} n x^{n-1} = \frac{1}{(1-x)^2}, \quad -1 < x < 1.$$

This provides a series expansion of $1/(1-x)^2$ within the interval $(-1, 1)$. By repeated differentiation it is easy to show that for $-1 < x < 1$,

$$\frac{1}{(1-x)^k} = \sum_{n=0}^{\infty} \binom{n+k-1}{n} x^n, \quad k = 1, 2, \dots$$

The radius of convergence of this series is $\rho = 1$, the same as for the original series.

EXAMPLE 5.4.5. Suppose we have the series

$$\sum_{n=1}^{\infty} \frac{2^n}{2n^2 + n} \left(\frac{x}{1-x} \right)^n,$$

which can be written as

$$\sum_{n=1}^{\infty} \frac{z^n}{2n^2 + n},$$

where $z = 2x/(1-x)$. This is a power series in z . By Theorem 5.4.2, the radius of convergence of this series is $\rho = 1$. We note that when $z = 1$ the series $\sum_{n=1}^{\infty} [1/(2n^2 + n)]$ is absolutely convergent. Thus by Theorem 5.4.5, the given series is uniformly convergent for $|z| \leq 1$, that is, for values of x satisfying

$$-\frac{1}{2} \leq \frac{x}{1-x} \leq \frac{1}{2},$$

or equivalently,

$$-1 \leq x \leq \frac{1}{3}.$$

5.5. SEQUENCES AND SERIES OF MATRICES

In Section 5.3 we considered sequences and series whose terms were scalar functions of x rather than being constant as was done in Sections 5.1 and 5.2. In this section we consider yet another extension, in which the terms of the series are matrices rather than scalars. We shall provide a brief discussion of this extension. The interested reader can find a more detailed study of this topic in Gantmacher (1959), Lancaster (1969), and Graybill (1983). As in Chapter 2, all matrix elements considered here are real.

For the purpose of our study of sequences and series of matrices we first need to introduce the norm of a matrix.

Definition 5.5.1. Let $\mathbf{A}$ be a matrix of order $m \times n$. A norm of $\mathbf{A}$, denoted by $\|\mathbf{A}\|$, is a real-valued function of $\mathbf{A}$ with the following properties:

1. $\|\mathbf{A}\| \geq 0$, and $\|\mathbf{A}\| = 0$ if and only if $\mathbf{A} = \mathbf{0}$.
2. $\|c\mathbf{A}\| = |c|\|\mathbf{A}\|$, where c is a scalar.
3. $\|\mathbf{A} + \mathbf{B}\| \leq \|\mathbf{A}\| + \|\mathbf{B}\|$, where $\mathbf{B}$ is any matrix of order $m \times n$.
4. $\|\mathbf{AC}\| \leq \|\mathbf{A}\|\|\mathbf{C}\|$, where $\mathbf{C}$ is any matrix for which the product $\mathbf{AC}$ is defined. $\square$

If $\mathbf{A} = (a_{ij})$, then examples of matrix norms that satisfy properties 1, 2, 3, and 4 include the following:

1. The Euclidean norm, $\|\mathbf{A}\|_2 = (\sum_{i=1}^m \sum_{j=1}^n a_{ij}^2)^{1/2}$.
2. The spectral norm, $\|\mathbf{A}\|_s = [e_{\max}(\mathbf{A}'\mathbf{A})]^{1/2}$, where $e_{\max}(\mathbf{A}'\mathbf{A})$ is the largest eigenvalue of $\mathbf{A}'\mathbf{A}$.

Definition 5.5.2. Let $\mathbf{A}_k = (a_{ijk})$ be matrices of orders $m \times n$ for $k \geq 1$. The sequence $\{\mathbf{A}_k\}_{k=1}^{\infty}$ is said to converge to the $m \times n$ matrix $\mathbf{A} = (a_{ij})$ if $\lim_{k \rightarrow \infty} a_{ijk} = a_{ij}$ for $i = 1, 2, \dots, m$; $j = 1, 2, \dots, n$. $\square$

For example, the sequence of matrices

$$\mathbf{A}_k = \begin{bmatrix} \frac{1}{k} & \frac{1}{k^2} - 1 & \frac{k-1}{k+1} \\ 2 & 0 & \frac{k^2}{2-k^2} \end{bmatrix}, \quad k = 1, 2, \dots,$$

converges to

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

as $k \rightarrow \infty$. The sequence

$$\mathbf{A}_k = \begin{bmatrix} \frac{1}{k} & k^2 - 2 \\ 1 & \frac{k}{1+k} \end{bmatrix}, \quad k = 1, 2, \dots,$$

does not converge, since $k^2 - 2$ goes to infinite as $k \rightarrow \infty$.

From Definition 5.5.2 it is easy to see that $\{\mathbf{A}_k\}_{k=1}^{\infty}$ converges to $\mathbf{A}$ if and only if

$$\lim_{k \rightarrow \infty} \|\mathbf{A}_k - \mathbf{A}\| = 0,$$

where $\|\cdot\|$ is any matrix norm.

Definition 5.5.3. Let $\{\mathbf{A}_k\}_{k=1}^{\infty}$ be a sequence of matrices of order $m \times n$. Then $\sum_{k=1}^{\infty} \mathbf{A}_k$ is called an infinite series (or just a series) of matrices. This series is said to converge to the $m \times n$ matrix $\mathbf{S} = (s_{ij})$ if and only if the series $\sum_{k=1}^{\infty} a_{ijk}$ converges for all $i = 1, 2, \dots, m$; $j = 1, 2, \dots, n$, where a_{ijk} is the (i, j) th element of $\mathbf{A}_k$, and

$$\sum_{k=1}^{\infty} a_{ijk} = s_{ij}, \quad i = 1, 2, \dots, m; j = 1, 2, \dots, n. \quad (5.54)$$

The series $\sum_{k=1}^{\infty} \mathbf{A}_k$ is divergent if at least one of the series in (5.54) is divergent. $\square$

From Definition 5.5.3 and Result 5.2.1 we conclude that $\sum_{k=1}^{\infty} \mathbf{A}_k$ diverges if $\lim_{k \rightarrow \infty} a_{ijk} \neq 0$ for at least one pair (i, j) , that is, if $\lim_{k \rightarrow \infty} \mathbf{A}_k \neq \mathbf{0}$.

A particular type of infinite series of matrices is the power series $\sum_{k=0}^{\infty} \alpha_k \mathbf{A}^k$, where $\mathbf{A}$ is a square matrix, α_k is a scalar ($k = 0, 1, \dots$), and $\mathbf{A}^0$ is by definition the identity matrix $\mathbf{I}$. For example, the power series

$$\mathbf{I} + \mathbf{A} + \frac{1}{2!} \mathbf{A}^2 + \frac{1}{3!} \mathbf{A}^3 + \dots + \frac{1}{k!} \mathbf{A}^k + \dots$$

represents an expansion of the exponential matrix function $\exp(\mathbf{A})$ (see Gantmacher, 1959).

Theorem 5.5.1. Let $\mathbf{A}$ be an $n \times n$ matrix. Then $\lim_{k \rightarrow \infty} \mathbf{A}^k = \mathbf{0}$ if $\|\mathbf{A}\| < 1$, where $\|\cdot\|$ is any matrix norm.

Proof. From property 4 in Definition 5.5.1 we can write

$$\|\mathbf{A}^k\| \leq \|\mathbf{A}\|^k, \quad k = 1, 2, \dots$$

Since $\|\mathbf{A}\| < 1$, then $\lim_{k \rightarrow \infty} \|\mathbf{A}^k\| = 0$, which implies that $\lim_{k \rightarrow \infty} \mathbf{A}^k = \mathbf{0}$ (why?). $\square$

Theorem 5.5.2. Let $\mathbf{A}$ be a symmetric matrix of order $n \times n$ such that $|\lambda_i| < 1$ for $i = 1, 2, \dots, n$, where λ_i is the i th eigenvalue of $\mathbf{A}$ (all the eigenvalues of $\mathbf{A}$ are real by Theorem 2.3.5). Then $\sum_{k=0}^{\infty} \mathbf{A}^k$ converges to $(\mathbf{I} - \mathbf{A})^{-1}$.

Proof. By the spectral decomposition theorem (Theorem 2.3.10) there exists an orthogonal matrix $\mathbf{P}$ such that $\mathbf{A} = \mathbf{P}\mathbf{\Lambda}\mathbf{P}'$, where $\mathbf{\Lambda}$ is a diagonal matrix whose diagonal elements are the eigenvalues of $\mathbf{A}$. Then

$$\mathbf{A}^k = \mathbf{P}\mathbf{\Lambda}^k\mathbf{P}', \quad k = 0, 1, 2, \dots$$

Since $|\lambda_i| < 1$ for all i , then $\mathbf{\Lambda}^k \rightarrow \mathbf{0}$ and hence $\mathbf{A}^k \rightarrow \mathbf{0}$ as $k \rightarrow \infty$. Furthermore, the matrix $\mathbf{I} - \mathbf{A}$ is nonsingular, since

$$\mathbf{I} - \mathbf{A} = \mathbf{P}(\mathbf{I} - \mathbf{\Lambda})\mathbf{P}'$$

and all the diagonal elements of $\mathbf{I} - \mathbf{\Lambda}$ are positive.

Now, for any nonnegative integer k we have the following identity:

$$(\mathbf{I} - \mathbf{A})(\mathbf{I} + \mathbf{A} + \mathbf{A}^2 + \dots + \mathbf{A}^k) = \mathbf{I} - \mathbf{A}^{k+1}.$$

Hence,

$$\mathbf{I} + \mathbf{A} + \dots + \mathbf{A}^k = (\mathbf{I} - \mathbf{A})^{-1}(\mathbf{I} - \mathbf{A}^{k+1}).$$

By letting k go to infinity we get

$$\sum_{k=0}^{\infty} \mathbf{A}^k = (\mathbf{I} - \mathbf{A})^{-1},$$

since $\lim_{k \rightarrow \infty} \mathbf{A}^{k+1} = \mathbf{0}$. $\square$

Theorem 5.5.3. Let $\mathbf{A}$ be a symmetric $n \times n$ matrix and λ be any eigenvalue of $\mathbf{A}$. Then $|\lambda| \leq \|\mathbf{A}\|$, where $\|\mathbf{A}\|$ is any matrix norm of $\mathbf{A}$.

Proof. We have that $\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$, where $\mathbf{v}$ is an eigenvector of $\mathbf{A}$ for the eigenvalue λ . If $\|\mathbf{A}\|$ is any matrix norm of $\mathbf{A}$, then

$$\|\lambda\mathbf{v}\| = |\lambda|\|\mathbf{v}\| = \|\mathbf{A}\mathbf{v}\| \leq \|\mathbf{A}\|\|\mathbf{v}\|.$$

Since $\mathbf{v} \neq \mathbf{0}$, we conclude that

$$|\lambda| \leq \|\mathbf{A}\|. \quad \square$$

Corollary 5.5.1. Let $\mathbf{A}$ be a symmetric matrix of order $n \times n$ such that $\|\mathbf{A}\| < 1$, where $\|\mathbf{A}\|$ is any matrix norm of $\mathbf{A}$. Then $\sum_{k=0}^{\infty} \mathbf{A}^k$ converges to $(\mathbf{I} - \mathbf{A})^{-1}$.

Proof. This result follows from Theorem 5.5.2, since for $i = 1, 2, \dots, n$, $|\lambda_i| \leq \|\mathbf{A}\| < 1$. $\square$

5.6. APPLICATIONS IN STATISTICS

Sequences and series have many useful applications in statistics. Some of these applications will be discussed in this section.

5.6.1. Moments of a Discrete Distribution

Perhaps one of the most visible applications of infinite series in statistics is in the study of the distribution of a discrete random variable that can assume a countable number of values. Under certain conditions, this distribution can be completely determined by its moments. By definition, the moments of a distribution are a set of descriptive constants that are useful for measuring its properties.

Let X be a discrete random variable that takes on the values $x_0, x_1, \dots, x_n, \dots$, with probabilities $p(n)$, $n \geq 0$. Then, by definition, the k th central moment of X , denoted by μ_k , is

$$\mu_k = E[(X - \mu)^k] = \sum_{n=0}^{\infty} (x_n - \mu)^k p(n), \quad k = 1, 2, \dots,$$

where $\mu = E(X) = \sum_{n=0}^{\infty} x_n p(n)$ is the mean of X . We note that $\mu_2 = \sigma^2$ is the variance of X . The k th noncentral moment of X is given by the series

$$\mu'_k = E(X^k) = \sum_{n=0}^{\infty} x_n^k p(n), \quad k = 1, 2, \dots \quad (5.55)$$

We note that $\mu'_1 = \mu$. If, for some integer N , $|x_n| \geq 1$ for $n > N$, and if the series in (5.55) converges absolutely, then so does the series for μ'_j ($j = 1, 2, \dots, k-1$). This follows from applying the comparison test:

$$|x_n|^j p(n) \leq |x_n|^k p(n) \quad \text{if } j < k \text{ and } n > N.$$

Examples of discrete random variables with a countable number of values include the Poisson (see Section 4.5.3) and the negative binomial. The latter random variable represents the number n of failures before the r th success when independent trials are performed, each of which has two probability outcomes, success or failure, with a constant probability p of success on each trial. Its probability mass function is therefore of the form

$$p(n) = \binom{n+r-1}{n} p^r (1-p)^n, \quad n = 0, 1, 2, \dots$$

By contrast, the Poisson random variable has the probability mass function

$$p(n) = \frac{e^{-\lambda} \lambda^n}{n!}, \quad n = 0, 1, 2, \dots,$$

where λ is the mean of X . We can verify that λ is the mean by writing

$$\begin{aligned} \mu &= \sum_{n=0}^{\infty} n \frac{e^{-\lambda} \lambda^n}{n!} \\ &= \lambda e^{-\lambda} \sum_{n=1}^{\infty} \frac{\lambda^{n-1}}{(n-1)!} \\ &= \lambda e^{-\lambda} \sum_{n=0}^{\infty} \frac{\lambda^n}{n!} \\ &= \lambda e^{-\lambda} (e^{\lambda}), \quad \text{by Maclaurin's expansion of } e^{\lambda} \\ &= \lambda. \end{aligned}$$

The second noncentral moment of the Poisson distribution is

$$\begin{aligned} \mu'_2 &= \sum_{n=0}^{\infty} n^2 \frac{e^{-\lambda} \lambda^n}{n!} \\ &= e^{-\lambda} \lambda \sum_{n=1}^{\infty} n \frac{\lambda^{n-1}}{(n-1)!} \\ &= e^{-\lambda} \lambda \sum_{n=1}^{\infty} (n-1+1) \frac{\lambda^{n-1}}{(n-1)!} \\ &= e^{-\lambda} \lambda \left[\lambda \sum_{n=2}^{\infty} \frac{\lambda^{n-2}}{(n-2)!} + \sum_{n=1}^{\infty} \frac{\lambda^{n-1}}{(n-1)!} \right] \\ &= e^{-\lambda} \lambda [\lambda e^{\lambda} + e^{\lambda}] \\ &= \lambda^2 + \lambda. \end{aligned}$$

In general, the k th noncentral moment of the Poisson distribution is given by the series

$$\mu'_k = \sum_{n=0}^{\infty} n^k \frac{e^{-\lambda} \lambda^n}{n!}, \quad k = 1, 2, \dots,$$

which converges for any k . This can be shown, for example, by the ratio test

$$\begin{aligned}\lim_{n \rightarrow \infty} \frac{a_{n+1}}{a_n} &= \lim_{n \rightarrow \infty} \left(\frac{n+1}{n} \right)^k \frac{\lambda}{n+1} \\ &= 0 < 1.\end{aligned}$$

Thus all the noncentral moments of the Poisson distribution exist.

Similarly, for the negative binomial distribution we have

$$\begin{aligned}\mu &= \sum_{n=0}^{\infty} n \binom{n+r-1}{n} p^r (1-p)^n \\ &= \frac{r(1-p)}{p} \quad (\text{why?}),\end{aligned}\tag{5.56}$$

$$\begin{aligned}\mu'_2 &= \sum_{n=0}^{\infty} n^2 \binom{n+r-1}{n} p^r (1-p)^n \\ &= \frac{r(1-p)(1+r-rp)}{p^2} \quad (\text{why?})\end{aligned}\tag{5.57}$$

and the k th noncentral moment,

$$\mu'_k = \sum_{n=0}^{\infty} n^k \binom{n+r-1}{n} p^r (1-p)^n, \quad k = 1, 2, \dots, \tag{5.58}$$

exists for any k , since, by the ratio test,

$$\begin{aligned}\lim_{n \rightarrow \infty} \frac{a_{n+1}}{a_n} &= \lim_{n \rightarrow \infty} \frac{\left(\frac{n+1}{n} \right)^k \binom{n+r}{n+1}}{\binom{n+r-1}{n}} (1-p) \\ &= (1-p) \lim_{n \rightarrow \infty} \left(\frac{n+1}{n} \right)^k \left(\frac{n+r}{n+1} \right) \\ &= 1-p < 1,\end{aligned}$$

which proves convergence of the series in (5.58).

A very important inequality that concerns the mean μ and variance σ^2 of any random variable X (not just the discrete ones) is Chebyshev's inequality,

namely,

$$P(|X - \mu| \geq r\sigma) \leq \frac{1}{r^2},$$

or equivalently,

$$P(|X - \mu| < r\sigma) \geq 1 - \frac{1}{r^2}, \quad (5.59)$$

where r is any positive number (see, for example, Lindgren, 1976, Section 2.3.2). The importance of this inequality stems from the fact that it is independent of the exact distribution of X and connects the variance of X with the distribution of its values. For example, inequality (5.59) states that at least $(1 - 1/r^2) \times 100\%$ of the values of X fall within $r\sigma$ from its mean, where $\sigma = \sqrt{\sigma^2}$ is the standard deviation of X .

Chebyshev's inequality is a special case of a more general inequality called Markov's inequality. If b is a nonzero constant and $h(x)$ is a nonnegative function, then

$$P[h(X) \geq b^2] \leq \frac{1}{b^2} E[h(X)],$$

provided that $E[h(X)]$ exists. Chebyshev's inequality follows from Markov's inequality by choosing $h(X) = (X - \mu)^2$.

Another important result that concerns the moments of a distribution is given by the following theorem, regarding what is known as the *Stieltjes moment problem*, which also applies to any random variable:

Theorem 5.6.1. Suppose that the moments μ'_k ($k = 1, 2, \dots$) of a random variable X exist, and the series

$$\sum_{k=1}^{\infty} \frac{\mu'_k}{k!} \tau^k \quad (5.60)$$

is absolutely convergent for some $\tau > 0$. Then these moments uniquely determine the cumulative distribution function $F(x)$ of X .

Proof. See, for example, Fisz (1963, Theorem 3.2.1). □

In particular, if

$$|\mu'_k| \leq M^k, \quad k = 1, 2, \dots,$$

for some constant M , then the series in (5.60) converges absolutely for any $\tau > 0$ by the comparison test. This is true because the series $\sum_{k=1}^{\infty} (M^k/k!) \tau^k$ converges (for example, by the ratio test) for any value of τ .

It should be noted that absolute convergence of the series in (5.60) is a sufficient condition for the unique determination of $F(x)$, but is not a necessary condition. This is shown in Rao (1973, page 106). Furthermore, if some moments of X fail to exist, then the remaining moments that do exist cannot determine $F(x)$ uniquely. The following counterexample is given in Fisz (1963, page 74):

Let X be a discrete random variable that takes on the values $x_n = 2^n/n^2$, $n \geq 1$, with probabilities $p(n) = 1/2^n$, $n \geq 1$. Then

$$\mu = E(X) = \sum_{n=1}^{\infty} \frac{1}{n^2},$$

which exists, because the series is convergent. However, μ'_2 does not exist, because

$$\mu'_2 = E(X^2) = \sum_{n=1}^{\infty} \frac{2^n}{n^4}$$

and this series is divergent, since $2^n/n^4 \rightarrow \infty$ as $n \rightarrow \infty$.

Now, let Y be another discrete random variable that takes on the value zero with probability $\frac{1}{2}$ and the values $y_n = 2^{n+1}/n^2$, $n \geq 1$, with probabilities $q(n) = 1/2^{n+1}$, $n \geq 1$. Then,

$$E(Y) = \sum_{n=1}^{\infty} \frac{1}{n^2} = E(X).$$

The second noncentral moment of Y does not exist, since

$$\mu'_2 = E(Y^2) = \sum_{n=1}^{\infty} \frac{2^{n+1}}{n^4},$$

and this series is divergent.

Since μ'_2 does not exist for both X and Y , none of their noncentral moments of order $k > 2$ exist either, as can be seen from applying the comparison test. Thus X and Y have the same first noncentral moments, but do not have noncentral moments of any order greater than 1. These two random variables have obviously different distributions.

5.6.2. Moment and Probability Generating Functions

Let X be a discrete random variable that takes on the values $x_0, x_1, x_2, \dots$ with probabilities $p(n)$, $n \geq 0$.

The Moment Generating Function of X

This function is defined as

$$\phi(t) = E(e^{tX}) = \sum_{n=0}^{\infty} e^{tx_n} p(n) \quad (5.61)$$

provided that the series converges. In particular, if $x_n = n$ for $n \geq 0$, then

$$\phi(t) = \sum_{n=0}^{\infty} e^{tn} p(n), \quad (5.62)$$

which is a power series in e^t . If ρ is the radius of convergence for this series, then by Theorem 5.4.4, $\phi(t)$ is a continuous function of t and has derivatives of all orders inside its interval of convergence. Since

$$\left. \frac{d^k \phi(t)}{dt^k} \right|_{t=0} = E(X^k) = \mu'_k, \quad k = 1, 2, \dots, \quad (5.63)$$

$\phi(t)$, when it exists, can be used to obtain all noncentral moments of X , which can completely determine the distribution of X by Theorem 5.6.1.

From (5.63), by using Maclaurin's expansion of $\phi(t)$, we can obtain an expression for this function as a power series in t :

$$\begin{aligned} \phi(t) &= \phi(0) + \sum_{n=1}^{\infty} \frac{t^n}{n!} \phi^{(n)}(0) \\ &= 1 + \sum_{n=1}^{\infty} \frac{\mu'_n}{n!} t^n. \end{aligned} \quad (5.64)$$

Let us now go back to the series in (5.62). If

$$\lim_{n \rightarrow \infty} \frac{p(n+1)}{p(n)} = p,$$

then by Theorem 5.4.2, the radius of convergence ρ is

$$\rho = \begin{cases} 1/p, & 0 < p < \infty, \\ 0, & p = \infty, \\ \infty, & p = 0. \end{cases}$$

Alternatively, if $\limsup_{n \rightarrow \infty} [p(n)]^{1/n} = q$, then

$$\rho = \begin{cases} 1/q, & 0 < q < \infty, \\ 0, & q = \infty, \\ \infty, & q = 0. \end{cases}$$

For example, for the Poisson distribution, where

$$p(n) = \frac{e^{-\lambda} \lambda^n}{n!}, \quad n = 0, 1, 2, \dots,$$

we have $\lim_{n \rightarrow \infty} [p(n+1)/p(n)] = \lim_{n \rightarrow \infty} [\lambda/(n+1)] = 0$. Hence, $\rho = \infty$, that is,

$$\phi(t) = \sum_{n=0}^{\infty} \frac{e^{-\lambda} \lambda^n}{n!} e^{tn}$$

converges uniformly for any value of t for which $e^t < \infty$, that is, $-\infty < t < \infty$. As a matter of fact, a closed-form expression for $\phi(t)$ can be found, since

$$\begin{aligned} \phi(t) &= e^{-\lambda} \sum_{n=0}^{\infty} \frac{(\lambda e^t)^n}{n!} \\ &= e^{-\lambda} \exp(\lambda e^t) \\ &= \exp(\lambda e^t - \lambda) \quad \text{for all } t. \end{aligned} \tag{5.65}$$

The k th noncentral moment of X is then given by

$$\mu'_k = \left. \frac{d^k \phi(t)}{dt^k} \right|_{t=0} = \left. \frac{d^k (\lambda e^t - \lambda)}{dt^k} \right|_{t=0}.$$

In particular, the first two noncentral moments are

$$\begin{aligned} \mu'_1 &= \mu = \lambda, \\ \mu'_2 &= \lambda + \lambda^2. \end{aligned}$$

This confirms our earlier finding concerning these two moments.

It should be noted that formula (5.63) is valid provided that there exists a $\delta > 0$ such that the neighborhood $N_\delta(0)$ is contained inside the interval of convergence. For example, let X have the probability mass function

$$p(n) = \frac{6}{\pi^2 n^2}, \quad n = 1, 2, \dots$$

Then

$$\lim_{n \rightarrow \infty} \frac{p(n+1)}{p(n)} = 1.$$

Hence, by Theorem 5.4.4, the series

$$\phi(t) = \sum_{n=1}^{\infty} \frac{6}{\pi^2 n^2} e^{tn}$$

converges uniformly for values of t satisfying $e^t \leq r$, where $r \leq 1$, or equivalently, for $t \leq \log r \leq 0$. If, however, $t > 0$, then the series diverges. Thus there does not exist a neighborhood $N_\delta(0)$ that is contained inside the interval of convergence for any $\delta > 0$. Consequently, formula (5.63) does not hold in this case.

From the moment generating function we can derive a series of constants that play a role similar to that of the moments. These constants are called cumulants. They have properties that are, in certain circumstances, more useful than those of the moments. Cumulants were originally defined and studied by Thiele (1903).

By definition, the cumulants of X , denoted by $\kappa_1, \kappa_2, \dots, \kappa_n, \dots$ are constants that satisfy the following identity in t :

$$\begin{aligned} \exp\left(\kappa_1 t + \frac{\kappa_2 t^2}{2!} + \dots + \frac{\kappa_n t^n}{n!} + \dots\right) \\ = 1 + \mu'_1 t + \frac{\mu'_2}{2!} t^2 + \dots + \frac{\mu'_n}{n!} t^n + \dots \end{aligned} \quad (5.66)$$

Using formula (5.64), this identity can be written as

$$\sum_{n=1}^{\infty} \frac{\kappa_n}{n!} t^n = \log \phi(t), \quad (5.67)$$

provided that $\phi(t)$ exists and is positive. By definition, the natural logarithm of the moment generating function of X is called the cumulant generating function.

Formula (5.66) can be used to express the noncentral moments in terms of the cumulants, and vice versa. Kendall and Stuart (1977, Section 3.14) give a general relationship that can be used for this purpose. For example,

$$\begin{aligned} \kappa_1 &= \mu'_1, \\ \kappa_2 &= \mu'_2 - \mu'^2_1, \\ \kappa_3 &= \mu'_3 - 3\mu'_1 \mu'_2 + 2\mu'^3_1. \end{aligned}$$

The cumulants have an interesting property in that they are, except for κ_1 , invariant to any constant shift c in X . That is, for $n = 2, 3, \dots$, κ_n is not changed if X is replaced by $X + c$. This follows from noting that

$$E[e^{(X+c)t}] = e^{ct} \phi(t),$$

which is the moment generating function of $X + c$. But

$$\log[e^{ct} \phi(t)] = ct + \log \phi(t).$$

By comparison with (5.67) we can then conclude that except for κ_1 , the cumulants of $X + c$ are the same as those of X . This contrasts sharply with the noncentral moments of X , which are not invariant to such a shift.

Another advantage of using cumulants is that they can be employed to obtain approximate expressions for the percentile points of the distribution of X (see Section 9.5.1).

EXAMPLE 5.6.1. Let X be a Poisson random variable whose moment generating function is given by formula (5.65). By applying (5.67) we get

$$\begin{aligned}\sum_{n=1}^{\infty} \frac{\kappa_n}{n!} t^n &= \log[\exp(\lambda e^t - \lambda)] \\ &= \lambda e^t - \lambda \\ &= \lambda \sum_{n=1}^{\infty} \frac{t^n}{n!}.\end{aligned}$$

Here, we have made use of Maclaurin's expansion of e^t . This series converges for any value of t . It follows that $\kappa_n = \lambda$ for $n = 1, 2, \dots$.

The Probability Generating Function

This is similar to the moment generating function. It is defined as

$$\psi(t) = E(t^X) = \sum_{n=0}^{\infty} t^{x_n} p(n).$$

In particular, if $x_n = n$ for $n \geq 0$, then

$$\psi(t) = \sum_{n=0}^{\infty} t^n p(n), \quad (5.68)$$

which is a power series in t . Within its interval of convergence, this series represents a continuous function with derivatives of all orders. We note that $\psi(0) = p(0)$ and that

$$\frac{1}{k!} \left. \frac{d^k \psi(t)}{dt^k} \right|_{t=0} = p(k), \quad k = 1, 2, \dots \quad (5.69)$$

Thus, the entire probability distribution of X is completely determined by $\psi(t)$.

The probability generating function is also useful in determining the

moments of X . This is accomplished by using the relation

$$\begin{aligned} \left. \frac{d^k \psi(t)}{dt^k} \right|_{t=1} &= \sum_{n=k}^{\infty} n(n-1) \cdots (n-k+1) p(n) \\ &= E[X(X-1) \cdots (X-k+1)]. \end{aligned} \quad (5.70)$$

The quantity on the right-hand side of (5.70) is called the k th factorial moment of X , which we denote by θ_k . The noncentral moments of X can be derived from the θ_k 's. For example,

$$\begin{aligned} \mu'_1 &= \theta_1, \\ \mu'_2 &= \theta_2 + \theta_1, \\ \mu'_3 &= \theta_3 + 3\theta_2 + \theta_1, \\ \mu'_4 &= \theta_4 + 6\theta_3 + 7\theta_2 + \theta_1. \end{aligned}$$

Obviously, formula (5.70) is valid provided that $t = 1$ belongs to the interval of convergence of the series in (5.68).

If a closed-form expression is available for the moment generating function, then a corresponding expression can be obtained for $\psi(t)$ by replacing e^t with t . For example, from formula (5.65), the probability generating function for the Poisson distribution is given by $\psi(t) = \exp(\lambda t - \lambda)$.

5.6.3. Some Limit Theorems

In Section 3.7 we defined convergence in probability of a sequence of random variables. In Section 4.5.1 convergence in distribution of the same sequence was introduced. In this section we introduce yet another type of convergence.

Definition 5.6.1. A sequence $\{X_n\}_{n=1}^{\infty}$ of random variables converges in quadratic mean to a random variable X if

$$\lim_{n \rightarrow \infty} E(X_n - X)^2 = 0.$$

This convergence is written symbolically as $X_n \xrightarrow{\text{q.m.}} X$. □

Convergence in quadratic mean implies convergence in probability. This follows directly from applying Markov's inequality: If $X_n \xrightarrow{\text{q.m.}} X$, then for any $\epsilon > 0$,

$$P(|X_n - X| \geq \epsilon) \leq \frac{1}{\epsilon^2} E(X_n - X)^2 \rightarrow 0$$

as $n \rightarrow \infty$. This shows that the sequence $\{X_n\}_{n=1}^{\infty}$ converges in probability to X .

5.6.3.1. The Weak Law of Large Numbers (Khinchine's Theorem)

Let $\{X_i\}_{i=1}^{\infty}$ be a sequence of independent and identically distributed random variables with a finite mean μ . Then $\bar{X}_n$ converges in probability to μ as $n \rightarrow \infty$, where $\bar{X}_n = (1/n)\sum_{i=1}^n X_i$ is the sample mean of a sample of size n .

Proof. See, for example, Lindgren (1976, Section 2.5.1) or Rao (1973, Section 2c.3). $\square$

Definition 5.6.2. A sequence $\{X_n\}_{n=1}^{\infty}$ of random variables converges strongly, or almost surely, to a random variable X , written symbolically as $X_n \xrightarrow{\text{a.s.}} X$, if for any $\epsilon > 0$,

$$\lim_{N \rightarrow \infty} P\left(\sup_{n \geq N} |X_n - X| > \epsilon\right) = 0. \quad \square$$

Theorem 5.6.2. Let $\{X_n\}_{n=1}^{\infty}$ be a sequence of random variables. Then we have the following:

1. If $X_n \xrightarrow{\text{a.s.}} c$, where c is constant, then X_n converges in probability to c .
2. If $X_n \xrightarrow{\text{q.m.}} c$, and the series $\sum_{n=1}^{\infty} E(X_n - c)^2$ converges, then $X_n \xrightarrow{\text{a.s.}} c$.

5.6.3.2. The Strong Law of Large Numbers (Kolmogorov's Theorem)

Let $\{X_n\}_{n=1}^{\infty}$ be a sequence of independent random variables such that $E(X_n) = \mu_n$ and $\text{Var}(X_n) = \sigma_n^2$, $n = 1, 2, \dots$. If the series $\sum_{n=1}^{\infty} \sigma_n^2/n^2$ converges, then $\bar{X}_n \xrightarrow{\text{a.s.}} \bar{\mu}_n$, where $\bar{\mu}_n = (1/n)\sum_{i=1}^n \mu_i$.

Proof. See Rao (1973, Section 2c.3). $\square$

5.6.3.3. The Continuity Theorem for Probability Generating Functions

See Feller (1968, page 280).

Suppose that for every $k \geq 1$, the sequence $\{p_k(n)\}_{n=0}^{\infty}$ represents a discrete probability distribution. Let $\psi_k(t) = \sum_{n=0}^{\infty} t^n p_k(n)$ be the corresponding probability generating function ($k = 1, 2, \dots$). In order for a limit

$$q_n = \lim_{k \rightarrow \infty} p_k(n)$$

to exist for every $n = 0, 1, \dots$, it is necessary and sufficient that the limit

$$\psi(t) = \lim_{k \rightarrow \infty} \psi_k(t)$$

exist for every t in the open interval $(0, 1)$. In this case,

$$\psi(t) = \sum_{n=0}^{\infty} t^n q_n.$$

This theorem implies that a sequence of discrete probability distributions converges if and only if the corresponding probability generating functions converge. It is important here to point out that the q_n 's may not form a discrete probability distribution (because they may not sum to 1). The function $\psi(t)$ may not therefore be a probability generating function.

5.6.4. Power Series and Logarithmic Series Distributions

The power series distribution, which was introduced by Kosambi (1949), represents a family of discrete distributions, such as the binomial, Poisson, and negative binomial. Its probability mass function is given by

$$p(n) = \frac{a_n \theta^n}{f(\theta)}, \quad n = 0, 1, 2, \dots,$$

where $a_n \geq 0$, $\theta > 0$, and $f(\theta)$ is the function

$$f(\theta) = \sum_{n=0}^{\infty} a_n \theta^n. \quad (5.71)$$

This function is defined provided that θ falls inside the interval of convergence of the series in (5.71).

For example, for the Poisson distribution, $\theta = \lambda$, where λ is the mean, $a_n = 1/n!$ for $n = 0, 1, 2, \dots$, and $f(\theta) = e^\lambda$. For the negative binomial, $\theta = 1 - p$ and $a_n = \binom{n+r-1}{n}$, $n = 0, 1, 2, \dots$, where n = number of failures, r = number of successes, and p = probability of success on each trial, and thus

$$f(\theta) = \sum_{n=0}^{\infty} \binom{n+r-1}{n} (1-p)^n = \frac{1}{p^r}.$$

A special case of the power series distribution is the logarithmic series distribution. It was first introduced by Fisher, Corbet, and Williams (1943) while studying abundance and diversity for insect trap data. The probability mass function for this distribution is

$$p(n) = -\frac{\theta^n}{n \log(1 - \theta)}, \quad n = 1, 2, \dots,$$

where $0 < \theta < 1$.

The logarithmic series distribution is useful in the analysis of various kinds of data. A description of some of its applications can be found, for example, in Johnson and Kotz (1969, Chapter 7).

5.6.5. Poisson Approximation to Power Series Distributions

See Pérez-Abreu (1991).

The Poisson distribution can provide an approximation to the distribution of the sum of random variables having power series distributions. This is based on the following theorem:

Theorem 5.6.3. For each $k \geq 1$, let $X_1, X_2, \dots, X_k$ be independent non-negative integer-valued random variables with a common power series distribution

$$p_k(n) = a_n \theta_k^n / f(\theta_k), \quad n = 0, 1, \dots,$$

where $a_n \geq 0$ ($n = 0, 1, \dots$) are independent of k and

$$f(\theta_k) = \sum_{n=0}^{\infty} a_n \theta_k^n, \quad \theta_k > 0.$$

Let $a_0 > 0$, $\lambda > 0$ be fixed and $S_k = \sum_{i=1}^k X_i$. If $k\theta_k \rightarrow \lambda$ as $k \rightarrow \infty$, then

$$\lim_{k \rightarrow \infty} P(S_k = n) = e^{-\lambda_0} \lambda_0^n / n!, \quad n = 0, 1, \dots,$$

where $\lambda_0 = \lambda a_1 / a_0$.

Proof. See Pérez-Abreu (1991, page 43). $\square$

By using this theorem we can obtain the well-known Poisson approximation to the binomial and the negative binomial distributions as shown below.

EXAMPLE 5.6.2 (The Binomial Distribution). For each $k \geq 1$, let $X_1, \dots, X_k$ be a sequence of independent Bernoulli random variables with success probability p_k . Let $S_k = \sum_{i=1}^k X_i$. Suppose that $kp_k \rightarrow \lambda > 0$ as $k \rightarrow \infty$. Then, for each $n = 0, 1, \dots$,

$$\begin{aligned} \lim_{k \rightarrow \infty} P(S_k = n) &= \lim_{k \rightarrow \infty} \binom{k}{n} p_k^n (1 - p_k)^{k-n} \\ &= e^{-\lambda} \lambda^n / n!. \end{aligned}$$

This follows from the fact that the Bernoulli distribution with success probability p_k is a power series distribution with $\theta_k = p_k / (1 - p_k)$ and $f(\theta_k) =$

$1 + \theta_k$. Since $a_0 = a_1 = 1$, and $k\theta_k \rightarrow \lambda$ as $k \rightarrow \infty$, we get from applying Theorem 5.6.3 that

$$\lim_{k \rightarrow \infty} P(S_k = n) = e^{-\lambda} \lambda^n / n!.$$

EXAMPLE 5.6.3 (The Negative Binomial Distribution). We recall that a random variable Y has the negative binomial distribution if it represents the number of failures n (in repeated trials) before the k th success ($k \geq 1$). Let p_k denote the probability of success on a single trial. Let $X_1, X_2, \dots, X_k$ be random variables defined as

X_1 = number of failures occurring before the 1st success,

X_2 = number of failures occurring between the 1st success
and the 2nd success,

$\vdots$

X_k = number of failures occurring between the $(k - 1)$ st
success and the k th success.

Such random variables have what is known as the geometric distribution. It is a special case of the negative binomial distribution with $k = 1$. The common probability distribution of the X_i 's is

$$P(X_i = n) = p_k (1 - p_k)^n, \quad n = 0, 1, \dots; \quad i = 1, 2, \dots, k.$$

This is a power series distribution with $a_n = 1$ ($n = 0, 1, \dots$), $\theta_k = 1 - p_k$, and

$$f(\theta_k) = \sum_{n=0}^{\infty} (1 - p_k)^n = \frac{1}{1 - (1 - p_k)} = \frac{1}{p_k}.$$

It is easy to see that $X_1, X_2, \dots, X_k$ are independent and that $Y = S_k = \sum_{i=1}^k X_i$.

Let us now assume that $k(1 - p_k) \rightarrow \lambda > 0$ as $k \rightarrow \infty$. Then from Theorem 5.6.3 we obtain the following result:

$$\lim_{k \rightarrow \infty} P(S_k = n) = e^{-\lambda} \lambda^n / n!, \quad n = 0, 1, \dots.$$

5.6.6. A Ridge Regression Application

Consider the linear model

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon},$$

where $\mathbf{y}$ is a vector of n response values, $\mathbf{X}$ is an $n \times p$ matrix of rank p , $\boldsymbol{\beta}$ is a vector of p unknown parameters, and $\boldsymbol{\epsilon}$ is a random error vector such that $E(\boldsymbol{\epsilon}) = \mathbf{0}$ and $\text{Var}(\boldsymbol{\epsilon}) = \sigma^2 \mathbf{I}_n$. All variables in this model are corrected for their means and scaled to unit length, so that $\mathbf{X}'\mathbf{X}$ and $\mathbf{X}'\mathbf{y}$ are in correlation form.

We recall from Section 2.4.2 that if the columns of $\mathbf{X}$ are multicollinear, then the least-squares estimator of $\boldsymbol{\beta}$, namely $\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$, is an unreliable estimator due to large variances associated with its elements. There are several methods that can be used to combat multicollinearity. A review of such methods can be found in Ofir and Khuri (1986). Ridge regression is one of the most popular of these methods. This method, which was developed by Hoerl and Kennard (1970a, b), is based on adding a positive constant k to the diagonal elements of $\mathbf{X}'\mathbf{X}$. This leads to a biased estimator $\boldsymbol{\beta}^*$ of $\boldsymbol{\beta}$ called the ridge regression estimator and is given by

$$\boldsymbol{\beta}^* = (\mathbf{X}'\mathbf{X} + k\mathbf{I}_n)^{-1}\mathbf{X}'\mathbf{y}.$$

The elements of $\boldsymbol{\beta}^*$ can have substantially smaller variances than the corresponding elements of $\hat{\boldsymbol{\beta}}$ (see, for example, Montgomery and Peck, 1982, Section 8.5.3).

Draper and Herzberg (1987) showed that the ridge regression residual sum of squares can be represented as a power series in k . More specifically, consider the vector of predicted responses,

$$\begin{aligned}\hat{\mathbf{y}}_k &= \mathbf{X}\boldsymbol{\beta}^* \\ &= \mathbf{X}(\mathbf{X}'\mathbf{X} + k\mathbf{I}_n)^{-1}\mathbf{X}'\mathbf{y},\end{aligned}\tag{5.72}$$

which is based on using $\boldsymbol{\beta}^*$. Formula (5.72) can be written as

$$\hat{\mathbf{y}}_k = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1} \left[\mathbf{I}_n + k(\mathbf{X}'\mathbf{X})^{-1} \right]^{-1} \mathbf{X}'\mathbf{y}.\tag{5.73}$$

From Theorem 5.5.2, if all the eigenvalues of $k(\mathbf{X}'\mathbf{X})^{-1}$ are less than one in absolute value, then

$$\left[\mathbf{I}_n + k(\mathbf{X}'\mathbf{X})^{-1} \right]^{-1} = \sum_{i=0}^{\infty} (-1)^i k^i (\mathbf{X}'\mathbf{X})^{-i}.\tag{5.74}$$

From (5.73) and (5.74) we get

$$\hat{\mathbf{y}}_k = (\mathbf{H}_1 - k\mathbf{H}_2 + k^2\mathbf{H}_3 - k^3\mathbf{H}_4 + \cdots)\mathbf{y},$$

where $\mathbf{H}_i = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-i}\mathbf{X}'$, $i \geq 1$. Thus the ridge regression residual sum of squares, which is the sum of squares of deviations of the elements of $\mathbf{y}$ from the corresponding elements of $\hat{\mathbf{y}}_k$, is

$$(\mathbf{y} - \hat{\mathbf{y}}_k)'(\mathbf{y} - \hat{\mathbf{y}}_k) = \mathbf{y}'\mathbf{Q}\mathbf{y},$$

where

$$\mathbf{Q} = (\mathbf{I}_n - \mathbf{H}_1 + k\mathbf{H}_2 - k^2\mathbf{H}_3 + k^3\mathbf{H}_4 - \cdots)^2.$$

It can be shown (see Exercise 5.32) that

$$\mathbf{y}'\mathbf{Q}\mathbf{y} = SS_E + \sum_{i=3}^{\infty} (i-2)(-k)^{i-1}S_i, \quad (5.75)$$

where SS_E is the usual least-squares residual sum of squares, which can be obtained when $k = 0$, that is,

$$SS_E = \mathbf{y}'[\mathbf{I}_n - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']\mathbf{y},$$

and $S_i = \mathbf{y}'\mathbf{H}_i\mathbf{y}$, $i \geq 3$. The terms to the right of (5.75), other than SS_E , are bias sums of squares induced by the presence of a nonzero k . Draper and Herzberg (1987) demonstrated by means of an example that the series in (5.75) may diverge or else converge very slowly, depending on the value of k .

FURTHER READING AND ANNOTATED BIBLIOGRAPHY

- Apostol, T. M. (1964). *Mathematical Analysis*. Addison-Wesley, Reading, Massachusetts. (Infinite series are discussed in Chap. 12.)
- Boyer, C. B. (1968). *A History of Mathematics*. Wiley, New York.
- Draper, N. R., and A. M. Herzberg (1987). "A ridge-regression sidelight." *Amer. Statist.*, **41**, 282–283.
- Draper, N. R., and H. Smith (1981). *Applied Regression Analysis*, 2nd ed. Wiley, New York. (Chap. 6 discusses ridge regression in addition to the various statistical procedures for selecting variables in a regression model.)
- Feller, W. (1968). *An Introduction to Probability Theory and Its Applications*, Vol. I, 3rd ed. Wiley, New York.
- Fisher, R. A., and E. A. Cornish (1960). "The percentile points of distribution having known cumulants." *Technometrics*, **2**, 209–225.
- Fisher, R. A., A. S. Corbet, and C. B. Williams (1943). "The relation between the number of species and the number of individuals in a random sample of an animal population." *J. Anim. Ecology*, **12**, 42–58.
- Fisz, M. (1963). *Probability Theory and Mathematical Statistics*, 3rd ed. Wiley, New York. (Chap. 5 deals almost exclusively with limit distributions for sums of independent random variables.)
- Fulks, W. (1978). *Advanced Calculus*, 3rd ed. Wiley, New York. (Chap. 2 discusses limits of sequences; Chap. 13 deals with infinite series of constant terms; Chap. 14 deals with sequences and series of functions; Chap. 15 provides a study of power series.)
- Gantmacher, F. R. (1959). *The Theory of Matrices*, Vol. I. Chelsea, New York.

- Graybill, F. A. (1983). *Matrices with Applications in Statistics*, 2nd ed. Wadsworth, Belmont, California. (Chap. 5 includes a section on sequences and series of matrices.)
- Hirschman, I. I., Jr. (1962). *Infinite Series*. Holt, Rinehart and Winston, New York. (This book is designed to be used in applied courses beyond the advanced calculus level. It emphasizes applications of the theory of infinite series.)
- Hoerl, A. E., and R. W. Kennard (1970a). "Ridge regression: Biased estimation for non-orthogonal problems." *Technometrics*, **12**, 55–67.
- Hoerl, A. E., and R. W. Kennard (1970b). "Ridge regression: Applications to non-orthogonal problems." *Technometrics*, **12**, 69–82; Correction. **12**, 723.
- Hogg, R. V., and A. T. Craig (1965). *Introduction to Mathematical Statistics*, 2nd ed. Macmillan, New York.
- Hyslop, J. M. (1954). *Infinite Series*, 5th ed. Oliver and Boyd, Edinburgh. (This book presents a concise treatment of the theory of infinite series. It provides the basic elements of this theory in a clear and easy-to-follow manner.)
- Johnson, N. L., and S. Kotz (1969). *Discrete Distributions*. Houghton Mifflin, Boston. (Chaps. 1 and 2 contain discussions concerning moments, cumulants, generating functions, and power series distributions.)
- Kendall, M., and A. Stuart (1977). *The Advanced Theory of Statistics*, Vol. 1, 4th ed. Macmillan, New York. (Moments, cumulants, and moment generating functions are discussed in Chap. 3.)
- Knopp, K. (1951). *Theory and Application of Infinite Series*. Blackie and Son, London. (This reference book provides a detailed and comprehensive study of the theory of infinite series. It contains many interesting examples.)
- Kosambi, D. D. (1949). "Characteristic properties of series distributions." *Proc. Nat. Inst. Sci. India*, **15**, 109–113.
- Lancaster, P. (1969). *Theory of Matrices*. Academic Press, New York. (Chap. 5 discusses functions of matrices in addition to sequences and series involving matrix terms.)
- Lindgren, B. W. (1976). *Statistical Theory*, 3rd ed. Macmillan, New York. (Chap. 2 contains a section on moments of a distribution and a proof of Markov's inequality.)
- Montgomery, D. C., and E. A. Peck (1982). *Introduction to Linear Regression Analysis*. Wiley, New York. (Chap. 8 discusses the effect of multicollinearity and the methods for dealing with it including ridge regression.)
- Nurcombe, J. R. (1979). "A sequence of convergence tests." *Amer. Math. Monthly*, **86**, 679–681.
- Ofir, C., and A. I. Khuri (1986). "Multicollinearity in marketing models: Diagnostics and remedial measures." *Internat. J. Res. Market.*, **3**, 181–205. (This is a review article that surveys the problem of multicollinearity in linear models and the various remedial measures for dealing with it.)
- Pérez-Abreu, V. (1991). "Poisson approximation to power series distributions." *Amer. Statist.*, **45**, 42–45.
- Pye, W. C., and P. G. Webster (1989). "A note on Raabe's test extended." *Math. Comput. Ed.*, **23**, 125–128.

- Rao, C. R. (1973). *Linear Statistical Inference and Its Applications*, 2nd ed. Wiley, New York. (Chap. 2 contains a section on limit theorems in statistics, including the weak and strong laws of large numbers.)
- Rudin, W. (1964). *Principles of Mathematical Analysis*, 2nd ed. McGraw-Hill, New York. (Sequences and series of scalar constants are discussed in Chap. 3; sequences and series of functions are studied in Chap. 7.)
- Thiele, T. N. (1903). *Theory of Observations*. Layton, London. Reprinted in *Ann. Math. Statist.*, **2**, 165–307 (1931).
- Wilks, S. S. (1962). *Mathematical Statistics*. Wiley, New York. (Chap. 4 discusses different types of convergence of random variables; Chap. 5 presents several results concerning the moments of a distribution.)
- Withers, C. S. (1984). “Asymptotic expansions for distributions and quantiles with power series cumulants.” *J. Roy. Statist. Soc. Ser. B*, **46**, 389–396.

EXERCISES

In Mathematics

- 5.1. Suppose that $\{a_n\}_{n=1}^{\infty}$ is a bounded sequence of positive terms.
- Define $b_n = \max\{a_1, a_2, \dots, a_n\}$, $n = 1, 2, \dots$. Show that the sequence $\{b_n\}_{n=1}^{\infty}$ converges, and identify its limit.
 - Suppose further that $a_n \rightarrow c$ as $n \rightarrow \infty$, where $c > 0$. Show that $c_n \rightarrow c$, where $\{c_n\}_{n=1}^{\infty}$ is the sequence of geometric means, $c_n = (\prod_{i=1}^n a_i)^{1/n}$.
- 5.2. Suppose that $\{a_n\}_{n=1}^{\infty}$ and $\{b_n\}_{n=1}^{\infty}$ are any two Cauchy sequences. Let $d_n = |a_n - b_n|$, $n = 1, 2, \dots$. Show that the sequence $\{d_n\}_{n=1}^{\infty}$ converges.
- 5.3. Prove Theorem 5.1.3.
- 5.4. Show that if $\{a_n\}_{n=1}^{\infty}$ is a bounded sequence, then the set E of all its subsequential limits is also bounded.
- 5.5. Suppose that $a_n \rightarrow c$ as $n \rightarrow \infty$ and that $\{a_i\}_{i=1}^{\infty}$ is a sequence of positive terms for which $\sum_{i=1}^n \alpha_i \rightarrow \infty$ as $n \rightarrow \infty$.
- Show that

$$\frac{\sum_{i=1}^n \alpha_i a_i}{\sum_{i=1}^n \alpha_i} \rightarrow c \quad \text{as } n \rightarrow \infty.$$

In particular, if $\alpha_i = 1$ for all i , then

$$\frac{1}{n} \sum_{i=1}^n a_i \rightarrow c \quad \text{as } n \rightarrow \infty.$$

- Show that the converse of the special case in (a) does not always

hold by giving a counterexample of a sequence $\{a_n\}_{n=1}^{\infty}$ that does not converge, yet $(\sum_{i=1}^n a_i)/n$ converges as $n \rightarrow \infty$.

5.6. Let $\{a_n\}_{n=1}^{\infty}$ be a sequence of positive terms such that

$$\frac{a_{n+1}}{a_n} \rightarrow b \quad \text{as } n \rightarrow \infty,$$

where $0 < b < 1$. Show that there exist constants c and r such that $0 < r < 1$ and $c > 0$ for which $a_n < cr^n$ for sufficiently large values of n .

5.7. Suppose that we have the sequence $\{a_n\}_{n=1}^{\infty}$, where $a_1 = 1$ and

$$a_{n+1} = \frac{a_n(3b + a_n^2)}{3a_n^2 + b}, \quad b > 0, \quad n = 1, 2, \dots$$

Show that the sequence converges, and find its limit.

5.8. Show that the sequence $\{a_n\}_{n=1}^{\infty}$ converges, and find its limit, where $a_1 = 1$ and

$$a_{n+1} = (2 + a_n)^{1/2}, \quad n = 1, 2, \dots$$

5.9. Let $\{a_n\}_{n=1}^{\infty}$ be a sequence and $s_n = \sum_{i=1}^n a_i$.

(a) Show that $\limsup_{n \rightarrow \infty} (s_n/n) \leq \limsup_{n \rightarrow \infty} a_n$.

(b) If s_n/n converges as $n \rightarrow \infty$, then show that $a_n/n \rightarrow 0$ as $n \rightarrow \infty$.

5.10. Show that the sequence $\{a_n\}_{n=1}^{\infty}$, where $a_n = \sum_{i=1}^n (1/i)$, is not a Cauchy sequence and is therefore divergent.

5.11. Suppose that the sequence $\{a_n\}_{n=1}^{\infty}$ satisfies the following condition: There is an r , $0 < r < 1$, such that

$$|a_{n+1} - a_n| < br^n, \quad n = 1, 2, \dots,$$

where b is a positive constant. Show that this sequence converges.

5.12. Show that if $a_n \geq 0$ for all n , then $\sum_{n=1}^{\infty} a_n$ converges if and only if $\{s_n\}_{n=1}^{\infty}$ is a bounded sequence, where s_n is the n th partial sum of the series.

5.13. Show that the series $\sum_{n=1}^{\infty} [1/(3n-1)(3n+2)]$ converges to $\frac{1}{6}$.

5.14. Show that the series $\sum_{n=1}^{\infty} (n^{1/n} - 1)^p$ is divergent for $p \leq 1$.

5.15. Let $\sum_{n=1}^{\infty} a_n$ be a divergent series of positive terms.

(a) If $\{a_n\}_{n=1}^{\infty}$ is a bounded sequence, then show that $\sum_{n=1}^{\infty} [a_n/(1 + a_n)]$ diverges.

(b) Show that (a) is true even if $\{a_n\}_{n=1}^{\infty}$ is not a bounded sequence.

5.16. Let $\sum_{n=1}^{\infty} a_n$ be a divergent series of positive terms. Show that

$$\frac{a_n}{s_n^2} \leq \frac{1}{s_{n-1}} - \frac{1}{s_n}, \quad n = 2, 3, \dots,$$

where s_n is the n th partial sum of the series; then deduce that $\sum_{n=1}^{\infty} (a_n/s_n^2)$ converges.

5.17. Let $\sum_{n=1}^{\infty} a_n$ be a convergent series of positive terms. Let $r_n = \sum_{i=n}^{\infty} a_i$. Show that for $m < n$,

$$\sum_{i=m}^n \frac{a_i}{r_i} > 1 - \frac{r_n}{r_m},$$

and deduce that $\sum_{n=1}^{\infty} (a_n/r_n)$ diverges.

5.18. Given the two series $\sum_{n=1}^{\infty} (1/n)$, $\sum_{n=1}^{\infty} (1/n^2)$. Show that

$$\lim_{n \rightarrow \infty} \left(\frac{1}{n} \right)^{1/n} = 1,$$

$$\lim_{n \rightarrow \infty} \left(\frac{1}{n^2} \right)^{1/n} = 1.$$

5.19. Test for convergence of the series $\sum_{n=1}^{\infty} a_n$, where

(a) $a_n = (n^{1/n} - 1)^n,$

(b) $a_n = \frac{\log(1+n)}{\log(1+e^{n^2})},$

(c) $a_n = \frac{1 \times 3 \times 5 \times \cdots \times (2n-1)}{2 \times 4 \times 6 \times \cdots \times 2n} \cdot \frac{1}{2n+1},$

(d) $a_n = \sqrt{n + \sqrt{n}} - \sqrt{n},$

(e) $a_n = \frac{(-1)^n 4^n}{n^n},$

(f) $a_n = \sin \left[\left(n + \frac{1}{n} \right) \pi \right].$

5.20. Determine the values of x for which each of the following series converges uniformly:

- (a)
$$\sum_{n=1}^{\infty} \frac{n+2}{3^n} x^{2n},$$
- (b)
$$\sum_{n=1}^{\infty} \frac{10^n}{n} x^n,$$
- (c)
$$\sum_{n=1}^{\infty} (n+1)^2 x^n,$$
- (d)
$$\sum_{n=1}^{\infty} \frac{\cos(nx)}{n(n^2+1)}.$$

5.21. Consider the series

$$\sum_{n=1}^{\infty} a_n = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n}.$$

Let $\sum_{n=1}^{\infty} b_n$ be a certain rearrangement of $\sum_{n=1}^{\infty} a_n$ given by

$$1 + \frac{1}{3} - \frac{1}{2} + \frac{1}{5} + \frac{1}{7} - \frac{1}{4} + \frac{1}{9} + \frac{1}{11} - \frac{1}{6} + \cdots,$$

where two positive terms are followed by one negative. Show that the sum of the original series is less than $\frac{10}{12}$, whereas that of the rearranged series (which is convergent) exceeds $\frac{11}{12}$.

5.22. Consider Cauchy's product of $\sum_{n=0}^{\infty} a_n$ with itself, where

$$a_n = \frac{(-1)^n}{\sqrt{n+1}}, \quad n = 0, 1, 2, \dots$$

Show that this product is divergent. [*Hint:* Show that the n th term of this product does not go to zero as $n \rightarrow \infty$.]

5.23. Consider the sequence of functions $\{f_n(x)\}_{n=1}^{\infty}$, where for $n = 1, 2, \dots$

$$f_n(x) = \frac{nx}{1+nx^2}, \quad x \geq 0.$$

Find the limit of this sequence, and determine whether or not the convergence is uniform on $[0, \infty)$.

5.24. Consider the series $\sum_{n=1}^{\infty} (1/n^x)$.

- (a) Show that this series converges uniformly on $[1 + \delta, \infty)$, where δ is any positive number. [Note: The function represented by this series is known as *Riemann's ζ -function* and is denoted by $\zeta(x)$.]
 (b) Is $\zeta(x)$ differentiable on $[1 + \delta, \infty)$? If so, give a series expansion for $\zeta'(x)$.

In Statistics

5.25. Prove formulas (5.56) and (5.57).

5.26. Find a series expansion for the moment generating function of the negative binomial distribution. For what values of t does this series converge uniformly? In this case, can formula (5.63) be applied to obtain an expression for the k th noncentral moment ($k = 1, 2, \dots$) of this distribution? Why or why not?

5.27. Find the first three cumulants of the negative binomial distribution.

5.28. Show that the moments μ'_n ($n = 1, 2, \dots$) of a random variable X determine the cumulative distribution functions of X uniquely if

$$\limsup_{n \rightarrow \infty} \frac{|\mu'_n|^{1/n}}{n} \text{ is finite.}$$

[Hint: Use the fact that $n! \sim \sqrt{2\pi} n^{n+1/2} e^{-n}$ as $n \rightarrow \infty$.]

5.29. Find the moment generating function of the logarithmic series distribution, and deduce that the mean and variance of this distribution are given by

$$\begin{aligned} \mu &= \alpha\theta/(1 - \theta), \\ \sigma^2 &= \mu \left(\frac{1}{1 - \theta} - \mu \right), \end{aligned}$$

where $\alpha = -1/\log(1 - \theta)$.

5.30. Let $\{X_n\}_{n=1}^{\infty}$ be a sequence of binomial random variables where the probability mass function of X_n ($n = 1, 2, \dots$) is given by

$$p_n(k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, 1, 2, \dots, n,$$

where $0 < p < 1$. Further, let the random variable Y_n be defined as

$$Y_n = \frac{X_n}{n} - p.$$

- (a) Show that $E(X_n) = np$ and $\text{Var}(X_n) = np(1 - p)$.
- (b) Apply Chebyshev's inequality to show that

$$P(|Y_n| \geq \epsilon) \leq \frac{p(1-p)}{n\epsilon^2},$$

where $\epsilon > 0$.

- (c) Deduce from (b) that Y_n converges in probability to zero. [*Note:* This result is known as *Bernoulli's law of large numbers*.]
- 5.31.** Let $X_1, X_2, \dots, X_n$ be a sequence of independent Bernoulli random variables with success probability p_n . Let $S_n = \sum_{i=1}^n X_i$. Suppose that $np_n \rightarrow \mu > 0$ as $n \rightarrow \infty$.
- (a) Give an expression for $\phi_n(t)$, the moment generating function of S_n .
 - (b) Show that

$$\lim_{n \rightarrow \infty} \phi_n(t) = \exp(\mu e^t - \mu),$$

which is the moment generating function of a Poisson distribution with mean μ .

- 5.32.** Prove formula (5.75).