CHAPTER 1

Taylor Approximation

Taylor's theorem plays a vital role in numerical analysis. The theorem enables us to approximate a given sufficiently smooth function by a polynomial called the *Taylor polynomial* and provides us with an expression for the error involved in the approximation called the *truncation error* or the *remainder term*.

In Section 1.1, we first define Taylor's polynomial for a given function at a given point and prove Taylor's theorem. In Section 1.2, we define the truncation error in approximating a sufficiently smooth function at a point by Taylor's polynomial and introduce the concept of estimating the truncation error.

1.1 Taylor's Theorem

Let f be a real-valued function defined on an interval I. We say $f \in C^n(I)$ if f is n-times continuously differentiable at every point in I. Also, we say $f \in C^\infty(I)$ if f is continuously differentiable of any order at every point in I.

The most important result used very frequently in numerical analysis, especially in error analysis of numerical methods, is the Taylor's expansion of a function in a neighborhood of a point $a \in \mathbb{R}$. In this section, we define the Taylor's polynomial and prove an important theorem called the *Taylor's theorem*. The idea of the proof of this theorem is similar to the one used in proving the mean value theorem, where we construct a function and apply Rolle's theorem several times to it.

Definition 1.1.1 [Taylor's Polynomial].

Let f be n-times continuously differentiable at a given point a. The ***Taylor's polynomial*** of degree n for the function f about the point a, denoted by T_n, is defined by

$$T_n(x) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k, \ x \in \mathbb{R}. \tag{1.1}$$

Theorem 1.1.2 [Taylor's Theorem].

Let f be $(n+1)$-times continuously differentiable function on an open interval containing the points a and x. Then there exists a number ξ between a and x such that

$$f(x) = T_n(x) + \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-a)^{n+1}, \tag{1.2}$$

where T_n is the Taylor's polynomial of degree n for f about the point a given by (1.1) and the second term on the right hand side is called the ***remainder term***.

Proof.

Let us assume $x > a$ and prove the theorem. The proof is similar if $x < a$.

Define $g(t)$ by

$$g(t) = f(t) - T_n(t) - A(t-a)^{n+1}$$

and choose A so that $g(x) = 0$, which gives

$$A = \frac{f(x) - T_n(x)}{(x-a)^{n+1}}.$$

Note that

$$g^{(k)}(a) = 0 \text{ for } k = 0, 1, \dots n.$$

Also, observe that the function g is continuous on $[a, x]$ and differentiable in (a, x).

Apply Rolle's theorem to g on $[a, x]$ (after verifying all the hypotheses of Rolle's theorem) to get

$$a < c_1 < x \text{ satisfying } g'(c_1) = 0.$$

Again apply Rolle's theorem to g' on $[a, c_1]$ to get

$$a < c_2 < c_1 \text{ satisfying } g''(c_2) = 0.$$

In turn apply Rolle's theorem to $g^{(2)}$, $g^{(3)}$, $\dots$, $g^{(n)}$ on intervals $[a, c_2]$, $[a, c_3]$, $\dots$, $[a, c_n]$, respectively.

At the last step, we get

$$a < c_{n+1} < c_n \text{ satisfying } g^{(n+1)}(c_{n+1}) = 0.$$

But

$$g^{(n+1)}(c_{n+1}) = f^{(n+1)}(c_{n+1}) - A(n+1)!,$$

which gives

$$A = \frac{f^{(n+1)}(c_{n+1})}{(n+1)!}.$$

Equating both values of A, we get

$$f(x) = T_n(x) + \frac{f^{(n+1)}(c_{n+1})}{(n+1)!}(x-a)^{n+1}.$$

This completes the proof.

Observe that the mean value theorem (Theorem A.3.7) is a particular case of the Taylor's theorem.

Remark 1.1.3.

The representation (1.2) is called the ***Taylor's formula*** for the function f about the point a.

The Taylor's theorem helps us to obtain an approximate value of a sufficiently smooth function in a small neighborhood of a given point a when the value of f and all its derivatives up to a sufficient order is known at the point a. For instance, if we know $f(a)$, $f'(a)$, $\cdots$, $f^{(n)}(a)$, and we seek an approximate value of $f(a+h)$ for some real number h, then the Taylor's theorem can be used to get

$$f(a+h) \approx f(a) + f'(a)h + \frac{f''(a)}{2!}h^2 + \cdots + \frac{f^{(n)}(a)}{n!}h^n.$$

Note here that we have not added the remainder term and therefore used the approximation symbol $\approx$. Observe that the remainder term

$$R_n(h) = \frac{f^{(n+1)}(\xi)}{(n+1)!}h^{n+1}$$

is not known since it involves the evaluation of $f^{(n+1)}$ at some unknown value ξ lying between a and $a+h$. Also, observe that as $h \to 0$, the remainder term approaches to zero, provided $f^{(n+1)}$ is bounded. This means that for smaller values of h, the Taylor's polynomial gives a good approximation of $f(a+h)$.

1.2 Truncation Error and its Estimate

The error involved in approximating $f(x)$ by the Taylor's polynomial $T_n(x)$ is given by $f(x) - T_n(x)$ and is called the ***Truncation error***. Using (1.2), we can write the truncation error as

$$f(x) - T_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-a)^{n+1},$$

for some ξ between x and a.

As remarked above, the truncation error cannot be obtained explicitly because it involves an unknown parameter ξ. However, we can estimate the truncation error under a smoothness condition on f.

Let f be an $(n+1)$-times continuously differentiable function defined on a closed and bounded interval $I := [\alpha, \beta]$ (*i.e.,* $f \in C^{n+1}(I)$). Then, there exists real number M_{n+1} such that

$$|f^{(n+1)}(\xi)| \leq M_{n+1}, \text{ for all } \xi \in I.$$

Therefore, for any fixed points $a, x \in I$, we can write

$$\left| \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-a)^{n+1} \right| \leq \frac{M_{n+1}}{(n+1)!}\left|x-a\right|^{n+1}.$$

We can further get an estimate of the remainder term that is independent of x as

$$\left| \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-a)^{n+1} \right| \leq \frac{M_{n+1}}{(n+1)!}|\beta-a|^{n+1},$$

which holds for all $x \in I$. Observe that the right hand side of the above estimate is a fixed number. We refer to such estimates as ***truncation error estimates*** or ***remainder estimates***.

Example 1.2.1.

A second degree polynomial approximation to

$$f(x) = \sqrt{x+1}, \quad x \in [-1, \infty)$$

using the Taylor's formula about $a = 0$ is given by

$$f(x) \approx 1 + \frac{x}{2} - \frac{x^2}{8},$$

where the remainder term is neglected and hence what we obtained here is only an approximate representation of f.

The truncation error is obtained using the remainder term in the formula (1.2) with $n = 2$ and is given by

$$f(x) - T_2(x) = \frac{x^3}{16(\sqrt{1+\xi}\,)^5},$$

for some point ξ between 0 and x.

Note that we cannot obtain a remainder estimate in the present example as f''' is not bounded in $[-1, \infty)$. However, for any $0 < \delta < 1$, if we restrict the domain of f to $[-\delta, \infty)$, then we can obtain the remainder estimate for a fixed $x \in [-\delta, \infty)$ as

$$|f(x) - T_2(x)| \leq \frac{|x|^3}{16(\sqrt{1-\delta}\,)^5}.$$

Further, if we restrict the domain of f to $[-\delta, b]$ for some real number $b > 0$, then we get the remainder estimate independent of x as

$$|f(x) - T_2(x)| \leq \frac{b^3}{16(\sqrt{1-\delta}\,)^5}.$$

1.3 Taylor's Series

Since the remainder term in (1.2) represents the error in Taylor's approximation of a given (sufficiently smooth) function, it is very important to write the remainder term in the form that gives the best possible information about the behaviour of the error. This is particularly important in numerical analysis where the interest is to understand how fast the error goes to zero as we improve the approximation by increasing the degree of the polynomial.

Definition 1.3.1 [Taylor's Series].

Let f be a C^∞-function in a neighborhood of a point a. The power series

$$\sum_{k=0}^{\infty} \frac{f^{(k)}(a)}{k!}(x-a)^k$$

is called the ***Taylor's series*** of f about the point a.

The question now is when this series converges and what is the limit of this series. These questions are answered in the following theorem[1] whose proof is omitted for this course.

Theorem 1.3.2.

Let $f \in C^\infty(I)$ and let $a \in I$. Assume that there exists an open neighborhood (interval) $N_a \subset I$ of the point a and there exists a constant M (may depend on a) such that

$$\left|f^{(k)}(x)\right| \le M^k,$$

for all $x \in N_a$ and $k = 0, 1, 2, \cdots$. Then for each $x \in I$, we have

$$f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(a)}{k!}(x-a)^k.$$

The following example illustrates the procedure for approximating a given function by its Taylor's polynomial and obtaining the corresponding truncation error.

Example 1.3.3.

As another example, let us approximate the function $f(x) = \cos(x)$ by a polynomial using Taylor's theorem about the point $a = 0$. First, let us take the Taylor's series expansion

$$\begin{aligned} f(x) &= \cos(0) - \sin(0)x - \frac{\cos(0)}{2!}x^2 + \frac{\sin(0)}{3!}x^3 + \cdots \\ &= \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!}x^{2k}. \end{aligned}$$

Now, we truncate this infinite series to get an approximate representation of $f(x)$ in

[1] Apostol, T. M., Mathematical Analysis, Narosa, 1974.

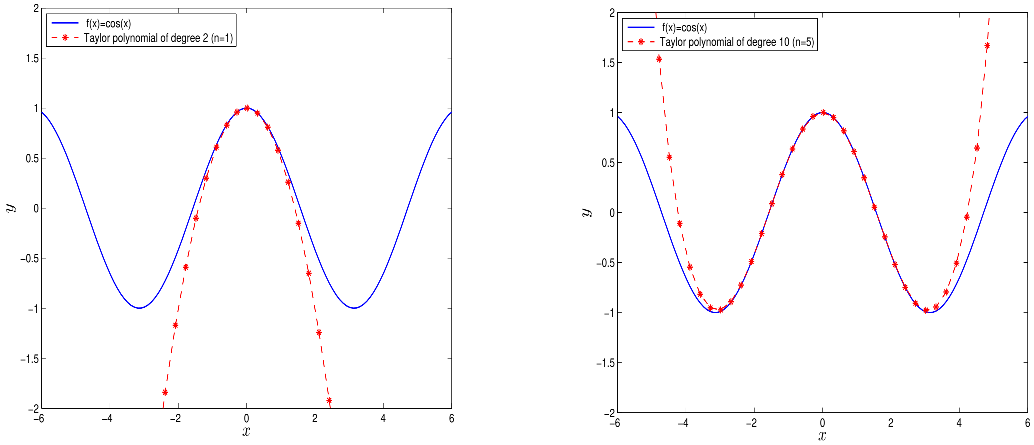

Figure 1.1: Comparison between the graph of $f(x) = \cos(x)$ and the Taylor polynomial of degree 2 and 10 about the point $a = 0$.

a sufficiently small neighborhood of $a = 0$ as

$$f(x) \approx \sum_{k=0}^{n} \frac{(-1)^k}{(2k)!} x^{2k},$$

which is the Taylor's polynomial of degree $2n$ for the function $f(x) = \cos(x)$ about the point $a = 0$. The truncated series is

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

Looking at the leading term of the truncated series, the truncation error can be taken as

$$f(x) - T_{2n}(x) = (-1)^{n+1} \frac{\cos(\xi)}{(2(n+1))!} x^{2(n+1)},$$

where ξ lies between 0 and x. It is important to observe in this particular example that for a given n, we get the Taylor polynomial of degree $2n$.

Figure 1.1 shows the comparison between the Taylor polynomial (red dot and dash line) of degree 2 ($n = 1$) and degree 10 ($n = 5$) for $f(x) = \cos(x)$ about $a = 0$ and the graph of $\cos(x)$ (blue solid line). We observe that for $n = 1$, Taylor polynomial gives a good approximation in a small neighborhood of $a = 0$. But sufficiently away from 0, this polynomial deviates significantly from the actual graph of $f(x) = \cos(x)$. Whereas, for $n = 5$, we get a good approximation in a relatively larger neighborhood of $a = 0$.

1.4 Exercises

1. Find the Taylor's polynomial of degree 2 for the function

$$f(x) = \sqrt{x+1}$$

   about the point $a = 1$. Also find the remainder.

2. Use Taylor's polynomial of degree 2 about $a = 0$ to evaluate approximately the value of the function $f(x) = e^x$ at $x = 0.5$. Obtain the remainder $R_2(0.5)$ in terms of the unknown ξ and show that ξ lie in the interval $(0, 0.5)$ by explicitly computing it.

3. Obtain Taylor polynomial of degree 3 and 5 for the function $f(x) = \sin(x)$ about the point $a = 0$. Give the remainder term in both the cases.

CHAPTER 2

Order of Convergence

In Section A.1, we defined convergent sequences and discussed some conditions under which a given sequence of real numbers converges. The definition and the discussed conditions never tell us how fast the sequence converges to the limit. Even if we know that a sequence of approximations converge to the exact one (limit), it is important in numerical analysis to know how fast the sequence of approximate values converge to the exact value. In this section, we introduce two important notations called *big Oh* and *little oh*, which are basic tools for the study of speed of convergence. We end this chapter by defining the notion of order of convergence.

2.1 Big Oh and Little oh Notations

The notions of big Oh and little oh are well understood through the following examples.

Example 2.1.1.

Consider the two sequences $\{n\}$ and $\{n^2\}$ both of which are unbounded and tend to infinity as $n \to \infty$. However we feel that the sequence $\{n\}$ grows 'slowly' compared to the sequence $\{n^2\}$.

Consider also the sequences $\{1/n\}$ and $\{1/n^2\}$ both of which decrease to zero as $n \to \infty$. However we feel that the sequence $\{1/n^2\}$ decreases more rapidly compared to the sequence $\{1/n\}$.

The above examples motivate us to develop tools that compare two sequences $\{a_n\}$ and $\{b_n\}$. Landau has introduced the concepts of Big Oh and Little oh for comparing two sequences that we will define below.

Definition 2.1.2 [Big Oh and Little oh].

Let $\{a_n\}$ and $\{b_n\}$ be sequences of real numbers. Then

1. the sequence $\{a_n\}$ is said to be ***Big Oh*** of $\{b_n\}$, and write $a_n = O(b_n)$, if there exists a real number C and a natural number N such that

$$|a_n| \leq C\,|b_n|, \quad \text{for all } n \geq N.$$

2. the sequence $\{a_n\}$ is said to be ***Little oh*** (sometimes said to be ***small oh***) of $\{b_n\}$, and write $a_n = o(b_n)$, if for every $\epsilon > 0$ there exists a natural number N such that

$$|a_n| \leq \epsilon\,|b_n|, \quad \text{for all } n \geq N.$$

Remark 2.1.3.

1. If $b_n \neq 0$ for every n, then we have $a_n = O(b_n)$ if and only if the sequence $\left\{\dfrac{a_n}{b_n}\right\}$ is bounded. That is, there exists a constant C such that

$$\left|\frac{a_n}{b_n}\right| \leq C.$$

2. If $b_n \neq 0$ for every n, then we have $a_n = o(b_n)$ if and only if the sequence $\left\{\dfrac{a_n}{b_n}\right\}$ converges to 0. That is,

$$\lim_{n\to\infty} \frac{a_n}{b_n} = 0.$$

3. For any pair of sequences $\{a_n\}$ and $\{b_n\}$ such that $a_n = o(b_n)$, it follows that $a_n = O(b_n)$.

   The converse is not true. Consider the sequences $a_n = n$ and $b_n = 2n + 3$, for which $a_n = O(b_n)$ holds but $a_n = o(b_n)$ does not hold.

4. Let $\{a_n\}$ and $\{b_n\}$ be two sequences that converge to 0. Then $a_n = O(b_n)$ means the sequence $\{a_n\}$ tends to 0 at least as fast as the sequence $\{b_n\}$; and $a_n = o(b_n)$ means the sequence $\{a_n\}$ tends to 0 faster than the sequence $\{b_n\}$.

5. If the sequence $\{a_n\}$ is bounded, then $a_n = O(b_n)$ where $b_n = 1$ for every n. In this case $a_n = O(1)$.

The Big Oh and Little oh notations may be adapted to functions as follows.

Definition 2.1.4 [Big Oh and Little oh for Functions].

Let $x_0 \in \mathbb{R}$. Let f and g be functions defined in an interval containing x_0. Then

1. the function f is said to be ***Big Oh*** of g as $x \to x_0$, and write $f(x) = O\big(g(x)\big)$, if there exists a real number C and a real number $\delta > 0$ such that
$$|f(x)| \le C\,|g(x)|, \quad \text{whenever } |x - x_0| \le \delta.$$

2. the function f is said to be ***Little Oh*** (also, ***Small oh***) of g as $x \to x_0$, and write $f(x) = o\big(g(x)\big)$, if for every $\epsilon > 0$ there exists a real number δ such that
$$|f(x)| \le \epsilon\,|g(x)|, \quad \text{whenever } |x - x_0| \le \delta.$$

Observe that the remarks similar to the ones discussed in Remark 2.1.3 hold in the case of functions as well.

Example 2.1.5.

The Taylor's formula for $f(x) = \cos(x)$ about the point $a = 0$ is
$$\cos(x) = \sum_{k=0}^{n} \frac{(-1)^k}{(2k)!} x^{2k} + (-1)^{n+1} \frac{\cos(\xi)}{(2(n+1))!} x^{2(n+1)}$$
where ξ lies between x and 0.

Let us denote the remainder term (truncation error) as (for a fixed n)
$$g(x) = (-1)^{n+1} \frac{\cos(\xi)}{(2(n+1))!} x^{2(n+1)}.$$

Clearly, $g(x) \to 0$ as $x \to 0$. The question now is

'How fast does $g(x) \to 0$ as $x \to 0$?'

The answer is

'As fast as $x^{2(n+1)} \to 0$ as $x \to 0$.'

That is,
$$g(x) = O\left(x^{2(n+1)}\right) \text{ as } x \to 0.$$

In this case, we may say that the sequence
$$\left\{ \sum_{k=0}^{n} \frac{(-1)^k}{(2k)!} x^{2k} \right\}$$
converges to $\cos(x)$ with order $2(n+1)$.

2.2 Order of Convergence

In the oh notations, we compared two sequences (or two functions) to understand how fast a sequence approaches the limit. We can also measure the rapidity of a convergent sequence $\{a_n\}$ in approaching the limit by looking at how the terms of the sequence get closer to the limit as n increases.

Definition 2.2.1 [Order of Convergence].

Let $\{a_n\}$ be a sequence such that $\lim_{n\to\infty} a_n = a$.

1. We say that the order of convergence is atleast ***linear*** if there exists a constant $c < 1$ and a natural number N such that
$$|a_{n+1} - a| \leq c\,|a_n - a| \quad \text{for all } n \geq N.$$
2. We say that the order of convergence is atleast ***superlinear*** if there exists a sequence $\{\epsilon_n\}$ that converges to 0, and a natural number N such that
$$|a_{n+1} - a| \leq \epsilon_n\,|a_n - a| \quad \text{for all } n \geq N.$$
3. We say that the order of convergence is atleast ***quadratic*** if there exists a constant C (not necessarily less than 1), and a natural number N such that
$$|a_{n+1} - a| \leq C\,|a_n - a|^2 \quad \text{for all } n \geq N.$$
4. Let $\alpha \in \mathbb{R}_+$. We say that the order of convergence is atleast α if there exists a constant C (not necessarily less than 1), and a natural number N such that
$$|a_{n+1} - a| \leq C\,|a_n - a|^\alpha \quad \text{for all } n \geq N.$$

Remark 2.2.2.

Another equivalent way to define the order of convergence of a convergent sequence is as follows:

We say that a sequence $\{a_n\}$ converges to a limit a with order of convergence $\alpha > 0$ if there exists a constant $\lambda > 0$ such that
$$\lim_{n\to\infty} \frac{|a_{n+1} - a|}{|a_n - a|^\alpha} = \lambda.$$

The number λ is called the *asymptotic error constant*.

2.3 Exercises

1. Prove or disprove:
   i) $2n^2 + 3n + 4 = o(n)$ as $n \to \infty$.
   ii) $\frac{n+1}{n^2} = o(\frac{1}{n})$ as $n \to \infty$.
   iii) $\frac{n+1}{n^2} = O(\frac{1}{n})$ as $n \to \infty$.
   iv) $\frac{n+1}{\sqrt{n}} = o(1)$ as $n \to \infty$.
   v) $\frac{1}{\ln n} = o(\frac{1}{n})$ as $n \to \infty$.
   vi) $\frac{1}{n \ln n} = o(\frac{1}{n})$ as $n \to \infty$.
   vii) $\frac{e^n}{n^5} = O(\frac{1}{n})$ as $n \to \infty$.

2. Prove or disprove:
   i) $e^x - 1 = O(x^2)$ as $x \to 0$.
   ii) $x^{-2} = O(\cot x)$ as $x \to 0$.
   iii) $\cot x = o(x^{-1})$ as $x \to 0$.
   iv) For $r > 0$, $x^r = O(e^x)$ as $x \to \infty$.
   v) For $r > 0$, $\ln x = O(x^r)$ as $x \to \infty$.

APPENDIX A

Mathematical Preliminaries

This chapter reviews some of the concepts and results from calculus that are frequently used in this course. We recall important definitions and theorems, and outline proofs of certain theorems. The readers are assumed to be familiar with a first course in calculus.

In Section A.1, we introduce sequences of real numbers, recall the definition of convergence of a sequence, and state the sandwich theorem. Further, we discuss the concept of limit and continuity of a real function in Section A.2. We state the sandwich theorem for continuous functions and the intermediate value theorem. These theorems play vital role in convergence analysis and finding initial guesses to some iterative methods for solving nonlinear equations. In Section A.3 we define the notion of derivative of a function, and prove Rolle's theorem and mean-value theorem for derivatives. The mean-value theorem for integration is discussed in Section A.4. These two theorems are crucially used in devising methods for numerical integration and differentiation.

A.1 Sequences of Real Numbers

Let $a, b \in \mathbb{R}$ be such that $a < b$. We use the notations $[a, b]$ and (a, b) for the closed and the open intervals, respectively, and are defined by

$$[a, b] = \{\, x \in \mathbb{R} \,:\, a \le x \le b \,\}$$

and

$$(a, b) = \{\, x \in \mathbb{R} \,:\, a < x < b \,\}.$$

Definition A.1.1 [Sequence].

A ***sequence*** of real numbers is an ordered list of real numbers

$$a_1, a_2, \ldots, a_n, a_{n+1}, \ldots.$$

In other words, a *sequence* is a function that associates the real number a_n for each natural number n. The notation $\{a_n\}$ is used to denote the sequence, *i.e.*

$$\{a_n\} := a_1, a_2, \ldots, a_n, a_{n+1}, \ldots.$$

Note

We may also use the notation $\{a_n\}_{n=1}^{\infty}$ to denote a sequence. This notation is particularly useful when we define a sequence where the index starts from some other integer, say, with $n = 0$ instead of $n = 1$. If we simply use the notation $\{a_n\}$, it means $n = 1, 2, \ldots$.

The concept of convergence of a sequence plays an important role in the numerical analysis. For instance, an iterative procedure is often constructed to approximate a solution x of an equation. This procedure produces a sequence of approximations, which is expected to convergence to the exact solution x.

Definition A.1.2 [Convergence of a Sequence].

Let $\{a_n\}$ be a sequence of real numbers and let a be a real number. The sequence $\{a_n\}$ is said to ***converge*** to a, written as

$$\lim_{n\to\infty} a_n = a \quad (\text{or } a_n \to a \text{ as } n \to \infty),$$

if for every $\epsilon > 0$, there exists a natural number N such that

$$|a_n - a| < \epsilon \quad \text{whenever} \quad n \geq N.$$

The real number a is called the ***limit*** of the sequence $\{a_n\}$.

The following result is very useful in computing the limit of a sequence sandwiched between two sequences having a common limit.

Theorem A.1.3 [Sandwich Theorem].

Let $\{a_n\}, \{b_n\}, \{c_n\}$ be sequences of real numbers such that

1. there exists an $n_0 \in \mathbb{N}$ such that for every $n \geq n_0$, the sequences satisfy the inequalities $a_n \leq b_n \leq c_n$ and
2. $\lim_{n\to\infty} a_n = \lim_{n\to\infty} c_n = a$, for some real number a.

Then the sequence $\{b_n\}$ also converges and $\lim_{n\to\infty} b_n = a$.

Definition A.1.4 [Bounded Sequence].

A sequence $\{a_n\}$ is said to be a ***bounded sequence*** if there exists a real number M such that

$$|a_n| \leq M \quad \text{for every} \quad n \in \mathbb{N}.$$

Note that a bounded sequence, in general, need not converge, for instance $\{(-1)^n\}$. However, it is always possible to find a convergent subsequence.

Theorem A.1.5 [Bolzano-Weierstrass theorem].

Every bounded sequence $\{a_n\}$ has a convergent subsequence $\{a_{n_k}\}$.

Definition A.1.6 [Monotonic Sequences].

A sequence $\{a_n\}$ of real numbers is said to be

1. an ***increasing sequence*** if $a_n \leq a_{n+1}$, for every $n \in \mathbb{N}$.
2. a ***strictly increasing sequence*** if $a_n < a_{n+1}$, for every $n \in \mathbb{N}$.
3. a ***decreasing sequence*** if $a_n \geq a_{n+1}$, for every $n \in \mathbb{N}$.
4. a ***strictly decreasing sequence*** if $a_n > a_{n+1}$, for every $n \in \mathbb{N}$.

A sequence $\{a_n\}$ is said to be a (strictly) ***monotonic sequence*** if it is either (strictly) increasing or (strictly) decreasing.

The following theorem shows the advantage of working with monotonic sequences.

Theorem A.1.7.

Bounded monotonic sequences always converge.

The following result gives the algebra of limits of sequences.

Theorem A.1.8.

Let $\{a_n\}$ and $\{b_n\}$ be two sequences. Assume that $\lim\limits_{n\to\infty} a_n$ and $\lim\limits_{n\to\infty} b_n$ exist. Then

1. $\lim\limits_{n\to\infty} (a_n + b_n) = \lim\limits_{n\to\infty} a_n + \lim\limits_{n\to\infty} b_n$.
2. $\lim\limits_{n\to\infty} c\, a_n = c \lim\limits_{n\to\infty} a_n$, for any number c.
3. $\lim\limits_{n\to\infty} a_n b_n = \lim\limits_{n\to\infty} a_n \lim\limits_{n\to\infty} b_n$.
4. $\lim\limits_{n\to\infty} \dfrac{1}{a_n} = \dfrac{1}{\lim\limits_{n\to\infty} a_n}$, provided $\lim\limits_{n\to\infty} a_n \neq 0$.

A.2 Limits and Continuity

In the previous section, we introduced the concept of limit for a sequences of real numbers. We now define the "limit" in the context of functions.

Definition A.2.1 [Limit of a Function].

1. Let f be a function defined on the left side (or both sides) of a, except possibly at a itself. Then, we say "the ***left-hand limit*** of $f(x)$ as x approaches a, equals l" and denote

$$\lim_{x\to a-} f(x) = l,$$

if we can make the values of $f(x)$ arbitrarily close to l (as close to l as we like) by taking x to be sufficiently close to a and x less than a.

2. Let f be a function defined on the right side (or both sides) of a, except possibly at a itself. Then, we say "the ***right-hand limit*** of $f(x)$ as x approaches a, equals r" and denote

$$\lim_{x\to a+} f(x) = r,$$

if we can make the values of $f(x)$ arbitrarily close to r (as close to r as we like) by taking x to be sufficiently close to a and x greater than a.

3. Let f be a function defined on both sides of a, except possibly at a itself. Then,

we say“the ***limit*** of $f(x)$ as x approaches a, equals L” and denote

$$\lim_{x\to a} f(x) = L,$$

if we can make the values of $f(x)$ arbitrarily close to L (as close to L as we like) by taking x to be sufficiently close to a (on either side of a) but not equal to a.

Remark A.2.2.

Note that in each of the above definitions the value of the function f at the point a does not play any role. In fact, the function f need not be defined at a.

In the previous section, we have seen some limit laws in the context of sequences. Similar limit laws also hold for limits of functions. We have the following result, often referred to as “the limit laws” or as “algebra of limits”.

Theorem A.2.3.

Let f, g be two functions defined on both sides of a, except possibly at a itself. Assume that $\lim\limits_{x\to a} f(x)$ and $\lim\limits_{x\to a} g(x)$ exist. Then

1. $\lim\limits_{x\to a} (f(x) + g(x)) = \lim\limits_{x\to a} f(x) + \lim\limits_{x\to a} g(x)$.
2. $\lim\limits_{x\to a} c\, f(x) = c \lim\limits_{x\to a} f(x)$, for any number c.
3. $\lim\limits_{x\to a} f(x)g(x) = \lim\limits_{x\to a} f(x) \lim\limits_{x\to a} g(x)$.
4. $\lim\limits_{x\to a} \dfrac{1}{g(x)} = \dfrac{1}{\lim\limits_{x\to a} g(x)}$, provided $\lim\limits_{x\to a} g(x) \neq 0$.

Remark A.2.4.

Polynomials, rational functions, all trigonometric functions wherever they are defined, have property called direct substitution property:

$$\lim_{x\to a} f(x) = f(a).$$

The following theorem is often used to estimate limits of functions.

Theorem A.2.5.

If $f(x) \leq g(x)$ when x is in an interval containing a (except possibly at a) and the limits of f and g both exist as x approaches a, then

$$\lim_{x\to a} f(x) \leq \lim_{x\to a} g(x).$$

The following theorem is often used to compute limits of functions.

Theorem A.2.6 [Sandwich Theorem].

Let f, g, and h be given functions such that

1. $f(x) \leq g(x) \leq h(x)$ when x is in an interval containing a (except possibly at a) and
2. $\lim\limits_{x\to a} f(x) = \lim\limits_{x\to a} h(x) = L$,

then

$$\lim_{x\to a} g(x) = L.$$

We will now give a rigorous definition of the limit of a function. Similar definitions can be written down for left-hand and right-hand limits of functions.

Definition A.2.7.

Let f be a function defined on some open interval that contains a, except possibly at a itself. Then we say that the ***limit*** of $f(x)$ as x approaches a is L and we write

$$\lim_{x\to a} f(x) = L.$$

if for every $\epsilon > 0$ there is a number $\delta > 0$ such that

$$|f(x) - L| < \epsilon \ \text{ whenever } \ 0 < |x - a| < \delta.$$

Definition A.2.8 [Continuity].

A function f is

1. ***continuous from the right*** at a if

$$\lim_{x\to a+} f(x) = f(a).$$

2. ***continuous from the left*** at a if
$$\lim_{x \to a-} f(x) = f(a).$$

3. ***continuous*** at a if
$$\lim_{x \to a} f(x) = f(a).$$

A function f is said to be ***continuous on an open interval*** if f is continuous at every point in the interval. If f is defined on a closed interval $[a, b]$, then f is said to be continuous at a if f is continuous from the right at a and similarly, f is said to be continuous at b if f is continuous from left at b.

Remark A.2.9.

Note that the definition for continuity of a function f at a, means the following three conditions are satisfied:

1. The function f must be defined at a. i.e., a is in the domain of f,
2. $\lim_{x \to a} f(x)$ exists, and
3. $\lim_{x \to a} f(x) = f(a)$.

Equivalently, for any given $\epsilon > 0$, there exists a $\delta > 0$ such that
$$|f(x) - f(a)| < \epsilon \text{ whenever } |x - a| < \delta.$$

Theorem A.2.10.

If f and g are continuous at a, then the functions $f + g$, $f - g$, cg (c is a constant), fg, f/g (provided $g(a) \neq 0$), $f \circ g$ (composition of f and g, whenever it makes sense) are all continuous.

Thus polynomials, rational functions, trigonometric functions are all continuous on their respective domains.

Theorem A.2.11 [Intermediate Value Theorem].

Suppose that f is continuous on the closed interval $[a, b]$ and let N be any number between $f(a)$ and $f(b)$, where $f(a) \neq f(b)$. Then there exists a point $c \in (a, b)$ such that
$$f(c) = N.$$

A.3 Differentiation

We next give the basic definition of a differentiable function.

Definition A.3.1 [Derivative].

The ***derivative*** of a function f at a point a, denoted by $f'(a)$, is defined as

$$f'(a) = \lim_{h\to 0} \frac{f(a+h) - f(a)}{h}, \tag{1.1}$$

if this limit exists. We say f is *differentiable* at a. A function f is said to be ***differentiable*** on (c, d) if f is differentiable at every point in (c, d).

There are alternate ways of defining the derivative of a function, which leads to different difference formulae.

Remark A.3.2.

The derivative of a function f at a point $x = a$ can also be given by

$$f'(a) = \lim_{h\to 0} \frac{f(a) - f(a-h)}{h}, \tag{1.2}$$

and

$$f'(a) = \lim_{h\to 0} \frac{f(a+h) - f(a-h)}{2h}, \tag{1.3}$$

provided the limits exist.

If we write $x = a + h$, then $h = x - a$ and $h \to 0$ if and only if $x \to a$. Thus, formula (1.1) can equivalently be written as

$$f'(a) = \lim_{x\to a} \frac{f(x) - f(a)}{x - a}.$$

Remark A.3.3 [Interpretation].

Take the graph of f, draw the line joining the points $(a, f(a))$ and $(x, f(x))$. Take its slope and take the limit of these slopes as $x \to a$. Then the point $(x, f(x))$ tends to $(a, f(a))$. The limit is nothing but the slope of the tangent line at $(a, f(a))$ to the curve $y = f(x)$. This geometric interpretation will be very useful in describing the Newton-Raphson method in the context of solving nonlinear equations.

Theorem A.3.4.

If f is differentiable at a, then f is continuous at a.

Proof.

Using the identities

$$f(x) = \frac{f(x) - f(a)}{x - a}(x - a) + f(a)$$

and taking limit as $x \to a$ yields the desired result.

The converse of Theorem A.3.4 is not true. For, the function $f(x) = |x|$ is continuous at $x = 0$ but is not differentiable there.

Theorem A.3.5.

Suppose f is differentiable at a. Then there exists a function ϕ such that

$$f(x) = f(a) + (x - a)f'(a) + (x - a)\phi(x),$$

and $\lim_{x\to a} \phi(x) = 0$.

Proof.

Define ϕ by

$$\phi(x) = \frac{f(x) - f(a)}{x - a} - f'(a).$$

Since f is differentiable at a, the result follows on taking limits on both sides of the last equation as $x \to a$.

Theorem A.3.6 [Rolle's Theorem].

Let f be a function that satisfies the following three hypotheses:

1. f is continuous on the closed interval $[a, b]$.
2. f is differentiable on the open interval (a, b).
3. $f(a) = f(b)$.

Then there is a number c in the open interval (a, b) such that $f'(c) = 0$.

Proof.

If f is a constant function *i.e.,* $f(x) = f(a)$ for every $x \in [a, b]$, clearly such a c exists. If f is not a constant, then at least one of the following holds.

Case 1: The graph of f goes above the line $y = f(a)$ *i.e.,* $f(x) > f(a)$ for some $x \in (a, b)$.
Case 2: The graph of f goes below the line $y = f(a)$ *i.e.,* $f(x) < f(a)$ for some $x \in (a, b)$.

In case (1), *i.e.,* if the graph of f goes above the line $y = f(a)$, then the global maximum cannot be at a or b. Therefore, it must lie in the open interval (a, b). Denote that point by c. That is, global maximum on $[a, b]$ is actually a local maximum, and hence $f'(c) = 0$. A similar argument can be given in case (2), to show the existence of local minimum in (a, b). This completes the proof of Rolle's theorem.

The mean value theorem plays an important role in obtaining error estimates for certain numerical methods.

Theorem A.3.7 [Mean Value Theorem].

Let f be a function that satisfies the following hypotheses:

1. f is continuous on the closed interval $[a, b]$.
2. f is differentiable on the open interval (a, b).

Then there is a number c in the open interval (a, b) such that

$$f'(c) = \frac{f(b) - f(a)}{b - a}.$$

or, equivalently,

$$f(b) - f(a) = f'(c)(b - a).$$

Proof.

Define ϕ on $[a, b]$ by

$$\phi(x) = f(x) - f(a) - \frac{f(b) - f(a)}{b - a}(x - a).$$

By Rolle's theorem there exists a $c \in (a, b)$ such that $\phi'(c) = 0$ and hence the proof is complete.

A.4 Integration

In Theorem A.3.7, we have discussed the mean value property for the derivative of a function. We now discuss the mean value theorems for integration.

Theorem A.4.1 [Mean Value Theorem for Integrals].

If f is continuous on $[a, b]$, then there exists a number c in $[a, b]$ such that

$$\int_a^b f(x)\,dx = f(c)(b-a).$$

Proof.

Let m and M be minimum and maximum values of f in the interval $[a, b]$, respectively. Then,

$$m(b-a) \leq \int_a^b f(x)\,dx \leq M(b-a).$$

Since f is continuous, the result follows from the intermediate value theorem.

Recall the average value of a function f on the interval $[a, b]$ is defined by

$$\frac{1}{b-a}\int_a^b f(x)\,dx.$$

Observe that the first mean value theorem for integrals asserts that the average of an integrable function f on an interval $[a, b]$ belongs to the range of the function f.

Remark A.4.2 [Interpretation].

Let f be a function on $[a, b]$ with $f > 0$. Draw the graph of f and find the area under the graph lying between the ordinates $x = a$ and $x = b$. Also, look at a rectangle with base as the interval $[a, b]$ with height $f(c)$ and compute its area. Both values are the same.

The Theorem A.4.1 is often referred to as the *first mean value theorem for integrals*. We now state the second mean value theorem for integrals, which is a general form of Theorem A.4.1.

Theorem A.4.3 [Second Mean Value Theorem for Integrals].

Let f and g be continuous on $[a, b]$, and let $g(x) \geq 0$ for all $x \in \mathbb{R}$. Then there exists a number $c \in [a, b]$ such that

$$\int_a^b f(x)g(x)\,dx = f(c)\int_a^b g(x)\,dx.$$

The proof of this theorem is left as an exercise.

A.5 Exercises

Sequences of Real Numbers

1. Let L be a real number and let $\{a_n\}$ be a sequence of real numbers. If there exists a positive integer N such that

$$|a_n - L| \leq \mu |a_{n-1} - L|,$$

   for all $n \geq N$ and for some fixed $\mu \in (0, 1)$, then show that $a_n \to L$ as $n \to \infty$.

2. Consider the sequences $\{a_n\}$ and $\{b_n\}$, where

$$a_n = \frac{1}{n}, \quad b_n = \frac{1}{n^2}, \quad n = 1, 2, \cdots.$$

   Clearly, both the sequences converge to zero. For the given $\epsilon = 10^{-2}$, obtain the smallest positive integers N_a and N_b such that

$$|a_n| < \epsilon \text{ whenever } n \geq N_a, \text{ and } |b_n| < \epsilon \text{ whenever } n \geq N_b.$$

   For any $\epsilon > 0$, show that $N_a > N_b$.

3. Let $\{x_n\}$ and $\{y_n\}$ be two sequences such that $x_n, y_n \in [a, b]$ and $x_n < y_n$ for each $n = 1, 2, \cdots$. If $x_n \to b$ as $n \to \infty$, then show that the sequence $\{y_n\}$ converges. Find the limit of the sequence $\{y_n\}$.

4. Let $I_n = \left[\dfrac{n-2}{2n}, \dfrac{n+2}{2n}\right]$, $n = 1, 2, \cdots$ and $\{a_n\}$ be a sequence with a_n is chosen arbitrarily in I_n for each $n = 1, 2, \cdots$. Show that $a_n \to \dfrac{1}{2}$ as $n \to \infty$.

Limits and Continuity

5. Let f be a real-valued function such that $f(x) \geq \sin(x)$ for all $x \in \mathbb{R}$. If $\lim_{x\to 0} f(x) = L$ exists, then show that $L \geq 0$.

6. Let P and Q be polynomials. Find
$$\lim_{x\to\infty} \frac{P(x)}{Q(x)} \quad \text{and} \quad \lim_{x\to 0} \frac{P(x)}{Q(x)}$$
in each of the following cases.
   i) The degree of P is less than the degree of Q.
   ii) The degree of P is greater than the degree of Q.
   iii) The agree of P is equal to the degree of Q.

7. Show that the equation $\sin x + x^2 = 1$ has at least one solution in the interval $[0, 1]$.

8. Let $f(x)$ be continuous on $[a, b]$, let $x_1, \cdots, x_n$ be points in $[a, b]$, and let $g_1, \cdots$, g_n be real numbers having same sign. Show that
$$\sum_{i=1}^{n} f(x_i) g_i = f(\xi) \sum_{i=1}^{n} g_i, \quad \text{for some } \xi \in [a, b].$$

9. Let $f : [0, 1] \to [0, 1]$ be a continuous function. Prove that the equation $f(x) = x$ has at least one solution lying in the interval $[0, 1]$ (Note: A solution of this equation is called a *fixed point* of the function f).

10. Show that the equation $f(x) = x$, where
$$f(x) = \sin\left(\frac{\pi x + 1}{2}\right), \quad x \in [-1, 1]$$
has at least one solution in $[-1, 1]$.

Differentiation

11. Let $c \in (a, b)$ and $f : (a, b) \to \mathbb{R}$ be differentiable at c. If c is a local extremum (maximum or minimum) of f, then show that $f'(c) = 0$.

12. Suppose f is differentiable in an open interval (a, b). Prove the following statements
(a) If $f'(x) \geq 0$ for all $x \in (a, b)$, then f is non-decreasing.
(b) If $f'(x) = 0$ for all $x \in (a, b)$, then f is constant.
(c) If $f'(x) \leq 0$ for all $x \in (a, b)$, then f is non-increasing.

13. Let $f : [a, b] \to \mathbb{R}$ be given by $f(x) = x^2$. Find a point c specified by the mean-value theorem for derivatives. Verify that this point lies in the interval (a, b).

Integration

14. Let $g : [0, 1] \to \mathbb{R}$ be a continuous function. Show that there exists a $c \in (0, 1)$ such that

$$\int\limits_0^1 x^2(1-x)^2 g(x)dx = \frac{1}{30} g(\xi).$$

15. If n is a positive integer, show that

$$\int\limits_{\sqrt{n\pi}}^{\sqrt{(n+1)\pi}} \sin(t^2)\, dt = \frac{(-1)^n}{c},$$

where $\sqrt{n\pi} \leq c \leq \sqrt{(n+1)\pi}$.