Arindama Singh

MA1102 Classnotes
Series & Matrices

Contents

Syllabus

Series: Sequences of real numbers, Series, ratio and root test, improper integral, integral test, alternating series, absolute and conditional convergence, power series, radius and interval of convergence of power series, term by term differentiation and integration of power series, Taylor's formulas, Taylor series, periodic functions and Fourier series, convergence of Fourier series, functions of any period, even and odd functions, half-range expansions.

Matrices: Matrix operations, special types of matrices, matrices as linear transformations, linear independence, basis and dimension, rank of a matrix, nullity of a matrix, elementary operations, inverse of a matrix, orthogonalization, determinant, existence- uniqueness of solutions of a linear system, Gaussian elimination, Gauss-Jordan elimination, Eigenvalues, eigenvectors, eigenvalues of special types of matrices, similarity of matrices, basis of eigenvectors, diagonalization.

Texts:

1. G.B. Thomas Jr., M.D. Weir and J.R. Hass, Thomas Calculus, Pearson Education, 2009.
2. E. Kreyszig, Advanced Engineering Mathematics, 10th Ed., John Willey & Sons, 2010.

References:

1. J. Hefferon, Linear Algebra, http://joshua.smcvt.edu/linearalgebra, 2014.
2. S. Lang, Introduction to Linear Algebra, 2nd Ed., Springer-Verlag, 1986.
3. M.T. Nair, Calculus of One Variable, Ane Books, 2014.
4. N. Piskunov, Differential and Integral Calculus Vol. 1-2, Mir Publishers, 1974.
5. A. Singh, Introduction to Matrix Theory, Ane Books, 2018.
6. G. Strang, Linear Algebra and its Applications, Cengage Learning, 4 th Ed., 2006.

1

Series of Numbers

1.1 Preliminaries

We use the following notation:

$\varnothing$ = the empty set.

$\mathbb{N} = \{1, 2, 3, \ldots\}$, the set of natural numbers.

$\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$, the set of integers.

$\mathbb{Q} = \left\{\frac{p}{q} : p \in \mathbb{Z},\ q \in \mathbb{N}\right\}$, the set of rational numbers.

$\mathbb{R}$ = the set of real numbers.

$\mathbb{R}_+$ = the set of all positive real numbers.

As we know, $\mathbb{N} \subsetneq \mathbb{Z} \subsetneq \mathbb{Q} \subsetneq \mathbb{R}$. The numbers in $\mathbb{R} \setminus \mathbb{Q}$ is the set of irrational numbers. Examples are $\sqrt{2}$, $3.10110111011110\cdots$ etc.

Along with the usual laws of $+$, $\cdot$, $<$, $\mathbb{R}$ satisfies the **completeness property**:

> Every nonempty subset of $\mathbb{R}$ having an upper bound has a least upper bound (**lub**) in $\mathbb{R}$.

Explanation: Let A be a nonempty subset of $\mathbb{R}$. A real number u is called an upper bound of A if each element of A is less than or equal to u. An upper bound ℓ of A is called a least upper bound if all upper bounds of A are greater than or equal to ℓ.

Notice that $\mathbb{Q}$ does not satisfy the completeness property. For example, the nonempty set $A = \{x \in \mathbb{Q} : x^2 < 2\}$ has an upper bound, say, 2. But its least upper bound is $\sqrt{2}$, which is not in $\mathbb{Q}$.

Similar to lub, we have the notion of glb, the greatest lower bound of a subset of $\mathbb{R}$. Let A be a nonempty subset of $\mathbb{R}$. A real number v is called a lower bound of A if each element of A is greater than or equal to v. A lower bound m of A is called a greatest lower bound if all lower bounds of A are less than or equal to m. The completeness property of $\mathbb{R}$ implies that

> Every nonempty subset of $\mathbb{R}$ having a lower bound has a greatest lower bound (**glb**) in $\mathbb{R}$.

The lub acts as a maximum of a nonempty set and the glb acts as a minimum of the set. In fact, when the lub$(A) \in A$, this lub is defined as the **maximum of** A and is denoted as $\max(A)$. Similarly, if the glb$(A) \in A$, this glb is defined as the **minimum of** A and is denoted by $\min(A)$.

A consequence of the completeness property of $\mathbb{R}$ is the **Archimedean property**:

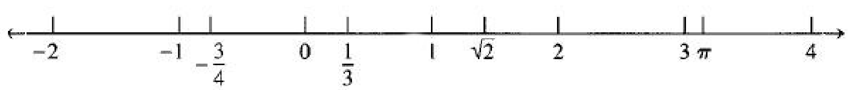
If $a > 0$ and $b > 0$, then there exists an $n \in \mathbb{N}$ such that $na \geq b$.

From the Archimedean property, it follows that corresponding to each real number x, there exists a unique integer n such that $n \leq x < n + 1$. This integer n is called the **integral part** of x, and is denoted by $[x]$. That is,

$$[x] = \text{ the largest integer less than or equal to } x.$$

Using the Archimedean property it can be proved that both $\mathbb{Q}$ and $\mathbb{R} \setminus \mathbb{Q}$ are **dense** in $\mathbb{R}$. That is, if $x < y$ are real numbers then there exist a rational number a and an irrational number b such that $x < a < y$ and $x < b < y$.

We may not explicitly use these properties of $\mathbb{R}$ but some theorems, whose proofs we will omit, can be proved using these properties. These properties allow $\mathbb{R}$ to be visualized as a number line:



Let $a, b \in \mathbb{R},\ a < b$.

$[a, b] = \{x \in \mathbb{R} : a \leq x \leq b\}$, the closed interval $[a, b]$.

$(a, b] = \{x \in \mathbb{R} : a < x \leq b\}$, the semi-open interval $(a, b]$.

$[a, b) = \{x \in \mathbb{R} : a \leq x < b\}$, the semi-open interval $[a, b)$.

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

$(-\infty, b] = \{x \in \mathbb{R} : x \leq b\}$, the closed infinite interval $(-\infty, b]$.

$(-\infty, b) = \{x \in \mathbb{R} : x < b\}$, the open infinite interval $(-\infty, b)$.

$[a, \infty) = \{x \in \mathbb{R} : x \geq a\}$, the closed infinite interval $[a, \infty)$.

$(a, \infty) = \{x \in \mathbb{R} : x \leq b\}$, the open infinite interval (a, ∞).

$(-\infty, \infty) = \mathbb{R}$, both open and closed infinite interval.

We also write $\mathbb{R}_+$ for $(0, \infty)$ and $\mathbb{R}_-$ for $(-\infty, 0)$. These are, respectively, the set of all positive real numbers, and the set of all negative real numbers.

A **neighborhood** of a point c is an open interval $(c - \delta, c + \delta)$ for some $\delta > 0$.

The **absolute value** of $x \in \mathbb{R}$ is defined as $|x| = \begin{cases} x & \text{if } x \geq 0 \\ -x & \text{if } x < 0. \end{cases}$

Thus $|x| = \sqrt{x^2}$. And $|-a| = a$ for $a \geq 0$. If $x, y \in \mathbb{R}$, then $|x-y|$ is the distance between real numbers x and y. Moreover,

$$|-x| = |x|,\ |xy| = |x|\,|y|,\ \left|\frac{x}{y}\right| = \frac{|x|}{|y|} \text{ if } y \neq 0,\ |x+y| \leq |x|+|y|,\ \big|\,|x|-|y|\,\big| \leq |x-y|.$$

Let $x \in \mathbb{R}$ and let $a > 0$. Then the following are true:

1. $|x| = a$ iff $x = \pm a$.
2. $|x| < a$ iff $-a < x < a$ iff $x \in (-a, a)$.
3. $|x| \leq a$ iff $-a \leq x \leq a$ iff $x \in [-a, a]$.
4. $|x| > a$ iff $-a < x$ or $x > a$ iff $x \in (-\infty, -a) \cup (a, \infty)$ iff $x \in \mathbb{R} \setminus [-a, a]$.
5. $|x| \geq a$ iff $-a \leq x$ or $x \geq a$ iff $x \in (-\infty, -a] \cup [a, \infty)$ iff $x \in \mathbb{R} \setminus (-a, a)$.

Therefore, for $a \in \mathbb{R}, \delta > 0$,

$$|x-a| < \delta \text{ iff } a-\delta < x < a+\delta \text{ iff } x \in (a-\delta, a+\delta).$$

The following statement is useful in proving equality using an inequality:

Let $a, b \in \mathbb{R}$. If for each $\epsilon > 0$, $|a-b| < \epsilon$, then $a = b$.

1.2 Sequences

The infinite sum $100 + 10 + 1 + 1/10 + 1/100 + \cdots$ is equal to the decimal number $111.111\cdots$, whereas the infinite sum $1 + 2 + 3 + 4 + \cdots$ is not a number. For the first sum, we rather take the partial sums

$$100,\ 100+10,\ 100+10+1,\ 100+10+1+1/10,\ \ldots$$

which are numbers and ask whether the sequence of these numbers approximates certain real number? We may ask a similar question about the second sum.

As another example, consider approximating $\sqrt{2}$ by the usual division procedure. We get the sequence

$$1,\ 1.4,\ 1.41,\ 1.414,\ 1.4142,\ 1.41421,\ 1.414213,\ 1.4142135,\ 1.41421356,\ \ldots$$

Does it approximate $\sqrt{2}$?

In general, we define a **sequence**, specifically, a **sequence of real numbers** as a function $f : \mathbb{N} \to \mathbb{R}$. The values of the function are $f(1),\ f(2),\ f(3),\ \ldots$ These are called the **terms of the sequence**. The nth term of the sequence is $f(n)$. Writing $f(n)$ as x_n, we write the sequence in many ways such as

$$(x_n), \quad (x_n)_{n=1}^{\infty}, \quad \{x_n\}_{n=1}^{\infty}, \quad \{x_n\}, \quad \text{or as } \ (x_1, x_2, x_3, \ldots)$$

showing explicitly its terms. For example, $x_n = n$ defines the sequence

$$f : \mathbb{N} \to \mathbb{R} \text{ with } f(n) = n,$$

that is, the sequence is $(1, 2, 3, 4, \ldots)$, the sequence of natural numbers. Informally, we say "the sequence $x_n = n$."

The sequence $x_n = 1/n$ is the sequence $(1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots)$; formally, $(1/n)$.

The sequence $x_n = 1/n^2$ is the sequence $(1/n^2)$, also $(1, \frac{1}{4}, \frac{1}{9}, \frac{1}{16}, \ldots)$.

The constant sequence (c) for a given real number c is the constant function $f : \mathbb{N} \to \mathbb{R}$, where $f(n) = c$ for each $n \in \mathbb{N}$. It is $(c, c, c, \ldots)$.

A sequence is an infinite list of real numbers; it is ordered like natural numbers, and unlike a set of numbers where there is no order.

There are sequences which *approximate* a real number and there are sequences which do not approximate any real number. For example, $(1/n)$ approximates the real number 0, whereas (n) approximates no real number. Also the sequence $(1, -1, 1, -1, 1, -1, \ldots)$, which may be written as $\big((-1)^{n-1}\big)$, approximates no real number. We would say that the sequence $(1/n)$ converges to 0 and the other two sequences diverge. The sequence (n) diverges to ∞ and the sequence $\big((-1)^{n-1}\big)$ diverges.

Look at the sequence $(1/n)$ closely. We feel that eventually, it will approximate 0. It means that whatever tolerance I fix, there is a term in the sequence after which every term of the sequence away from 0 is within that tolerance. What does it mean?

Suppose I am satisfied with an approximation to 0 within the tolerance 5. Then, I see that the terms of the sequence, starting with 1 and then $1/2, 1/3, \ldots$, all of them are within 5 units away from 0. In fact, $|1/n - 0| < 5$ for all n. Now, you see, bigger the tolerance, it is easier to fix a *tail* of the sequence satisfying the tolerance condition. Suppose I fix my tolerance as $1/5$. Then I see that the sixth term onwards, all the terms of the sequence are within $1/5$ units away from 0. That is, $|1/n - 0| < 1/5$ for all $n \ge 6$. If I fix my tolerance as 10^{-10}, then we see that $|1/n - 0| < 10^{-10}$ for all $n \ge 10^{10} + 1$.

This leads to the formal definition of *convergence of a sequence*.

Let (x_n) be a sequence. Let $a \in \mathbb{R}$. We say that (x_n) **converges to** a iff for each $\epsilon > 0$, there exists an $m \in \mathbb{N}$ such that for all natural numbers $n > m$, $|x_n - a| < \epsilon$.

(1.1) ***Example***

Show that the sequence $(1/n)$ converges to 0.

Let $\epsilon > 0$. Take $m = [1/\epsilon]$, the natural number such that $m \le 1/\epsilon < m + 1$. Then $1/(m+1) < \epsilon$. Moreover, if $n > m$, then $1/n \le 1/(m+1) < \epsilon$. That is, for any given $\epsilon > 0$, there exists an m, (we have defined it here) such that for every $n > m$, we see that $|1/n - 0| < \epsilon$. Therefore, $(1/n)$ converges to 0. ∎

Notice that in (1.1), we could have resorted to the Archimedean property directly and chosen any natural number $m > 1/\epsilon$.

Now that $(1/n)$ converges to 0, the sequence whose first 1000 terms are like (n) and 1001st term onward, it is like $(1/n)$ also converges to 0. Because, for any given $\epsilon > 0$, we choose our m as $[1/\epsilon]+1000$. That is, convergence behavior of a sequence does not change if first finite number of terms are changed.

For a constant sequence $x_n = c$, suppose $\epsilon > 0$ is given. We see that for each $n \in \mathbb{N}$, $|x_n - c| = 0 < \epsilon$. Therefore, the constant sequence (c) converges to c.

Sometimes, it is easier to use the condition $|x_n - a| < \epsilon$ as $a - \epsilon < x_n < a + \epsilon$.

A sequence thus converges to a iff each neighborhood of a contains a tail of the sequence.

We say that a sequence (x_n) **converges** iff it converges to some a. Thus to say that (x_n) **diverges** means that the sequence does not converge to any real number what so ever.

(1.2) ***Example***

Show that the sequence $\{(-1)^n\}$ diverges.

It means that whatever real number r we choose, it is not a limit of the sequence $-1, 1, -1, 1, -1, \ldots$. To see this, we consider three cases:

Case 1: $r = 1$. Let $\epsilon = 1/2$. If the sequence converges to 1, then we have an $m \in \mathbb{N}$ such that both $(m+1)$st term and $(m+2)$nd term are no more than $1/2$ away from 1. Now, one of $m+1$ and $m+2$ is odd; write this odd number as k. In that case, $x_k = -1$. Then it follows that $|x_k - 1| = |-1-1| < 1/2$, a contradiction.

Case 2: $r = -1$. Similar to Case 1. Consider an even number greater than m.

Case 3: $r \ne 1$, $r \ne -1$. Let $\epsilon = \frac{1}{2}\min\{|r-1|, |r+1|\}$. That is, to whichever point 1 or -1 the real number r is closer, take ϵ as half of that distance. Then neither $|r-1| < \epsilon$ nor $|r-(-1)| < \epsilon$. That is, no term of the sequence is within a distance ϵ from r. So, the sequence does not converge to r. ∎

As you see, proving that a sequence does not converge is comparatively difficult. To show that a sequence converges the definition demands that we first guess what could be its limit; then the definition helps in verifying that our guess is correct or not. Also, notice that when x_n converges to a, the m in the definition may depend on the given ϵ. Thus $\{x_n\}$ does not converge means that corresponding to any real number a we get an $\epsilon > 0$ such that there are infinitely many terms of the sequence away from a by at least ϵ, that is, $|x_n - a| \ge \epsilon$ for infinitely many n's.

There can be *non-convergence* in a way unlike the sequence $\{(-1)^n\}$. The terms of the sequence may grow indefinitely taking positive values or may diminish indefinitely taking negative values. In the first case, whatever natural number you choose, there are infinitely many terms of the sequence which are bigger than the chosen natural number. We say, it surpasses each natural number and in the second case, it remains smaller than each negative integer. These correspond to the two special cases of divergence.

Let (x_n) be a sequence. We say that (x_n) **diverges to** ∞ iff for every $r > 0$, there exists an $m \in \mathbb{N}$ such that for all natural numbers $n > m$, $x_n > r$.

We call an open interval (r, ∞) a neighborhood of ∞. A sequence thus diverges to ∞ implies the following:

1. Each neighborhood of ∞ contains a tail of the sequence.
2. Every tail of the sequence contains arbitrarily large positive numbers.

We say that (x_n) **diverges to** $-\infty$ iff for every $r > 0$, there exists an $m \in \mathbb{N}$ such that for all natural numbers $n > m$, $x_n < -r$.

Calling an open interval $(-\infty, s)$ a neighborhood of $-\infty$, we see that a sequence diverges to $-\infty$ implies the following:

1. Each neighborhood of $-\infty$ contains a tail of the sequence.
2. Every tail of the sequence contains arbitrarily small negative numbers.

We use a unified notation for convergence to a real number and divergence to $\pm\infty$. Let (x_n) be a sequence. When x_n converges to a real number r, we say that the **limit of** (x_n) **is** r; and when x_n diverges to $\pm\infty$, we say that the **limit of** (x_n) **is** $\pm\infty$. For $\ell \in \mathbb{R} \cup \{-\infty, \infty\}$, we write the phrase "the limit of x_n is ℓ" in any one of the following manner:

$$\lim_{n\to\infty} x_n = \ell, \quad \lim x_n = \ell, \quad x_n \to \ell \text{ as } n \to \infty, \quad x_n \to \ell.$$

(1.3) ***Example***

Show that (a) $\lim \sqrt{n} = \infty$; (b) $\lim \log(1/n) = -\infty$.

(a) Let $r > 0$. Choose an $m > r^2$. Let $n > m$. Then $\sqrt{n} > \sqrt{m} > r$. Therefore, $\lim \sqrt{n} = \infty$.

(b) Let $r > 0$. Choose a natural number $m > e^r$. Let $n > m$. Then $1/n < 1/m < e^{-r}$. Since $\log x$ is an increasing function, we have $\log(1/n) < \log e^{-r} = -r$. Therefore, $\log(1/n) \to -\infty$. □

Using the definitions of limit of a sequence many useful results can be shown. In addition, using the completeness property of $\mathbb{R}$ some more results about sequences can be proved.

1.3 Results on Sequences

It is of fundamental importance that if you obtain the limit of a sequence by some method, then by following another method, you would not get a different limit.

(1.4) *Theorem*

Limit of a sequence is unique.

Proof. Let (x_n) be a sequence. Suppose that $\lim x_n = \ell$ and also that $\lim x_n = s$. We consider the following exhaustive cases.

Case 1: $\ell \in \mathbb{R}$ and $s \in \mathbb{R}$. On the contrary, suppose that $s \neq \ell$; that is, $|s - \ell| > 0$. Choose $\epsilon = |s - \ell|/2$. We have natural numbers k and m such that for every $n \geq k$ and $n \geq m$,

$$|x_n - \ell| < \epsilon \quad \text{and} \quad |x_n - s| < \epsilon.$$

Fix one such n, say $M > \max\{k, m\}$. Both the above inequalities hold for $n = M$. Then

$$|s - \ell| = |s - x_M + x_M - \ell| \leq |x_M - s| + |x_M - \ell| < 2\epsilon = |s - \ell|.$$

So, $|s - \ell| < |s - \ell|$, a contradiction.

Case 2: $\ell \in \mathbb{R}$ and $s = \infty$. Since the sequence converges to ℓ, for $\epsilon = 1$, there exists a natural number k such that for every $n \geq k$, we have $|x_n - \ell| < 1$. Since the sequence diverges to ∞, we have $m \in \mathbb{N}$ such that for every $n > m$, $x_n > \ell + 1$. Now, fix an $M > \max\{k, m\}$. Then both of the above hold for this $n = M$. So, $x_M < \ell + 1$ and $x_M > \ell + 1$. This is a contradiction.

Case 3: $\ell \in \mathbb{R}$ and $s = -\infty$. It is similar to Case 2.

Case 4: $\ell = \infty$, $s = -\infty$. Again choose an M so that x_M is both greater than 1 and also less than -1 leading to a contradiction. ∎

Sometimes it is helpful in determining whether a sequence converges, even if we are not able to find its limit. Essentially, there are two results which help us to do this. In stating these results, the following terminology will be used.

We say that a sequence (x_n) is **bounded** iff there exists a positive real number k such that for each $n \in \mathbb{N}$, $|x_n| \le k$; that is, when the whole sequence is contained in an interval of finite length.

We also say that (x_n) is **bounded below** iff there exists an $m \in \mathbb{R}$ such that $x_n \ge m$ for each n; and the sequence (x_n) is called **bounded above** iff there exists an $M \in \mathbb{R}$ such that each $x_n \le M$ for each n.

Clearly, a sequence is bounded iff it is both bounded below and bounded above.

The sequence $\big((-1)^n/n\big)$ is bounded and it converges to 0.

Divergent sequences can be bounded or unbounded. For example, $\big((-1)^n\big)$ is a bounded sequence whereas (n) and $(-n)$ are unbounded sequences. The sequence $\big((-1)^n\big)$ diverges, but it neither diverges to ∞ nor to $-\infty$. The sequence (n) diverges to ∞; the sequence $(-n)$ diverges to $-\infty$.

The sequence $\big((-1)^n \log n\big)$ is unbounded; it diverges; but it neither diverges to ∞ nor to $-\infty$.

A sequence (x_n) is called **increasing** iff $x_n \le x_{n+1}$ for each n. Similarly, (x_n) is called **decreasing** iff $x_n \ge x_{n+1}$ for each n. A sequence which is either increasing or decreasing is called a **monotonic** sequence.

A sequence (x_n) is called a **Cauchy sequence** iff for each $\epsilon > 0$, there exists an $M \in \mathbb{N}$ such that for all natural numbers n and m with $n > m > M$, we have $|x_n - x_m| < \epsilon$. It follows that for all $n > m$, if $\lim |x_n - x_m| \to 0$ as $m \to \infty$, then (x_n) is a Cauchy sequence.

Let (x_n) be a sequence. Let $k_1 < k_2 < k_3 < \cdots$ be an increasing sequence of indices. The sequence $\big(x_{k_n}\big)$ for $n = 1, 2, 3, \ldots$, is called a **subsequence** of the sequence (x_n).

For example, $(1, 4, 9, 16, \ldots)$ is a subsequence of the sequence $1, 2, 3, 4, \ldots$.

Including the two criteria for convergence of a sequence, one by Cauchy and the other by Weirstrass, we mention some other important results.

(1.5) *Theorem*

(1) *Each convergent sequence is bounded.*

(2) Algebra of Limits*: Suppose* $\lim x_n = a$ *and* $\lim y_n = b$. *Then the following are true:*

   (a) Sum*:* $\lim (x_n + y_n) = a + b$.

   (b) Difference*:* $\lim (x_n - y_n) = a - b$.

   (c) Constant Multiple*:* $\lim (cx_n) = ca$ *for any real number* c.

(d) Product*:* $\lim(x_n y_n) = ab$.
(e) Division*:* $\lim(x_n/y_n) = a/b$, *provided no* y_n *is* 0 *and* $b \neq 0$.
(f) Domination*: If for each* $n \in \mathbb{N}$, $x_n \leq y_n$, *then* $a \leq b$.

(3) Sandwich Theorem*: Let* (x_n), (y_n) *and* (z_n) *be sequences. Suppose there exists* $m \in \mathbb{N}$ *such that for all* $n > m$, *we have* $x_n \leq y_n \leq z_n$. *If* $x_n \to \ell$ *and* $z_n \to \ell$, *then* $y_n \to \ell$.

(4) Weirstrass Criterion*: A bounded monotonic sequence converges. Specifically,*

(a) *an increasing sequence which is bounded above converges to its* lub*;*
(b) *a decreasing sequence which is bounded below converges to its* glb.

(5) Cauchy Criterion*: A sequence* (x_n) *converges iff it is a Cauchy sequence.*

(6) Limits of functions to Limits of sequences*: Let* $m \in \mathbb{N}$. *Let* $f(x)$ *be a function whose domain includes* $[m, \infty)$. *Let* (x_n) *be a sequence such that* $x_n = f(n)$ *for all* $n \geq m$. *If* $\lim_{x\to\infty} f(x) = \ell$, *then* $\lim_{n\to\infty} x_n = \ell$.

(7) Limits of sequences to Limits of functions*: Let* $a < c < b$. *Let* $f : D \to \mathbb{R}$ *be a function where* D *contains* $(a, c) \cup (c, b)$. *Let* $\ell \in \mathbb{R}$. *Then* $\lim_{x\to c} f(x) = \ell$ *iff for each non-constant sequence* (x_n) *converging to* c, *the sequence of functional values* $\big(f(x_n)\big)$ *converges to* ℓ.

(8) Subsequence Criterion*: Let* (x_n) *be a sequence.*

(a) *If* $x_n \to \ell$, *then every subsequence of* (x_n) *converges to* ℓ.
(b) *If* $x_{2n} \to \ell$ *and* $x_{2n+1} \to \ell$, *then* $x_n \to \ell$.
(c) *If* (x_n) *is bounded, then it has a convergent subsequence.*

(9) Continuity*: Let* $f(x)$ *be a continuous real valued function whose domain contains each term of a convergent sequence* x_n *and also its limit. Then* $\lim_{n\to\infty} f(x_n) = f\Big(\lim_{n\to\infty} x_n\Big)$.

For sequences (x_n) and (y_n), we write their sum as $(x_n + y_n)$ and product as $(x_n y_n)$. Sum of two divergent sequences may converge; similarly, product of two divergent sequences may converge. For example, $(1, 0, 1, 0, 1, \ldots)$ and $(0, 1, 0, 1, 0, \ldots)$ are divergent but their sum $(1, 1, 1, 1, \ldots)$ is convergent and their product $(0, 0, 0, \ldots)$ is convergent. Also, $\big((-1)^n\big)$ diverges but $\big((-1)^n(-1)^n\big)$ converges; whereas (n) diverges, $\big(1/n^2\big)$ converges and their product $(1/n)$ converges.

The condition $\lim y_n \neq 0$ is important in the division rule. The sequences $(1/n)$ and $\big(1/n^2\big)$ are convergent but their product $(1/n)/(1/n^2) = n$ does not converge to $\lim(1/n)/\lim(1/n^2)$, which is an indeterminate.

When we say that $\infty + \infty = \infty$, what we mean is if (x_n) and (y_n) are any sequences such that $\lim x_n = \infty$ and $\lim y_n = \infty$, then $\lim(x_n + y_n) = \infty$. Similarly, other equalities concerning these two special symbols $\pm\infty$ can be shown. We note them down here:

Let $r > 0$ be any real number. As usual, addition and multiplication are commutative, and

$$\infty + \pm r = \infty \pm 0 = \infty + \infty = \infty, \quad -\infty \pm r = -\infty \pm 0 = -\infty,$$
$$r \cdot \infty = (-r) \cdot (-\infty) = \frac{\infty}{r} = \frac{-\infty}{-r} = \infty \cdot \infty = (-\infty) \cdot (-\infty) = \infty,$$
$$r \cdot (-\infty) = (-r) \cdot \infty = \frac{\infty}{-r} = \frac{-\infty}{r} = \infty \cdot (-\infty) = -\infty.$$

The indeterminate forms are:

$$\infty - \infty, \quad 0 \cdot (\pm\infty), \quad \frac{\pm\infty}{\pm\infty}.$$

The reason for these expressions to be indeterminate follows the same principle. For example, $\infty - \infty$ is interpreted as if (x_n) diverges to ∞ and (y_n) diverges to $-\infty$, then $(x_n + y_n)$ may diverge to ∞, or to $-\infty$, or converge to any real number, or neither. I leave it to you for supplying appropriate examples for each of these scenario. Similarly, other forms above are indeterminate.

One consequence of the *constant multiple* rule is that every *nonzero multiple* of a divergent sequence diverges. For if x_n diverges and $c \neq 0$ but cx_n converges, then $x_n = (1/c)cx_n$ would converge!

The domination result implies that if a sequence has only positive terms, its limit cannot be negative. Notice that if all but an initial finite number of terms of a sequence are positive, then also its limit cannot be negative. Similarly, the limit of a sequence of negative terms (leaving some first finite number of terms) cannot be positive. Moreover, the domination statement includes the case of divergence to ∞. Specifically, if $x_n \to \infty$ and for each n, $x_n \le y_n$, then $y_n \to \infty$.

For an application of the Sandwich theorem, consider a sequence (x_n).

If $|x_n| \to 0$, then $x_n \to 0$. Reason: $-|x_n| \le x_n \le |x_n|$ and Sandwitch theorem.

Also, if $x_n \to 0$, then $|x_n| \to 0$. Reason: $\big| |x_n| - 0 \big| \le |x_n - 0|$.

That is, (x_n) converges to 0 iff $(|x_n|)$ converges to 0.

In general, if (x_n) converges to ℓ, then $(|x_n|)$ converges to $|\ell|$. It follows from the inequality

$$\big| |x_n| - |\ell| \big| \le |x_n - \ell|.$$

However, even if $(|x_n|)$ converges, (x_n) may not converge. For instance, take $x_n = (-1)^n$.

Remark 1.6 The Weirstrass criterion can be proved as follows using the completeness principle of real numbers. For (1.5-4a), let (x_n) be an increasing sequence which is bounded above. Then the set of terms of the sequence has an lub, say, s. That is, for each $n \in \mathbb{N}$, $x_n \leq s$, and for each $\epsilon > 0$ there is a term of the sequence, say, x_k such that $s - \epsilon < x_k \leq s$. Since (x_n) is an increasing sequence, $s - \epsilon < x_k \leq x_{k+1} \leq x_{k+2} \leq \cdots \leq s$. So, the tail $x_k, x_{k+1}, x_{k+2}, \ldots$ is contained in $(s - \epsilon, s + \epsilon)$. Therefore, (x_n) converges to s.

(1.7) *Example*

(1) $\displaystyle\lim_{n\to\infty} \frac{5n^2 + 2n + 7}{n^2 - 11n + 5} = \lim_{n\to\infty} \frac{5 + \frac{2}{n} + \frac{7}{n^2}}{1 - \frac{11}{n} + \frac{5}{n^2}} = \frac{5}{1} = 5.$

the operation of division in the limit is applicable since the limit of the denominator is nonzero.

(2) Since $\dfrac{-1}{n} \leq \dfrac{\cos n}{n} \leq \dfrac{1}{n}$, Sandwich theorem implies that $\dfrac{\cos n}{n} \to 0.$

(3) As $n < \sqrt{n^2 + 1} + n$, $0 < \sqrt{n^2 + 1} - n = \dfrac{1}{\sqrt{n^2 + 1} + n} < \dfrac{1}{n}.$

By Sandwich theorem, $\displaystyle\lim_{n\to\infty} \sqrt{n^2 + 1} - n = 0.$

(4) Let $p > 0$. Show that $\lim (1/n^p) = 0$.

Let $\epsilon > 0$. Using Archimedean Property, take $m \in \mathbb{N}$ so that $(1/\epsilon)^{1/p} < m$. Now, $1/\epsilon < m^p$; so $1/m^p < \epsilon$. If $n > m$, then $\dfrac{1}{n^p} < \dfrac{1}{m^p} < \epsilon$.

(5) Let $x > 0$. Show that $\lim(x^{1/n}) = 1$.

The function $f(t) = x^t$ is continuous at each $t \in [0, \infty)$.
So, $\lim(x^{1/n}) = x^{\lim(1/n)} = x^0 = 1$.

(6) Show that if $|x| < 1$, then $\lim x^n = 0$.

Write $|x| = \dfrac{1}{1 + r}$ for some $r > 0$. By the Binomial theorem,

$$(1 + r)^n \geq 1 + nr > nr.$$

So,

$$0 < |x|^n = (1 + r)^{-n} < \frac{1}{nr}.$$

By Sandwich theorem, $\lim |x|^n = 0$. Now, $-|x|^n \leq x^n \leq |x|^n$. Again, by Sandwich theorem, $\lim x^n = 0$.

(7) Show that $\lim(n^{1/n}) = 1$.

Let $x_n = n^{1/n} - 1$. We see that $x_n \geq 0$. Using Binomial theorem for $n \geq 2$,

$$x_n + 1 = n^{1/n} \Rightarrow n = (x_n + 1)^n \geq 1 + nx_n + \frac{n(n-1)}{2} x_n^2.$$

Hence
$$n \geq \frac{n(n-1)}{2} x_n^2 \Rightarrow 0 \leq x_n \leq \frac{\sqrt{2}}{\sqrt{n-1}}.$$
Apply Sandwich theorem to conclude that $\lim x_n = 0$.

□

(1.8) ***Example***

Let $x_n = 1 + \frac{1}{2} + \frac{1}{3} + \cdots \frac{1}{n}$ for $n \in \mathbb{N}$. Does the sequence (x_n) converge?
$$x_{2n} - x_n = \frac{1}{n+1} + \frac{1}{n+2} + \cdots + \frac{1}{2n} \geq \frac{1}{2n} \cdot n = \frac{1}{2}.$$
Hence, as $n \to \infty$, $|x_{2n} - x_n|$ does not converge to 0. That is, (x_n) is not a Cauchy sequence; so it does not converge. □

(1.9) ***Example***

Define a sequence (x_n) by $x_1 = 1,\ x_2 = 2$, and $x_{k+2} = (x_{k+1} + x_k)/2$ for $k \geq 1$. Does the sequence converge?

$x_{n+2} - x_{n+1} = (x_n - x_{n+1})/2$. Thus,
$$|x_{n+2} - x_{n+1}| = \frac{1}{2}|x_{n+1} - x_n| = \cdots = \frac{1}{2^n}|x_2 - x_1| = \frac{1}{2^n}.$$
If $n > m$, then
$$0 \leq |x_n - x_m| \leq |x_n - x_{n-1}| + \cdots + |x_{m+1} - x_m| \leq \frac{1}{2^{n-2}} + \frac{1}{2^{n-3}} + \cdots + \frac{1}{2^{m-1}}$$
$$= \frac{1}{2^{m-1}} \cdot \frac{1 - (1/2)^{n-m}}{1 - 1/2} < \frac{1}{2^{m-1}} \cdot \frac{1}{1/2} = \frac{1}{2^{m-2}}.$$
As $m \to \infty$, by Sandwich theorem, $\lim |x_n - x_m| = 0$. Therefore, (x_n) is a Cauchy sequence; so it converges.

Aliter: (x_{2n}) is a decreasing sequence bounded below by 1. So, it converges to some real number a.

(x_{2n-1}) is an increasing sequence bounded above by 2. So, it converges to some real number b.

Since $2x_{2m+1} = x_{2m-1} + x_{2m}$ for each m, taking the limit, we have $2b = b + a$. That is, $a = b$. Therefore, the sequence converges. □

(1.10) ***Example***

Let (x_n) be the sequence given by $x_1 = 2$ and $x_{k+1} = \dfrac{x_k}{2} + \dfrac{1}{x_k}$ for $k \geq 1$. Does (x_n) converge?

Notice that it is a sequence of positive terms. If (x_n) converges to ℓ, then taking the limit on both the sides of the recursive formula, we find that $\ell = \frac{\ell}{2} + \frac{1}{\ell}$. It implies that $2\ell^2 = \ell^2 + 2$. That is, $\ell = \pm\sqrt{2}$. We see that $x_1 > 0$ and then the recursive formula says that each term is positive. So, a possible limit for (x_n) is $\ell = \sqrt{2}$. Moreover, first few terms say that the sequence may be a decreasing sequence. Thus, if at all a limit exists, it must be the greatest lower bound of the sequence. We guess that each term of the sequence is at least $\sqrt{2}$ and the sequence is a monotonically decreasing sequence. We must prove both.

(a) Observe that $x_1 = 2 \geq \sqrt{2}$. For $n \geq 1$,

$$x_{n+1} = \frac{x_n}{2} + \frac{1}{x_n} \geq \sqrt{2} \text{ if } x_n^2 - 2\sqrt{2}x_n + 2 \geq 0 \text{ if } (x_n - \sqrt{2})^2 \geq 0,$$

which is always true, Therefore, $x_n \geq \sqrt{2}$ for each n.

(b) Now, for the decreasing nature of the sequence,

$$x_{n+1} \leq x_n \ \text{ if } \ \frac{x_n}{2} + \frac{1}{x_n} \leq x_n \ \text{ if } \ \frac{1}{x_n} \leq \frac{1}{x_n} \ \text{ if } \ 2 \leq x_n^2 \ \text{ if } \ x_n \geq \sqrt{2},$$

which we already proved in (a).

Hence (x_n) is monotonically decreasing and bonded below by $\sqrt{2}$. Therefore, it converges. Moreover, as our earlier calculation shows, $\lim x_n = \sqrt{2}$. ∎

(1.11) *Example*

Show that the sequence (t_n) with $t_n = \left(1 + \frac{1}{n}\right)^n$ converges.

Using the Binomial theorem, we obtain

$$\begin{aligned}
t_n = \left(1 + \frac{1}{n}\right)^n &= 1 + n \cdot \frac{1}{n} + \frac{n(n-1)}{2!} \cdot \frac{1}{n^2} + \cdots + \frac{1}{n^n} \\
&= 1 + 1 + \frac{1}{2!}\left(1 - \frac{1}{n}\right) + \frac{1}{3!}\left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right) + \cdots + \frac{1}{n!}\left(1 - \frac{1}{n}\right)\cdots\left(1 - \frac{n-1}{n}\right) \\
&< 1 + 1 + \frac{1}{2!}\left(1 - \frac{1}{n+1}\right) + \cdots + \frac{1}{n!}\left(1 - \frac{1}{n+1}\right)\cdots\left(1 - \frac{n-1}{n+1}\right) \\
&\quad + \frac{1}{(n+1)!}\left(1 - \frac{1}{n+1}\right)\cdots\left(1 - \frac{n}{n+1}\right) \\
&= 1 + (n+1) \cdot \frac{1}{n+1} + \frac{(n+1)(n)}{2!} \cdot \frac{1}{(n+1)^2} + \cdots + \frac{1}{(n+1)^{n+1}} \\
&= \left(1 + \frac{1}{n+1}\right)^{n+1} = t_{n+1}.
\end{aligned}$$

Thus, (t_n) is an increasing sequence. Next, since $k! \geq 2^{k-1}$, we have $1/k! \leq 1/2^{k-1}$ for $k \in \mathbb{N}$. Using this, we obtain

$$t_n = 1 + 1 + \frac{1}{2!}\Big(1 - \frac{1}{n}\Big) + \frac{1}{3!}\Big(1 - \frac{1}{n}\Big)\Big(1 - \frac{2}{n}\Big) + \cdots + \frac{1}{n!}\Big(1 - \frac{1}{n}\Big) \cdots \Big(1 - \frac{n-1}{n}\Big)$$
$$\leq 1 + 1 + \frac{1}{2!} + \frac{1}{3!} + \cdots + \frac{1}{n!} \leq 1 + 1 + \frac{1}{2} + \frac{1}{2^2} + \cdots + \frac{1}{2^{n-1}} < 1 + 2 = 3.$$

Since t_n is an increasing sequence having an upper bound, it converges. □

Since $\big((1 + 1/n)^n\big)$ is a convergent sequence, its limit is a real number. We denote the limit of this sequence as e. Since each term of the sequence lies between 2 and 3, we conclude that $2 \leq e \leq 3$.

1.4 Series

A series is an infinite sum of numbers. As it is, two numbers can be added; so by induction, a finite of them can also be added. For an infinite sum to be meaningful, we look at the sequence of partial sums. Let (x_n) be a sequence. The series $x_1 + x_2 + \cdots + x_n + \cdots$ is meaningful when another sequence, namely,

$$x_1,\ x_1 + x_2,\ x_1 + x_2 + x_3,\ \ldots,\ \sum_{k=1}^{n} x_k,\ \ldots$$

is convergent. The infinite sum itself is denoted by $\sum_{n=1}^{\infty} x_n$ and also by $\sum x_n$.

We say that the series $\sum x_n$ is convergent iff the sequence (s_n) is convergent, where the nth **partial sum** s_n is given by $s_n = \sum_{k=1}^{n} x_k$.

Thus we may define convergence of a series as follows:

We say that the series $\sum x_n$ **converges to** $\ell \in \mathbb{R}$ iff for each $\epsilon > 0$, there exists an $m \in \mathbb{N}$ such that for each natural number $n > m$, $|\sum_{k=1}^{n} x_k - \ell| < \epsilon$. In this case, we also say that the **series sums to** ℓ, and write $\sum x_n = \ell$.

Further, we say that a series **converges** iff it converges to some $\ell \in \mathbb{R}$.

A series is said to be **divergent** iff it is not convergent.

Similar to convergence, if the sequence of partial sums (s_n) diverges to $\pm\infty$, we say that the series $\sum x_n$ diverges to $\pm\infty$.

That is, the series $\sum x_n$ **diverges to** ∞ iff for each $r > 0$, there exists $m \in \mathbb{N}$ such that for each natural number $n > m$, $\sum_{k=1}^{n} x_k > r$. We write it as $\sum x_n = \infty$, and say that the series sums to ∞.

Similarly, the series $\sum x_n$ **diverges to** $-\infty$ iff for each $r > 0$, there exists $m \in \mathbb{N}$ such that for each natural number $n > m$, $\sum_{k=1}^{n} x_k < -r$. We then write $\sum x_n = -\infty$, and say that the series sums to $-\infty$.

Notice that 'converges to a real number' and 'diverges to $\pm\infty$' both are written the same way. There can be series which diverge but neither to ∞ nor to $-\infty$. Further, if a series sums to ℓ, then it cannot sum to s where $s \neq \ell$, due to the uniqueness of limit of a sequence.

The series $-1 - 2 - 3 - 4 - \cdots - n - \cdots$ diverges to $-\infty$.

The series $1 - 1 + 1 - 1 + \cdots$ diverges. It neither diverges to ∞ nor to $-\infty$. Because, the sequence of partial sums here is $1, 0, 1, 0, 1, 0, 1, \ldots$.

The series $\displaystyle\sum_{n=1}^{\infty} \frac{1}{2^n}$ converges to 1. Because, if (s_n) is the sequence of partial sums, then

$$s_n = \sum_{k=1}^{n} \frac{1}{2^k} = \frac{1}{2} \cdot \frac{1 - (1/2)^n}{1 - 1/2} = 1 - \frac{1}{2^n} \to 1 \text{ as } n \to \infty.$$

A general scenario is discussed in the next example.

(1.12) *Example*

Let $a \neq 0$ and let $r \in \mathbb{R}$. Consider the **geometric series**

$$\sum_{n=1}^{\infty} ar^{n-1} = a + ar + ar^2 + ar^3 + \cdots .$$

The nth partial sum of the geometric series is

$$s_n = a + ar + ar^2 + ar^3 + \cdots ar^{n-1} = \frac{a(1 - r^n)}{1 - r}.$$

(a) If $|r| < 1$, then $r^n \to 0$. The geometric series converges to $\displaystyle\lim_{n\to\infty} s_n = \frac{a}{1 - r}$.
Therefore, $\displaystyle\sum_{n=0}^{\infty} ar^n = \sum_{n=1}^{\infty} ar^{n-1} = \frac{a}{1 - r}$ for $|r| < 1$.

(b) If $r = -1$ or $|r| > 1$, then r^n diverges; so the geometric series $\sum ar^{n-1}$ diverges. And, for $r = 1$, the geometric series $1 + 1 + \cdots$ diverges. $\Box$

(1.13) *Example*

The series $1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \cdots$ diverges to ∞. To see this, Write $s_n = \sum_{k=1}^{n} \frac{1}{k}$, the partial sum of the series up to n terms. Then, use (1.8).

For another solution, let r be any positive real number. Choose a natural number $m > 2r$ and take $n = 2^m$. Then

$$
\begin{aligned}
s_n = \sum_{k=1}^{n} \frac{1}{k} &\ge 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{2^m - 1} \\
&= 1 + \Big(\frac{1}{2} + \frac{1}{3}\Big) + \Big(\frac{1}{4} + \frac{1}{5} + \frac{1}{6} + \frac{1}{7}\Big) + \cdots + \Big(\sum_{k=2^{m-1}}^{2^m - 1} \frac{1}{k}\Big) \\
&\ge 1 + \Big(\frac{1}{4} + \frac{1}{4}\Big) + \Big(\frac{1}{8} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8}\Big) + \cdots + \Big(\sum_{k=2^{m-1}}^{2^m - 1} \frac{1}{2^m}\Big) \\
&= 1 + \frac{1}{2} + \frac{1}{2} + \cdots + \frac{1}{2} = 1 + \frac{m-1}{2} > 1 + \frac{2r-1}{2} = \frac{2r+1}{2} > r.
\end{aligned}
$$

That is, corresponding to any $r > 0$ there exists $n \in \mathbb{N}$ so that $s_n > r$. ($n = 2^m$, where $m > 2r$ is a natural number.) Therefore, the series diverges to ∞. This is called the **harmonic series**. ∎

(1.14) *Example*

Does the series $\sum \frac{1}{n(n+1)}$ converge?

Since $\frac{1}{k(k+1)} = \frac{1}{k} - \frac{1}{k+1}$, we have

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

Since $1/(n+1) \to 0$, the series converges to 1. Thus $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n(n+1)} = 1.$ ∎

(1.15) *Example*

The sum of the convergent series $\displaystyle\sum_{n=1}^{\infty} \frac{3^n - 4}{6^n}$ can be computed as follows:

$$
\sum_{n=1}^{\infty} \frac{3^n - 4}{6^n} = \sum_{n=0}^{\infty} \frac{1/2}{2^n} - \frac{4}{6} \sum_{n=0}^{\infty} \frac{1}{6^n} = \frac{1/2}{1 - 1/2} - \frac{4/6}{1 - 1/6} = 1 - \frac{4}{5} = \frac{1}{5}.
$$
∎

The following result sometimes helps in ascertaining that a given series diverges.

(1.16) *Theorem*

If a series $\sum a_n$ converges, then the sequence (a_n) converges to 0.

Proof. Let s_n denote the partial sum $\sum_{k=1}^{n} a_k$. Then $a_n = s_n - s_{n-1}$. If the series converges, say, to ℓ, then $\lim s_n = \ell = \lim s_{n-1}$. It follows that $\lim a_n = 0$.

It says that if $\lim a_n$ does not exist, or if $\lim a_n$ exists but is not equal to 0, then the series $\sum a_n$ diverges. ∎

(1.17) *Example*

(1) The series $\sum_{n=1}^{\infty} \frac{-n}{3n+1}$ diverges because $\lim_{n\to\infty} \frac{-n}{3n+1} = -\frac{1}{3} \neq 0$.

(2) The series $\sum(-1)^n$ diverges because $\lim(-1)^n$ does not exist. □

Notice what (1.16) does not say. The harmonic series diverges even though $\lim \frac{1}{n} = 0$.

(1.18) *Theorem*

If both the series $\sum a_n$ and $\sum b_n$ converge, then the series $\sum(a_n + b_n)$, $\sum(a_n - b_n)$ and $\sum ka_n$ converge; where k is any real number.

If $\sum a_n$ converges and $\sum b_n$ diverges, then $\sum(a_n + b_n)$ and $\sum(a_n - b_n)$ diverge.

Further, if $\sum a_n$ diverges and $k \neq 0$, then $\sum ka_n$ diverges.

Proofs of the statements in (1.18) follow from (1.5-2).

Notice that sum of two divergent series can converge. For example, both $\sum(1/n)$ and $\sum(-1/n)$ diverge but their sum $\sum 0$ converges.

Since deleting a finite number of terms of a sequence does not alter its convergence, omitting a finite number of terms or adding a finite number of terms to a convergent (divergent) series implies the convergence (divergence) of the new series. Of course, the sum of the convergent series will be affected. For example,

$$\sum_{n=3}^{\infty} \left(\frac{1}{2^n}\right) = \sum_{n=1}^{\infty} \left(\frac{1}{2^n}\right) - \frac{1}{2} - \frac{1}{4}.$$

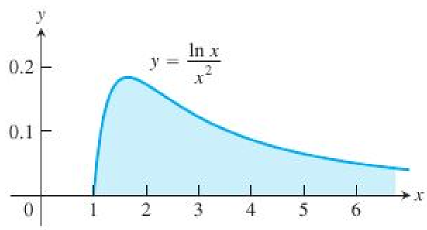
However,

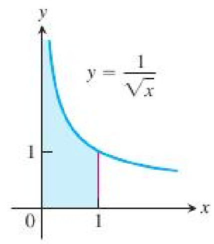
$$\sum_{n=3}^{\infty} \left(\frac{1}{2^{n-2}}\right) = \sum_{n=1}^{\infty} \left(\frac{1}{2^n}\right).$$

This is called **re-indexing** the series. As long as we preserve the order of the terms of the series, we can re-index without affecting its convergence and sum.

1.5 Improper Integrals

In the definite integral $\int_a^b f(x)dx$ we required that both a, b are finite and also the range of $f(x)$ is a subset of some finite interval. However, there are functions which violate one or both of these requirements, and yet, the area under the curves and above the x-axis remain bounded.





Such integrals are called **Improper Integrals**. Suppose $f(x)$ is continuous on $[0, \infty)$. It makes sense to write

$$\int_0^\infty f(x)dx = \lim_{b\to\infty} \int_0^b f(x)dx$$

provided that the limit exists. In such a case, we say that the improper integral $\int_0^\infty f(x)dx$ **converges** and its *value* is given by the limit. We say that the improper integral **diverges** iff it is not convergent. Obviously, we are interested in computing the value of an improper integral, in which case, the integral is required to converge. Integrals of the type $\int_a^b f(x)\,dx$ can become improper when $f(x)$ is not continuous at a point in the interval $[a, b]$. Here are the possible types of improper integrals.

1. If $f(x)$ is continuous on $[a, \infty)$, then $\displaystyle\int_a^\infty f(x)\,dx = \lim_{b\to\infty} \int_a^b f(x)\,dx.$
2. If $f(x)$ is continuous on $(-\infty, b]$, then $\displaystyle\int_{-\infty}^b f(x)\,dx = \lim_{a\to-\infty} \int_a^b f(x)\,dx.$
3. If $f(x)$ is continuous on $(-\infty, \infty)$, then
   $\displaystyle\int_{-\infty}^\infty f(x)\,dx = \int_{-\infty}^c f(x)\,dx + \int_c^\infty f(x)\,dx$, for any $c \in \mathbb{R}$.
4. If $f(x)$ is continuous on $(a, b]$ and discontinuous at $x = a$, then
   $\displaystyle\int_a^b f(x)\,dx = \lim_{t\to a+} \int_t^b f(x)\,dx.$
5. If $f(x)$ is continuous on $[a, b)$ and discontinuous at $x = b$, then
   $\displaystyle\int_a^b f(x)\,dx = \lim_{t\to b-} \int_a^t f(x)\,dx.$

6. If $f(x)$ is continuous on $[a, c) \cup (c, b]$ and discontinuous at $x = c$, then
$$\int_a^b f(x)\,dx = \int_a^c f(x)\,dx + \int_c^b f(x)\,dx.$$

In each case, if the limit of the concerned integral is finite, then we say that the improper integral (on the left) **converges**, else, the improper integral **diverges**; the finite value as obtained from the limit is the *value* of the improper integral. A convergent improper integral converges to its value. Two important sub-cases of divergent improper integrals are when the limit of the concerned integral is ∞ or $-\infty$. In these cases, we say that the improper integral **diverges to** ∞ **or to** $-\infty$ as is the case.

(1.19) *Example*

Is the area under the curve $y = (\log x)/x^2$ for $x \geq 1$ finite?

The question is whether $\displaystyle\int_1^\infty \frac{\log x}{x^2}\,dx$ converges?

Let $b > 1$. Integrating by parts,

$$\int_1^b \frac{\log x}{x^2}\,dx = \left[\log x\left(-\frac{1}{x}\right)\right]_1^b - \int_1^b \frac{1}{x}\left(\frac{-1}{x}\right)dx = -\frac{\log b}{b} - \frac{1}{b} + 1.$$

$$\lim_{b\to\infty} \int_1^b \frac{\log x}{x^2}\,dx = \lim_{b\to\infty}\left[-\frac{\log b}{b} - \frac{1}{b} + 1\right] = 1.$$

Therefore, the improper integral $\displaystyle\int_1^\infty \frac{\log x}{x^2}\,dx$ converges to 1. That is, the required area is finite and it is equal to 1. ∎

(1.20) *Example*

Is $\displaystyle\int_{-\infty}^\infty \frac{1}{1+x^2}\,dx$ convergent?

$$\int_a^b \frac{1}{1+x^2}\,dx = \tan^{-1} b - \tan^{-1} a.$$

So,

$$\int_{-\infty}^0 \frac{1}{1+x^2}\,dx = \lim_{a\to-\infty} \int_a^0 \frac{1}{1+x^2}\,dx = \lim_{a\to-\infty} (-\tan^{-1} a) = -\left(-\frac{\pi}{2}\right) = \frac{\pi}{2}.$$

$$\int_0^\infty \frac{1}{1+x^2}\,dx = \lim_{b\to\infty} \int_0^b \frac{1}{1+x^2}\,dx = \lim_{b\to\infty} (\tan^{-1} b) = \frac{\pi}{2}.$$

Therefore,

$$\int_{-\infty}^{\infty} \frac{1}{1+x^2}\,dx = \int_{-\infty}^{0} \frac{1}{1+x^2}\,dx + \int_{0}^{\infty} \frac{1}{1+x^2}\,dx$$

is convergent and its value is $\pi/2 + \pi/2 = \pi$. □

(1.21) ***Example***

Consider evaluating $\displaystyle\int_0^3 \frac{dx}{x-1}$. Overlooking the point $x = 1$, where the integrand is not defined, we may compute

$$\int_0^3 \frac{dx}{x-1} = \log|x-1|\Big]_0^3 = \log 2 - \log 1 = \log 2.$$

However, it is an improper integral and its value, if exists, must be computed as follows:

$$\int_0^3 \frac{dx}{x-1} = \lim_{b\to 1-}\int_0^b \frac{dx}{x-1} + \lim_{a\to 1+}\int_a^3 \frac{dx}{x-1}.$$

The integral converges provided both the limits are finite. However,

$$\lim_{b\to 1-}\int_0^b \frac{dx}{x-1} = \lim_{b\to 1-}\big(\log|b-1| - \log|-1|\big) = \lim_{b\to 1-}\log(1-b) = -\infty.$$

Therefore, $\displaystyle\int_0^3 \frac{dx}{x-1}$ does not converge. □

(1.22) ***Example***

Evaluate $\displaystyle\int_0^3 \frac{dx}{(x-1)^{2/3}}$.

The integrand is not defined at $x = 1$. We consider it as an improper integral.

$$\int_0^3 \frac{dx}{(x-1)^{2/3}} = \lim_{b\to 1-}\int_0^b \frac{dx}{(x-1)^{2/3}} + \lim_{a\to 1+}\int_a^3 \frac{dx}{(x-1)^{2/3}}.$$

$$\lim_{b\to 1-}\int_0^b \frac{dx}{(x-1)^{2/3}} = \lim_{b\to 1-} 3(x-1)^{1/3}\Big|_0^b = \lim_{b\to 1-}\big(3(b-1)^{1/3} - 3(-1)^{1/3}\big) = 3.$$

$$\lim_{a\to 1+}\int_a^3 \frac{dx}{(x-1)^{2/3}} = \lim_{a\to 1+} 3(x-1)^{1/3}\Big|_a^3 = \lim_{a\to 1+}\big(3(3-1)^{1/3} - 3(a-1)^{1/3}\big) = 3(2)^{1/3}.$$

Hence $\displaystyle\int_0^3 \frac{dx}{(x-1)^{2/3}} = 3(1+2^{1/3})$. □

In the above example, had we not noticed that the integrand has discontinuity in the interior, we would have ended up at a wrong computation such as

$$\int_0^3 \frac{dx}{(x-1)^{2/3}} = 3(x-1)^{1/3}\Big|_0^3 = 3(2^{1/3} - (-1)^{1/3}),$$

even though the answer happens to be correct here.

(1.23) ***Example***

For what values of $p \in \mathbb{R}$, the improper integral $\displaystyle\int_1^\infty \frac{dx}{x^p}$ converges? What is its value, when it converges?

Case 1: $p = 1$.

$$\int_1^b \frac{dx}{x^p} = \int_1^b \frac{dx}{x} = \log b - \log 1 = \log b.$$

Since $\lim_{b\to\infty} \log b = \infty$, the improper integral diverges to ∞.

Case 2: $p < 1$.

$$\int_1^b \frac{dx}{x^p} = \frac{-x^{-p+1}}{-p+1}\Big|_1^b = \frac{1}{1-p}(b^{1-p} - 1).$$

Since $\lim_{b\to\infty} b^{1-p} = \infty$, the improper integral diverges to ∞.

Case 3: $p > 1$.

$$\int_1^b \frac{dx}{x^p} = \frac{1}{1-p}(b^{1-p} - 1) = \frac{1}{1-p}\Big(\frac{1}{b^{p-1}} - 1\Big).$$

Since $\displaystyle\lim_{b\to\infty} \frac{1}{b^{p-1}} = 0$, we have

$$\int_1^\infty \frac{dx}{x^p} = \lim_{b\to\infty}\int_1^b \frac{dx}{x^p} = \lim_{b\to\infty} \frac{1}{1-p}\Big(\frac{1}{b^{p-1}} - 1\Big) = \frac{1}{p-1}.$$

Hence, the improper integral $\displaystyle\int_1^\infty \frac{dx}{x^p}$ converges to $\displaystyle\frac{1}{p-1}$ for $p > 1$ and diverges to ∞ for $p \le 1$. ∎

(1.24) ***Example***

For what values of $p \in \mathbb{R}$, the improper integral $\displaystyle\int_0^1 \frac{dx}{x^p}$ converges?

Notice that for $p \le 0$, the integral is not an improper integral, and its value is $1/(1-p)$. We consider the rest of the cases as follows.

Case 1: $p = 1$.

$$\int_0^1 \frac{dx}{x^p} = \lim_{a\to 0+}\int_a^1 \frac{dx}{x} = \lim_{a\to 0+}[\log 1 - \log a] = \infty.$$

Therefore, the improper integral diverges to ∞.

Case 2: $0 < p < 1$.

$$\int_0^1 \frac{dx}{x^p} = \lim_{a\to 0+} \int_a^1 \frac{dx}{x^p} = \lim_{a\to 0+} \frac{1-a^{1-p}}{1-p} = \frac{1}{1-p}.$$

Therefore, the improper integral converges to $1/(1-p)$.

Case 3: $p > 1$.

$$\int_0^1 \frac{dx}{x^p} = \lim_{a\to 0+} \frac{1-a^{1-p}}{1-p} = \lim_{a\to 0+} \frac{1}{p-1}\Big(\frac{1}{a^{p-1}} - 1\Big) = \infty.$$

Hence the improper integral diverges to ∞.

Therefore, $\int_0^1 \frac{dx}{x^p} = \frac{1}{1-p}$ for $p < 1$; and $\int_0^1 \frac{dx}{x^p}$ diverges to ∞ for $p \ge 1$. □

1.6 Convergence Tests for Improper Integrals

Sometimes it is helpful to be sure that an improper integral converges, even if we are unable to evaluate it.

(1.25) ***Theorem*** (Comparison Test)

Let $f(x)$ and $g(x)$ be continuous functions on $[a, \infty)$. Suppose that $0 \le f(x) \le g(x)$ for all $x \ge a$.

(1) *If $\int_a^\infty g(x)\,dx$ converges, then $\int_a^\infty f(x)\,dx$ converges.*

(2) *If $\int_a^\infty f(x)\,dx$ diverges to ∞, then $\int_a^\infty g(x)\,dx$ diverges to ∞.*

Proof. Since $0 \le f(x) \le g(x)$ for all $x \ge a$,

$$\int_a^b f(x)dx \le \int_a^b g(x)dx.$$

As $\lim_{b\to\infty} \int_a^b g(x)dx = \ell$ for some $\ell \in \mathbb{R}$, $\lim_{b\to\infty} \int_a^b f(x)dx$ exists and the limit is less than or equal to ℓ. This proves (1). Proof of (2) is similar to that of (1). ∎

We also use a similar result stated below, without proof.

(1.26) ***Theorem*** (Limit Comparison Test)

Let $f(x)$ and $g(x)$ be continuous functions on $[a, \infty)$. If $\lim_{x\to\infty} \frac{f(x)}{g(x)} = L \ne 0$, then $\int_a^\infty f(x)dx$ and $\int_a^\infty g(x)dx$ either both converge, or both diverge.

In (1.25) we talk about non-negative functions. The reason is the following result, which we will not prove:

(1.27) *Theorem*

Let $f(x)$ be a continuous function on $[a,b)$, for $b \in \mathbb{R}$ or $b = \infty$. If the improper integral $\int_a^b |f(x)|\,dx$ converges, then the improper integral $\int_a^b f(x)\,dx$ also converges.

(1.28) *Example*

(1) $\displaystyle\int_1^\infty \frac{\sin^2 x}{x^2}\,dx$ converges because

$$0 \le \frac{\sin^2 x}{x^2} \le \frac{1}{x^2} \text{ for all } x \ge 1, \text{ and } \int_1^\infty \frac{dx}{x^2} \text{ converges.}$$

(2) $\displaystyle\int_2^\infty \frac{dx}{\sqrt{x^2-1}}$ diverges to ∞ because (Recall: $\displaystyle\lim_{x\to\infty} \log x = \infty$.)

$$\frac{1}{\sqrt{x^2-1}} \ge \frac{1}{x} \text{ for all } x \ge 2, \text{ and } \int_2^\infty \frac{dx}{x} \text{ diverges to } \infty.$$

(3) $\displaystyle\int_1^\infty \frac{dx}{1+x^2}$ converges or diverges?

Since $\displaystyle\lim_{x\to\infty}\left[\frac{1}{1+x^2}\Big/\frac{1}{x^2}\right] = \lim_{x\to\infty}\frac{x^2}{1+x^2} = 1$, the limit comparison test says that the given improper integral and $\displaystyle\int_1^\infty \frac{dx}{x^2}$ both converge or diverge together. The latter converges, so does the former. However, they may converge to different values.

$$\int_1^\infty \frac{dx}{1+x^2} = \lim_{b\to\infty}[\tan^{-1} b - \tan^{-1} 1] = \frac{\pi}{2} - \frac{\pi}{4} = \frac{\pi}{4}.$$

$$\int_1^\infty \frac{dx}{x^2} = \lim_{b\to\infty}\left(\frac{-1}{b} - \frac{-1}{1}\right) = 1.$$

(4) Does the improper integral $\displaystyle\int_1^\infty \frac{10^{10}dx}{e^x+1}$ converge?

$$\lim_{x\to\infty}\frac{10^{10}}{e^x+1}\Big/\frac{1}{e^x} = \lim_{x\to\infty}\frac{10^{10}e^x}{e^x+1} = 10^{10}.$$

Also, $e \ge 2$ implies that for all $x \ge 1$, $e^x \ge x^2$. So, $e^{-x} \le x^{-2}$.

Since $\displaystyle\int_1^\infty \frac{dx}{x^2}$ converges, $\displaystyle\int_1^\infty \frac{dx}{e^x}$ also converges.

By limit comparison test, the given improper integral converges.

□

(1.29) *Example*

Show that $\int_1^\infty \frac{\sin x}{x^p}\,dx$ converges for all $p > 0$.

For $p > 1$ and $x \ge 1$, $\left|\frac{\sin x}{x^p}\right| \le \frac{1}{x^p}$. Since $\int_1^\infty \frac{dx}{x^p}$ converges, $\int_1^\infty \left|\frac{\sin x}{x^p}\right| dx$ converges. By (1.27), $\int_1^\infty \frac{\sin x}{x^p}\,dx$ converges.

For $0 < p \le 1$, use integration by parts:

$$\int_1^b \frac{\sin x}{x^p}\,dx = -\frac{\cos b}{b^p} + \frac{\cos 1}{1^p} + p\int_1^b \frac{\cos x}{x^{p+1}}\,dx.$$

Taking the limit as $b \to \infty$, we see that the first term goes to 0; the second term is already a real number, the third term, an improper integral converges as in the case for $p > 1$ above.

Therefore, the given improper integral converges. □

(1.30) *Example*

Show that $\int_0^\infty \frac{\sin x}{x^p}\,dx$ converges for $0 < p \le 1$.

For $p = 1$, the integral $\int_0^1 \frac{\sin x}{x}\,dx$ is not an improper integral. Since $\frac{\sin x}{x}$ with its value at 0 as 1 is continuous on $[0, 1]$, this integral exists.

For $0 < p < 1$ and $0 < x \le 1$, since $\frac{\sin x}{x^p} \le \frac{1}{x^p}$ and $\int_0^1 \frac{dx}{x^p}$ converges due to (1.24), the improper integral $\int_0^1 \frac{\sin x}{x^p}\,dx$ converges.

Next, the improper integral $\int_1^\infty \frac{\sin x}{x}\,dx$ converges due to (1.29).

Hence $\int_0^\infty \frac{\sin x}{x^p}\,dx = \int_0^1 \frac{\sin x}{x}\,dx + \int_1^\infty \frac{\sin x}{x}\,dx$ converges. □

(1.31) *Example*

Show that $\Gamma(x) = \int_0^\infty e^{-t}t^{x-1}\,dt$ converges for each $x > 0$.

Fix $x > 0$. Since $\lim_{t\to\infty} e^{-t}t^{x+1} = 0$, there exists $t_0 \ge 1$ such that $0 < e^{-t}t^{x+1} < 1$ for $t > t_0$. That is,

$$0 < e^{-t}t^{x-1} < t^{-2} \quad \text{for} \quad t > t_0.$$

Since $\int_1^\infty t^{-2}\,dt$ is convergent, $\int_{t_0}^\infty t^{-2}\,dt$ is also convergent. By the comparison test,

$$\int_{t_0}^{\infty} e^{-t}t^{x-1}\,dt \text{ is convergent.}$$

The integral $\int_1^{t_0} e^{-t}t^{x-1}\,dt$ exists and is not an improper integral.

Next, we consider the improper integral $\int_0^1 e^{-t}t^{x-1}\,dt$. Let $0 < a < 1$.

For $a \le t \le 1$, we have $0 < e^{-t}t^{x-1} < t^{x-1}$. So,

$$\int_a^1 e^{-t}t^{x-1}\,dt < \int_a^1 t^{x-1}\,dt = \frac{1-a^x}{x} < \frac{1}{x}.$$

Taking the limit as $a \longrightarrow 0+$, we see that

$$\int_0^1 e^{-t}t^{x-1}\,dt \text{ is convergent,}$$

and its value is less than or equal to $1/x$. Therefore,

$$\int_0^\infty e^{-t}t^{x-1}\,dt = \int_0^1 e^{-t}t^{x-1}\,dt + \int_1^{t_0} e^{-t}t^{x-1}\,dt + \int_{t_0}^\infty e^{-t}t^{x-1}\,dt$$

and the integral is convergent. ∎

The function $\Gamma(x)$ is called the **gamma function**; it is defined on $(0,\infty)$. For $x > 0$, using integration by parts, we have

$$\Gamma(x+1) = \int_0^\infty t^x e^{-t}\,dt = \Big[t^x(-e^{-t})\Big]_0^\infty - \int_0^\infty x t^{x-1}(-e^{-t})\,dt = x\Gamma(x).$$

It thus follows that $\Gamma(n+1) = n!$ for any non-negative integer n. We take $0! = 1$.

(1.32) ***Example***

Test the convergence of $\displaystyle\int_{-\infty}^{\infty} e^{-t^2}\,dt$.

Since e^{-t^2} is continuous on $[-1,1]$, $\displaystyle\int_{-1}^{1} e^{-t^2}\,dt$ exists.

For $t > 1$, we have $t < t^2$. So, $0 < e^{-t^2} < e^{-t}$. Since $\displaystyle\int_1^\infty e^{-t}\,dt$ is convergent, by Comparison test, $\displaystyle\int_1^\infty e^{-t^2}\,dt$ is convergent.

Now, $\displaystyle\int_{-a}^{-1} e^{-t^2}\,dt = \int_a^1 e^{-t^2}\,d(-t) = \int_1^a e^{-t^2}dt$. Taking limit as $a \longrightarrow \infty$, we see that $\displaystyle\int_{-\infty}^{1} e^{-t^2}\,dt$ is convergent and its value is equal to $\displaystyle\int_1^\infty e^{-t^2}\,dt$.

Combining the three integrals above, we conclude that $\int_{-\infty}^{\infty} e^{-t^2}\,dt$ converges. □

The Gamma function takes other forms by substitution of the variable of integration. Substituting t by rt we have

$$\Gamma(x) = r^x \int_0^\infty e^{-rt} t^{x-1}\,dt \quad \text{for } 0 < r,\ 0 < x.$$

Substituting t by t^2, we have

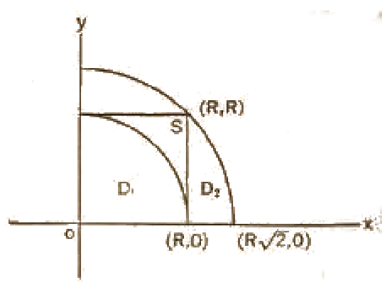
$$\Gamma(x) = 2\int_0^\infty e^{-t^2} t^{2x-1}\,dt \quad \text{for } 0 < x.$$

(1.33) ***Example***

Show that $\Gamma(\frac{1}{2}) = \sqrt{\pi}$.

$$\Gamma\left(\frac{1}{2}\right) = 2\int_0^\infty e^{-t^2}\,dt.$$

To evaluate this integral, consider the double integral of $e^{-x^2-y^2}$ over two circular sectors D_1 and D_2, and the square S as indicated below.



Since the integrand is positive, we have $\iint_{D_1} < \iint_S < \iint_{D_2}$.

Now, evaluate these integrals by converting them to iterated integrals as follows:

$$\int_0^R e^{-r^2} r\,dr \int_0^{\pi/2} d\theta < \int_0^R e^{-x^2}\,dx \int_0^R e^{-y^2}\,dy < \int_0^{R\sqrt{2}} e^{-r^2} r\,dr \int_0^{\pi/2} d\theta$$

$$\frac{\pi}{4}(1 - e^{-R^2}) < \left(\int_0^R e^{-x^2}\,dx\right)^2 < \frac{\pi}{4}(1 - e^{-2R^2})$$

Take the limit as $R \to \infty$ to obtain

$$\left(\int_0^\infty e^{-x^2}\,dx\right)^2 = \frac{\pi}{4}$$

From this, the result follows. ∎

(1.34) *Example*

Prove: $B(x,y) = \int_0^1 t^{x-1}(1-t)^{y-1}\,dt$ converges for $x > 0, y > 0$.

We write the integral as a sum of two integrals:

$$B(x,y) = \int_0^{1/2} t^{x-1}(1-t)^{y-1}\,dt + \int_{1/2}^1 t^{x-1}(1-t)^{y-1}\,dt$$

Setting $u = 1 - t$, the second integral looks like

$$\int_{1/2}^1 t^{x-1}(1-t)^{y-1}\,dt = \int_0^{1/2} u^{y-1}(1-u)^{x-1}\,dt$$

Therefore, it is enough to show that the first integral converges. Notice that here, $0 < t \leq 1/2$.

Case 1: $x \geq 1$.

For $0 < t < 1/2$, $1 - t > 0$. So, for all $y > 0$, the function $(1-t)^{y-1}$ is well defined, continuous, and bounded on $(0, 1/2]$. So is the function t^{x-1}. Therefore, the integral $\int_0^{1/2} t^{x-1}(1-t)^{y-1}\,dt$ exists and is not an improper integral.

Case 2: $0 < x < 1$.

Here, the function t^{x-1} is well defined and continuous on $(0, 1/2]$. By (1.24), the integral $\int_0^{1/2} t^{x-1}\,dt$ converges. Notice that $(1-t)^{y-1} \leq 1$ for $y - 1 \geq 0$, and $(1-t)^{y-1} \leq (1/2)^{y-1}$ for $y - 1 < 0$. So, there exists a constant c depending on the given value of y such that $t^{x-1}(1-t)^{y-1} \leq c\,t^{x-1}$ for $0 < t \leq 1/2$. We thus see that $\int_0^{1/2} t^{x-1}(1-t)^{y-1}\,dt$ converges. ∎

The function $B(x,y)$ for $x > 0,\ y > 0$ is called the **beta function**.

By setting t as $1 - t$, we see that $B(x,y) = B(y,x)$.

By substituting t with $\sin^2 t$, the Beta function can be written as

$$B(x,y) = 2\int_0^{\pi/2} (\sin t)^{2x-1}(\cos t)^{2y-1}\,dt, \quad \text{for } x > 0,\ y > 0.$$

Changing the variable t to $t/(1+t)$, the Beta function can be written as

$$B(x,y) = \int_0^\infty \frac{t^{x+1}}{(1+t)^{x+y}}\,dt \quad \text{for } x > 0,\ y > 0.$$

Again, using multiple integrals it can be shown that

$$B(x,y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)} \quad \text{for } x > 0,\ y > 0.$$

1.7 Tests of Convergence for Series

Recall that a function $f(x)$ is called a *decreasing* function iff $f(s) \geq f(t)$ for any $s \leq t$ in the domain of $f(x)$.

We connect the convergence of improper integrals to the convergence of series as follows.

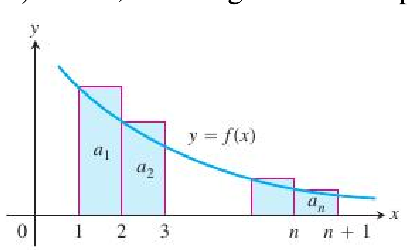
(1.35) *Theorem* (Integral Test)

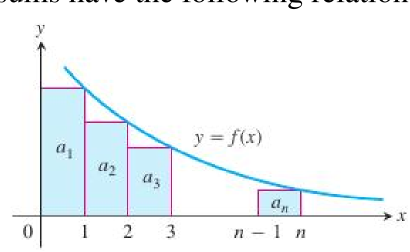
Let $\sum a_n$ be a series of positive terms. Let $f : [1,\infty) \to \mathbb{R}$ be a continuous, positive and decreasing function such that $a_n = f(n)$ for each $n \in \mathbb{N}$.

(1) *If $\int_1^\infty f(t)dt$ is convergent, then $\sum a_n$ is convergent.*

(2) *If $\int_1^\infty f(t)dt$ diverges to ∞, then $\sum a_n$ diverges to ∞.*

Proof. Since $f(t)$ is positive and decreasing, in any closed interval $[k, k+1]$ for $k \in \mathbb{N}$, the maximum value of $f(t)$ is $f(k)$ and the minimum value of $f(t)$ is $f(k+1)$. Thus, the integrals and the partial sums have the following relation:





$$\begin{aligned}
\int_1^{n+1} f(t)\,dt &= \int_1^2 f(t)\,dt + \int_2^3 f(t)\,dt + \cdots + \int_n^{n+1} f(t)\,dt \\
&\le \int_1^2 f(1)\,dt + \int_2^3 f(2)\,dt + \cdots + \int_n^{n+1} f(n)\,dt \\
&= f(1) + f(2) + \cdots f(n) \\
&= f(1) + \int_1^2 f(2)\,dt + \int_2^3 f(3)\,dt + \cdots + \int_{n-1}^n f(n)\,dt \\
&\le f(1) + \int_1^2 f(t)\,dt + \int_2^3 f(t)\,dt + \cdots + \int_{n-1}^n f(n)\,dt \\
&= f(1) + \int_1^n f(t)\,dt.
\end{aligned}$$

As $f(1) = a_1, \ldots, f(n) = a_n$, we obtain

$$\int_1^{n+1} f(t)\,dt \;\le\; a_1 + a_2 + \cdots + a_n \;\le\; a_1 + \int_1^n f(t)\,dt.$$

If $\displaystyle\int_1^n f(t)\,dt$ is finite, then the right hand inequality shows that the sequence of partial sums $a_1 + \cdots + a_n$ of the series $\sum a_n$ is an increasing sequence having an upper bound. Hence, this sequence converges; that is, the series $\sum a_n$ is convergent.

If $\displaystyle\int_1^n f(t)\,dt = \infty$, then the left hand inequality similarly shows that $\sum a_n$ diverges to ∞. ∎

We remark that in the above theorem, the hypothesis that $f(x)$ is positive is taken for convenience. Even if $f(x)$ is not positive, under the rest of the hypotheses the same conclusion can be obtained. The reason is, $f(x)$ is continuous and decreasing implies that if it is not positive throughout its domain $[1, \infty)$, then either it is negative on $[1, \infty)$ or there exists $a \in [1, \infty)$ such that $f(x)$ is positive on $[1, a)$ and negative on (a, ∞). Correspondingly, the sequence $(f(n))$ will have a tail consisting of negative terms. Then convergence of the improper integral will imply the convergence of the series. The second conclusion will be reformulated as "if the improper integral diverges to $\pm\infty$, then the series will diverge to $\pm\infty$".

Notice that when the series converges, the value of the integral can be different from the sum of the series. Moreover, Integral test assumes implicitly that (a_n) is a decreasing sequence. Further, the integral test is also applicable when the interval of integration is $[m, \infty)$ instead of $[1, \infty)$.

(1.36) *Example*

Show that $\sum_{n=1}^{\infty} \frac{1}{n^p}$ converges for $p > 1$ and diverges for $p \le 1$.

For $p = 1$, the series is the harmonic series; and it diverges. Suppose $p \neq 1$. Consider the function $f(t) = 1/t^p$ from $[1, \infty)$ to $\mathbb{R}$. This is a continuous, positive and decreasing function.

$$\int_1^\infty \frac{1}{t^p}\,dt = \lim_{b\to\infty} \frac{t^{-p+1}}{-p+1}\Big|_1^b = \frac{1}{1-p} \lim_{b\to\infty} \left(\frac{1}{b^{p-1}} - 1\right) = \begin{cases} \frac{1}{p-1} & \text{if } p > 1 \\ \infty & \text{if } p < 1. \end{cases}$$

Then the Integral test proves the statement. ∎

Note that for $p > 1$, the sum of the series $\sum n^{-p}$ need not be equal to $1/(p-1)$.

(1.37) *Example*

Does the series $\sum_{n=1}^{\infty} \frac{n+7}{n(n+3)\sqrt{n+5}}$ converge?

Let $a_n = \frac{n+7}{n(n+3)\sqrt{n+5}}$ and $b_n = \frac{1}{n^{3/2}}$. Then

$$\frac{a_n}{b_n} = \frac{\sqrt{n}(n+7)}{(n+3)\sqrt{n+5}} \to 1 \text{ as } n \to \infty.$$

Since $\sum \frac{1}{n^{3/2}}$ is convergent, Limit comparison test says that the given series is convergent. ∎

(1.38) *Example*

Show that the series $\sum_{n=2}^{\infty} \frac{1}{n(\log n)^\alpha}$ converges for $\alpha > 1$ and diverges to ∞ for $\alpha \le 1$.

The function $f(x) = \frac{1}{x(\log x)^\alpha}$ is continuous, positive, and decreasing on $[2, \infty)$. By the integral test, it converges when $\int_2^\infty \frac{1}{x(\log x)^\alpha}\,dx$ converges. Evaluating the integral, we have

$$\int_2^\infty \frac{1}{x(\log x)^\alpha}\,dx = \int_{\log 2}^\infty \frac{1}{t^\alpha}\,dt.$$

As in (1.36), we conclude that the series converges for $\alpha > 1$ and diverges to ∞ for $\alpha \le 1$. ∎

There are various ways to determine whether a series converges or not; occasionally, some information on its sum is also obtained.

(1.39) ***Theorem*** (Comparison Test)

Let $\sum a_n$ and $\sum b_n$ be series of non-negative terms. Suppose there exists $k > 0$ such that $0 \le a_n \le kb_n$ for each $n > m$ for some $m \in \mathbb{N}$.

(1) *If $\sum b_n$ converges, then $\sum a_n$ converges.*

(2) *If $\sum a_n$ diverges to ∞, then $\sum b_n$ diverges to ∞.*

Proof. (1) Consider all partial sums of the series having more than m terms. We see that

$$a_1 + \cdots + a_m + a_{m+1} + \cdots + a_n \le a_1 + \cdots + a_m + k \sum_{j=m+1}^{n} b_j.$$

Since $\sum b_n$ converges, so does $\sum_{j=m+1}^{n} b_j$. And then $\sum a_n$ is an increasing bounded sequence; so it converges.

(2) If $\sum b_n$ is convergent, then by (a), $\sum a_n$ would become convergent! ∎

Caution: The comparison test holds for series of non-negative terms.

(1.40) ***Theorem*** (Ratio Comparison Test)

Let $\sum a_n$ and $\sum b_n$ be series of non-negative terms. Suppose there exists $m \in \mathbb{N}$ such that for each $n > m$, $a_n > 0$, $b_n > 0$, and $\dfrac{a_{n+1}}{a_n} \le \dfrac{b_{n+1}}{b_n}$.

(1) *If $\sum b_n$ converges, then $\sum a_n$ converges.*

(2) *If $\sum a_n$ diverges to ∞, then $\sum b_n$ diverges to ∞.*

Proof. For $n > m$,

$$a_n = \frac{a_n}{a_{n-1}} \frac{a_{n-1}}{a_{n-2}} \cdots \frac{a_{m+2}}{a_{m+1}} a_{m+1} \le \frac{b_n}{b_{n-1}} \frac{b_{n-1}}{b_{n-2}} \cdots \frac{b_{m+2}}{b_{m+1}} a_{m+1} = \frac{a_{m+1}}{b_{m+1}} b_n.$$

By (1.39), if $\sum b_n$ converges, then $\sum a_n$ converges. This proves (1). And, (2) follows from (1) by contradiction. ∎

(1.41) ***Theorem*** (Limit Comparison Test)

Let $\sum a_n$ and $\sum b_n$ be series of non-negative terms. Suppose there exists $m \in \mathbb{N}$ such that for each $n > m$, $a_n > 0$, $b_n > 0$, and $\lim\limits_{n\to\infty} \dfrac{a_n}{b_n} = k$.

(1) *If $k > 0$ then $\sum b_n$ and $\sum a_n$ converge or diverge to ∞, together.*

(2) *If $k = 0$ and $\sum b_n$ converges, then $\sum a_n$ converges.*

(3) *If $k = \infty$ and $\sum b_n$ diverges to ∞ then $\sum a_n$ diverges to ∞.*

Proof. (1) Let $\epsilon = k/2 > 0$. The limit condition implies that there exists $M \in \mathbb{N}$ such that

$$\frac{k}{2} < \frac{a_n}{b_n} < \frac{3k}{2} \quad \text{for each } n > M > m.$$

By the Comparison test, the conclusion is obtained.

(2) Let $\epsilon = 1$. The limit condition implies that there exists $M \in \mathbb{N}$ such that

$$-1 < \frac{a_n}{b_n} < 1 \quad \text{for each } n > M > m.$$

Using the right hand inequality and the Comparison test we conclude that convergence of $\sum b_n$ implies the convergence of $\sum a_n$.

(3) If $k = \infty$, $\lim(b_n/a_n) = 0$. Use (2). ∎

(1.42) ***Example***

Do the series (a) $\displaystyle\sum_{n=1}^{\infty} \frac{\log n}{n^{3/2}}$ (b) $\displaystyle\sum_{n=1}^{\infty} \frac{1 + n \log n}{1 + n^2}$ converge?

(a) Take $a_n = \dfrac{\log n}{n^{3/2}}$ and $b_n = \dfrac{1}{n^{5/4}}$. Then

$$\lim_{n\to\infty} \frac{a_n}{b_n} = \lim_{n\to\infty} \frac{\log n}{n^{1/4}} = \lim_{x\to\infty} \frac{\log x}{x^{1/4}} = \lim_{x\to\infty} \frac{1/x}{(1/4)x^{-3/4}} = \lim_{x\to\infty} \frac{4}{x^{1/4}} = 0.$$

Since $\sum b_n$ converges, by the Limit comparison test, $\sum a_n$ converges.

(b) Take $a_n = \dfrac{1 + n \log n}{1 + n^2}$ and $b_n = \dfrac{1}{n}$. Then

$$\lim_{n\to\infty} \frac{a_n}{b_n} = \lim_{n\to\infty} \frac{n + n^2 \log n}{1 + n^2} = \infty.$$

As $\sum b_n$ diverges to ∞, by the Limit comparison test, $\sum a_n$ diverges to ∞. □

(1.43) ***Theorem*** (D' Alembert Ratio Test)

Let $\sum a_n$ *be a series of positive terms. Suppose* $\displaystyle\lim_{n\to\infty} \frac{a_{n+1}}{a_n} = \ell$.

(1) *If* $\ell < 1$, *then* $\sum a_n$ *converges.*

(2) *If* $\ell > 1$ *or* $\ell = \infty$, *then* $\sum a_n$ *diverges to* ∞.

(3) *If* $\ell = 1$, *then no conclusion is obtained.*

Proof. (1) Given that $\lim(a_{n+1}/a_n) = \ell < 1$. Choose δ such that $\ell < \delta < 1$. There exists $m \in \mathbb{N}$ such that for each $n > m$, $a_{n+1}/a_n < \delta$. Then

$$\frac{a_n}{a_{m+1}} = \frac{a_n}{a_{n-1}} \frac{a_{n-1}}{a_{n-2}} \cdots \frac{a_{m+2}}{a_{m+1}} < \delta^{n-m+1}.$$

Thus, $a_n < \delta^{n-m+1} a_{m+1}$. Consequently,

$$a_{m+1} + a_{m+2} + \cdots + a_n < a_{m+1}(1 + \delta + \delta^2 + \cdots \delta^{n-m+1}).$$

Since $\delta < 1$, this approaches a limit as $n \to \infty$. Therefore, the series

$$a_{m+1} + a_{m+2} + \cdots a_n + \cdots$$

converges. In that case, the series $\sum a_n = (a_1 + \cdots + a_m) + a_{m+1} + a_{m+2} + \cdots$ converges.

(2) Given that $\lim(a_{n+1}/a_n) = \ell > 1$. There exists $m \in \mathbb{N}$ such that for each $n > m$, $a_{n+1} > a_n$. Then

$$a_{m+1} + a_{m+2} + \cdots + a_n > a_{m+1}(n - m).$$

Since $a_{m+1} > 0$, this approaches ∞ as $n \to \infty$. Therefore, the series

$$a_{m+1} + a_{m+2} + \cdots a_n + \cdots$$

diverges to ∞. In that case, the series $\sum a_n = (a_1 + \cdots + a_m) + a_{m+1} + a_{m+2} + \cdots$ diverges to ∞. The other case of $\ell = \infty$ is similar.

(3) The series $\sum(1/n)$ diverges to ∞. Here, $\lim(a_{n+1}/a_n) = \lim(n/(n+1)) = 1$. But the series $\sum(1/n^2)$ is convergent although $\lim(a_{n+1}/a_n) = 1$. ∎

(1.44) ***Example***

Does the series $\displaystyle\sum_{n=1}^{\infty} \frac{n!}{n^n}$ converge?

Write $a_n = n!/(n^n)$. Then

$$\frac{a_{n+1}}{a_n} = \frac{(n+1)!n^n}{(n+1)^{n+1}\, n!} = \left(\frac{n}{n+1}\right)^n \to \frac{1}{e} < 1 \text{ as } n \to \infty.$$

By D' Alembert's ratio test, the series converges.

Then it follows that the sequence $\big(n!/n^n\big)$ converges to 0. □

(1.45) ***Example***

By Ratio test, it follows that the series

$$1 + 1 + \frac{1}{2!} + \frac{1}{3!} + \cdots + \frac{1}{n!} + \cdots$$

is convergent. In fact, this series converges to e. To see this, consider

$$s_n = 1 + 1 + \frac{1}{2!} + \cdots + \frac{1}{n!}, \quad t_n = \left(1 + \frac{1}{n}\right)^n.$$

By the Binomial theorem,

$$t_n = 1 + 1 + \frac{1}{2!}\Big(1 - \frac{1}{n}\Big) + \cdots + \frac{1}{n!}\Big[\Big(1 - \frac{1}{n}\Big)\Big(1 - \frac{2}{n}\Big)\cdots\Big(1 - \frac{n-1}{n}\Big)\Big] \le s_n.$$

Thus taking limit as $n \to \infty$, we have

$$e = \lim_{n\to\infty} t_n \le \lim_{n\to\infty} s_n.$$

Also, for $n > m$, where m is any fixed natural number,

$$t_n \ge 1 + 1 + \frac{1}{2!}\Big(1 - \frac{1}{n}\Big) + \cdots + \frac{1}{m!}\Big[\Big(1 - \frac{1}{n}\Big)\Big(1 - \frac{2}{n}\Big)\cdots\Big(1 - \frac{m-1}{n}\Big)\Big]$$

Taking limit as $n \to \infty$ we have

$$e = \lim_{n\to\infty} t_n \ge s_m.$$

Since m is arbitrary, taking the limit as $m \to \infty$, we have

$$e \ge \lim_{m\to\infty} s_m.$$

Therefore, $\lim_{m\to\infty} s_m = e$. That is, $\sum_{n=0}^{\infty} \frac{1}{n!} = e$. ∎

(1.46) ***Example***

Does the series $\sum_{n=1}^{\infty} \frac{4^n (n!)^2}{(2n)!}$ converge?

Let $a_n = \frac{4^n (n!)^2}{(2n)!}$. We have $\frac{a_{n+1}}{a_n} = \frac{4^{n+1}((n+1)!)^2}{(2(n+1))!}\,\frac{(2n)!}{4^n (n!)^2} = \frac{2(n+1)}{2n+1}$.

Since its limit is equal to 1, the Ratio test fails. However, $\frac{a_{n+1}}{a_n} > 1$. Since $a_1 = 2$, we see that each $a_n > 2$. That is, the sequence (a_n) does not converge to 0. Therefore, the series diverges. Since it is a series of positive terms, it diverges to ∞. ∎

(1.47) ***Theorem*** (Cauchy Root Test)

Let $\sum a_n$ be a series of positive terms. Suppose $\lim_{n\to\infty} (a_n)^{1/n} = \ell$.

(1) *If $\ell < 1$, then $\sum a_n$ converges.*

(2) *If $\ell > 1$ or $\ell = \infty$, then $\sum a_n$ diverges to ∞.*

(3) *If $\ell = 1$, then no conclusion is obtained.*

Proof. (1) Suppose $\ell < 1$. Choose δ such that $\ell < \delta < 1$. Due to the limit condition, there exists an $m \in \mathbb{N}$ such that for each $n > m$, $(a_n)^{1/n} < \delta$. That is, $a_n < \delta^n$. Since $0 < \delta < 1$, $\sum \delta^n$ converges. By Comparison test, $\sum a_n$ converges.
(2) Given that $\ell > 1$ or $\ell = \infty$, we see that $(a_n)^{1/n} > 1$ for infinitely many values of n. That is, the sequence (a_n) does not converge to 0. Therefore, $\sum a_n$ is divergent. It diverges to ∞ since it is a series of positive terms.
(3) Once again, for both the series $\sum(1/n)$ and $\sum(1/n^2)$, we see that $(a_n)^{1/n}$ has the limit 1. But one is divergent, the other is convergent. ∎

Remark 1.48 In fact, for a sequence (a_n) of positive terms if $\lim_{n\to\infty} \frac{a_{n+1}}{a_n}$ exists, then $\lim_{n\to\infty} (a_n)^{1/n}$ exists and the two limits are equal.

To see this, suppose $\lim_{n\to\infty} \frac{a_{n+1}}{a_n} = \ell$. Let $\epsilon > 0$. Then we have an $m \in \mathbb{N}$ such that for all $n > m$, $\ell - \epsilon < \frac{a_{n+1}}{a_n} < \ell + \epsilon$. Use the right side inequality first. For all such n, $a_n < (\ell + \epsilon)^{n-m} a_m$. Then

$$(a_n)^{1/n} < (\ell + \epsilon)\big((\ell + \epsilon)^{-m} a_m\big)^{1/n} \to \ell + \epsilon \text{ as } n \to \infty.$$

Therefore, $\lim (a_n)^{1/n} \le \ell + \epsilon$ for every $\epsilon > 0$. That is, $\lim (a_n)^{1/n} \le \ell$.
Similarly, the left side inequality gives $\lim (a_n)^{1/n} \ge \ell$.

Notice that this gives an alternative proof of (1.47).

(1.49) *Example*

Does the series $\sum_{n=0}^{\infty} 2^{(-1)^n - n} = 2 + \frac{1}{4} + \frac{1}{2} + \frac{1}{16} + \cdots$ converge?

Let $a_n = 2^{(-1)^n - n}$. Then

$$\frac{a_{n+1}}{a_n} = \begin{cases} 1/8 & \text{if } n \text{ even} \\ 2 & \text{if } n \text{ odd.} \end{cases}$$

Clearly, its limit does not exist. But

$$(a_n)^{1/n} = \begin{cases} 2^{1/n-1} & \text{if } n \text{ even} \\ 2^{-1/n-1} & \text{if } n \text{ odd} \end{cases}$$

This has limit $1/2 < 1$. Therefore, by Cauchy root test, the series converges. □

1.8 Alternating series

If the terms of a series have alternating signs, then the earlier tests are not applicable. For example, the methods discussed so far fail on deciding whether the series $\sum(-1)^n/n$ converges or not.

(1.50) ***Theorem*** (Leibniz Alternating Series Test)

Let (a_n) be a sequence of positive terms decreasing to 0; that is, for each n, $a_n \geq a_{n+1} > 0$, and $\lim_{n\to\infty} a_n = 0$. Then the series $\sum_{n=1}^{\infty}(-1)^{n+1}a_n$ converges, and its sum lies between $a_1 - a_2$ and a_1.

Proof. The partial sum upto $2n$ terms is

$$s_{2n} = (a_1-a_2)+(a_3-a_4)+\cdots+(a_{2n-1}-a_{2n}) = a_1-\big[(a_2-a_3)+\cdots+(a_{2n-2}-a_{2n-1})+a_{2n}\big].$$

It is a sum of n positive terms bounded above by a_1 and below by $a_1 - a_2$. Hence s_{2n} converges to some s with $a_1 - a_2 \leq s \leq a_1$.

The partial sum upto $2n + 1$ terms is $s_{2n+1} = s_{2n} + a_{2n+1}$. It converges to s as $\lim a_{2n+1} = 0$. Hence the series converges to some s with $a_1 - a_2 \leq s \leq a_1$. ∎

The bounds for s can be sharpened by taking $s_{2n} \leq s \leq s_{2n-1}$ for $n > 1$.

Leibniz test now implies that the series $1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} + \cdots$ is convergent to some s with $1/2 \leq s \leq 1$. By taking more terms, we can have different bounds such as

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} = \frac{7}{12} \leq s \leq 1 - \frac{1}{2} + \frac{1}{3} = \frac{10}{12}$$

In contrast, the harmonic series $1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \cdots$ diverges to ∞.

We say that the series $\sum a_n$ is **absolutely convergent** iff the series $\sum |a_n|$ is convergent.

An alternating series $\sum a_n$ is said to be **conditionally convergent** iff it is convergent but it is not absolutely convergent.

Thus for a series of non-negative terms, convergence and absolute convergence coincide. As we just saw, an alternating series may be convergent but not absolutely convergent.

The series $1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} + \cdots$ is a conditionally convergent series. It shows that the converse of the following theorem is not true.

(1.51) ***Theorem***

An absolutely convergent series is convergent.

Proof. Let $\sum a_n$ be an absolutely convergent series. Then $\sum |a_n|$ is convergent. Let $\epsilon > 0$. By Cauchy criterion, there exists an $n_0 \in \mathbb{N}$ such that for all $n > m > n_0$, we have

$$|a_m| + |a_{m+1}| + \cdots + |a_n| < \epsilon.$$

Now,

$$|a_m + a_{m+1} + \cdots + a_n| \leq |a_m| + |a_{m+1}| + \cdots + |a_n| < \epsilon.$$

Again, by Cauchy criterion, the series $\sum a_n$ is convergent. ∎

An absolutely convergent series can be rearranged in any way we like, but the sum remains the same. Whereas a rearrangement of the terms of a conditionally convergent series may lead to divergence or convergence to any other number. In fact, a conditionally convergent series can always be rearranged in a way so that the rearranged series converges to any desired number; we will not prove this fact.

(1.52) ***Example***

Do the series (a) $\displaystyle\sum_{n=1}^{\infty}(-1)^{n+1}\frac{1}{2^n}$ (b) $\displaystyle\sum_{n=1}^{\infty}\frac{\cos n}{n^2}$ converge?

(a) $\sum(2)^{-n}$ converges. Therefore, the given series converges absolutely; hence it converges.

(b) $\left|\frac{\cos n}{n^2}\right| \leq \frac{1}{n^2}$ and $\sum(n^{-2})$ converges. By comparision test, the given series converges absolutely; and hence it converges. □

(1.53) ***Example***

Discuss the convergence of the series $\displaystyle\sum_{n=1}^{\infty}\frac{(-1)^{n+1}}{n^p}$.

For $p > 1$, the series $\sum n^{-p}$ converges. Therefore, the given series converges absolutely for $p > 1$.

For $0 < p \leq 1$, by Leibniz test, the series converges. But $\sum n^{-p}$ does not converge. Therefore, the given series converges conditionally for $0 < p \leq 1$.

For $p \leq 0$, $\lim \frac{(-1)^{n+1}}{n^p} \neq 0$. Therefore, the given series diverges in this case. □

(1.54) ***Example***

Does the series $1 - \frac{1}{4} - \frac{1}{16} + \frac{1}{9} + \frac{1}{25} + \frac{1}{49} - \cdots$ converge?

Here, the series has been made up from the terms $1/n^2$ by taking first one term, next two negative terms of squares of next even numbers, then three positive terms which are squares of next three odd numbers, and so on. This is a rearrangement of the series

$$1 - \frac{1}{2^2} + \frac{1}{3^2} - \frac{1}{4^2} + \frac{1}{5^2} - \frac{1}{6^2} + \cdots$$

which is absolutely convergent (since $\sum(1/n^2)$ is convergent). Therefore, the given series is convergent and its sum is the same as that of the alternating series $\sum(-1)^{n+1}(1/n^2)$. $\Box$

1.9 Exercises for Chapter 1

1. Show the following:

   (a) $\lim_{n\to\infty} \frac{\ln n}{n} = 0.$ (b) $\lim_{n\to\infty} n^{1/n} = 1.$ (c) $\lim_{n\to\infty} x^n = 0$ for $|x| < 1.$

   (d) $\lim_{n\to\infty} \frac{n^p}{x^n} = 0$ for $x > 1.$ (e) $\lim_{n\to\infty} \frac{x^n}{n!} = 0$ (f) $\lim_{n\to\infty} \left(1 + \frac{x}{n}\right)^n = e^x$

2. Prove the following:

   (a) It is not possible that a series converges to a real number ℓ and also diverges to $-\infty$.
   (b) It is not possible that a series diverges to ∞ and also to $-\infty$.

3. Prove the following:

   (a) If both the series $\sum a_n$ and $\sum b_n$ converge, then the series $\sum(a_n + b_n)$, $\sum(a_n - b_n)$ and $\sum ka_n$ converge; where k is any real number.
   (b) If $\sum a_n$ converges and $\sum b_n$ diverges to $\pm\infty$, then $\sum(a_n + b_n)$ diverges to $\pm\infty$, and $\sum(a_n - b_n)$ diverges to $\mp\infty$.
   (c) If $\sum a_n$ diverges to $\pm\infty$, and $k > 0$, then $\sum ka_n$ diverges to $\pm\infty$.
   (d) If $\sum a_n$ diverges to $\pm\infty$, and $k < 0$, then $\sum ka_n$ diverges to $\mp\infty$.

4. Give examples for the following:

   (a) $\sum a_n$ and $\sum b_n$ both diverge, but $\sum(a_n + b_n)$ converges to a nonzero number.
   (b) $\sum a_n$ and $\sum b_n$ both diverge, and $\sum(a_n + b_n)$ diverges to ∞.
   (c) $\sum a_n$ and $\sum b_n$ both diverge, and $\sum(a_n + b_n)$ diverges to $-\infty$.

5. Show that the sequence 1, 1.1, 1.1011, 1.10110111, 1.1011011101111... converges.
6. Determine whether the following series converge:

   (a) $\displaystyle\sum_{n=1}^{\infty} \frac{-n}{3n+1}$ (b) $\displaystyle\sum_{n=1}^{\infty} \frac{\ln n}{n^{3/2}}$ (c) $\displaystyle\sum_{n=1}^{\infty} \frac{1+n\ln n}{1+n^2}$

   Ans: (a) converges (b) diverges (c) diverges to ∞.
7. Test for convergence the series $\frac{1}{3} + \left(\frac{2}{3}\right)^2 + \left(\frac{3}{7}\right)^3 + \cdots + \left(\frac{n}{2n+1}\right)^n + \cdots$.
   Ans: converges.
8. Is the integral $\displaystyle\int_{-\infty}^{\infty} \frac{1}{1+x^2}\,dx$ convergent? *Ans*: Yes.
9. Is the area under the curve $y = (\ln x)/x^2$ for $1 \le x < \infty$ finite? *Ans*: Yes.
10. Evaluate (a) $\displaystyle\int_0^3 \frac{dx}{(x-1)^{2/3}}$ (b) $\displaystyle\int_0^3 \frac{dx}{x-1}$

    Ans: (a) $3(1+2^{1/3})$ (b) does not converge.
11. Show that $\displaystyle\int_1^{\infty} \frac{\sin x}{x^p}\,dx$ converges for all $p > 0$.
12. Show that $\displaystyle\int_0^{\infty} \frac{\sin x}{x^p}\,dx$ converges for $0 < p \le 1$.
13. Show that the series $\displaystyle\sum_{n=2}^{\infty} \frac{1}{n(\ln n)^{\alpha}}$ converges for $\alpha > 1$ and diverges to ∞ for $\alpha \le 1$.
14. Does the series $\displaystyle\sum_{n=1}^{\infty} \frac{4^n (n!)^2}{(2n)!}$ converge? *Ans*: diverges to ∞.
15. Does the series $1 - \frac{1}{4} - \frac{1}{16} + \frac{1}{9} + \frac{1}{25} + \frac{1}{49} - \cdots$ converge? *Ans*: Yes.
16. Let (a_n) be a sequence of positive terms. Show that if $\sum_{n=1}^{\infty} \frac{a_n}{1+a_n}$ converges, then $\sum_{n=1}^{\infty} a_n$ converges.
17. Let (a_n) be a sequence of positive decreasing terms. Show that if $\sum_{n=1}^{\infty} a_n$ converges, then the sequence (na_n) converges to 0.

2

Series Representation of Functions

2.1 Power Series

A power series apparently is a generalization of a polynomial. A polynomial in x looks like

$$p(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n.$$

A power series is an infinite sum of the same form. The question is though a polynomial defines a function for $x \in \mathbb{R}$, when does a power series define a function? That is, for what values of x, a power series sums to a number?

Let $c \in \mathbb{R}$. A **power series about** $x = c$ is a series of the form

$$\sum_{n=0}^{\infty} a_n (x-c)^n = a_0 + a_1(x-c) + a_2(x-c)^2 + \cdots$$

The point c is called the **center** of the power series and the real numbers $a_0, a_1, \cdots, a_n, \cdots$ are its **coefficients**.

When $x = c$ and $n = 0$, we agree to read the term $a_0(x-c)^0$ as a_0. This will save space in writing a power series.

For example, the geometric series

$$1 + x + x^2 + \cdots + x^n + \cdots$$

is a power series about $x = 0$ with each coefficient as 1. We know that its sum is $\dfrac{1}{1-x}$ for $-1 < x < 1$. And we know that for $|x| \geq 1$, the geometric series does not converge. That is, the series defines a function from $(-1, 1)$ to $\mathbb{R}$ and it is not meaningful for other values of x.

(2.1) *Example*

Show that the following power series converges for $0 < x < 4$.

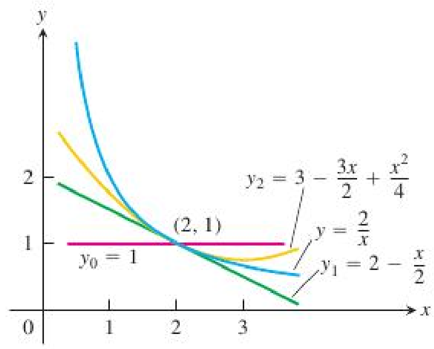
$$1 - \frac{1}{2}(x-2) + \frac{1}{4}(x-2)^2 + \cdots + \frac{(-1)^n}{2^n}(x-2)^n + \cdots$$

It is a geometric series with the ratio as $r = (-1/2)(x-2)$. Thus it converges for $|(-1/2)(x-2)| < 1$. Simplifying we get the constraint as $0 < x < 4$.

Notice that the power series sums to

$$\frac{1}{1-r} = \frac{1}{1 - \frac{-1}{2(x-2)}} = \frac{2}{x}.$$

Thus, the power series gives a series expansion of the function $\frac{2}{x}$ for $0 < x < 4$. Truncating the series to n terms give us polynomial approximations of $\frac{2}{x}$. ∎



A fundamental result for the power series is the following, which we state and prove for power series about the point 0. Results on power series about any point c can be obtained from this particular case in a similar manner or with the substitution $y = x - c$.

(2.2) ***Theorem*** (Convergence Theorem for Power Series)

Suppose the power series $\sum_{n=0}^{\infty} a_n x^n$ *is convergent for* $x = \alpha$ *and divergent for* $x = \beta$ *for some* $\alpha > 0,\ \beta > 0$. *Then the power series converges absolutely for all* x *with* $|x| < \alpha$; *and it diverges for all* x *with* $|x| > \beta$.

Proof. The power series converges for $x = \alpha$ means that $\sum a_n \alpha^n$ converges. Thus $\lim\limits_{n\to\infty} a_n \alpha^n = 0$. Then we have an $M \in \mathbb{N}$ such that for all $n > M$, $|a_n \alpha^n| < 1$. Let $x \in \mathbb{R}$ be such that $|x| < \alpha$. Write $t = |\frac{x}{\alpha}|$. We have

$$\text{for each } n > M,\ |a_n x^n| = |a_n \alpha^n|\, |\tfrac{x}{\alpha}|^n < t^n.$$

As $0 \le t < 1$, the geometric series $\sum_{n=m+1}^{\infty} t^n$ converges. By comparison test, $\sum_{n=m+1}^{\infty} |a_n x^n|$ converges. Adding to it some finite terms, it follows that $\sum_{n=0}^{\infty} |a_n x^n|$ converges. That is, the power series converges absolutely for all x with $|x| < \alpha$.

For the divergence part of the theorem, suppose, on the contrary that the power series converges for some $c > \beta$. By the convergence part, the series must converge for $x = \beta$, a contradiction. ∎

Notice that if the power series is about a point $x = c$, then we take $y = x - c$ and apply (2.2). Also, for $x = 0$, the power series $\sum a_n x^n$ always converges.

In view of (2.2), we introduce some terminology.

Consider the power series $\sum_{n=0}^{\infty} a_n (x-c)^n$. The real number

$$R = \text{lub}\{\alpha \ge 0 : \text{ the power series converges for all } x \text{ with } |x-c| < \alpha\}$$

is called the **radius of convergence** of the power series.

That is, R is such a non-negative number that the power series converges for all x with $|x-c| < R$ and it diverges for all x with $|x-c| > R$. Further, $R = \infty$ iff the power series converges for all $x \in \mathbb{R}$.

If the radius of convergence of the power series $\sum a_n (x-c)^n$ is R, then the **interval of convergence** of the power series is

$[c-R, c+R]$ iff it converges at both $x = c-R$ and $x = c+R$.

$(c-R, c+R)$ iff it diverges at both $x = c-R$ and $x = c+R$.

$[c-R, c+R)$ iff it converges at $x = c-R$ and diverges at $x = c+R$.

$(c-R, c+R]$ iff it diverges at $x = c-R$ and converges at $x = c+R$.

That is, the interval of convergence of the power series is the open interval $(c-R, c+R)$ along with the point(s) $c-R$ and/or $c+R$, wherever it is convergent. Due to (2.2) the power series converges everywhere inside the interval of convergence, it converges absolutely inside the open interval $(c-R, c+R)$, and it diverges everywhere beyond the interval of convergence.

Also, see that when $R = \infty$, the power series converges for all $x \in \mathbb{R}$, and when $R = 0$, the power series converges only at the point $x = c$, whence its sum is a_0.

To determine the interval of convergence, you must find the radius of convergence R, and then test for its convergence separately for the end-points $x = c - R$ and $x = c + R$.

The radius of convergence can be found out by ratio test and/or root test, or any other test.

(2.3) *Theorem*

The radius of convergence of the power series $\displaystyle\sum_{n=0}^{\infty} a_n (x-c)^n$ *is* $\displaystyle\lim_{n\to\infty} |a_n|^{-1/n}$ *provided this limit is either a real number or* ∞.

Proof. For the power series $\sum_{n=0}^{\infty} a_n(x-c)^n$, suppose $\lim_{n\to\infty} |a_n|^{1/n} = r \in \mathbb{R} \cup \{\infty\}$. We need to show the following:

(1) If $r \in \mathbb{R}$ and $r > 0$, then the radius of convergence of the power series is $1/r$.

(2) If $r = 0$, then the power series converges for all $x \in \mathbb{R}$.

(3) If $r = \infty$, then the power series converges only for $x = 0$.

(1) Suppose $\lim_{n\to\infty} |a_n|^{1/n} = r > 0$. Now,

$$|x-c| < \frac{1}{r} \Rightarrow |x-c| \lim_{n\to\infty} |a_n|^{1/n} < 1 \Rightarrow \lim_{n\to\infty} |a_n(x-c)^n|^{1/n} < 1.$$

By the root test, it follows that the series is convergent when $|x-c| < \frac{1}{r}$. Next,

$$|x-c| > \frac{1}{r} \Rightarrow |x-c| \lim_{n\to\infty} |a_n|^{1/n} > 1 \Rightarrow \lim_{n\to\infty} |a_n(x-c)^n|^{1/n} > 1.$$

Again, the root test implies that the series is divergent when $|x-c| > \frac{1}{r}$.

(2) If $r = 0$, then for any $x \in \mathbb{R}$, $\lim |a_n(x-c)^n|^{1/n} = |x-c| \lim |a_n|^{1/n} = 0$. By the root test, the series converges for each $x \in \mathbb{R}$.

(3) If $r = \infty$, then for any $x \neq c$, $\lim |a_n(x-c)^n| = \lim |x-c||a_n|^{1/n} = \infty$. By the root test, $\sum a_n(x-c)^n$ diverges for each $x \neq c$. ∎

Instead of the Root test, if we apply the Ratio test, then we obtain the following theorem.

(2.4) *Theorem*

The radius of convergence of the power series $\sum_{n=0}^{\infty} a_n(x-c)^n$ *is given by* $\lim_{n\to\infty} \left|\frac{a_n}{a_{n+1}}\right|$, *provided that this limit is either a real number or equal to* ∞.

Also, we sometimes need to use the method of substitution. In this connection, the following result is useful.

(2.5) *Theorem*

Let $R > 0$ *and let* $f : (-R, R) \to \mathbb{R}$ *be a continuous function. If* $\sum_{n=0}^{\infty} a_n x^n$ *converges absolutely for* $|x| < R$, *then* $\sum_{n=0}^{\infty} a_n(f(x))^n$ *converges absolutely for* $|f(x)| < R$.

(2.6) *Example*

For what values of x, do the following power series converge?

$$\text{(a)} \sum_{n=0}^{\infty} n!x^n \quad \text{(b)} \sum_{n=0}^{\infty} \frac{x^n}{n!} \quad \text{(c)} \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{2n+1} \quad \text{(d)} \sum_{n=0}^{\infty} (-1)^n \frac{x^{n+1}}{n+1}$$

(a) $a_n = n!$. Thus $\lim |a_n/a_{n+1}| = \lim 1/(n+1) = 0$. Hence $R = 0$. That is, the series is only convergent for $x = 0$.

(b) $a_n = 1/n!$. Thus $\lim |a_n/a_{n+1}| = \lim(n+1) = \infty$. Hence $R = \infty$. That is, the series is convergent for all $x \in \mathbb{R}$.

(c) Here, the power series in not in the form $\sum b_n x^n$. The series can be thought of as

$$x\Big(1 - \frac{x^2}{3} + \frac{x^4}{5} + \cdots\Big) = x \sum_{n=0}^{\infty} (-1)^n \frac{t^n}{2n+1} \quad \text{for } t = x^2.$$

Now, for the power series $\sum (-1)^n \dfrac{t^n}{2n+1}$, $a_n = (-1)^n/(2n+1)$.
Thus $\lim |a_n/a_{n+1}| = \lim \frac{2n+3}{2n+1} = 1$. Hence $R = 1$. That is, for $|t| = x^2 < 1$, the series converges and for $|t| = x^2 > 1$, the series diverges.

Alternatively, you can use the geometric series. That is, for any $x \in \mathbb{R}$, consider the series

$$x\Big(1 - \frac{x^2}{3} + \frac{x^4}{5} + \cdots\Big).$$

Write b_n for the nth coefficient of the power series. By the ratio test, the series converges if

$$\lim_{n\to\infty} \left|\frac{b_n}{b_{n+1}}\right| = \lim_{n\to\infty} \frac{2n+3}{2n+1} |x^2| = x^2 < 1.$$

That is, the power series converges for $-1 < x < 1$. Also, by the ratio test, the series diverges for $|x| > 1$.

What happens for $|x| = 1$?
For $x = -1$, the original power series is an alternating series; it converges due to Liebniz. Similarly, for $x = 1$, the alternating series also converges.
Hence the interval of convergence for the original power series (in x) is $[-1, 1]$.

(d) Consider the series in the form $x \displaystyle\sum_{n=0}^{\infty} (-1)^{n+1} \frac{x^n}{n+1}$.

For the series $\displaystyle\sum_{n=0}^{\infty} (-1)^{n+1} \frac{x^n}{n+1}$, $\quad a_n = \dfrac{(-1)^{n+1}}{n+1}$. Thus, $\displaystyle\lim_{n\to\infty} \frac{|a_n|}{|a_{n+1}|} = \lim_{n\to\infty} \frac{n+2}{n+1} = 1$.
Hence $R = 1$. That is, the series is convergent for all $x \in (-1, 1)$.

Here again, you can use directly the ratio test on the series for any fixed x as in (c).

For $x = -1$, the series is $-1 - \dfrac{1}{2} - \dfrac{1}{3} - \cdots$ Since the harmonic series is divergent, the power series at $x = -1$ is divergent.
For $x = 1$, the series is $1 - \dfrac{1}{2} + \dfrac{1}{3} - \dfrac{1}{4} + \cdots$ Since the alternating harmonic series is

convergent, the power series at $x = 1$ is convergent.
Therefore, the interval of convergence of the original power series is $(-1, 1]$. ∎

If R is the radius of convergence of a power series $\sum a_n(x-a)^n$, then the series defines a function $f(x)$ from the open interval $(a-R, a+R)$ to $\mathbb{R}$ by

$$f(x) = a_0 + a_1(x-a) + a_2(x-a)^2 + \cdots = \sum_{n=0}^{\infty} a_n(x-a)^n \quad \text{for } x \in (a-R, a+R).$$

This function can be differentiated and integrated term-by-term and it so happens that the new series obtained by such term-by-term differentiation or integration has the same radius of convergence and they define the derivative and the integral of $f(x)$. We state it without proof.

(2.7) *Theorem*

Let the power series $\sum_{n=0}^{\infty} a_n(x-c)^n$ have radius of convergence $R > 0$. Then the power series defines a function $f : (c-R, c+R) \to \mathbb{R}$. Further, $f'(x)$ and $\int f(x)dx$ exist as functions from $(c-R, c+R)$ to $\mathbb{R}$ and these are given by

$$f(x) = \sum_{n=0}^{\infty} a_n(x-c)^n, \quad f'(x) = \sum_{n=1}^{\infty} n a_n(x-c)^{n-1}, \quad \int f(x)dx = \sum_{n=0}^{\infty} a_n \frac{(x-c)^{n+1}}{n+1} + C,$$

where all the three power series converge for all $x \in (c-R, c+R)$.

Caution: Term by term differentiation may not work for series, which are not power series. For example, $\displaystyle\sum_{n=1}^{\infty} \frac{\sin(n!x)}{n^2}$ is convergent for all x. Differentiaing term by term, we obtain the series $\displaystyle\sum_{n=1}^{\infty} \frac{n!\cos(n!x)}{n^2}$, which diverges for all x.

Further, power series about the same point can be multiplied by using a generalization of multiplication of polynomials. We write the multiplication of power series about $x = 0$ for simplicity.

(2.8) *Theorem*

Let the power series $\sum a_n(x-c)^n$ and $\sum b_n(x-c)^n$ have the same radii of convergence $R > 0$. Then their product $\sum d_n(x-c)^n$ has the same radius of convergence R, where

$$d_n = \sum_{k=0}^{n} a_k b_{n-k} = a_0 b_n + a_1 b_{n-1} + \cdots + a_{n-1} b_1 + a_n b_0.$$

Moreover, the functions they define satisfy the following for $c - R < x < c + R$:

$$\textit{If } f(x) = \sum a_n(x-c)^n, \; g(x) = \sum b_n(x-c)^n, \; \textit{then } f(x)g(x) = \sum d_n(x-c)^n.$$

(2.9) ***Example***

Determine power series expansions of the following functions:

(a) $g(x) = \dfrac{2}{(x-1)^3}$ (b) $g(x) = \tan^{-1} x$ (c) $\log(1+x)$ (d) $\dfrac{\log(1+x)}{1-x}$

(a) For $-1 < x < 1$, $\dfrac{1}{1-x} = 1 + x + x^2 + x^3 + \cdots$.
Differentiating term by term, we have

$$\frac{1}{(1-x)^2} = 1 + 2x + 3x^2 + 4x^3 + \cdots$$

Differentiating once more, we get

$$\frac{2}{(1-x)^3} = 2 + 6x + 12x^2 + \cdots = \sum_{n=2}^{\infty} n(n-1)x^{n-2} \quad \text{for } -1 < x < 1.$$

(b) $\dfrac{1}{1+x^2} = 1 - x^2 + x^4 - x^6 + x^8 - \cdots$ for $|x^2| < 1$.
Integrating term by term we have

$$\tan^{-1} x + C = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \cdots \quad \text{for } -1 < x < 1.$$

Evaluating at $x = 0$, we see that $C = 0$. Hence the power series for $\tan^{-1} x$.
(c) For $-1 < x < 1$, $\dfrac{1}{1+x} = 1 - x + x^2 - x^3 + x^4 - \cdots$.
Integrating term by term and evaluating at $x = 0$, we obtain

$$\log(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \cdots \quad \text{for } -1 < x < 1.$$

(d) Using the results in (c) and the geometric series for $1/(1-x)$, we have

$$\frac{\log(1+x)}{1-x} = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}\, x^n}{n} \cdot \sum_{n=0}^{\infty} x^n \quad \text{for } -1 < x < 1.$$ □

For obtaining the product of the two power series, we need to write the first in the form $\sum a_n x^n$. (Notice that for the second series, each $b_n = 1$.) Here, the first series is

$$\log(1+x) = \sum_{n=0}^{\infty} a_n x^n, \quad \text{where } a_0 = 0 \text{ and } a_n = \frac{(-1)^{n-1}}{n} \text{ for } n \geq 1.$$

Thus the product above is $\dfrac{\log(1+x)}{1-x} = \displaystyle\sum_{n=0}^{\infty} c_n x^n$, where

$$c_n = a_0 b_n + a_1 b_{n-1} + \cdots + a_n b_0 = a_0 + a_1 + \cdots a_n = 1 - \frac{1}{2} + \frac{1}{3} - \cdots + \frac{(-1)^{n-1}}{n}.$$

The following theorem due to Abel shows that continuity is somehow inbuilt in a power series. It helps in evaluating a series at the endpoints of the interval of convergence.

(2.10) ***Theorem*** (Abel)

Let R be the radius of convergence of the power series $\sum_{n=0}^{\infty} a_n x^n$ whose sum is $f(x)$ for $|x| < R$.

(1) *If the power series converges at $x = -R$, then* $\sum_{n=0}^{\infty} a_n (-R)^n = \lim_{x \to -R+} f(x)$.

(2) *If the power series converges at $x = R$, then* $\sum_{n=0}^{\infty} a_n R^n = \lim_{x \to R-} f(x)$.

(2.11) ***Example***

We know that $\log(1+x) = \sum_{n=1}^{\infty} (-1)^{n+1} \frac{x^n}{n}$. The radius of convergence of this power series is 1. By Abel's theorem,

$$\lim_{x \to 1-} \log(1+x) = \sum_{n=1}^{\infty} (-1)^{n+1} \frac{1^n}{n} \Rightarrow \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} = \log 2. \qquad \square$$

Observe what Abel's theorem does not say. In general, if a function has a power series representation in the interval $|x| < R$ and the function has a limit as $x \to R-$, then the series need not converge at $x = R$. See the following example.

(2.12) ***Example***

Consider the function $f(x) = \frac{1}{1+x^2}$. It has the power series representation $\sum_{n=0}^{\infty} (-1)^n x^{2n}$ for $|x| < 1$. As $x \to 1-$, the function $f(x)$ has the limit $1/2$. However, at $x = 1$, the power series $\sum_{n=0}^{\infty} (-1)^n (1)^{2n} = 1 - 1 + 1 - 1 + 1 \cdots$ does not converge. $\square$

2.2 Taylor's formulas

For an elegant power series representation of smooth functions we require Taylor's formulas. It has two forms: differential form and integral form. The differential form is a generalization of the Mean Value Theorem for differentiable functions. However, we will first prove the integral form, and then deduce the differential form. In what follows, if $f : [a, b] \to \mathbb{R}$, then its derivative at $x = a$ is taken as the right hand derivative, and at $x = b$, the derivative is taken as the left hand derivative of $f(x)$.

(2.13) ***Theorem*** (Taylor's Formula in Integral Form)

Let $f(x)$ be an $(n+1)$-times continuously differentiable function on $[a,b]$. Let $x \in [a,b]$. Then

$$f(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n + R_n(x),$$

where $R_n(x) = \int_a^x \frac{(x-t)^n}{n!} f^{(n+1)}(t)\,dt$. An estimate for $R_n(x)$ is given by

$$\frac{m\,(x-a)^{n+1}}{(n+1)!} \le R_n(x) \le \frac{M\,(x-a)^{n+1}}{(n+1)!}$$

where $m \le f^{n+1}(x) \le M$ for $x \in [a,b]$.

Proof. We prove it by induction on n. For $n = 0$, we should show that

$$f(x) = f(a) + R_0(x) = f(a) + \int_a^x f'(t)\,dt.$$

But this follows from the Fundamental theorem of calculus. Now, suppose that Taylor's formula holds for $n = k$. That is, we have

$$f(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots + \frac{f^{(k)}(a)}{k!}(x-a)^k + R_k(x),$$

where $R_k(x) = \int_a^x \frac{(x-t)^k}{k!} f^{(k+1)}(t)\,dt$. We evaluate $R_k(x)$ using integration by parts with the first function as $f^{(k+1)}(t)$ and the second function as $(x-t)^k/k!$. Remember that the variable of integration is t and x is a fixed number. Then

$$\begin{aligned}
R_k(x) &= \Big[-f^{(k+1)}(t)\frac{(x-t)^{k+1}}{(k+1)!} \Big]_a^x + \int_a^x f^{(k+2)}(t)\frac{(x-t)^{k+1}}{(k+1)!}\,dt \\
&= f^{(k+1)}(a)\frac{(x-a)^{k+1}}{(k+1)!} + \int_a^x f^{(k+2)}(t)\frac{(x-t)^{k+1}}{(k+1)!}\,dt \\
&= \frac{f^{(k+1)}(a)}{(k+1)!}(x-a)^{k+1} + R_{k+1}(x).
\end{aligned}$$

This completes the proof of Taylor's formula. For the estimate of $R_n(x)$, Observe that

$$\int_a^x \frac{(x-t)^n}{n!}\,dt = \frac{-(x-t)^{n+1}}{(n+1)!}\Big|_a^x = \frac{(x-a)^{n+1}}{(n+1)!}.$$

Since $m \le f^{n+1}(x) \le M$, the estimate for $R_n(x)$ follows. ∎

To derive the differential form of Taylor's formula, we use the following result.

(2.14) *Theorem* (Weighted Mean Value Theorem)

Let $f(t)$ and $g(t)$ be continuous real valued functions on $[a,b]$, where $g(t)$ does not change sign in $[a,b]$. Then there exists $c \in [a,b]$ such that

$$\int_a^b f(t)g(t)\,dt = f(c)\int_a^b g(t)\,dt.$$

Proof. Without loss of generality assume that $g(t) \geq 0$. Since $f(t)$ is continuous, let $\alpha = \min f(t)$ and $\beta = \max f(t)$ in $[a,b]$. Then

$$\alpha\int_a^b g(t)\,dt \leq \int_a^b f(t)g(t)\,dt \leq \beta\int_a^b g(t)\,dt.$$

If $\int_a^b g(t)\,dt = 0$, then $\int_a^b f(t)g(t)\,dt = 0$. In this case, $\int_a^b f(t)g(t)\,dt = f(c)\int_a^b g(t)\,dt$. So, suppose that $\int_a^b g(t)\,dt \neq 0$. Then $\int_a^b g(t)\,dt > 0$. Consequently,

$$\alpha \leq \frac{\int_a^b f(t)g(t)\,dt}{\int_a^b g(t)\,dt} \leq \beta.$$

By the intermediate value theorem, there exists $c \in [a,b]$ such that

$$\frac{\int_a^b f(t)g(t)\,dt}{\int_a^b g(t)\,dt} = f(c).$$ ∎

(2.15) *Theorem* (Taylor's Formula in Differential Form)

Let $n \in \mathbb{N}$. Suppose that $f^{(n)}(x)$ is continuously differentiable on $[a,b]$. Then there exists $c \in [a,b]$ such that

$$f(x) = f(a)+f'(a)(x-a)+\frac{f''(a)}{2!}(x-a)^2+\cdots+\frac{f^{(n)}(a)}{n!}(x-a)^n+\frac{f^{(n+1)}(c)}{(n+1)!}(x-a)^{n+1}.$$

Proof. Let $x \in (a,b)$. The function $g(t) = (x-t)^n$ does not change sign in $[a,x]$. By the weighted mean value theorem, there exists $c \in [a,x]$ such that

$$\begin{aligned}\int_a^x (x-t)^n f^{(n+1)}(t)\,dt &= f^{(n+1)}(c)\int_a^x (x-t)^n\,dt = -f^{(n+1)}(c)\frac{(x-t)^{n+1}}{n+1}\bigg|_a^x \\ &= f^{(n+1)}(c)\frac{(x-a)^{n+1}}{n+1}.\end{aligned}$$

Using the Taylor's formula in integral form, we have

$$R_n(x) = \frac{1}{n!}\int_a^x (x-t)^n f^{(n+1)}(t)\,dt = \frac{1}{n!} f^{(n+1)}(c)\frac{(x-a)^{n+1}}{n+1} = \frac{f^{(n+1)}(c)}{(n+1)!}(x-a)^{n+1}.$$

∎

Remark 2.16 Taylor's formula in differential form can be proved directly by repeated use of the Mean value theorem, or Rolle's theorem. It is as follows.

For $x = a$, the formula holds. So, let $x \in (a, b]$. For any $t \in [a, x]$, let

$$p(t) = f(a) + f'(a)(t-a) + \frac{f''(a)}{2!}(t-a)^2 + \cdots + \frac{f^{(n)}(a)}{n!}(t-a)^n.$$

Here, we treat x as a certain point, not a variable; and t as a variable. Write

$$g(t) = f(t) - p(t) - \frac{f(x) - p(x)}{(x-a)^{n+1}}\,(t-a)^{n+1}.$$

We see that $g(a) = 0,\ g'(a) = 0,\ g''(a) = 0,\ \ldots, g^{(n)}(a) = 0,$ and $g(x) = 0.$

By Rolle's theorem, there exists $c_1 \in (a, x)$ such that $g'(c_1) = 0$. Since $g(a) = 0$, apply Rolle's theorem once more to get a $c_2 \in (a, c_1)$ such that $g''(c_2) = 0$.

Continuing this way, we get a $c_{n+1} \in (a, c_n)$ such that $g^{(n+1)}(c_{n+1}) = 0$.

Since $p(t)$ is a polynomial of degree at most n, $p^{(n+1)}(t) = 0$. Then

$$g^{(n+1)}(t) = f^{(n+1)}(t) - \frac{f(x) - p(x)}{(x-a)^{n+1}}\,(n+1)!.$$

Evaluating at $t = c_{n+1}$ we have $f^{(n+1)}(c_{n+1}) - \dfrac{f(x) - p(x)}{(x-a)^{n+1}}\,(n+1)! = 0$. That is,

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

Consequently, $g(t) = f(t) - p(t) - \dfrac{f^{(n+1)}(c_{n+1})}{(n+1)!}\,(t-a)^{n+1}$.

Evaluating it at $t = x$ and using the fact that $g(x) = 0$, we get

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

Since x is an arbitrary point in $(a, b]$, this completes the proof.

The polynomial

$$p(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n$$

in Taylor's formula is called as **Taylor's polynomial of order** n. Notice that the degree of the Taylor's polynomial may be less than or equal to n. The expression given for $f(x)$ there is called **Taylor's formula** for $f(x)$. Taylor's polynomial is an approximation to $f(x)$ with **the error**

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

How good $f(x)$ is approximated by $p(x)$ depends on the smallness of the error $R_n(x)$. For example, if we use $p(x)$ of order 5 for approximating $\sin x$ at $x = 0$, then we get

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} + R_6(x), \text{ where } R_6(x) = \frac{\sin\theta}{6!}x^6.$$

Here, θ lies between 0 and x. The absolute error is bounded above by $|x|^6/6!$. However, if we take the Taylor's polynomial of order 6, then $p(x)$ is the same as in the above, but the absolute error is now $|x|^7/7!$. If x is near 0, this is smaller than the earlier bound.

Notice that if $f(x)$ is a polynomial of degree n, then Taylor's polynomial of order n is equal to the original polynomial.

As (2.6) shows, by clever manipulation of known series and functions, we may be able to have a series representation of some of them.

In general, we ask: Which functions can have a power series representation, and how to obtain a power series from such a given function?

2.3 Taylor series

Taylor's formulas (2.15 and 2.13) say that under suitable hypotheses a function can be written as

$$f(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n + R_n(x),$$

where

$$R_n(x) = \frac{f^{(n+1)}(c)}{(n+1)!}(x-a)^{n+1} \quad \text{OR} \quad R_n(x) = \int_a^x \frac{(x-t)^n}{n!} f^{(n+1)}(t)\,dt.$$

If $R_n(x)$ converges to 0 for all x in an interval around the point $x = a$, then the ensuing series on the right hand side would converge and then the function can be written in the form of a series. That is, under the conditions that $f(x)$ has derivatives of all order, and $R_n(x) \to 0$ for all x in an interval around $x = a$, the function $f(x)$ has a power series representation

$$f(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n + \cdots$$

Such a series is called the **Taylor series** expansion of the function $f(x)$. When $a = 0$, the Taylor series is called the **Maclaurin series**.

Conversely, if a function $f(x)$ has a power series expansion about $x = a$, then by repeated differentiation and evaluation at $x = a$ shows that the coefficients of the power series are precisely of the form $\frac{f^{(n)}(a)}{n!}$ as in the Taylor series.

(2.17) *Example*

Find the Taylor series expansion of the function $f(x) = 1/x$ at $x = 2$. In which interval around $x = 2$, the series converges?

We see that $f(x) = \frac{1}{x}$, $f(2) = \frac{1}{2}$; $\cdots$; $f^{(n)}(x) = \frac{(-1)^n n!}{x^{n+1}}$, $f^{(n)}(2) = \frac{(-1)^n n!}{2^{n+1}}$.

Hence the Taylor series for $f(x) = 1/x$ is

$$\frac{1}{2} - \frac{x-2}{2^2} + \frac{(x-2)^2}{2^3} - \cdots + (-1)^n \frac{(x-2)^n}{2^{n+1}} + \cdots$$

We now require the remainder term in the Taylor expansion. The absolute value of the remainder term in the differential form is (for any c, x in an interval around $x = 2$)

$$|R_n| = \left|\frac{f^{(n+1)}(c)}{(n+1)!}(x-2)^{n+1}\right| = \left|\frac{(x-2)^{n+1}}{c^{n+2}}\right|$$

Here, c lies between x and 2. Clearly, if x is near 2, $|R_n| \to 0$. Hence the Taylor series represents the function near $x = 2$.

However, a direct calculation can be done looking at the Taylor series so obtained. Here, the series is a geometric series with ratio $r = -(x-2)/2$. Hence it converges absolutely whenever

$$|r| < 1, \quad \text{i.e.,} \quad |x-2| < 2 \quad \text{i.e.,} \quad 0 < x < 4.$$

Thus the series represents the function $f(x) = 1/x$ for $0 < x < 4$. □

(2.18) *Example*

Consider the function $f(x) = e^x$. For its Maclaurin series, we find that

$$f(0) = 1,\ f'(0) = 1, \cdots, f^{(n)}(0) = 1, \cdots$$

Hence

$$e^x = 1 + x + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!} + \cdots$$

Using the integral form of the remainder,

$$|R_n(x)| = \left| \int_a^x \frac{(x-t)^n}{n!} f^{(n+1)}(t)\, dt \right| = \left| \int_0^x \frac{(x-t)^n}{n!} e^t\, dt \right| \to 0 \text{ as } n \to \infty.$$

Hence, e^x has the above power series expansion for each $x \in \mathbb{R}$.

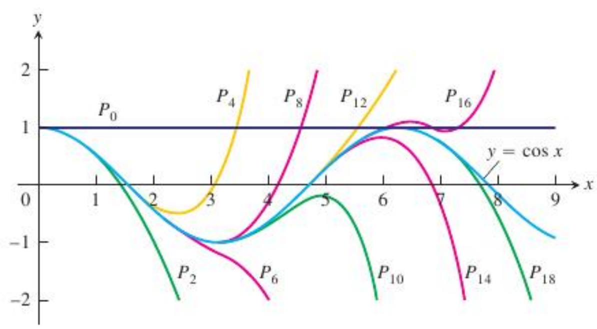
Directly, by the ratio test, this power series has the radius of convergence

$$R = \lim_{n\to\infty} \frac{a_n}{a_{n+1}} = \lim_{n\to\infty} \frac{(n+1)!}{n!} = \infty.$$

Therefore, for every $x \in \mathbb{R}$ the above series converges. ∎

(2.19) ***Example***

You can show that $\cos x = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n}}{(2n)!}$. The Taylor polynomials approximating $\cos x$ are $P_{2n}(x) = \sum_{k=n}^{\infty} \frac{(-1)^k x^{2k}}{(2k)!}$. The following picture shows how these polynomials approximate $\cos x$ for $0 \le x \le 9$.



In the above Maclaurin series expansion of $\cos x$, we have the absolute value of the remainder in the differential form as

$$|R_{2n}(x)| = \frac{|x|^{2n+1}}{(2n+1)!} \to 0 \text{ as } n \to \infty$$

for any $x \in \mathbb{R}$. Hence the series represents $\cos x$ for each $x \in \mathbb{R}$. ∎

(2.20) ***Example***

Let $m \in \mathbb{R}$. Show that, for $-1 < x < 1$,

$$(1+x)^m = 1 + \sum_{n=1}^{\infty} \binom{m}{n} x^n, \quad \text{where } \binom{m}{n} = \frac{m(m-1)\cdots(m-n+1)}{n!}.$$

To see this, find the derivatives of the given function:

$$f(x) = (1+x)^m, \quad f^{(n)}(x) = m(m-1)\cdots(m-n+1)x^{m-n}.$$

Then the Maclaurin series for $f(x)$ is the given series. For this series, we see that

$$\lim_{n\to\infty}\left|\frac{a_n}{a_{n+1}}\right| = \lim_{n\to\infty}\left|\frac{m(m-1)\cdots(m-n+1)}{n!}\,\frac{(n+1)!}{m(m-1)\cdots(m-n)}\right| = \lim_{n\to\infty}\left|\frac{n+1}{m-n}\right| = 1.$$

Alternatively, you can show that the remainder term in the Maclaurin series expansion goes to 0 as $n \to \infty$ for $-1 < x < 1$.

The series so obtained is called a **binomial series** expansion of $(1+x)^m$. Substituting values of m, we get series for different functions. Notice that when $m \in \mathbb{N}$, the binomial series terminates to give a polynomial and it represents $(1+x)^m$ for each $x \in \mathbb{R}$. $\square$

(2.21) *Example*

Consider the function $f(x) = \sqrt{1+x}$. With $m = 1/2$, the binomial series expansion gives

$$\sqrt{1+x} = \sum_{n=0}^{\infty} \frac{(-1)^{n+1}(2n)!}{2^{2n}(n!)^2(2n-1)} x^n = 1 + \frac{x}{2} - \frac{x^2}{8} + \frac{x^3}{16} - \cdots \quad \text{for } -1 < x < 1.$$

Using the estimate (We have not derived it.)

$$\frac{4^n}{\sqrt{\pi(n+1/3)}} \le \binom{2n}{n} \le \frac{4^n}{\sqrt{\pi(n+1/4)}}$$

it follows that at $x = 1$, the power series converges absolutely. Thus, By Abel's theorem, we obtain

$$\sum_{n=0}^{\infty} \frac{(-1)^{n+1}(2n)!}{2^{2n}(n!)^2(2n-1)} = \lim_{x\to 1-} f(x) = \sqrt{2}.$$

Similarly, the power series is absolutely convergent at $x = -1$. Therefore,

$$\sum_{n=0}^{\infty} \frac{(-1)^{n+1}(2n)!}{2^{2n}(n!)^2(2n-1)}(-1)^n = \lim_{x\to -1+} f(x) = 0.$$

It implies that $\displaystyle\sum_{n=1}^{\infty} \frac{(2n)!}{2^{2n}(n!)^2(2n-1)} = 1.$ $\square$

Remark 2.22 There exist functions which are n times differentiable for each $n \in \mathbb{N}$ but they are not represented by their Taylor series in any interval. For example, consider the following function:

$$f(x) = \begin{cases} e^{-1/x^2} & \text{if } x \neq 0 \\ 0 & \text{if } x = 0. \end{cases}$$

For each $n \in \mathbb{N}$, $f^{(n)}(0) = 0$; so, the Taylor series of $f(x)$ is the power series of which each coefficient is 0. Except at the point $x = 0$, the Taylor series does not match with the function. Here, notice that $R_n(x)$, which is equal to e^{-1/x^2} does not converge to 0 as $n \to \infty$. Thus, in order that a function is represented by its Taylor series in an interval, it is essential that the remainder term $R_n(x)$ must converge to 0 for all x in that interval as $n \to \infty$.

2.4 Fourier series

In the power series for $\sin x = x - x^3/3! + \cdots$, the periodicity of $\sin x$ is not obvious. Recall that $f(x)$ is called 2ℓ**-periodic** for $\ell > 0$ iff $f(x+2\ell) = f(x)$ for all $x \in \mathbb{R}$. For the time being, we restrict to 2π-periodic functions. We will see that a 2π-periodic function can be expanded in a series involving sines and cosines instead of powers of x.

A **trigonometric series** is of the form $\dfrac{1}{2}a_0 + \displaystyle\sum_{n=1}^{\infty}(a_n \cos nx + b_n \sin nx)$.

Since both cosine and sine functions are 2π-periodic, if the trigonometric series converges to a function $f(x)$, then necessarily $f(x)$ is also 2π-periodic. Thus,

$$f(0) = f(2\pi) = f(4\pi) = f(6\pi) = \cdots \quad \text{and} \quad f(-\pi) = f(\pi), \quad \text{etc.}$$

Moreover, if $f(x) = \frac{1}{2}a_0 + \sum_{n=1}^{\infty}(a_n \cos nx + b_n \sin nx)$, say, for all $x \in [-\pi, \pi]$, then the coefficients can be determined from $f(x)$. Towards this, multiply $f(t)$ by $\cos mt$ and integrate to obtain:

$$\int_{-\pi}^{\pi} f(t)\cos mt\, dt = \frac{1}{2}a_0 \int_{-\pi}^{\pi} \cos mt\, dt + \sum_{n=1}^{\infty} a_n \int_{-\pi}^{\pi} \cos nt \cos mt\, dt$$
$$+ \sum_{n=1}^{\infty} b_n \int_{-\pi}^{\pi} \sin nt \cos mt\, dt.$$

For $m, n = 0, 1, 2, 3, \ldots,$

$$\int_{-\pi}^{\pi} \cos nt \cos mt \, dt = \begin{cases} 0 & \text{if } n \neq m \\ \pi & \text{if } n = m > 0 \\ 2\pi & \text{if } n = m = 0 \end{cases} \quad \text{and} \quad \int_{-\pi}^{\pi} \sin nt \cos mt \, dt = 0.$$

Thus, we obtain

$$\int_{-\pi}^{\pi} f(t) \cos mt \, dt = \pi a_m, \quad \text{for all } m = 0, 1, 2, 3, \cdots$$

Similarly, by multiplying $f(t)$ by $\sin mt$ and integrating, and using the fact that

$$\int_{-\pi}^{\pi} \sin nt \sin mt \, dt = \begin{cases} 0 & \text{if } n \neq m \\ \pi & \text{if } n = m > 0 \\ 0 & \text{if } n = m = 0 \end{cases}$$

we obtain

$$\int_{-\pi}^{\pi} f(t) \sin mt \, dt = \pi b_m, \quad \text{for all } m = 1, 2, 3, \cdots$$

Let $f : \mathbb{R} \to \mathbb{R}$ be a 2π-periodic function integrable on $[-\pi, \pi]$. Write

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \cos nt \, dt, \quad \text{for } n = 0, 1, 2, 3, \ldots;$$

$$b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \sin nt \, dt \quad \text{for } n = 1, 2, 3, \ldots.$$

Then the trigonometric series

$$\frac{1}{2} a_0 + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx)$$

is called the **Fourier series** of $f(x)$.

To state and understand a relevant result about when the Fourier series of a function $f(x)$ converges at a point, we require the following notation and terminology.

Let $f : \mathbb{R} \to \mathbb{R}$ be a function. Let $c \in \mathbb{R}$. Write

$$f(c+) = \lim_{h \to 0+} f(c + h), \; f(c-) = \lim_{h \to 0+} f(c - h).$$

We use the following terminology:

> $f(x)$ has a **finite jump** at $x = c$ iff $f(c+)$ exists, $f(c-)$ exists, and $f(c+) \neq f(c-)$.

$f(x)$ is **piecewise continuous** iff on any finite interval $f(x)$ is continuous except for at most a finite number of finite jumps.

the **right hand slope** of $f(x)$ is equal to $\lim_{h\to 0+} \dfrac{f(c+h)-f(c+)}{h}$.

the **left hand slope** of $f(x)$ is equal to $\lim_{h\to 0-} \dfrac{f(c+h)-f(c-)}{h}$.

$f(x)$ is **piecewise smooth** iff $f(x)$ is piecewise continuous and $f(x)$ has both left hand slope and right hand slope at every point.

(2.23) ***Theorem*** (Convergence of Fourier Series)

Let $f : \mathbb{R} \to \mathbb{R}$ be a 2π-periodic piecewise smooth function. Then the Fourier series of $f(x)$ converges at each $x \in \mathbb{R}$. Further, at any point $c \in \mathbb{R}$, the following statements hold:

(1) *If $f(x)$ is continuous at c, then the Fourier series sums to $f(c)$.*

(2) *If $f(x)$ is not continuous at c, then the Fourier series sums to $\frac{1}{2}[f(c+)+f(c-)]$.*

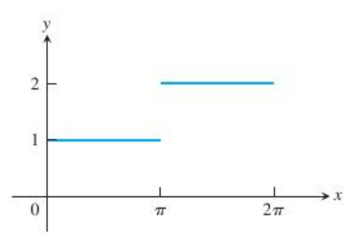
We observe that a 2π-periodic function $f : \mathbb{R} \to \mathbb{R}$, which is bounded and piecewise monotonic on $[-\pi, \pi]$ is piecewise smooth. Similarly, a 2π-periodic function $f : \mathbb{R} \to \mathbb{R}$, which has left and right derivatives at each point and is piecewise continuous on $[-\pi, \pi]$, is also piecewise smooth. Thus, the Fourier series for these types of functions converge.

Fourier series can represent functions which cannot be represented by a Taylor series, or a conventional power series; for example, a step function.

(2.24) ***Example***

Find the Fourier series of the function $f(x)$ given by the following which is extended to $\mathbb{R}$ with periodicity 2π:

$$f(x) = \begin{cases} 1 & \text{if } 0 \le x < \pi \\ 2 & \text{if } \pi \le x < 2\pi \end{cases}$$



Due to periodic extension, we can rewrite the function $f(x)$ on $[-\pi, \pi)$ as

$$f(x) = \begin{cases} 2 & \text{if } -\pi \le x < 0 \\ 1 & \text{if } 0 \le x < \pi \end{cases}$$

Then the coefficients of the Fourier series are computed as follows:

$$a_0 = \frac{1}{\pi}\int_{-\pi}^{0} f(t)\,dt + \frac{1}{\pi}\int_{0}^{\pi} f(t)\,dt = 3.$$

$$a_n = \frac{1}{\pi}\int_{-\pi}^{0} 2\cos nt\,dt + \frac{1}{\pi}\int_{0}^{\pi} \cos nt\,dt = 0.$$

$$b_n = \frac{1}{\pi}\int_{-\pi}^{0} 2\sin nt\,dt + \frac{1}{\pi}\int_{0}^{\pi} \sin nt\,dt = \frac{(-1)^n - 1}{n\pi}.$$

Notice that $b_1 = -\frac{2}{\pi},\ b_2 = 0,\ b_3 = -\frac{2}{3\pi},\ b_4 = 0, \ldots$. Therefore,

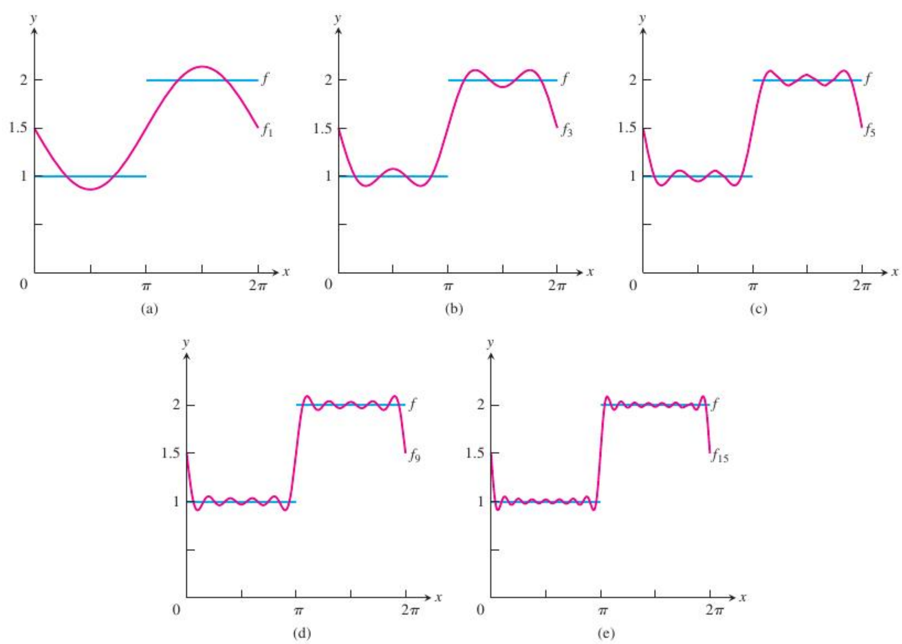
$$f(x) = \frac{3}{2} - \frac{2}{\pi}\Big(\sin x + \frac{\sin 3x}{3} + \frac{\sin 5x}{5} + \cdots\Big).$$

Here, the last expression for $f(x)$ holds for all $x \in \mathbb{R}$, where ever $f(x)$ is continuous; in particular, for $x \in [-\pi, \pi)$ except at $x = 0$. Notice that $x = 0$ is a point of discontinuity of $f(x)$. By the convergence theorem, the Fourier series at $x = 0$ sums to $\dfrac{f(0+) + f(0-)}{2}$, which is equal to $\dfrac{3}{2}$. ∎

Once we have a series representation of a function, we should see how the **partial sums** of the series approximate the function. In the above example, let us write

$$f_m(x) = \frac{1}{2}a_0 + \sum_{n=1}^{m}(a_n \cos nx + b_n \sin nx).$$

The approximations $f_1(x),\ f_3(x),\ f_5(x), f_9(x)$ and $f_{15}(x)$ to $f(x)$ are shown in the figure below.



Remark 2.25 In the above example, we found that

$$f(x) = \frac{3}{2} - \frac{2}{\pi}\Big(\sin x + \frac{\sin 3x}{3} + \frac{\sin 5x}{5} + \cdots\Big).$$

In particular, at $x = \pi$, the function $f(x)$ is not differentiable. However, each term on the right hand side is differentiable at $x = \pi$. It means that each term of a Fourier series can be differentiable at a point, but the sum of the series is not differentiable at that point. Indeed, if we differentiate each term of the series and sum the terms after evaluation at π, we get the divergent series $\frac{2}{\pi}(-1 - 1 - 1 - \cdots)$.

(2.26) *Example*

Show that the Fourier series for $f(x) = x^2$ defined on $[0, 2\pi)$ is given by

$$\frac{4\pi^2}{6} + \sum_{n=1}^{\infty}\Big(\frac{4}{n^2}\cos nx - \frac{4\pi}{n}\sin nx\Big).$$

Extend $f(x)$ to $\mathbb{R}$ by periodicity 2π. We thus have $f(2\pi) = f(0) = 0$. Then

$$f(-\pi) = f(-\pi{+}2\pi) = f(\pi) = \pi^2,\ f(-\pi/2) = f(-\pi/2{+}2\pi) = f(3\pi/2) = (3\pi/2)^2.$$

Thus the function $f(x)$ on $[-\pi, \pi)$ is defined by

$$f(x) = \begin{cases} (x + 2\pi)^2 & \text{if } -\pi \le x < 0 \\ x^2 & \text{if } 0 \le x < \pi. \end{cases}$$

Notice that $f(x)$ is neither odd nor even. The coefficients of the Fourier series for $f(x)$ are

$$a_0 = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\,dt = \frac{1}{\pi}\int_0^{2\pi} t^2\,dt = \frac{8\pi^2}{3}.$$

$$a_n = \frac{1}{\pi}\int_0^{2\pi} t^2\cos nt\,dt = \frac{4}{n^2} \quad \text{for } n = 1, 2, 3, \ldots$$

$$b_n = \frac{1}{\pi}\int_0^{2\pi} t^2\sin nt\,dt = -\frac{4\pi}{n} \quad \text{for } n = 1, 2, 3, \ldots$$

Hence the Fourier series for $f(x)$ is as claimed.

As per the extension of $f(x)$ to $\mathbb{R}$, we see that in the interval $(2k\pi, 2(k+1)\pi)$, the function is defined by $f(x) = (x - 2k\pi)^2$. Thus it has discontinuities at the points $x = 0,\ \pm 2\pi,\ \pm 4\pi,\ \ldots$ At such a point $x = 2k\pi$, the series converges to the average value of the left and right side limits, i.e., the series when evaluated at $2k\pi$ yields the value

$$\frac{1}{2}\Big[\lim_{x\to 2k\pi-} f(x) + \lim_{x\to 2k\pi+} f(x)\Big] = \frac{1}{2}\Big[\lim_{x\to 2k\pi-}(x-2k\pi)^2 + \lim_{x\to 2k\pi+}(x-2(k+1)\pi)^2\Big] = 2\pi^2.$$

∎

2.5 Odd and even functions

If $f(x)$ is an odd function, that is, if $f(-x) = -f(x)$, then the Fourier coefficients for $f(x)$ are

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\cos nt\,dt = 0 \quad \text{for } n = 0, 1, 2, \ldots.$$

$$b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\sin nt\,dt = \frac{2}{\pi}\int_{0}^{\pi} f(t)\sin nt\,dt \quad \text{for } n = 1, 2, 3, \ldots.$$

In this case, the Fourier series for $f(x)$ is given by

$$\sum_{n=1}^{\infty} b_n \sin nx, \quad b_n = \frac{2}{\pi}\int_{0}^{\pi} f(t)\sin nt\,dt \quad \text{for } n = 1, 2, 3, \ldots.$$

Similarly, if $f(x)$ is an even function, that is, $f(-x) = f(x)$, then its Fourier series is given by

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} a_n, \quad a_n = \frac{2}{\pi}\int_{0}^{\pi} f(t)\cos nt\,dt \quad \text{for } n = 0, 1, 2, \ldots.$$

(2.27) ***Example***

Show that $x^2 = \dfrac{\pi^2}{3} + 4\sum_{n=1}^{\infty}(-1)^n \dfrac{\cos nx}{n^2}$ for all $x \in [-\pi, \pi]$.

Let $f(x)$ be the periodic extension of the function $x \mapsto x^2$ on $[-\pi, \pi)$ to $\mathbb{R}$. Since $\pi^2 = (-\pi)^2$, such an extension with period 2π exists, and it is continuous. The extension of $f(x) = x^2$ to $\mathbb{R}$ is not the function x^2. For instance, in the interval $[\pi, 3\pi]$, its extension looks like $f(x) = (x - 2\pi)^2$. With this understanding, we go for the Fourier series expansion of $f(x) = x^2$ in the interval $[-\pi, \pi]$. We also see that $f(x)$ is an even function. Its Fourier series is a cosine series. The coefficients of the series are as follows:

$$a_0 = \frac{2}{\pi}\int_{0}^{\pi} t^2\,dt = \frac{2}{3}\pi^2.$$

$$a_n = \frac{2}{\pi}\int_{0}^{\pi} t^2 \cos nt\,dt = \frac{4}{n^2}(-1)^n \quad \text{for } n = 1, 2, 3, \ldots$$

Therefore,

$$f(x) = x^2 = \frac{\pi^2}{3} + 4\sum_{n=1}^{\infty}(-1)^n \frac{\cos nx}{n^2} \quad \text{for all } x \in [-\pi, \pi].$$

In particular, by taking $x = 0$ and $x = \pi$, we have

$$\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n^2} = \frac{\pi^2}{12}, \qquad \sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}.$$

Due to the periodic extension of $f(x)$ to $\mathbb{R}$, we see that

$$(x - 2\pi)^2 = \frac{\pi^2}{3} + 4 \sum_{n=1}^{\infty} (-1)^n \frac{\cos nx}{n^2} \quad \text{for all } x \in [\pi, 3\pi].$$

It also follows that the same sum (of the series) is equal to $(x - 4\pi)^2$ for $x \in [3\pi, 5\pi]$, etc. ∎

(2.28) ***Example***

Show that for $0 < x < 2\pi$, $\quad \dfrac{1}{2}(\pi - x) = \displaystyle\sum_{n=1}^{\infty} \frac{\sin nx}{n}$.

Let $f(x) = x$ for $0 \le x < 2\pi$. Extend $f(x)$ to $\mathbb{R}$ by taking the periodicity as 2π. As in (2.27), $f(x)$ is not an odd function. For instance, $f(-\pi/2) = f(3\pi/2) = 3\pi/2 \neq f(\pi/2) = \pi/2$.

The coefficients of the Fourier series for $f(x)$ are as follows:

$$a_0 = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t)\, dt = \frac{1}{\pi} \int_0^{2\pi} t\, dt = 2\pi,$$

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \cos nt\, dt = \frac{1}{\pi} \int_0^{2\pi} t \cos nt\, dt = 0.$$

$$\begin{aligned} b_n &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \sin nt\, dt = \frac{1}{\pi} \int_0^{2\pi} t \sin nt\, dt \\ &= \frac{1}{\pi} \left[\frac{-n \cos nt}{n} \right]_0^{2\pi} + \frac{1}{n\pi} \int_0^{2\pi} \cos nt\, dt = -\frac{2}{\pi}. \end{aligned}$$

By the convergence theorem, $x = \pi - 2 \displaystyle\sum_{n=1}^{\infty} \frac{\sin nx}{n}$ for $0 < x < 2\pi$, which yields the required result. ∎

2.6 Half range Fourier series

Suppose a piecewise smooth function $f : (0, \pi) \to \mathbb{R}$ is given. To find its Fourier series, we need to extend it to $\mathbb{R}$ so that the extended function is 2π-periodic. Such an extension can be done in at least two ways.

1. Even Extension:

First, extend $f(x)$ to $(-\pi, \pi)$ by requiring that $f(x)$ is an even function. This requirement forces $f(-x) = f(x)$ for each $x \in (-\pi, \pi)$. Next, we extend this $f(x)$ which has now been defined on $(-\pi, \pi)$ to $\mathbb{R}$ with periodicity 2π. Notice that since $f(0)$, $f(-\pi)$ and $f(\pi)$ are not given, we may have to define them. The only constraint is that $f(x)$ is required to be even; so we must take $f(\pi) = f(-\pi)$. If we define these values in such a way that the extended function is continuous at these points, then the Fourier series will converge to these values at these points. Since the extended function is even and of period 2π, its Fourier series is a cosine series given by

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos nx, \quad a_n = \frac{2}{\pi}\int_0^{\pi} f(t)\cos nt\,dt, \quad n = 0, 1, 2, 3, \ldots, \quad x \in \mathbb{R}.$$

In this case, we say that the Fourier series is a **half-range cosine series** for $f(x)$.

In case, domain of f is $[0, \pi)$, $(0, \pi]$ or $[0, \pi]$, the values $f(0)$ and/or $f(\pi)$ may be already available, and we do not need to define those.

2. Odd Extension:

First, extend $f(x)$ from $(0, \pi)$ to $(-\pi, \pi)$ by requiring that $f(x)$ is an odd function. This requirement forces $f(-x) = -f(x)$ for each $x \in (-\pi, \pi)$. In particular, the extended function $f(x)$ will satisfy $f(0) = f(-0) = -f(0)$ leading to $f(0) = 0$. Next, we extend this $f(x)$ which has now been defined on $(-\pi, \pi)$ to $\mathbb{R}$ with periodicity 2π. This will force $f(-\pi) = f(\pi)$. Again, the requirement that $f(x)$ is odd implies that $f(-\pi) = -f(\pi) = f(\pi)$ leading to $f(-\pi) = f(\pi) = 0$. Thus, the odd extension of $f(x)$ with periodicity 2π will satisfy $f(k\pi) = 0$ for all integers k. By the convergence theorem, the Fourier series of the extended odd function $f(x)$ will be equal to $f(x)$ at each $x \neq k\pi$.

In addition, if f is already defined on $[0, \pi]$ with $f(0) = f(\pi) = 0$, then the Fourier series of the odd extension of $f(x)$ with period 2π will represent the function $f(x)$ at all points $x \in \mathbb{R}$.

The Fourier series expansion of this extended $f(x)$ is a sine series given by

$$f(x) = \sum_{n=1}^{\infty} b_n \sin nx, \quad b_n = \frac{2}{\pi}\int_0^{\pi} f(t)\sin nt\,dt, \quad n = 1, 2, 3, \ldots, \quad x \in \mathbb{R}.$$

In this case, we say that the Fourier series is a **half-range sine series** for $f(x)$.

(2.29) *Example*

Find half-range cosine and sine series for $f(x) = \begin{cases} x & \text{if } 0 \le x \le \pi/2 \\ \pi - x & \text{if } \pi/2 < x \le \pi. \end{cases}$

1. With an even extension, the Fourier coefficients are given by $a_0 = \pi/2$ and for $n = 1, 2, 3, \ldots$,

$$\begin{aligned} a_n &= \frac{2}{\pi}\int_0^{\pi} f(t)\cos nt\,dt = \frac{2}{\pi}\int_0^{\pi/2} t\cos nt\,dt + \frac{2}{\pi}\int_{\pi/2}^{\pi}(\pi - t)\cos nt\,dt \\ &= \frac{2}{\pi n^2}\Big(2\cos\frac{n\pi}{2} - 1 - \cos n\pi\Big). \end{aligned}$$

Notice that $a_n = 0$ when n is odd, and $a_n = 0$ when $n = 4k$ for any integer k. Thus

$$f(x) = \frac{\pi}{4} - \frac{2}{\pi}\Big(\frac{\cos 2x}{1^2} + \frac{\cos 6x}{3^2} + \frac{\cos 10x}{5^2} + \cdots\Big), \quad \text{for } x \in [0, \pi].$$

2. With an odd extension, the Fourier coefficients are given by $a_n = 0$ and for $n = 1, 2, 3, \ldots$,

$$\begin{aligned} b_n &= \frac{2}{\pi}\int_0^{\pi} f(t)\sin nt\,dt = \frac{2}{\pi}\int_0^{\pi/2} t\sin nt\,dt + \frac{2}{\pi}\int_{\pi/2}^{\pi}(\pi - t)\sin nt\,dt \\ &= \frac{2}{\pi}\Big[-\frac{t}{n}\cos nt + \frac{1}{n^2}\sin nt\Big]_0^{\pi/2} + \frac{2}{\pi}\Big[\frac{t-\pi}{n}\cos nt - \frac{1}{n^2}\sin nt\Big]_{\pi/2}^{\pi} \\ &= \frac{2}{\pi}\Big(-\frac{\pi}{2n}\cos\frac{n\pi}{2} + \frac{1}{n^2}\sin\frac{n\pi}{2}\Big) + \frac{2}{\pi}\Big(\frac{\pi}{2n}\cos\frac{2\pi}{n} + \frac{1}{n^2}\sin\frac{n\pi}{2}\Big) \\ &= \frac{4}{\pi n^2}\sin\frac{n\pi}{2} = \begin{cases} \frac{4}{\pi n^2}(-1)^{(n-1)/2} & \text{for } n \text{ odd} \\ 0 & \text{for } n \text{ even.} \end{cases} \end{aligned}$$

Here, $f(0) = f(\pi) = 0$. Thus $f(x) = \dfrac{4}{\pi}\Big(\dfrac{\sin x}{1^2} - \dfrac{\sin 3x}{3^2} + \dfrac{\sin 5x}{5^2} - \cdots\Big), \quad$ for $x \in [0, \pi]$.

Justify why do we write the equalities $f(x) = \cdots$ in both the cases. ∎

(2.30) *Example*

Find the sine series expansion of $\cos x$ in $[0, \pi]$.

We work with the odd extension of $\cos x$ with period 2π to $\mathbb{R}$. Observe that the odd extension $f(x)$ of $\cos x$ has the following values in $[-\pi, \pi]$:

$$f(x) = \begin{cases} -\cos x & \text{if } -\pi < x < 0 \\ \cos x & \text{if } 0 < x < \pi \\ 0 & \text{if } x = -\pi, 0, \pi. \end{cases}$$

The Fourier coefficients are given by $a_n = 0$ and

$$b_n = \frac{2}{\pi}\int_0^{\pi} \cos t \sin nt\, dt = \begin{cases} 0 & \text{for } n \text{ odd} \\ \frac{4n}{\pi(n^2-1)} & \text{for } n \text{ even.} \end{cases}$$

Therefore, the Fourier sine expansion of $\cos x$ is $\displaystyle\frac{8}{\pi}\sum_{n=1}^{\infty}\frac{n\,\sin(2nx)}{4n^2-1}$ in $[0, \pi]$.
At $x = 0$ and $x = \pi$, the Fourier series sums to $0 = f(0) = f(\pi)$, but different from $\cos 0$ and $\cos \pi$. At other points, the series represents $\cos x$. That is,

$$\cos x = \frac{8}{\pi}\sum_{n=1}^{\infty}\frac{n\,\sin(2nx)}{4n^2-1} \quad \text{for } x \in (0, \pi).$$

Similarly, you can verify that $\displaystyle\sin x = \frac{2}{\pi} - \frac{4}{\pi}\sum_{n=1}^{\infty}\frac{\cos(2nx)}{4n^2-1}$ for $x \in (0, \pi)$. □

There is another approach to construct a Fourier series for a function with domain as $(0, \pi)$, which we discuss in the last section.

2.7 Functions defined on $(-\ell, \ell)$

Suppose $f : (-\ell, \ell) \to \mathbb{R}$ is given. We first extend $f(x)$ to a 2ℓ-periodic function on $\mathbb{R}$. While constructing such an extension, we may have to define $f(-\ell)$. Especially, we take $f(-\ell) = \lim_{x \to -\ell+} f(x)$ so that the extended function is continuous at $x = -\ell$. Of course, if originally f is a function given on $[-\ell, \ell)$, then this definition of $f(-\ell)$ is not necessary. Next, we change the independent variable by taking $x = \ell y/\pi$, equivalently, $y = \pi x/\ell$. That is, we define the function $h(y) = f(\ell y/\pi)$. Since $f(x)$ is a 2ℓ-periodic function, $h(y)$ is a 2π-periodic function. Then, we construct the Fourier series for $h(y)$ and in this Fourier series substitute $y = \pi x/\ell$ for obtaining the Fourier series for $f(x)$. Now, the Fourier series for $h(y)$ is given by

$$\frac{a_0}{2} + \sum_{n=1}^{\infty}\Big(a_n \cos ny + b_n \sin ny\Big),$$

where the Fourier coefficients are (In a_n, $n = 0, 1, 2, 3, \ldots$; and in b_n, $n = 1, 2, 3, \ldots$.)

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f\Big(\frac{\ell}{\pi}s\Big)\cos ns\, ds, \quad b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f\Big(\frac{\ell}{\pi}s\Big)\cos ns\, ds.$$

Substituting $t = \frac{\ell}{\pi}s, \quad ds = \frac{\pi}{\ell}dt$, we have

$$a_n = \frac{1}{\ell}\int_{-\ell}^{\ell} f(t)\cos\left(\frac{n\pi t}{\ell}\right)dt, \quad b_n = \frac{1}{\ell}\int_{-\ell}^{\ell} f(t)\sin\left(\frac{n\pi t}{\ell}\right)dt.$$

Substituting $y = \pi x/\ell$ in the above Fourier series, we obtain the Fourier series for $f(x)$, which is

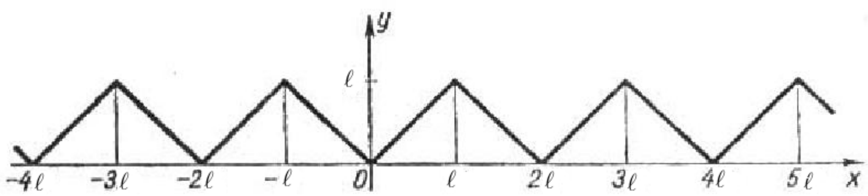
$$\frac{a_0}{2} + \sum_{n=1}^{\infty}\left[a_n\cos\left(\frac{n\pi x}{\ell}\right) + b_n\sin\left(\frac{n\pi x}{\ell}\right)\right].$$

In general, such a Fourier series need neither be a sine series nor a cosine series. Sometimes, this series is called a *full range* Fourier series.

(2.31) *Example*

Construct the Fourier series for $f(x) = |x|$ for $x \in [-\ell, \ell)$ for a given $\ell > 0$.

We extend the given function to $f : \mathbb{R} \to \mathbb{R}$ with period 2ℓ. It is shown in the following figure:



Notice that the function $f : \mathbb{R} \to \mathbb{R}$ is not $|x|$; it is $|x|$ on $[-\ell, \ell)$. Due to its period as 2ℓ, it is $|x - 2\ell|$ on $[\ell, 3\ell)$ etc. However, it is an even function; so its Fourier series is a cosine function.

The Fourier coefficients are

$$b_n = 0, \quad a_0 = \frac{2}{\ell}\int_0^{\ell}|t|\,dt = \frac{2}{\ell}\int_0^{\ell} t\,dt = \ell,$$

$$a_n = \frac{2}{\ell}\int_0^{\ell} t\cos\left(\frac{n\pi t}{\ell}\right)dt = \begin{cases} 0 & \text{for } n \text{ even} \\ -\frac{4\ell}{n^2\pi^2} & \text{for } n \text{ odd} \end{cases}$$

Therefore the (full range) Fourier series for $f(x)$ shows that in $[-\ell, \ell]$,

$$|x| = \frac{\ell}{2} - \frac{4\ell}{\pi^2}\left[\frac{\cos(\pi x/\ell)}{1^2} + \frac{\cos(3\pi x/\ell)}{3^2} + \cdots + \frac{\cos((2n+1)\pi x/\ell)}{(2n+1)^2} + \cdots\right].$$

Notice that the Fourier series represents the function at the end-points $x = -\ell$ and $x = \ell$ also since the function is continuous at these points. ∎

2.8 Functions defined on $(0, \ell)$

Suppose that $f : (0, \ell) \to \mathbb{R}$ is given for which we require a Fourier series. As in the case of functions defined on $(0, \pi)$, we have essentially two approaches to the problem. One, by an even extension, and the other, by an odd extension.

1. Even extension: In this approach, we extend $f(x)$ to $(-\ell, \ell)$ so that the extended function is an even function. Then, we find its Fourier series following the idea of the last section. That is, we first impose $f(-x) = f(x)$ for $x \in (0, \ell)$ and define $f(0)$ (anyway we like) so that f on $(-\ell, \ell)$ is an even function. Of course, if $f : [0, \ell) \to \mathbb{R}$, then we do not have to define $f(0)$. Next, we extend this $f(x)$ to $\mathbb{R}$ by making it periodic with period 2ℓ. Then, we find the Fourier series for the function $h(y) = f(\ell y/\pi)$, and then substitute $y = \pi x/\ell$ in this Fourier series to obtain the Fourier series for $f(x)$. Since $f(x)$ has been first extended to an even function of period 2ℓ, the function $h(y)$ is even and it has period 2π so that the final Fourier series is a cosine series. In this case, the Fourier series for $h(y)$ is

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} \Big(a_n \cos ny \Big), \quad a_n = \frac{2}{\pi} \int_0^{\pi} f\Big(\frac{\ell}{\pi} s\Big) \cos ns \, ds, \quad n = 0, 1, 2, 3, \ldots.$$

Substituting $t = \frac{\ell}{\pi} s$, $ds = \frac{\pi}{\ell} dt$, and $y = \pi x/\ell$ in the above Fourier series, we obtain the Fourier series for $f(x)$. It is

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} \Big(a_n \cos \frac{n\pi x}{\ell} \Big), \quad a_n = \frac{2}{\ell} \int_0^{\ell} f(t) \cos \Big(\frac{n\pi t}{\ell} \Big) \, dt, \quad n = 0, 1, 2, 3, \ldots.$$

The Fourier series sums to the function $f : \mathbb{R} \to \mathbb{R}$, where $f(x)$ is an even function of period 2ℓ and it agrees with the given function $f(x)$ on $(0, \ell)$. In particular, if $f(x)$ is piecewise continuous on $(0, \ell)$, then at $x = 0$, the Fourier series sums to $\lim_{x \to 0+} f(x)$ and at $x = \ell$, the Fourier series sums to $\lim_{x \to \ell-} f(x)$.

The Fourier series so obtained is called the **half-range cosine series** as earlier.

(2.32) *Example*

Find the half-range cosine series for $f(x) = 2x - 1$ for $0 < x < 1$ and show that

$$\frac{1}{1^2} + \frac{1}{3^2} + \frac{1}{5^2} + \cdots = \frac{\pi^2}{8}.$$

Here, the half-range cosine series is

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos \Big(\frac{n\pi x}{1} \Big),$$

where $a_0 = \frac{2}{1}\int_0^1 (2x-1)dx = 0$ and

$$a_n = \frac{2}{1}\int_0^1 (2x-1)\cos\left(\frac{n\pi x}{1}\right)dx = \frac{4}{n^2\pi^2}\left[(-1)^n - 1\right] = \begin{cases} 0 & \text{for } n \text{ even} \\ \frac{-8}{n^2\pi^2} & \text{for } n \text{ odd.} \end{cases}$$

Notice that $f(x)$ is continuous on $0 < x < 1$. Hence,

$$2x - 1 = -\frac{8}{\pi^2}\left[\frac{\cos \pi x}{1^2} + \frac{\cos 3\pi x}{3^2} + \frac{\cos 5\pi x}{5^2} + \cdots\right] \quad \text{for } 0 < x < 1.$$

At $x = 0$, the Fourier series sums to $\lim\limits_{x\to 0+}(2x-1) = -1$. Hence,

$$\frac{1}{1^2} + \frac{1}{3^2} + \frac{1}{5^2} + \cdots = \frac{\pi^2}{8}. \qquad \square$$

2. Odd extension: In this approach, we extend $f(x)$ to $(-\ell, \ell)$ so that the extended function is an odd function. Then, we find its Fourier series following the idea of the last section. That is, we first impose $f(-x) = -f(x)$ for $x \in (0, \ell)$ so that f on $(-\ell, \ell)$ is an odd function. Notice that it also requires setting $f(0) = f(-\ell) = 0$. Next, we extend this $f(x)$ to $\mathbb{R}$ by making it periodic with period 2ℓ. Then, we find the Fourier series for the function $h(y) = f(\ell y/\pi)$, and then substitute $y = \pi x/\ell$ in this Fourier series to obtain the Fourier series for $f(x)$. Since $f(x)$ has been first extended to an odd function of period 2ℓ, the function $h(y)$ is odd and it has period 2π so that its Fourier series is a sine series. In this case, the Fourier series for $h(y)$ is

$$\sum_{n=1}^{\infty}\Big(b_n \sin ny\Big), \quad b_n = \frac{2}{\pi}\int_0^{\pi} f\Big(\frac{\ell}{\pi}s\Big)\sin ns\, ds, \quad n = 1, 2, 3, \ldots.$$

Substituting $t = \frac{\ell}{\pi}s,\ ds = \frac{\pi}{\ell}dt,$ and $y = \pi x/\ell$ in the Fourier series, we obtain the Fourier series for $f(x)$. It is

$$\sum_{n=1}^{\infty} b_n \sin\left(\frac{n\pi x}{\ell}\right), \quad b_n = \frac{2}{\ell}\int_0^{\ell} f(t)\sin\left(\frac{n\pi t}{\ell}\right)dt, \quad n = 1, 2, 3, \ldots.$$

The Fourier series sums to the function $f : \mathbb{R} \to \mathbb{R}$, where $f(x)$ is an odd function of period 2ℓ and it agrees with the given function $f(x)$ on $(0, \ell)$. In particular, at $x = 0$ and at $x = \ell$, the Fourier series sums to 0.

The series so obtained is called a **half-range sine series**.

(2.33) ***Example***

Find the half-range sine series for $f(x) = \begin{cases} 1 & \text{if } 0 < x < 1 \\ 0 & \text{if } 1 \le x < 2. \end{cases}$

Here, $\ell = 2$. Thus, the Fourier sine series is $\sum_{n=1}^{\infty} b_n \sin\left(\frac{n\pi x}{2}\right)$, where

$$b_n = \frac{2}{2}\int_0^2 f(x)\sin\left(\frac{n\pi t}{2}\right)dt = \int_0^1 \sin\left(\frac{n\pi t}{2}\right)dt = \frac{2}{n\pi}\left[1 - \cos\left(\frac{n\pi}{2}\right)\right].$$ □

2.9 Functions defined on (a, b)

We now consider determining Fourier series for functions $f : (a, b) \to \mathbb{R}$. In applications, we do not usually encounter such a case. We go ahead for theoretical interest. Since we know how to construct Fourier series for a function with domain as $(0, \pi)$ or as $(-\pi, \pi)$, we have the following three approaches to the problem.

1. In the *first approach*, we define the continuous bijection $g : (0, \pi) \to (a, b)$ given by

$$g(y) = a + \frac{b-a}{\pi}\, y \quad \text{for } y \in (0, \pi).$$

Now, the composition $f \circ g$ is a function from $(0, \pi)$ to $\mathbb{R}$. Next, we take an even extension of $f \circ g$ with periodicity 2π; and call this extended function as h. We then construct the Fourier series for $h(y) = f(g(y))$. Finally, we substitute

$$y = g^{-1}(x) = \frac{\pi(x-a)}{b-a} \quad \text{for } x \in (a, b)$$

in the Fourier series for $h(y)$. This gives the **half-range cosine series** for $f(x)$ on (a, b).

2. In the *second approach*, we use the same continuous bijection $g : (0, \pi) \to (a, b)$ given by

$$g(y) = a + \frac{b-a}{\pi}\, y \quad \text{for } y \in (0, \pi).$$

The composition map $f \circ g$ is a function from $(0, \pi)$ to $\mathbb{R}$. Next, we take an odd extension of $f \circ g$ with periodicity 2π; and call this extended function as h. We then construct the Fourier series for $h(y)$. Finally, we substitute

$$y = g^{-1}(x) = \frac{\pi(x-a)}{b-a} \quad \text{for } x \in (a, b)$$

in the Fourier series for $h(y)$. This gives the **half-range sine series** for $f(x)$ on (a, b).

3. In the *third approach*, we define the continuous bijection $g : (-\pi, \pi) \to (a, b)$ given by

$$g(y) = \frac{a+b}{2} + \frac{b-a}{2\pi}y \quad \text{for } y \in (-\pi, \pi).$$

The composition $f \circ g$ is a function from $(-\pi, \pi)$ to $\mathbb{R}$. Next, we extend $f \circ g$ with periodicity 2π; and call this extended function as h. We then construct the Fourier series for $h(y)$. Finally, in the Fourier series for $h(y)$, we substitute

$$y = g^{-1}(x) = \frac{\pi}{b-a}(2x - a - b)$$

to obtain the Fourier series for $f(x)$ on (a, b).

In general, such a Fourier series need neither be a sine series nor a cosine series. Sometimes, this series is called a *full range* Fourier series.

Observe that if a function f has domain $(0, \pi)$, then we can scale it to length 2π by using the third approach. That is, we take the function $g : (-\pi, \pi) \to (0, \pi)$ as

$$g(y) = \frac{\pi + y}{2} \quad \text{for } y \in (0, \pi).$$

Then, we find the Fourier series for $h(y) = f(g(y))$. Finally, in the Fourier series for $h(y)$, we substitute $y = g^{-1}(x) = 2x - \pi$ to obtain the Fourier series for $f(x)$. In general, such a Fourier series may involve both sine and cosine terms.

(2.34) *Example*

We consider Example 2.29 once again to illustrate the third approach when the function is initially defined on $[0, \pi]$. There, we had

$$f(x) = \begin{cases} x & \text{if } 0 \le x \le \pi/2 \\ \pi - x & \text{if } \pi/2 < x \le \pi. \end{cases}$$

Here, $f(x)$ is also given at the end-points. We use the same formula for $g(x)$ at the end-points also. That is, we define $g : [-\pi, \pi] \to [0, \pi]$ by

$$x = g(y) = \frac{\pi + y}{2}, \quad h(y) = f\Big(\frac{\pi + y}{2}\Big) = \begin{cases} (\pi + y)/2 & \text{if } -\pi \le y \le 0 \\ (\pi - y)/2 & \text{if } 0 < y \le \pi. \end{cases}$$

Notice that the function $h(y)$ happens to be an even function with $h(-\pi) = h(\pi) = 0$. Then the Fourier coefficients are given by $b_n = 0$ and

$$a_0 = \frac{1}{\pi}\int_{-\pi}^{0} \frac{\pi + t}{2}\, dt + \frac{1}{\pi}\int_{0}^{\pi} \frac{\pi - t}{2}\, dt = \frac{\pi}{2}.$$

$$a_n = \frac{1}{\pi}\int_{-\pi}^{0}\frac{\pi+t}{2}\cos nt\,dt + \frac{1}{\pi}\int_{0}^{\pi}\frac{\pi-t}{2}\cos nt\,dt = \begin{cases}\frac{2}{\pi n^2} & \text{for } n \text{ odd}\\ 0 & \text{for } n \text{ even.}\end{cases}$$

The Fourier series for $h(y)$ is given by

$$\frac{\pi}{4} + \sum_{n \text{ odd}} \frac{2}{\pi n^2}\cos ny.$$

Using $y = g^{-1}(x) = 2x - \pi$, we have the Fourier series for $f(x)$. Also $f(x)$ is continuous at $x = 0$ and $x = \pi$. Therefore,

$$f(x) = \frac{\pi}{4} + \frac{2}{\pi}\sum_{n=1}^{\infty}\frac{1}{(2n+1)^2}\cos\big((2n+1)(2x-\pi)\big) \quad \text{for } x \in [0, \pi].$$

Notice that it is the same series we obtained earlier in Example 2.29(1) by even extension. This is so because $h(y)$ happens to be an even function. ∎

A Fun Problem: Show that the nth partial sum of the Fourier series for $f(x)$ can be written as the following integral:

$$s_n(x) = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x+t)\frac{\sin(2n+1)t/2}{2\sin t/2}\,dt.$$

We know that $s_n(x) = \frac{a_0}{2} + \sum_{k=1}^{n}(a_k\cos kx + b_k\sin kx)$, where

$$a_k = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\cos kt\,dt, \quad b_k = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\sin kt\,dt.$$

Substituting these values in the expression for $s_n(x)$, we have

$$\begin{aligned}
s_n(x) &= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)\,dt + \frac{1}{\pi}\sum_{k=1}^{n}\Big[\int_{-\pi}^{\pi} f(t)\cos kx\,\cos kt\,dt + \int_{-\pi}^{\pi} f(t)\sin kx\,\sin kt\,dt\Big]\\
&= \frac{1}{\pi}\int_{-\pi}^{\pi}\Big[\frac{f(t)}{2} + \sum_{k=1}^{n}\{f(t)\cos kx\,\cos kt + f(t)\sin kx\,\sin kt\}\Big]\,dt\\
&= \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\Big[\frac{1}{2} + \sum_{k=1}^{n}\cos k(t-x)\Big]\,dt := \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\sigma_n(t-x)\,dt.
\end{aligned}$$

The expression $\sigma_n(z)$ for $z = t - x$ can be re-written as follows:

$$\sigma_n(z) = \frac{1}{2} + \cos z + \cos 2z + \cdots + \cos nz.$$

Thus

$$\begin{aligned}
&2\sigma_n(z)\cos z \\
&= \cos z + 2\cos z\cos z + 2\cos z\cos 2z + \cdots + 2\cos z\cos nz \\
&= \cos z + [1+\cos 2z] + [\cos z + \cos 3z] + \cdots + [\cos(n-1)z + \cos(n+1)z] \\
&= 1 + 2\cos z + 2\cos 2z + \cdots + 2\cos(n-1)z + 2\cos nz + 2\cos(n+1)z \\
&= 2\sigma_n(z) - \cos nz + \cos(n+1)z.
\end{aligned}$$

This gives

$$\sigma_n(z) = \frac{\cos nz - \cos(n+1)z}{2(1-\cos z)} = \frac{\sin(2n+1)z/2}{2\sin z/2}.$$

Substituting $\sigma_n(z)$ with $z = t - x$, we have

$$s_n(x) = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\frac{\sin(2n+1)(t-x)/2}{2\sin(t-x)/2}\,dt.$$

Since the integrand is 2π-periodic, the value of the integral remains same on any interval of length 2π. Thus

$$s_n(x) = \frac{1}{\pi}\int_{x-\pi}^{x+\pi} f(t)\frac{\sin(2n+1)(t-x)/2}{2\sin(t-x)/2}\,dt.$$

Introduce a new variable $y = t - x$, i.e., $t = x + y$. And then write the integral in terms of t instead of y to obtain

$$s_n(x) = \int_{-\pi}^{\pi} f(x+y)\frac{\sin(2n+1)y/2}{2\sin y/2}\,dy = \int_{-\pi}^{\pi} f(x+t)\frac{\sin(2n+1)t/2}{2\sin t/2}\,dt.$$

This integral is called the **Dirichlet Integral**. In particular, taking $f(x) = 1$, we see that $a_0 = 2$, $a_k = 0$ and $b_k = 0$ for $k \in \mathbb{N}$; and then we get the identity

$$\frac{1}{\pi}\int_{-\pi}^{\pi} \frac{\sin\big((2n+1)t/2\big)}{2\sin\big(t/2\big)}\,dt = 1 \quad \text{for each } n \in \mathbb{N}.$$

2.10 Exercises for Chapter 2

1. Determine the interval of convergence for each of the following power series:

   (a) $\displaystyle\sum_{n=1}^{\infty}\frac{x^n}{n}$ (b) $\displaystyle\sum_{n=1}^{\infty}\frac{x^n}{n^2}$ (c) $\displaystyle\sum_{n=0}^{\infty}(-1)^n\frac{x^{n+1}}{n+1}$

   Ans: (a) $[-1, 1)$ (b) $[-1, 1]$ (c) $(-1, 1]$.

2. Determine the interval of convergence of the series $\dfrac{2x}{1} - \dfrac{(2x)^2}{2} + \dfrac{(2x)^3}{3} - \cdots$.
   Ans: $(-1/2, 1/2]$.
3. Determine power series expansion of the following functions:
   (a) $\ln(1+x)$ (b) $\dfrac{\ln(1+x)}{1-x}$
4. The function $\dfrac{1}{1-x}$ has the power series representation $\sum_{n=0}^{\infty} x^n$ with interval of convergence $(-1, 1)$. Prove that the function has power series representation around any $c \neq 1$.
5. Find the sum of the alternating harmonic series $\displaystyle\sum_{n=0}^{\infty} \frac{(-1)^n}{n+1}$. *Ans*: $\ln 2$.
6. Give an approximation scheme for $\displaystyle\int_0^a \frac{\sin x}{x}\,dx$ where $a > 0$.
   Ans: $a + \sum_{n=1}^{\infty} (-a)^{2n+1}/[(2n+1)^2(2n)!]$.
7. Show that $\sin^{-1} x = x + \dfrac{1}{2}\cdot\dfrac{1}{3}x^3 + \dfrac{1\cdot 3}{2\cdot 4}\cdot\dfrac{1}{5}x^5 + \dfrac{1\cdot 3\cdot 5}{2\cdot 4\cdot 6}\cdot\dfrac{1}{7}x^7 + \cdots$ for $-1 < x < 1$.
   Then, deduce that $1 + \dfrac{1}{2}\dfrac{1}{3} + \dfrac{1\cdot 3}{2\cdot 4}\cdot\dfrac{1}{5} + \dfrac{1\cdot 3\cdot 5}{2\cdot 4\cdot 6}\cdot\dfrac{1}{7} + \cdots = \dfrac{\pi}{2}$.
8. Find the Fourier series of $f(x)$ given by: $f(x) = 0$ for $-\pi \le x < 0$; and $f(x) = 1$ for $0 \le x \le \pi$. Say also how the Fourier series represents $f(x)$. Hence give a series expansion of $\pi/4$.
   Ans: $1/2 + (2/\pi)\sum_{n=0}^{\infty}(2n+1)^{-1}\sin[(2n+1)x]$.
9. Considering the fourier series for $|x|$, deduce that $\dfrac{\pi^2}{8} = \displaystyle\sum_{n=0}^{\infty} \frac{1}{(2n+1)^2}$.
10. Considering the fourier series for x, deduce that $\dfrac{\pi}{4} = \displaystyle\sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1}$.
11. Considering the fourier series for $f(x)$ given by: $f(x) = -1$, for $-\pi \le x < 0$ and $f(x) = 1$ for $0 \le x \le \pi$ deduce that $\dfrac{\pi}{4} = \displaystyle\sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1}$.
12. Considering $f(x) = x^2$, show that for each $x \in [0, \pi]$,
    $$\frac{\pi^2}{6} + 2\sum_{n=1}^{\infty} \frac{(-1)^n \cos nx}{n^2} = \sum_{n=1}^{\infty} \frac{n^2\pi^2(-1)^{n+1} + 2(-1)^n - 2}{n^3\pi} \sin nx.$$
13. Represent the function $f(x) = 1 - |x|$ for $-1 \le x \le 1$ as a cosine series.
    Ans: $1/2 + (4/\pi^2)\sum_{n=0}^{\infty}(2n+1)^{-2}\cos[(2n+1)\pi x]$.

3

Basic Matrix Operations

3.1 Addition and multiplication

A **matrix** is a rectangular array of symbols. For us these symbols are real numbers or, in general, complex numbers. The individual numbers in the array are called the **entries** of the matrix. The number of rows and the number of columns in any matrix are necessarily positive integers. A matrix with m rows and n columns is called an $m \times n$ matrix and it may be written as

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix},$$

or as $A = [a_{ij}]$ for short with $a_{ij} \in \mathbb{F}$ for $i = 1, \ldots, m$ and $j = 1, \ldots, n$. The number a_{ij} which occurs at the entry in ith row and jth column is referred to as the (i, j)-th entry of the matrix $[a_{ij}]$.

As usual, $\mathbb{R}$ denotes the set of all real numbers and $\mathbb{C}$ denotes the set of all complex numbers. We will write $\mathbb{F}$ for either $\mathbb{R}$ or $\mathbb{C}$. The numbers in $\mathbb{F}$ will also be referred to as **scalars**. Thus each entry of a matrix is a scalar.

The set of all $m \times n$ matrices with entries from $\mathbb{F}$ will be denoted by $\mathbb{F}^{m\times n}$.

A **row vector** of size n is a matrix in $\mathbb{F}^{1\times n}$. A typical row vector is written as $[a_1 \ \cdots \ a_n]$. Similarly, a **column vector** of size n is a matrix in $\mathbb{F}^{n\times 1}$. A typical column vector is written as

$$\begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} \quad \text{or as} \quad \begin{bmatrix} a_1 & \cdots & a_n \end{bmatrix}^T$$

for saving space. We will write both $\mathbb{F}^{1\times n}$ and $\mathbb{F}^{n\times 1}$ as $\mathbb{F}^n$. The elements of $\mathbb{F}^n$, called **vectors** will be written as $(a_1, \ldots, a_n)$.

So, $(a_1, \ldots, a_n)$ is either the row vector $\begin{bmatrix} a_1 & \cdots & a_n \end{bmatrix}$ or the column vector $\begin{bmatrix} a_1 & \cdots & a_n \end{bmatrix}^T$.

Any matrix in $\mathbb{F}^{m\times n}$ is said to have its **size** as $m \times n$. If $m = n$, the rectangular array becomes a square array with n rows and n columns; and the matrix is called a square matrix of **order** n.

Two matrices of the same size are considered **equal** when their corresponding entries are equal, i.e., if $A = [a_{ij}] \in \mathbb{F}^{m\times n}$ and $B = [b_{ij}] \in \mathbb{F}^{m\times n}$, then

$$A = B \quad \text{iff} \quad a_{ij} = b_{ij}$$

for each $i \in \{1, \ldots, m\}$ and for each $j \in \{1, \ldots, n\}$. Thus matrices of different sizes are unequal.

Sum of two matrices of the same size is a matrix whose entries are obtained by adding the corresponding entries in the given two matrices. That is, if $A = [a_{ij}] \in \mathbb{F}^{m\times n}$ and $B = [b_{ij}] \in \mathbb{F}^{m\times n}$, then

$$A + B = [a_{ij} + b_{ij}] \in \mathbb{F}^{m\times n}.$$

For example,

$$\begin{bmatrix}1&2&3\\2&3&1\end{bmatrix} + \begin{bmatrix}3&1&2\\2&1&3\end{bmatrix} = \begin{bmatrix}4&3&5\\4&4&4\end{bmatrix}.$$

The + here is called **addition** as usual. We informally say that matrices are added entry-wise. Matrices of different sizes can never be added. It follows that

$$A + B = B + A$$

whenever + is defined. Similarly, matrices can be **multiplied by a scalar** entry-wise. If $\alpha \in \mathbb{F}$ and $A = [a_{ij}] \in \mathbb{F}^{m\times n}$, then

$$\alpha\, A = [\alpha\, a_{ij}] \in \mathbb{F}^{m\times n}.$$

We write the **zero matrix** in $\mathbb{F}^{m\times n}$, all entries of which are 0, as 0. Thus,

$$A + 0 = 0 + A = A$$

for all matrices $A \in \mathbb{F}^{m\times n}$, with an implicit understanding that $0 \in \mathbb{F}^{m\times n}$. For $A = [a_{ij}]$, the matrix $-A \in \mathbb{F}^{m\times n}$ is taken as one whose (ij)th entry is $-a_{ij}$. Thus

$$-A = (-1)A \quad \text{and} \quad A + (-A) = -A + A = 0.$$

We also abbreviate $A + (-B)$ to $A - B$, as usual.

For example,

$$3\begin{bmatrix}1&2&3\\2&3&1\end{bmatrix} - \begin{bmatrix}3&1&2\\2&1&3\end{bmatrix} = \begin{bmatrix}0&5&7\\4&8&0\end{bmatrix}.$$

The addition and scalar multiplication as defined above satisfy the following properties:

Let $A, B, C \in \mathbb{F}^{m\times n}$. Let $\alpha, \beta \in \mathbb{F}$.

1. $A + B = B + A$.
2. $(A + B) + C = A + (B + C)$.
3. $A + 0 = 0 + A = A$.
4. $A + (-A) = (-A) + A = 0$.
5. $\alpha(\beta A) = (\alpha\beta)A$.
6. $\alpha(A + B) = \alpha A + \alpha B$.
7. $(\alpha + \beta)A = \alpha A + \beta A$.
8. $1\,A = A$.

Notice that whatever we discuss here for matrices apply to row vectors and column vectors, in particular. But remember that a row vector cannot be added to a column vector unless both are of size 1×1.

Another operation that we have on matrices is **multiplication of matrices**, which is a bit involved. Let $A = [a_{ik}] \in \mathbb{F}^{m\times n}$ and $B = [b_{kj}] \in \mathbb{F}^{n\times r}$. Then their **product** AB is a matrix $[c_{ij}] \in \mathbb{F}^{m\times r}$, where the entries are given by

$$c_{ij} = a_{i1}b_{1j} + \cdots + a_{in}b_{nj} = \sum_{k=1}^{n} a_{ik}b_{kj}.$$

Notice that the matrix product AB is defined *only when the number of columns in A is equal to the number of rows in B.*

A particular case might be helpful. Suppose A is a row vector in $\mathbb{F}^{1\times n}$ and B is a column vector in $\mathbb{F}^{n\times 1}$. Then their product $AB \in \mathbb{F}^{1\times 1}$; it is a matrix of size 1×1. Often we will identify such matrices with numbers. The product now looks like:

$$\begin{bmatrix} a_1 & \cdots & a_n \end{bmatrix} \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} = \begin{bmatrix} a_1b_1 + \cdots + a_nb_n \end{bmatrix}$$

This is helpful in visualizing the general case, which looks like

$$\begin{bmatrix} a_{11} & & a_{1k} & & a_{1n} \\ & & & & \\ a_{i1} & \cdots & a_{ik} & \cdots & a_{in} \\ & & & & \\ a_{m1} & & a_{mk} & & a_{mn} \end{bmatrix} \begin{bmatrix} b_{11} & b_{1j} & b_{1r} \\ & \vdots & \\ b_{\ell 1} & b_{\ell j} & b_{\ell r} \\ & \vdots & \\ b_{n1} & b_{nj} & b_{nr} \end{bmatrix} = \begin{bmatrix} c_{11} & c_{1j} & c_{1r} \\ & & \\ c_{i1} & c_{ij} & c_{ir} \\ & & \\ c_{m1} & c_{mj} & c_{mr} \end{bmatrix}$$

The ith row of A multiplied with the jth column of B gives the (ij)th entry in AB. Thus to get AB, you have to multiply all m rows of A with all r columns of B.

(3.1) ***Example***

In the following, call the first matrix as A and the second as B. The equality shows how the right hand side AB is computed.

$$\begin{bmatrix} 3 & 5 & -1 \\ 4 & 0 & 2 \\ -6 & -3 & 2 \end{bmatrix} \begin{bmatrix} 2 & -2 & 3 & 1 \\ 5 & 0 & 7 & 8 \\ 9 & -4 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 22 & -2 & 43 & 42 \\ 26 & -16 & 14 & 6 \\ -9 & 4 & -37 & -28 \end{bmatrix}.$$

Further, look at the kth column of AB, the columns of A and the entries in the kth column of B. We have

$$\begin{bmatrix} 22 \\ 26 \\ -9 \end{bmatrix} = 2 \begin{bmatrix} 3 \\ 4 \\ -6 \end{bmatrix} + 5 \begin{bmatrix} 5 \\ 0 \\ -3 \end{bmatrix} + 9 \begin{bmatrix} -1 \\ 2 \\ 2 \end{bmatrix}.$$

$$\begin{bmatrix} -2 \\ -16 \\ 4 \end{bmatrix} = -2 \begin{bmatrix} 3 \\ 4 \\ -6 \end{bmatrix} + 0 \begin{bmatrix} 5 \\ 0 \\ -3 \end{bmatrix} - 4 \begin{bmatrix} -1 \\ 2 \\ 2 \end{bmatrix}.$$

$$\begin{bmatrix} 43 \\ 14 \\ -37 \end{bmatrix} = 3 \begin{bmatrix} 3 \\ 4 \\ -6 \end{bmatrix} + 7 \begin{bmatrix} 5 \\ 0 \\ -3 \end{bmatrix} + 1 \begin{bmatrix} -1 \\ 2 \\ 2 \end{bmatrix}.$$

$$\begin{bmatrix} 42 \\ 6 \\ -28 \end{bmatrix} = 1 \begin{bmatrix} 3 \\ 4 \\ -6 \end{bmatrix} + 8 \begin{bmatrix} 5 \\ 0 \\ -3 \end{bmatrix} + 1 \begin{bmatrix} -1 \\ 2 \\ 2 \end{bmatrix}.$$

In general, suppose that the columns of A are $C_1, \ldots, C_n$ and that the kth column of B has entries $b_{1k}, \ldots, b_{nk}$ in that order; then, the kth column of AB is given by $b_{1k}C_1 + \cdots + b_{nk}C_n$. $\square$

If $u \in \mathbb{F}^{1\times n}$ and $v \in \mathbb{F}^{n\times 1}$, then $uv \in \mathbb{F}^{1\times 1}$, which we identify with a scalar; but $vu \in \mathbb{F}^{n\times n}$.

$$\begin{bmatrix} 3 & 6 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix} = \begin{bmatrix} 19 \end{bmatrix}, \quad \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix} \begin{bmatrix} 3 & 6 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 6 & 1 \\ 6 & 12 & 2 \\ 12 & 24 & 4 \end{bmatrix}.$$

It shows clearly that matrix multiplication is not commutative. Commutativity can break down due to various reasons. First of all when AB is defined, BA may not be defined. Second, even when both AB and BA are defined, they may not be of the same size; and third, even when they are of the same size, they need not be equal. For example,

$$\begin{bmatrix}1 & 2\\ 2 & 3\end{bmatrix}\begin{bmatrix}0 & 1\\ 2 & 3\end{bmatrix} = \begin{bmatrix}4 & 7\\ 6 & 11\end{bmatrix} \quad \text{but} \quad \begin{bmatrix}0 & 1\\ 2 & 3\end{bmatrix}\begin{bmatrix}1 & 2\\ 2 & 3\end{bmatrix} = \begin{bmatrix}2 & 3\\ 8 & 13\end{bmatrix}.$$

It does not mean that AB is never equal to BA. There can be some particular matrices A and B both in $\mathbb{F}^{n\times n}$ such that $AB = BA$. An extreme case is $AI = IA$, where I is the **identity matrix** defined by $I = [\delta_{ij}]$, where Kronecker's delta is defined as follows:

$$\delta_{ij} = \begin{cases}1 & \text{if } i = j\\ 0 & \text{if } i \neq j\end{cases} \qquad \text{for } i, j \in \mathbb{N}.$$

In fact, I serves as the identity of multiplication. I looks like

$$\begin{bmatrix}1 & 0 & \cdots & 0 & 0\\ 0 & 1 & \cdots & 0 & 0\\ & & \vdots & & \\ 0 & 0 & \cdots & 1 & 0\\ 0 & 0 & \cdots & 0 & 1\end{bmatrix} = \begin{bmatrix}1 & & & & \\ & 1 & & & \\ & & \ddots & & \\ & & & 1 & \\ & & & & 1\end{bmatrix}.$$

We often do not write the zero entries for better visibility of some pattern.

Unlike numbers, product of two nonzero matrices can be a zero matrix. For example,

$$\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}\begin{bmatrix}0 & 0\\ 0 & 1\end{bmatrix} = \begin{bmatrix}0 & 0\\ 0 & 0\end{bmatrix}.$$

It is easy to verify the following properties of matrix multiplication:

1. If $A \in \mathbb{F}^{m\times n}$, $B \in \mathbb{F}^{n\times r}$ and $C \in \mathbb{F}^{r\times p}$, then $(AB)C = A(BC)$.
2. If $A, B \in \mathbb{F}^{m\times n}$ and $C \in \mathbb{F}^{n\times r}$, then $(A + B)C = AB + AC$.
3. If $A \in \mathbb{F}^{m\times n}$ and $B, C \in \mathbb{F}^{n\times r}$, then $A(B + C) = AB + AC$.
4. If $\alpha \in \mathbb{F}$, $A \in \mathbb{F}^{m\times n}$ and $B \in \mathbb{F}^{n\times r}$, then $\alpha(AB) = (\alpha A)B = A(\alpha B)$.

You can see matrix multiplication in a block form. Suppose for two matrices A and B, their product is well defined. By looking at smaller matrices in A and in B, we can write their product as follows:

$$A = \begin{bmatrix}A_1 & A_2\\ A_3 & A_4\end{bmatrix},\ B = \begin{bmatrix}B_1 & B_2\\ B_3 & B_4\end{bmatrix} \;\Rightarrow\; AB = \begin{bmatrix}A_1B_1 + A_2B_3 & A_1B_2 + A_2B_4\\ A_3B_1 + A_4B_3 & A_3B_2 + A_4B_4\end{bmatrix}$$

provided that the blocks A_i and B_j are such that all the products involved in the above are well-defined.

Powers of square matrices can be defined inductively by taking

$$A^0 = I \quad \text{and} \quad A^n = AA^{n-1} \text{ for } n \in \mathbb{N}.$$

A square matrix A of order m is called **invertible** iff there exists a matrix B of order m such that

$$AB = I = BA.$$

Such a matrix B is called an **inverse** of A. If C is another inverse of A, then

$$C = CI = C(AB) = (CA)B = IB = B.$$

Therefore, an inverse of a matrix is unique and is denoted by A^{-1}. We talk of invertibility of square matrices only; and all square matrices are not invertible. For example, I is invertible but 0 is not. If $AB = 0$ for nonzero square matrices A and B, then neither A nor B is invertible. (Show it.) Invertible matrices play a crucial role in solving linear systems uniquely. We will come back to the issue later.

It is easy to verify that if $A, B \in \mathbb{F}^{n\times n}$ are invertible matrices, then $(AB)^{-1} = B^{-1}A^{-1}$.

Remark 3.2 If A and B are square matrices of the same order, then $AB = I$ implies that $BA = I$. This fact will be proved later. It means that a square matrix A of order n is invertible iff there exists a square matrix B of order n such that $AB = I$ iff there exists a square matrix C of order n such that $CA = I$.

3.2 Transpose and adjoint

We consider another operation on matrices. Given a matrix $A \in \mathbb{F}^{m\times n}$, its **transpose** is a matrix in $\mathbb{F}^{n\times m}$, which is denoted by A^T, and is defined by

$$\text{the } (i, j)\text{th entry of } A^T = \text{the } (j, i)\text{th entry of} A.$$

That is, the ith column of A^T is the column vector $\begin{bmatrix} a_{i1} & \cdots & a_{in} \end{bmatrix}^T$. The rows of A are the columns of A^T and the columns of A are the rows of A^T. For example,

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{bmatrix} \Rightarrow A^T = \begin{bmatrix} 1 & 2 \\ 2 & 3 \\ 3 & 1 \end{bmatrix}.$$

In particular, if $u = \begin{bmatrix} a_1 & \cdots & a_m \end{bmatrix}$ is a row vector, then its transpose is

$$u^T = \begin{bmatrix} a_1 \\ \vdots \\ a_m \end{bmatrix},$$

which is a column vector. Similarly, the transpose of a column vector is a row vector. Notice that the transpose notation goes well with our style of writing a column vector as the transpose of a row vector. If you write A as a row of column vectors, then you can express A^T as a column of row vectors, as in the following:

$$A = \begin{bmatrix} C_1 \cdots C_n \end{bmatrix} \Rightarrow A^T = \begin{bmatrix} C_1^T \\ \vdots \\ C_n^T \end{bmatrix}, \quad A = \begin{bmatrix} R_1 \\ \vdots \\ R_m \end{bmatrix} \Rightarrow A^T = \begin{bmatrix} R_1^T \cdots R_m^T \end{bmatrix}.$$

The following are some of the properties of this operation of transpose.

1. $(A^T)^T = A$.
2. $(A + B)^T = A^T + B^T$.
3. $(\alpha A)^T = \alpha A^T$.
4. $(AB)^T = B^T A^T$.
5. If A is invertible, then A^T is invertible, and $(A^T)^{-1} = (A^{-1})^T$.

In the above properties, we assume that the operations are allowed, that is, in (2), A and B must be of the same size. Similarly, in (4), the number of columns in A must be equal to the number of rows in B; and in (5), A must be a square matrix.

It is easy to see all the above properties, except perhaps the fourth one. For this, let $A \in \mathbb{F}^{m\times n}$ and $B \in \mathbb{F}^{n\times r}$. Now, the (j, i)th entry in $(AB)^T$ is the (i, j)th entry in AB; and it is given by

$$a_{i1}b_{j1} + \cdots + a_{in}b_{jn}.$$

On the other side, the (j, i)th entry in $B^T A^T$ is obtained by multiplying the jth row of B^T with the ith column of A^T. This is same as multiplying the entries in the jth column of B with the corresponding entries in the ith row of A, and then taking the sum. Thus it is

$$b_{j1}a_{i1} + \cdots + b_{jn}a_{in}.$$

This is the same as computed earlier.

We write $\overline{\alpha}$ for the **complex conjugate** of a scalar α. That is, $\overline{b + ic} = b - ic$ for $b, c \in \mathbb{R}$. Thus, if $a_{ij} \in \mathbb{R}$, then $\overline{a}_{ij} = a_{ij}$.

Close to the operations of transpose of a matrix is the adjoint. Let $A = [a_{ij}] \in \mathbb{F}^{m\times n}$. The **adjoint** of A is denoted as A^*, and is defined by

$$\text{the } (i, j)\text{th entry of } A^* = \text{the complex conjugate of } (j, i)\text{th entry of} A.$$

The adjoint of A is also called the **conjugate transpose** of A.

When A has only real entries, $A^* = A^T$. The ith column of A^* is the column vector $\begin{bmatrix}\overline{a}_{i1}, & \cdots, & \overline{a}_{in}\end{bmatrix}^T$. For example,

$$A = \begin{bmatrix}1&2&3\\2&3&1\end{bmatrix} \Rightarrow A^* = \begin{bmatrix}1&2\\2&3\\3&1\end{bmatrix}, \quad B = \begin{bmatrix}1+i&2&3\\2&3&1-i\end{bmatrix} \Rightarrow B^* = \begin{bmatrix}1-i&2\\2&3\\3&1+i\end{bmatrix}.$$

Similar to the transpose, the adjoint satisfies the following properties:

1. $(A^*)^* = A$.
2. $(A + B)^* = A^* + B^*$.
3. $(\alpha A)^* = \overline{\alpha} A^*$.
4. $(AB)^* = B^* A^*$.
5. If A is invertible, then A^* is invertible, and $(A^*)^{-1} = (A^{-1})^*$.

Here also, in (2), the matrices A and B must be of the same size, and in (4), the number of columns in A must be equal to the number of rows in B.

3.3 Special types of matrices

Recall that the *zero matrix* is a matrix each entry of which is 0. We write 0 for all zero matrices of all sizes. The size is to be understood from the context.

Let $A = [a_{ij}] \in \mathbb{F}^{n\times n}$ be a square matrix of order n. The entries a_{ii} are called as the **diagonal entries** of A. The first diagonal entry is a_{11}, and the last diagonal entry is a_{nn}. The entries of A, which are not the diagonal entries, are called as **off-diagonal entries** of A; they are a_{ij} for $i \neq j$. In the following matrix, the diagonal entries are shown in red:

$$\begin{bmatrix}1&2&3\\2&3&4\\3&4&0\end{bmatrix}.$$

Here, 1 is the first diagonal entry, 3 is the second diagonal entry and 0 is the third and the last diagonal entry.

If all off-diagonal entries of A are 0, then A is said to be a **diagonal matrix**. Only a square matrix can be a diagonal matrix. There is a way to generalize this notion to any matrix, but we do not require it. Notice that the diagonal entries in a diagonal matrix need not all be nonzero. For example, the zero matrix of order n is also a diagonal matrix. The following is a diagonal matrix. We follow the convention of not showing the off-diagonal entries in a diagonal matrix.

$$\begin{bmatrix} 1 & & \\ & 3 & \\ & & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

We also write a diagonal matrix with diagonal entries $d_1, \ldots, d_n$ as $\operatorname{diag}(d_1, \ldots, d_n)$. Thus the above diagonal matrix is also written as

$$\operatorname{diag}(1, 3, 0).$$

Recall that the *identity matrix* is a square matrix of which each diagonal entry is 1 and each off-diagonal entry is 0. Obviously,

$$I^T = I^* = I^{-1} = \operatorname{diag}(1, \ldots, 1) = I.$$

When identity matrices of different orders are used in a context, we will use the notation I_m for the identity matrix of order m. If $A \in \mathbb{F}^{m \times n}$, then $AI_n = A$ and $I_m A = A$.

We write e_i for a column vector whose ith component is 1 and all other components 0. That is, the jth component of e_i is δ_{ij}. In $\mathbb{F}^{n \times 1}$, there are then n distinct column vectors

$$e_1, \ldots, e_n.$$

The e_is are referred to as the **standard basis vectors**. These are the columns of the identity matrix of order n, in that order; that is, e_i is the ith column of I. The transposes of these e_is are the rows of I. That is, the ith row of I is e_i^T. Thus

$$I = \begin{bmatrix} e_1 & \cdots & e_n \end{bmatrix} = \begin{bmatrix} e_1^T \\ \vdots \\ e_n^T \end{bmatrix}.$$

A **scalar matrix** is a square matrix having all diagonal entries equal, and all off-diagonal entries as 0. That is, a scalar matrix is of the form αI, for some scalar α. The following is a scalar matrix:

$$\operatorname{diag}(3, 3, 3, 3) = \begin{bmatrix} 3 & & & \\ & 3 & & \\ & & 3 & \\ & & & 3 \end{bmatrix}$$

If $A, B \in \mathbb{F}^{m\times m}$ and A is a scalar matrix, then $AB = BA$. Conversely, if $A \in \mathbb{F}^{m\times m}$ is such that $AB = BA$ for all $B \in \mathbb{F}^{m\times m}$, then A must be a scalar matrix. This fact is not obvious, and you should try to prove it.

A matrix $A \in \mathbb{F}^{m\times n}$ is said to be **upper triangular** iff all entries below the diagonal are zero. That is, $A = [a_{ij}]$ is upper triangular when $a_{ij} = 0$ for $i > j$. Similarly, a matrix is called **lower triangular** iff all its entries above the diagonal are zero. Both upper triangular and lower triangular matrices are referred to as **triangular** matrices. In the following, L is a lower triangular matrix, and U is an upper triangular matrix, both of order 3.

$$L = \begin{bmatrix} 1 & & \\ 2 & 3 & \\ 3 & 4 & 5 \end{bmatrix}, \quad U = \begin{bmatrix} 1 & 2 & 3 \\ & 3 & 4 \\ & & 5 \end{bmatrix}.$$

A diagonal matrix is both upper triangular and lower triangular. Transpose of a lower triangular matrix is an upper triangular matrix and vice versa.

A square matrix A is called **hermitian** iff $A^* = A$. And A is called **skew hermitian** iff $A^* = -A$. A hermitian matrix with real entries satisfies $A^T = A$; and accordingly, such a matrix is called a **real symmetric** matrix. In general, A is called a **symmetric** matrix iff $A^T = A$. We also say that A is **skew symmetric** iff $A^T = -A$. In the following, B is symmetric, C is skew-symmetric, D is hermitian, and E is skew-hermitian. B is also hermitian and C is also skew-hermitian.

$$B = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 5 \end{bmatrix}, \quad C = \begin{bmatrix} 0 & 2 & -3 \\ -2 & 0 & 4 \\ 3 & -4 & 0 \end{bmatrix}, \quad D = \begin{bmatrix} 1 & -2i & 3 \\ 2i & 3 & 4 \\ 3 & 4 & 5 \end{bmatrix}, \quad E = \begin{bmatrix} 0 & 2+i & 3 \\ -2+i & i & 4i \\ -3 & 4i & 0 \end{bmatrix}.$$

Notice that a skew-symmetric matrix must have a zero diagonal, and the diagonal entries of a skew-hermitian matrix must be 0 or purely imaginary. Reason:

$$a_{ii} = -\overline{a}_{ii} \Rightarrow 2\mathrm{Re}(a_{ii}) = 0.$$

Let A be a square matrix. Since $A + A^T$ is symmetric and $A - A^T$ is skew symmetric, every square matrix can be written as a sum of a symmetric matrix and a skew symmetric matrix:

$$A = \frac{1}{2}(A + A^T) + \frac{1}{2}(A - A^T).$$

Similar rewriting is possible with hermitian and skew hermitian matrices:

$$A = \frac{1}{2}(A + A^*) + \frac{1}{2}(A - A^*).$$

A square matrix A is called **unitary** iff $A^*A = I = AA^*$. In addition, if A is real, then it is called an orthogonal matrix. That is, an **orthogonal matrix** is a matrix with real entries satisfying $A^TA = I = AA^T$. Notice that a square matrix is unitary iff it is invertible and its inverse is equal to its adjoint. Similarly, a real matrix is orthogonal iff it is invertible and its inverse is its transpose. In the following, B is a unitary matrix of order 2, and C is an orthogonal matrix (also unitary) of order 3:

$$B = \frac{1}{2}\begin{bmatrix} 1+i & 1-i \\ 1-i & 1+i \end{bmatrix}, \quad C = \frac{1}{3}\begin{bmatrix} 2 & 1 & 2 \\ -2 & 2 & 1 \\ 1 & 2 & -2 \end{bmatrix}.$$

The following are examples of orthogonal 2×2 matrices, for any fixed $\theta \in \mathbb{F}$:

$$O_1 := \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}, \quad O_2 := \begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix}.$$

O_1 is said to be a *rotation by an angle* θ and O_2 is called a *reflection by an angle* $\theta/2$ along the x-axis. Can you say why are they so called?

A square matrix A is called **normal** iff $A^*A = AA^*$. All hermitian matrices and all real symmetric matrices are normal matrices. For example,

$$\begin{bmatrix} 1+i & 1+i \\ -1-i & 1+i \end{bmatrix}$$

is a normal matrix; verify this. Also see that this matrix is neither hermitian nor skew-hermitian. In fact, a matrix is normal iff it is in the form $B + iC$, where B, C are hermitian and $BC = CB$. Can you prove this fact?

3.4 Linear independence

We look at row and column vectors and study some of their properties.

If $v = \begin{bmatrix} a_1 & \cdots & a_n \end{bmatrix}^T \in \mathbb{F}^{n\times 1}$, then we can express v in terms of the standard basis vectors $e_1, \ldots, e_n$ as $v = a_1e_1 + \cdots + a_ne_n$. We now generalize the notions involved.

Let $v_1, \ldots, v_m, v \in \mathbb{F}^n$. We say that v is a **linear combination** of the vectors $v_1, \ldots, v_m$ if there exist scalars $\alpha_1, \ldots, \alpha_m \in \mathbb{F}$ such that

$$v = \alpha_1 v_1 + \cdots \alpha_m v_m.$$

For example, in $\mathbb{F}^{2\times 1}$, one linear combination of $v_1 = \begin{bmatrix} 1 & 1 \end{bmatrix}^T$ and $v_2 = \begin{bmatrix} 1 & -1 \end{bmatrix}^T$ is as follows:

$$2\begin{bmatrix} 1 \\ 1 \end{bmatrix} + 1\begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}.$$

Is $\begin{bmatrix}4 & 2\end{bmatrix}^T$ a linear combination of v_1 and v_2? Yes, since

$$\begin{bmatrix}4\\-1\end{bmatrix} = 1\begin{bmatrix}1\\1\end{bmatrix} + 3\begin{bmatrix}1\\-1\end{bmatrix}.$$

In fact, every vector in $\mathbb{F}^{2\times 1}$ is a linear combination of v_1 and v_2. Reason:

$$\begin{bmatrix}a\\b\end{bmatrix} = \frac{a+b}{2}\begin{bmatrix}1\\1\end{bmatrix} + \frac{a-b}{2}\begin{bmatrix}1\\-1\end{bmatrix}.$$

However, every vector in $\mathbb{F}^{2\times 1}$ is not a linear combination of $\begin{bmatrix}1 & 1\end{bmatrix}^T$ and $\begin{bmatrix}2 & 2\end{bmatrix}^T$. Reason? Any linear combination of these two vectors is a multiple of $\begin{bmatrix}1 & 1\end{bmatrix}^T$. Then $\begin{bmatrix}1 & 0\end{bmatrix}^T$ is not a linear combination of these two vectors.

Remark 3.3 As we have seen in (3.1), the columns of AB are linear combinations of columns of A. The scalars in these linear combinations are the entries of corresponding columns of B.

The vectors $v_1, \ldots, v_m$ in $\mathbb{F}^n$ are called **linearly dependent** iff at least one of them is a linear combination of others. The vectors are called **linearly independent** iff none of them is a linear combination of others.

For example, $(1,1),\ (1,-1),\ (4,1)$ are linearly dependent vectors whereas $(1,1)$, $(1,-1)$ are linearly independent vectors.

Linear independence of a list of vectors can be characterized the following way:

(3.4) ***Theorem***

The vectors $v_1, \ldots, v_m \in \mathbb{F}^n$ are linearly independent iff for $\alpha_1, \ldots, \alpha_m \in \mathbb{F}$,

$$\textit{if } \alpha_1 v_1 + \cdots \alpha_m v_m = 0 \textit{ then } \alpha_1 = \cdots = \alpha_m = 0.$$

Notice that if $\alpha_1 = \cdots = \alpha_m = 0$, then obviously, $\alpha_1 v_1 + \cdots + \alpha_m v_m = 0$. But the above characterization demands its converse. It says that if you start with a linear combination and equate it to 0, then you must be able to derive that each coefficient in that linear combination is 0. That is, the only way the 0 vector can be written as a linear combination of the list of vectors $v_1, \ldots, v_m$ is the trivial linear combination, where each coefficient is 0. The condition given in the theorem is false when we have scalars not all zero such that the linear combination becomes 0. In fact, we prove this statement:

$v_1, \ldots, v_m$ are linearly dependent

iff $\alpha_1 v_1 + \cdots + \alpha_m v_m = 0$ for scalars $\alpha_1, \ldots, \alpha_m$ not all zero.

Proof. If the vectors $v_1, \ldots, v_m$ are linearly dependent then one of them is a linear combination of others. That is, we have an $i \in \{1, \ldots, m\}$ such that

$$v_i = \alpha_1 v_1 + \cdots + \alpha_{i-1} v_{i-1} + \alpha_{i+1} v_{i+1} + \cdots + \alpha_m v_m.$$

Then

$$\alpha_1 v_1 + \cdots + \alpha_{i-1} v_{i-1} + (-1) v_i + \alpha_{i+1} v_{i+1} + \cdots + \alpha_m v_m = 0.$$

Here, we see that a linear combination becomes zero, where at least one of the coefficients, that is, the ith one is nonzero.

Conversely, suppose we have scalars $\alpha_1, \ldots, \alpha_m$ not all zero such that

$$\alpha_1 v_1 + \cdots + \alpha_m v_m = 0.$$

Suppose $\alpha_j \neq 0$. Then

$$v_j = -\frac{\alpha_1}{\alpha_j} v_1 - \cdots - \frac{\alpha_{j-1}}{\alpha_j} v_{j-1} - \frac{\alpha_{j+1}}{\alpha_j} v_{j+1} - \cdots - \frac{\alpha_m}{\alpha_j} v_m.$$

That is, $v_1, \ldots, v_n$ are linearly dependent. ∎

(3.5) *Example*

Are the vectors $(1, 1, 1)$, $(2, 1, 1)$, $(3, 1, 0)$ linearly independent?

We start with an arbitrary linear combination and equate it to the zero vector. Solve the resulting linear equations to determine whether all the coefficients are necessarily 0 or not. So, let

$$a(1, 1, 1) + b(2, 1, 1) + c(3, 1, 0) = (0, 0, 0).$$

Comparing the components, we have

$$a + 2b + 3c = 0, \;\; a + b + c = 0, \;\; a + b = 0.$$

The last two equations imply that $c = 0$. Substituting in the first, we see that $a + 2b = 0$. This and the equation $a + b = 0$ give $b = 0$. Then it follows that $a = 0$.

We conclude that the given vectors are linearly independent. □

Caution: Be careful with the direction of implication here. Your work-out must be in the form

$$\sum_{i=1}^{m} \alpha_i v_i = 0 \Rightarrow \cdots \Rightarrow \text{ each } \alpha_i = 0.$$

To see how linear independence is helpful, consider the following system of linear equations:

$$\begin{aligned} x_1 + 2x_2 - 3x_3 &= 2 \\ 2x_1 - x_2 + 2x_3 &= 3 \\ 4x_1 + 3x_2 - 4x_3 &= 7 \end{aligned}$$

Here, we find that the third equation is redundant, since 2 times the first plus the second gives the third. That is, the third one linearly depends on the first two. (You can of course choose any other equation here as linearly depending on other two, but that is not important.) Now, take the row vectors of coefficients of the unknowns and the right hand side, as in the following:

$$v_1 = (1,\ 2,\ -3,\ 2), \quad v_2 = (2,\ -1,\ 2,\ 3), \quad v_3 = (4,\ 3,\ -4,\ 7).$$

We see that $v_3 = 2v_1 + v_2$, as it should be. Here, the list of vectors v_1, v_2, v_3 is linearly dependent. But the list v_1, v_2 is linearly independent. Thus, solving the given system of linear equations is the same thing as solving the system with only first two equations. For solving linear systems, it is of primary importance to find out which equations linearly depend on others. Once determined, such equations can be thrown away, and the rest can be solved.

A question: can you find four linearly independent vectors in $\mathbb{R}^{1\times 3}$?

3.5 Determinant

There are two important quantities associated with a square matrix. One is the trace and the other is the determinant.

The sum of all diagonal entries of a square matrix is called the **trace** of the matrix. That is, if $A = [a_{ij}] \in \mathbb{F}^{m\times m}$, then

$$\operatorname{tr}(A) = a_{11} + \cdots + a_{mm} = \sum_{k=1}^{m} a_{kk}.$$

In addition to $\operatorname{tr}(I_m) = m,\ \operatorname{tr}(0) = 0$, the trace satisfies the following properties:

Let $A, B \in \mathbb{F}^{m\times m}$. Let $\beta \in \mathbb{F}$.

1. $\operatorname{tr}(\beta A) = \beta \operatorname{tr}(A)$.
2. $\operatorname{tr}(A^T) = \operatorname{tr}(A)$ and $\operatorname{tr}(A^*) = \overline{\operatorname{tr}(A)}$.
3. $\operatorname{tr}(A + B) = \operatorname{tr}(A) + \operatorname{tr}(B)$ and $\operatorname{tr}(AB) = \operatorname{tr}(BA)$.
4. $\operatorname{tr}(A^*A) = 0$ iff $\operatorname{tr}(AA^*) = 0$ iff $A = 0$.

The last one follows from the observation that $\operatorname{tr}(A^*A) = \sum_{i=1}^{m} \sum_{j=1}^{m} |a_{ij}|^2 = \operatorname{tr}(AA^*)$.

The second quantity, called the **determinant** of a square matrix $A = [a_{ij}] \in \mathbb{F}^{n\times n}$, written as $\det(A)$, is defined inductively as follows:

If $n = 1$, then $\det(A) = a_{11}$.
If $n > 1$, then $\det(A) = \sum_{j=1}^{n} (-1)^{1+j}\, a_{1j} \det(A_{1j})$

where the matrix $A_{1j} \in \mathbb{F}^{(n-1)\times(n-1)}$ is obtained from A by deleting the first row and the jth column of A.

When $A = [a_{ij}]$ is written showing all its entries, we also write $\det(A)$ by replacing the two big closing brackets [and] by two vertical bars | and |. For a 2×2 matrix, its determinant is seen as follows:

$$\det \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = (-1)^{1+1} a_{11} \det[a_{22}] + (-1)^{1+2} a_{12} \det[a_{21}] = a_{11}a_{22} - a_{12}a_{21}.$$

Similarly, for a 3×3 matrix, we need to compute three 2×2 determinants. For example,

$$\begin{aligned} \det \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 2 \end{bmatrix} &= \begin{vmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 2 \end{vmatrix} \\ &= (-1)^{1+1} \times 1 \times \begin{vmatrix} 3 & 1 \\ 1 & 2 \end{vmatrix} + (-1)^{1+2} \times 2 \times \begin{vmatrix} 2 & 1 \\ 3 & 2 \end{vmatrix} + (-1)^{1+3} \times 3 \times \begin{vmatrix} 2 & 3 \\ 3 & 1 \end{vmatrix} \\ &= 1 \times \begin{vmatrix} 3 & 1 \\ 1 & 2 \end{vmatrix} - 2 \times \begin{vmatrix} 2 & 1 \\ 3 & 2 \end{vmatrix} + 3 \times \begin{vmatrix} 2 & 3 \\ 3 & 1 \end{vmatrix} \\ &= (3 \times 2 - 1 \times 1) - 2 \times (2 \times 2 - 1 \times 3) + 3 \times (2 \times 1 - 3 \times 3) \\ &= 5 - 2 \times 1 + 3 \times (-7) = -18. \end{aligned}$$

To see the determinant geometrically, consider a 2×2 matrix $A = [a_{ij}]$ with real entries. Let u be the vector with initial point at $(0, 0)$ and end-point at (a_{11}, a_{12}). Similarly, let v be the vector starting from the origin and ending at the point (a_{21}, a_{22}). Their sum $u + v$ is the vector whose initial point is the origin and end-point is $(a_{11} + a_{21}, a_{21} + a_{22})$. Denote by Δ, the area of the parallelogram with one vertex at the origin, and other vertices at the end-points of vectors u, v and $u + v$.

Writing the acute angle between the vectors u and v as θ, we have

$$\begin{aligned}\Delta^2 &= |u|^2|v|^2\sin^2\theta = |u|^2|v|^2(1-\cos^2\theta) = |u|^2|v|^2\left(1-\frac{(u\cdot v)^2}{|u|^2|v|^2}\right)\\ &= |u|^2|v|^2 - (u\cdot v)^2 = (a_{11}^2+a_{12}^2)(a_{21}^2+a_{22}^2) - (a_{11}a_{21}+a_{12}a_{22})^2\\ &= (a_{11}a_{22}-a_{12}a_{21})^2 = (\det(A))^2.\end{aligned}$$

That is, the absolute value of $\det(A)$ is the area of the parallelogram whose sides are represented by the row vectors of A. In $\mathbb{R}^{1\times 3}$, similarly, you can show that the absolute value of $\det(A)$ is the volume of the parallelepiped whose sides are represented by the row vectors of A.

For a lower triangular matrix, its determinant is the product of its diagonal entries.

$$\begin{vmatrix} a_{11} & & & \\ a_{12} & a_{22} & & \\ a_{13} & a_{23} & a_{33} & \\ & & \ddots & \\ a_{n1} & \cdots & & a_{nn}\end{vmatrix} = a_{11}\begin{vmatrix} a_{22} & & \\ a_{23} & a_{33} & \\ & \ddots & \\ a_{n1} & \cdots & a_{nn}\end{vmatrix} = \cdots = a_{11}a_{22}\cdots a_{nn}.$$

The determinant of any triangular matrix (upper or lower), is the product of its diagonal entries. In particular, the determinant of a diagonal matrix is also the product of its diagonal entries. Thus, if I is the identity matrix of order n, then $\det(I) = 1$ and $\det(-I) = (-1)^n$.

Our definition of determinant expands the determinant in the first row. In fact, the same result may be obtained by expanding it in any other row, or even any other column. Along with this, some more properties of the determinant is listed in the following theorem. We introduce some terminology to state the theorem.

Let $A \in \mathbb{F}^{n\times n}$.

Write the sub-matrix obtained from A by deleting the ith row and the jth column as A_{ij}. The (i, j)th **co-factor** of A is $(-1)^{i+j}\det(A_{ij})$; it is denoted by $C_{ij}(A)$. Sometimes, when the matrix A is fixed in a context, we write $C_{ij}(A)$ as C_{ij}.

The **adjugate** of A is the $n\times n$ matrix obtained by taking transpose of the matrix whose (i, j)th entry is $C_{ij}(A)$; it is denoted by $\mathrm{adj}(A)$. That is, $\mathrm{adj}(A) \in \mathbb{F}^{n\times n}$ is the matrix whose (i, j)th entry is the (j, i)th co-factor $C_{ji}(A)$.

Denote by $A_j(x)$ the matrix obtained from A by replacing the jth column of A by the (new) column vector $x \in \mathbb{F}^{n\times 1}$.

Some important facts about the determinant are listed below without proof.

(3.6) *Theorem*

Let $A \in \mathbb{F}^{n\times n}$. Let $i, j, k \in \{1, \ldots, n\}$. Then the following statements are true.

(1) $\det(A) = \sum_i a_{ij} C_{ij}(A) = \sum_i a_{ij}(-1)^{i+j}\det(A_{ij})$ *for any fixed* j.

(2) *For any* $j \in \{1, \ldots, n\}$, $\det(\,A_j(x+y)\,) = \det(\,A_j(x)\,) + \det(\,A_j(y)\,)$.

(3) *For any* $\alpha \in \mathbb{F}$, $\det(\,A_j(\alpha x)\,) = \alpha \det(\,A_j(x)\,)$.

(4) *For* $A \in \mathbb{F}^{n\times n}$, *let* $B \in \mathbb{F}^{n\times n}$ *be the matrix obtained from* A *by interchanging the* j*th and the* k*th columns, where* $j \neq k$. *Then* $\det(B) = -\det(A)$.

(5) *If a column of* A *is replaced by that column plus a scalar multiple of another column, then determinant does not change.*

(6) *Columns of* A *are linearly dependent iff* $\det(A) = 0$.

(7) $\det(A) = \sum_j a_{ij}(-1)^{i+j} A_{ij}$ *for any fixed* i.

(8) *All of* (2)-(6) *are true for rows instead of columns.*

(9) *If* A *is a triangular matrix, then* $\det(A)$ *is equal to the product of the diagonal entries of* A.

(10) $\det(AB) = \det(A)\det(B)$ *for any matrix* $B \in \mathbb{F}^{n\times n}$.

(11) *If* A *is invertible, then* $\det(A) \neq 0$ *and* $\det(A^{-1}) = (\det(A))^{-1}$.

(12) *If* $B = P^{-1}AP$ *for some invertible matrix* P, *then* $\det(A) = \det(B)$.

(13) A *is invertible iff columns of* A *are linearly independent iff rows of* A *are linearly independent iff* $\det(A) \neq 0$.

(14) $\det(A^T) = \det(A)$.

(15) $A\,\mathrm{adj}(A) = \mathrm{adj}(A)A = \det(A)\,I$.

Proof. (1) Construct a matrix $C \in \mathbb{F}^{(n+1)\times(n+1)}$ by taking its first row and first column as $e_1 \in \mathbb{F}^{(n+1)\times 1}$ and filling up the rest with the entries of A. In block form, it looks like:

$$C = \begin{bmatrix} 1 & \\ & A \end{bmatrix}.$$

Omitted entries are all 0. Then $\det(C) = \det(A)$. Now, exchange the first row and the $(i+1)$th rows in C. Call the matrix so obtained as D. Then

$$\det(C) = -\det(D) = -\sum_i a_{ij}(-1)^{i+1+j}\det(D_{ij}),$$

where $D_{ij} \in \mathbb{F}^{n\times n}$ are the minors in D. The ith row of D_{ij} is $e_i \in \mathbb{F}^{1\times n}$. To compute $\det(D_{ij})$, exchange the first and the ith rows in D_{ij}. Then $\det(D_{ij}) = -\det(A_{ij})$, where A_{ij} are the minors in A. Therefore,

$$\det(A) = \det(C) = -\sum_i a_{ij}(-1)^{i+1+j}\det(D_{ij}) = \sum_i a_{ij}(-1)^{i+j}\det(A_{ij}).$$

(2)-(4) These properties can be seen by direct calculations.

(5)-(7) These properties follow from (2)-(4).

(8) Similar to the proofs of (2)-(6).

(9)-(10) Use induction on the order of the matrix.

(11) Let A be invertible. Now, $1 = \det(I) = \det(AA^{-1}) = \det(A)\det(A^{-1})$ shows that $\det(A) \neq 0$ and $\det(A^{-1}) = (\det(A))^{-1}$.

(12) It follows from (10)-(11).

(13) If the columns of A are linearly independent, then A is invertible. By (11), $\det(A) \neq 0$. So, suppose the columns of A are linearly dependent. Say, the kth column of A is a linear combination of the columns 1 through $k-1$. Let B be the matrix identical to A except at the kth column; the kth column of B is the kth column of A minus that linear combination of the columns 1 through $k-1$ of A. Due to (5), $\det(A) = \det(B) = 0$, as the kth column of B has all entries 0.

(14) Consider the case $j = 1$. This property asserts that a determinant can be expanded in its first column. For $n = 1, 2$ it is easy to verify that expansion of a determinant can be made in the first column. Suppose that determinants of all matrices of order less than or equal to $m-1$ can be expanded in their first columns. Let $A \in \mathbb{F}^{m\times m}$. Now, expanding in the first row,

$$\det(A) = \sum_i a_{1j}(-1)^{1+j}\det(A_{1j}).$$

The minors $\det(A_{ij})$ can be expanded in their first column. That is,

$$\det(A_{1j}) = \sum_{i=2}^{m}(-1)^{i-1+1}a_{i1}B_{ij},$$

where B_{ij} denotes the determinant of the $(n-2)\times(n-2)$ matrix obtained from A by deleting the first and the ith rows, and deleting the first and the jth columns. Thus the only term in $\det(A)$ involving $a_{1j}ai1$ is $(-1)^{i+j+1}a_{1j}a_{i1}Bij$.

By the inductive assumption, $\det(A_{1j}) = \sum_{j=2}^{n}(-1)^{i+j+1}a_{1j}B_{ij}$. Therefore, the only term in $\det(A) = \sum_i a_{1j}(-1)^{1+j}\det(A_{1j})$ involving $a_{1j}a_{i1}$ is $(-1)^{i+j+1}a_{1j}a_{i1}Bij$. For $j \neq 1$, a proof similar to the proof of (1) above can be constructed.

(15) Consider $A\,\mathrm{adj}(A)$. Due to (1), the jth diagonal entry in this product is

$$\sum_{i=1}^{n} a_{ij}C_{ij}(A) = \det(A).$$

For the non-diagonal entries, let $i \neq j$. Construct a matrix B which is identical to A except at the jth row. The jth row of B is the same as the ith row of A. Now, look at the co-factors $C_{kj}(B)$. Such a co-factor is obtained by deleting the kth row and the jth column of B and then taking the determinant of the resulting $(n-1)\times(n-1)$ matrix. This is same as $C_{kj}(A)$. Hence, the (i,k)th entry in $A\,\mathrm{adj}(A)$ is

$$\sum_{k=1}^{n} a_{ik}C_{ik}(A) = \sum_{k=1}^{n} a_{ik}C_{ik}(B) = \det(B) = 0,$$

since the ith and the kth rows are equal in B. Therefore, $A\,\mathrm{adj}(A) = \det(A)\,I$.
The product formula $\mathrm{adj}(A)A = \det(A)\,I$ is proved similarly. ∎

From (6) it follows that if some column of A is the zero vector, then $\det(A) = 0$. Also, if some column of A is a scalar multiple of another column, then $\det(A) = 0$. Similar conditions on rows (instead of columns) imply that $\det(A) = 0$.

Using the above properties, the *computational complexity* for evaluating a determinant can be reduced drastically. The trick is to bring a matrix to a triangular form using the row operations mentioned in (3.6-3,4,5). Next, compute the product of the diagonal entries in the upper triangular form to get the determinant.

(3.7) *Example*

$$\begin{vmatrix} 1 & 0 & 0 & 1 \\ -1 & 1 & 0 & 1 \\ -1 & -1 & 1 & 1 \\ -1 & -1 & -1 & 1 \end{vmatrix} \overset{R1}{=} \begin{vmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & -1 & 1 & 2 \\ 0 & -1 & -1 & 2 \end{vmatrix} \overset{R2}{=} \begin{vmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & -1 & 4 \end{vmatrix} \overset{R3}{=} \begin{vmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 8 \end{vmatrix} = 8.$$

Here, $R1$ replaces the second row with second plus the first row, then replaces the third row with the third plus the first row, and the fourth row with the fourth plus the first row. $R2$ replaces the third and the fourth rows with the third plus the second, and the fourth plus the second, respectively. Finally, $R3$ replaces the fourth row with the fourth plus the third row.

Finally, the upper triangular matrix has the required determinant. □

(3.8) *Theorem*

Let $A \in \mathbb{C}^{n\times n}$.

(1) *If A is hermitian, then* $\det(A) \in \mathbb{R}$.

(2) *If A is unitary, then* $|\det(A)| = 1$.

(3) *If A is orthogonal, then* $\det(A) = \pm 1$.

Proof. (1) Let A be hermitian. Then $A = A^*$. It implies

$$\det(A) = \det(A^*) = \det(\overline{A}) = \overline{\det(A)}.$$

Hence, $\det(A) \in \mathbb{R}$.

(2) Let A be unitary. Then $A^*A = AA^* = I$. Now,

$$1 = \det(I) = \det(A^*A) = \det(\overline{A})\det(A) = \overline{\det(A)}\det(A) = |\det(A)|^2.$$

Hence $|\det(A)| = 1$.

(3) Let A be an orthogonal matrix. That is, $A \in \mathbb{R}^{n\times n}$ and A is unitary. Then $\det(A) \in \mathbb{R}$ and by (2), $|\det(A)| = 1$. That is, $\det(A) = \pm 1$. ∎

3.6 Exercises for Chapter 3

1. Show that given any $n \in \mathbb{N}$ there exist matrices $A, B \in \mathbb{R}^{n\times n}$ such that $AB \neq BA$.
2. Let $A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}$. Compute A^n. *Ans*: $\begin{bmatrix} 1 & n+1 & n(n+1) \\ 0 & 1 & 2(n+1) \\ 0 & 0 & 1 \end{bmatrix}$.
3. Let $A \in \mathbb{F}^{m\times n}$; $B \in \mathbb{F}^{n\times k}$; $A_1, \ldots, A_m$ be the rows of A; $B_1, \ldots, B_k$ be the columns of B. Show that
   (a) $A_1B, \ldots, A_mB$ are the rows of AB.
   (b) $AB_1, \ldots, AB_k$ are the columns of AB.
4. Let $A \in \mathbb{F}^{n\times n}$; I be the identity matrix of order n. Find the inverse of the $2n \times 2n$ matrix $\begin{bmatrix} I & A \\ 0 & I \end{bmatrix}$. *Ans*: $\begin{bmatrix} I & -A \\ 0 & I \end{bmatrix}$.
5. If A is a hermitian (symmetric) invertible matrix, then show that A^{-1} is hermitian (symmetric).
6. If A is a lower (upper) triangular invertible matrix, then A^{-1} is lower (upper) triangular.
7. Show that each orthogonal 2×2 matrix is either a reflection or a rotation.
8. Let $u, v, w \in \mathbb{F}^{n\times 1}$. Show that $\{u + v, v + w, w + u\}$ is linearly independent iff $\{u, v, w\}$ is linearly independent.
9. Find linearly independent vectors from $U = \{(a, b, c) : 2a + 3b - 4c = 0\}$ so that the set of all linear combinations of which is exactly U.
   Ans: $(2, 0, 1)$, $(0, 4, 3)$.

10. Determine linearly independent vectors so that the set of linear combinations of which is $U = \{(a, b, c, d, e) \in \mathbb{R}^5 : a = c = e, b + d = 0\}$.
Ans: $(1, 0, 1, 0, 1)$, $(0, 1, 0, -1, 0)$.

11. Calculate the determinants of the following matrices:
(a) $\begin{bmatrix} 2 & 3 & 4 \\ 0 & 1 & 2 \\ 2 & 0 & 1 \end{bmatrix}$ (b) $\begin{bmatrix} 1 & -1 & 2 & -2 \\ 2 & -2 & 1 & -1 \\ -2 & 1 & -1 & 2 \\ -1 & 2 & -2 & 1 \end{bmatrix}$ (c) $\begin{bmatrix} 2 & 3 & 1 & 2 \\ 0 & 2 & 1 & 3 \\ 0 & 2 & 1 & -2 \\ -2 & 1 & -1 & 1 \end{bmatrix}$.
Ans: (a) 6 (b) 0 (c) 40.

12. Let $a_1, \ldots, a_n \in \mathbb{R}$. Let $A \in \mathbb{R}^{n \times n}$ have the (i, j)th entry as a_j^{i-1}. Show that $\det(A) = \Pi_{i<j}(a_i - a_j)$. [$A$ is called a Vandermonde's matrix.]

4

Row Reduced Echelon Form

4.1 Elementary row operations

While solving a system of linear equations, we add and subtract equations, multiply an equation with a nonzero constant, and exchange two equations. These heuristics give rise to the row operations on a matrix.

There are three kinds of **Elementary Row Operations** for a matrix $A \in \mathbb{F}^{m\times n}$:

Exchange of two rows.

Multiplication of a row by a nonzero constant.

Adding to a row a nonzero multiple of another row.

(4.1) ***Example***

See the following computation on the first matrix. We get the second matrix from the first by adding to the third row (-3) times the first row. In symbols, we write this operation as $R_3 \leftarrow R_3 - 3R_1$. Similarly, the third matrix is obtained from the second by adding to the second row (-2) the first row. We write this operation as $R_2 \leftarrow R_2 - 2R_1$.

$$\begin{bmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \\ 3 & 3 & 3 \end{bmatrix} \xrightarrow{R_3 \leftarrow R_3 - 3R_1} \begin{bmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \\ 0 & 0 & 0 \end{bmatrix} \xrightarrow{R_2 \leftarrow R_2 - 2R_1} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

□

We will write $A \xrightarrow{O} B$ to mean that the matrix B is obtained from A by using the row operation O. Specifically, we will use the following notation for the elementary row operations (We assume that α is a nonzero scalar.):

$R_i \leftrightarrow R_j$: The ith row and the jth row are exchanged.

$R_i \leftarrow \alpha R_i$: The ith row is multiplied by a scalar α.

$R_i \leftarrow R_i + \alpha R_j$: To the ith row α times the jth row is added.

The elementary row operations can be captured by matrix multiplication. For this purpose, we define the three types of **elementary matrices** $E[i,j]$, $E_\alpha[i]$ and $E_\alpha[i,j]$ as in the following:

> $E[i,j]$ is obtained from I by exchanging the ith and jth rows.
>
> $E_\alpha[i]$ is obtained from I by multiplying α to the ith row.
>
> $E_\alpha[i,j]$ is obtained from I by adding to its ith row α times the jth row.

The context will specify the order of these elementary matrices.

The following are some elementary matrices of order 3:

$$E[1,2] = \begin{bmatrix} 0&1&0\\1&0&0\\0&0&1 \end{bmatrix}, \quad E_{-3}[2] = \begin{bmatrix} 1&0&0\\0&-3&0\\0&0&1 \end{bmatrix}, \quad E_{-3}[3,1] = \begin{bmatrix} 1&0&0\\0&1&0\\-3&0&1 \end{bmatrix}.$$

To use these elementary matrices, consider a 3×3 matrix, say, $A = \begin{bmatrix} 1&1&1\\2&2&2\\3&3&3 \end{bmatrix}$. Now,

$$E[1,2]A = \begin{bmatrix} 0&1&0\\1&0&0\\0&0&1 \end{bmatrix}\begin{bmatrix} 1&1&1\\2&2&2\\3&3&3 \end{bmatrix} = \begin{bmatrix} 2&2&2\\1&1&1\\3&3&3 \end{bmatrix}, \quad A \xrightarrow{R_1\leftrightarrow R_2} \begin{bmatrix} 2&2&2\\1&1&1\\3&3&3 \end{bmatrix}.$$

$$E_{-3}[2]A = \begin{bmatrix} 1&0&0\\0&-3&0\\0&0&1 \end{bmatrix}\begin{bmatrix} 1&1&1\\2&2&2\\3&3&3 \end{bmatrix} = \begin{bmatrix} 1&1&1\\-6&-6&-6\\3&3&3 \end{bmatrix}, \ A \xrightarrow{R_2\leftarrow -3R_2} \begin{bmatrix} 1&1&1\\-6&-6&-6\\3&3&3 \end{bmatrix}.$$

$$E_{-3}[3,1]A = \begin{bmatrix} 1&0&0\\0&1&0\\-3&0&1 \end{bmatrix}\begin{bmatrix} 1&1&1\\2&2&2\\3&3&3 \end{bmatrix} = \begin{bmatrix} 1&1&1\\2&2&2\\0&0&0 \end{bmatrix}, \quad A \xrightarrow{R_3\leftarrow R_3-3R_1} \begin{bmatrix} 1&1&1\\2&2&2\\0&0&0 \end{bmatrix}.$$

We find that the way $E[1,2]$, $E_{-3}[2]$ and $E_{-3}[3,1]$ have been obtained from I, in the same way $E[1,2]A$, $E_{-3}[2]A$ and $E_{-3}[3,1]A$ have been obtained from A. This is true in general.

Let $A \in \mathbb{F}^{m\times n}$. Consider $E[i,j]$, $E_\alpha[i]$, $E_\alpha[i,j] \in \mathbb{F}^{m\times m}$ for $\alpha \neq 0$. The following may be verified:

1. $A \xrightarrow{R_i\leftrightarrow R_j} E[i,j]A$.

   That is, $E[i,j]A$ is the matrix obtained from A by exchanging the ith and the jth rows.

2. $A \xrightarrow{R_i\leftarrow \alpha R_i} E_\alpha[i]A$.

   That is, $E_\alpha[i]A$ is the matrix obtained from A by multiplying α to the ith row.

3. $A \xrightarrow{R_i \leftarrow R_i + \alpha R_j} E_\alpha[i,j]A.$

   That is, $E_\alpha[i,j]A$ is the matrix obtained from A by adding to the ith α times the jth row.

For example, the computation $\begin{bmatrix}1&1&1\\2&2&2\\3&3&3\end{bmatrix} \xrightarrow{R_3 \leftarrow R_3 - 3R_1} \begin{bmatrix}1&1&1\\2&2&2\\0&0&0\end{bmatrix} \xrightarrow{R_2 \leftarrow R_2 - 2R_1} \begin{bmatrix}1&1&1\\0&0&0\\0&0&0\end{bmatrix}$

shows that $\begin{bmatrix}1&1&1\\2&2&2\\0&0&0\end{bmatrix} = E_{-3}[3,1]\begin{bmatrix}1&1&1\\2&2&2\\3&3&3\end{bmatrix}$, and $\begin{bmatrix}1&1&1\\0&0&0\\0&0&0\end{bmatrix} = E_{-2}[2,1]\begin{bmatrix}1&1&1\\2&2&2\\0&0&0\end{bmatrix}$.

Often we will apply elementary operations in a sequence. In this way, the above operations could be shown as

$$\begin{bmatrix}1&1&1\\2&2&2\\3&3&3\end{bmatrix} \xrightarrow{O} \begin{bmatrix}1&1&1\\0&0&0\\0&0&0\end{bmatrix}, \quad \text{where } O = R_3 \leftarrow R_3 - 3R_1,\ R_2 \leftarrow R_2 - 2R_1.$$

In this case, $\begin{bmatrix}1&1&1\\0&0&0\\0&0&0\end{bmatrix} = E_{-2}[2,1]\,E_{-3}[3,1]\begin{bmatrix}1&1&1\\2&2&2\\3&3&3\end{bmatrix}$.

Notice that the elementary matrices are multiplied in the reverse order corresponding to the elementary row operations.

4.2 Row reduced echelon form

Elementary row operations can be used to convert a matrix to a nice form, which we discuss next.

The first nonzero entry (from left) in a nonzero row of a matrix is called a **pivot**. We denote a pivot in a row by putting a box around it. A column where a pivot occurs is called a **pivotal column**.

A matrix $A \in \mathbb{F}^{m\times n}$ is said to be in **row reduced echelon form (RREF)** iff the following conditions are satisfied:

1. Each pivot is equal to 1.
2. The column index of the pivot in any nonzero row R is smaller than the column index of the pivot in any row below R.
3. In a pivotal column, all entries other than the pivot, are zero.

4. All zero rows are at the bottom.

(4.2) *Example*

The matrix $\begin{bmatrix} \boxed{1} & 2 & 0 & 0 \\ 0 & 0 & \boxed{1} & 0 \\ 0 & 0 & 0 & \boxed{1} \end{bmatrix}$ is in row reduced echelon form whereas the matrices

$$\begin{bmatrix} 0 & \boxed{1} & 3 & 0 \\ 0 & 0 & 0 & \boxed{2} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & \boxed{1} & 3 & 1 \\ 0 & 0 & 0 & \boxed{1} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & \boxed{1} & 3 & 0 \\ 0 & 0 & 0 & \boxed{1} \\ 0 & 0 & 0 & \boxed{1} \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & \boxed{1} & 0 & 0 \\ 0 & 0 & \boxed{1} & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \boxed{1} \end{bmatrix}$$

are not in row reduced echelon form. ▯

Any matrix can be brought to a row reduced echelon form by using elementary row operations. We give an algorithm to achieve this.

Reduction to RREF

1. Set the work region R as the whole matrix A.
2. If all entries in R are 0, then stop.
3. If there are nonzero entries in R, then find the leftmost nonzero column. Mark it as the pivotal column.
4. Find the topmost nonzero entry in the pivotal column. Suppose it is α. Box it; it is a pivot.
5. If the pivot is not on the top row of R, then exchange the row of A which contains the top row of R with the row where the pivot is.
6. If $\alpha \neq 1$, then replace the top row of R in A by $1/\alpha$ times that row.
7. Make all entries, except the pivot, in the pivotal column as zero by replacing each row above and below the top row of R using elementary row operations in A with that row and the top row of R.
8. Find the sub-matrix to the right and below the pivot. If no such sub-matrix exists, then stop. Else, reset the work region R to this sub-matrix, and go to Step 2.

We will refer to the output of the above reduction algorithm as *the row reduced echelon form* or, **the RREF** of a given matrix.

(4.3) ***Example***

$$A = \begin{bmatrix} \boxed{1} & 1 & 2 & 0 \\ 3 & 5 & 7 & 1 \\ 1 & 5 & 4 & 5 \\ 2 & 8 & 7 & 9 \end{bmatrix} \xrightarrow{O1} \begin{bmatrix} \boxed{1} & 1 & 2 & 0 \\ 0 & 2 & 1 & 1 \\ 0 & 4 & 2 & 5 \\ 0 & 6 & 3 & 9 \end{bmatrix} \xrightarrow{R_2 \leftarrow 1/2 R_2} \begin{bmatrix} \boxed{1} & 1 & 2 & 0 \\ 0 & \boxed{1} & \frac{1}{2} & \frac{1}{2} \\ 0 & 4 & 2 & 5 \\ 0 & 6 & 3 & 9 \end{bmatrix}$$

$$\xrightarrow{O2} \begin{bmatrix} \boxed{1} & 0 & \frac{3}{2} & -\frac{1}{2} \\ 0 & \boxed{1} & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 6 \end{bmatrix} \xrightarrow{R_3 \leftarrow 1/3 R_3} \begin{bmatrix} \boxed{1} & 0 & \frac{3}{2} & -\frac{1}{2} \\ 0 & \boxed{1} & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 0 & \boxed{1} \\ 0 & 0 & 0 & 6 \end{bmatrix} \xrightarrow{O3} \begin{bmatrix} \boxed{1} & 0 & \frac{3}{2} & 0 \\ 0 & \boxed{1} & \frac{1}{2} & 0 \\ 0 & 0 & 0 & \boxed{1} \\ 0 & 0 & 0 & 0 \end{bmatrix} = B$$

Here, $O1 = R_2 \leftarrow R_2 - 3R_1,\ R_3 \leftarrow R_3 - R_1,\ R_4 \leftarrow R_4 - 2R_1$;
$O2 = R_2 \leftarrow R_2 - R_1,\ R_3 \leftarrow R_3 - 4R_2,\ R_4 \leftarrow R_4 - 6R_2$; and
$O3 = R_1 \leftarrow R_1 + 1/2R_3,\ R_2 \leftarrow R_2 - 1/2R_3,\ R_4 \leftarrow R_4 - 6R_3$.
Notice that

$$B = E_{-6}[4,3]\, E_{-1/2}[2,3]\, E_{1/2}[1,3]\, E_{1/3}[3]\, E_{-6}[4,2]\, E_{-4}[3,2]\, E_{-1}[2,1] E_{1/2}[2]$$
$$E_{-2}[4,1]\, E_{-1}[3,1]\, E_{-3}[2,1]\, A.$$

The products are in reverse order. ∎

(4.4) ***Example***

Consider $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$ Then

$$\text{RREF}(A)\,\text{RREF}(B) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \neq \text{RREF}(AB) = AB = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Thus, RREF of a product need not be equal to the product of RREFs. ∎

Observe that if a square matrix A is invertible, then $AA^{-1} = I$ implies that A does not have a zero row.

(4.5) ***Theorem***

A square matrix is invertible iff it is a product of elementary matrices.

Proof. $E[i, j]$ is its own inverse, $E_{1/\alpha}[i]$ is the inverse of $E_\alpha[i]$, and $E_{-\alpha}[i, j]$ is the inverse of $E_\alpha[i, j]$. So, product of elementary matrices is invertible.

Conversely, suppose that A is invertible. Let EA^{-1} be the RREF of A^{-1}. If EA^{-1} has a zero row, then $EA^{-1}A$ also has a zero row. That is, E has a zero row, which is impossible since E is invertible. So, EA^{-1} does not have a zero row. Then each row in the square matrix EA^{-1} has a pivot. But the only square matrix in RREF having a pivot at each row is the identity matrix. Therefore, $EA^{-1} = I$. That is, $A = E$, a product of elementary matrices. ∎

Using the above theorem, we can show that the row reduced form of a matrix does not depend on the algorithm we use.

(4.6) *Theorem*

Let $A \in \mathbb{F}^{m\times n}$. There exists a unique matrix in $\mathbb{F}^{m\times n}$ in row reduced echelon form obtained from A by elementary row operations.

Proof. Suppose $B, C \in \mathbb{F}^{m\times n}$ are matrices in RREF such that each has been obtained from A by elementary row operations. Recall that elementary matrices are invertible and their inverses are also elementary matrices. Then $B = E_1 A$ and $C = E_2 A$ for some invertible matrices $E_1, E_2 \in \mathbb{F}^{m\times m}$. Now, $B = E_1 A = E_1(E_2)^{-1}C$. Write $E = E_1(E_2)^{-1}$ to have $B = EC$, where E is invertible.

We consider a particular case first, when $n = 1$. Here, B and C are column vectors in RREF. Thus, they can be zero vectors or e_1. Since $B = EC$, where E is invertible, it cannot happen that one is the zero vector and the other is e_1. Thus, either both are zero vectors or both are e_1. In either case, $B = C$.

For $n > 1$, assume, on the contrary, that $B \neq C$. Then there exists a column index, say $k \geq 1$, such that the first $k-1$ columns of B coincide with the first $k-1$ columns of C, respectively; and the kth column of B is not equal to the kth column of C. Let u be the kth column of B, and let v be the kth column of C. We have $u = Ev$ and $u \neq v$.

Suppose the pivotal columns that appear within the first $k-1$ columns in C are $e_1, \ldots, e_j$. Then $e_1, \ldots, e_j$ are also the pivotal columns in B that appear within the first $k-1$ columns. Since $B = EC$, we have $C = E^{-1}B$; and consequently,

$$e_1 = Ee_1 = E^{-1}e_1, \ \ldots, \ e_j = Ee_j = E^{-1}e_j.$$

The column vector u may be a pivotal column in B or a non-pivotal column in B. Similarly, v may be pivotal or non-pivotal in C. If both u and v are pivotal columns, then both are equal to e_{j+1}. This contradicts $u \neq v$. So, assume that u is non-pivotal in B or v is non-pivotal in C.

If u is non-pivotal in B, then $u = \alpha_1 e_1 + \cdots + \alpha_j e_j$ for some scalars $\alpha_1, \ldots, \alpha_j$. (See it.) Then

$$v = E^{-1}u = \alpha_1 E^{-1}e_1 + \cdots + \alpha_j E^{-1}e_j = \alpha_1 e_1 + \cdots + \alpha_j e_j = u.$$

This contradicts $u \neq v$.

If v is a non-pivotal column in C, then $v = \beta_1 e_1 + \cdots + \beta_j e_j$ for some scalars $\beta_1, \ldots, \beta_j$. Then

$$u = Ev = \beta_1 E e_1 + \cdots + \beta_j E e_j = \beta_1 e_1 + \cdots + \beta_j e_j = v.$$

Here also, $u = v$, which is a contradiction.

Therefore, $B = C$. ∎

The number of pivots in the RREF of a matrix A is called the **rank** of A, and it is denoted by $\text{rank}(A)$. Since RREF of a matrix is unique, rank is well-defined.

For instance, in (4.3), $\text{rank}(A) = \text{rank}(B) = 3$.

Suppose B is a matrix in RREF. If B is invertible, then its RREF does not have a zero row. So, the RREF is equal to I. But B is already in RREF. So, $B = I$. Conversely, if $B = I$, then it is in invertible, and also it is in RREF. Therefore, a matrix in RREF is invertible iff it is equal to I.

(4.7) ***Theorem***

A square matrix is invertible iff its rank is equal to its order.

Proof. Let A be a square matrix of order n. Let B be the RREF of A. Then $B = EA$, where E is invertible.

Let A be invertible. Then B is invertible. Since B is in RREF, $B = I$. So, $\text{rank}(A) = n$.

Conversely, suppose $\text{rank}(A) = n$. Then B has n number of pivots. Thus $B = I$. In that case, $A = E^{-1}B = E^{-1}$; and A is invertible. ∎

4.3 Determining rank & linear independence

Let A be an $m \times n$ matrix. Recall that $\text{rank}(A)$ is the number of pivots in the RREF of A. If $\text{rank}(A) = r$, then there are r number of linearly independent columns in A and other columns are linear combinations of these r columns. The linearly independent columns correspond to the pivotal columns in the RREF of A. Also, there exist r number of linearly independent rows of A such that other rows are linear combinations of these r rows. The linearly independent rows correspond to the nonzero rows in the RREF of A.

(4.8) ***Example***

Consider $A = \begin{bmatrix} 1 & 1 & 1 & 2 & 1 \\ 1 & 2 & 1 & 1 & 1 \\ 3 & 5 & 3 & 4 & 3 \\ -1 & 0 & -1 & -3 & -1 \end{bmatrix}$.

Here, $row(3) = row(1) + 2row(2)$ and $row(4) = row(2) - 2row(1)$. And, $row(2)$ is not a scalar multiple of $row(1)$, that is, $row(1)$, $row(2)$ are linearly independent; and all other rows are linear combinations of $row(1)$ and $row(2)$.

Also, $col(3) = col(5) = col(1)$, $col(4) = 3\,col(1) - col(2)$, and $col(1)$, $col(2)$ are linearly independent.

Verify that the RREF of A is given by $\begin{bmatrix} 1 & 0 & 1 & 3 & 1 \\ 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$.

As expected, $\text{rank}(A) = 2$. ∎

The connection between linear combinations, linear dependence and linear independence of rows and columns of a matrix, and its RREF may be stated as follows.

(4.9) Observation In the RREF of A suppose $R_{i1}, \ldots, R_{ir}$ are the rows of A which have become the nonzero rows in the RREF, and other rows have become the zero rows. Also, suppose $C_{j1}, \ldots, C_{jr}$ for $j1 < \cdots < jr$, are the columns of A which have become the pivotal columns in the RREF, other columns being non-pivotal. Then the following are true:

1. All rows of A other than $R_{i1}, \ldots, R_{ir}$ are linear combinations of $R_{i1}, \ldots, R_{ir}$.
2. The columns $C_{j1}, \ldots, C_{jr}$ have respectively become $e_1, \ldots, e_r$ in the RREF.
3. All columns of A other than $C_{j1}, \ldots, C_{jr}$ are linear combinations of $C_{j1}, \ldots, C_{jr}$.
4. If $e_1, \ldots, e_k$ are all the pivotal columns in the RREF that occur to the left of a non-pivotal column, then the non-pivotal column is in the form $(a_1, \ldots, a_k, 0, \ldots, 0)^T$. Further, if a column C in A has become this non-pivotal column in the RREF, then $C = a_1 C_{j1} + \cdots + a_k C_{jk}$.
5. If A is a square matrix, then A is invertible iff its RREF is I.

As the above observation shows, elementary operations can be used to determine linear dependence or independence of a finite set of vectors in $\mathbb{F}^{1\times n}$. Suppose that you are given with m number of vectors from $\mathbb{F}^{1\times n}$, say,

$$u_1 = (u_{11}, \ldots, u_{1n}), \ldots, u_m = (u_{m1}, \ldots, u_{mn}).$$

We form the matrix A with rows as $u_1, \ldots, u_m$. We then reduce A to its RREF, say, B. If there are r number of nonzero rows in B, then the rows corresponding to those rows in A are linearly independent, and the other rows (which have become the zero rows in B) are linear combinations of those r rows.

(4.10) *Example*

From among the vectors $(1, 2, 2, 1)$, $(2, 1, 0, -1)$, $(4, 5, 4, 1)$, $(5, 4, 2, -1)$, find linearly independent vectors; and point out which are the linear combinations of these independent ones. □

We form a matrix with the given vectors as rows and then bring the matrix to its RREF.

$$\begin{bmatrix} \boxed{1} & 2 & 2 & 1 \\ 2 & 1 & 0 & -1 \\ 4 & 5 & 4 & 1 \\ 5 & 4 & 2 & -1 \end{bmatrix} \xrightarrow{O1} \begin{bmatrix} \boxed{1} & 2 & 2 & 1 \\ 0 & \boxed{-3} & -4 & -3 \\ 0 & -3 & -4 & -3 \\ 0 & -6 & -8 & -6 \end{bmatrix} \xrightarrow{O2} \begin{bmatrix} \boxed{1} & 0 & -2/3 & -1 \\ 0 & \boxed{1} & 4/3 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Here, $O1 = R_2 \leftarrow R_2 - 2R_1,\ R_3 \leftarrow R_3 - 4R_1,\ R_4 \leftarrow R_4 - 5R_1$ and
$O2 = R_2 \leftarrow -3R_2,\ R_1 \leftarrow R_1 - 2R_2,\ R_3 \leftarrow R_3 + 3R_2,\ R_4 \leftarrow R_4 + 6R_2.$

No row exchanges have been applied in this reduction, and the nonzero rows are the first and the second rows. Therefore, the linearly independent vectors are $(1, 2, 2, 1)$, $(0, 1, 4/3, 1)$; and the third and the fourth are linear combinations of these.

The same method can be used in $\mathbb{F}^{n\times 1}$. Just use the transpose of the columns, form a matrix, and continue with row reductions. Finally, take the transposes of the nonzero rows in the RREF.

Notice that compared to Gram-Schmidt, reduction to RREF is more efficient way of extracting a linearly independent set retaining the span.

(4.11) *Example*

Suppose A and B are $n \times n$ matrices satisfying $AB = I$. Show that both A and B are invertible, and $BA = I$.

Let EA be the RREF of A, where E is a suitable product of elementary matrices. If A is not invertible, then EA has a zero row, so that EAB also has a zero row. Since

$EAB = E$ does not have a zero row, we conclude that A is invertible. Consequently, $B = A^{-1}$ is also invertible, and $BA = I$. ∎

However, if A and B are not square matrices, then $AB = I$ need not imply that $BA = I$. For instance,

$$\begin{bmatrix} 2 & 0 & -1 \\ 1 & 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 2 & 0 & -1 \\ 1 & 1 & -1 \end{bmatrix} = \begin{bmatrix} 3 & 1 & -2 \\ 1 & 1 & -1 \\ 4 & 2 & -3 \end{bmatrix}.$$

You have seen earlier that there do not exist three linearly independent vectors in $\mathbb{R}^2$. With the help of rank, now we can see why does it happen.

(4.12) *Theorem*

Let $u_1, \ldots, u_k,\ v_1, \ldots, v_m \in \mathbb{F}^n$. Suppose $v_1, \ldots, v_m \in \operatorname{span}(u_1, \ldots, u_k)$ and $m > k$. Then $v_1, \ldots, v_m$ are linearly dependent.

Proof. Consider all vectors as row vectors. Form the matrix A by taking its rows as $u_1, \ldots, u_k, v_1, \ldots, v_m$ in that order. Now, $r = \operatorname{rank}(A) \le k$. Similarly, construct the matrix B by taking its rows as $v_1, \ldots, v_m, u_1, \ldots, u_k$, in that order. Now, A and B have the same RREF since one is obtained from the other by re-ordering the rows. Therefore, $\operatorname{rank}(B) = \operatorname{rank}(A) = r \le k$. Since $m > k \ge r$, out of $v_1, \ldots, v_m$ at most r vectors can be linearly independent.So, $v_1, \ldots, v_m$ are linearly dependent. ∎

The following theorem is a corollary to the above.

(4.13) *Theorem*

Let $v_1, \ldots, v_n \in \mathbb{F}^m$. Then there exists a unique $r \le n$ such that some r of these vectors are linearly independent and other $n - r$ vectors are linear combinations of these r vectors.

To see further connection between these notions, let $u_1, \ldots, u_r, u \in \mathbb{F}^{m\times 1}$. Let $a_1, \ldots, a_r \in \mathbb{F}$ and let $P \in \mathbb{F}^{m\times m}$ be invertible. We see that

$$u = a_1 u_1 + \cdots + a_r u_r \quad \text{iff} \quad Pu = a_1 P u_1 + \cdots a_r P u_r.$$

Taking $u = 0$, we see that the vectors $u_1, \ldots, u_r$ are linearly independent iff $Pu_1, \ldots, Pu_r$ are linearly independent.

Now, if $A \in \mathbb{F}^{m\times n}$, then its columns are vectors in $\mathbb{F}^{m\times 1}$. The above equation implies that if there exist r number of columns in A which are linearly independent and other columns are linear combinations of these r columns, then the same is true for the matrix PA.

Similarly, let $Q \in \mathbb{F}^{n\times n}$ be invertible. If there exist r number of rows of A which are linearly independent and other rows are linear combinations of these r rows, then the same is true for the matrix AQ.

These facts along with the last theorem can be used to prove the following theorem.

(4.14) ***Theorem***

Let $A \in \mathbb{F}^{m\times n}$. Then

$$\begin{aligned}\operatorname{rank}(A) &= \textit{the maximum number of linearly independent rows in } A\\ &= \textit{the maximum number of linearly independent columns in } A\\ &= \operatorname{rank}(A^T)\\ &= \operatorname{rank}(PAQ) \ \textit{ for invertible matrices } P \in \mathbb{F}^{m\times m} \textit{ and } Q \in \mathbb{F}^{n\times n}.\end{aligned}$$

4.4 Computing inverse of a matrix

Let $A \in \mathbb{F}^{n\times n}$. If A is invertible, then using Property (15) of the determinant, its inverse can be computed. However, computation of determinant is easier when elementary row operations are used. This suggests that we use elementary row operations directly for computing the inverse of a given matrix.

Observe that when A is an invertible matrix of order n, its RREF has exactly n pivots. The entries in each pivotal column above and below the pivot are 0. The pivots are each equal to 1. Therefore, such a row reduced echelon matrix is nothing but I, the identity matrix of order n.

Now, look at the sequence of elementary matrices corresponding to the elementary operations used in this row reduction of A. The product of these elementary matrices is A^{-1}, since this product times A is I, which is the row reduced form of A. Now, if we use the same elementary operations on I, then the result will be $A^{-1}I = A^{-1}$. Thus we obtain a procedure to compute the inverse of a matrix A provided it is invertible.

The work will be easier if we write the matrix A and the identity matrix I side by side and apply the elementary operations on both of them simultaneously. For this purpose, we introduce the notion of an augmented matrix.

If $A \in \mathbb{F}^{m\times n}$ and $B \in \mathbb{F}^{m\times k}$, then the matrix $[A|B] \in \mathbb{F}^{m\times(n+k)}$ obtained from A and B by writing first all the columns of A and then the columns of B, in that order, is called an **augmented matrix**.

The vertical bar shows the separation of columns of A and of B, though, conceptually unnecessary.

For computing the inverse of a matrix, start with the augmented matrix $[A|I]$. Then we reduce $[A|I]$ to its RREF. If the A-portion in the RREF is I, then the I-portion in the RREF gives A^{-1}. If the A-portion in the RREF contains a zero row, then A is not invertible. Notice that if a zero row has appeared during the RREF conversion, then we need not proceed towards the RREF; the matrix A is not invertible.

(4.15) *Example*

For illustration, consider the following square matrices:

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

We want to find the inverses of the matrices, if at all they are invertible.

Augment A with an identity matrix to get

$$\left[\begin{array}{cccc|cccc} \boxed{1} & -1 & 2 & 0 & 1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 2 & 0 & 1 & 0 & 0 \\ 2 & 1 & -1 & -2 & 0 & 0 & 1 & 0 \\ 1 & -2 & 4 & 2 & 0 & 0 & 0 & 1 \end{array}\right].$$

Use elementary row operations. Since $a_{11} = 1$, we leave $row(1)$ untouched. To zero-out the other entries in the first column, we use the sequence of elementary row operations $R_2 \leftarrow R_2 + R_1$, $R_3 \leftarrow R_3 - 2R_1$, $R_4 \leftarrow R_4 - R_1$ to obtain

$$\left[\begin{array}{cccc|cccc} \boxed{1} & -1 & 2 & 0 & 1 & 0 & 0 & 0 \\ 0 & -1 & 2 & 2 & 1 & 1 & 0 & 0 \\ 0 & 3 & -5 & -2 & -2 & 0 & 1 & 0 \\ 0 & -1 & 2 & 2 & -1 & 0 & 0 & 1 \end{array}\right].$$

The pivot is -1 in $(2,2)$ position. Use $R_2 \leftarrow -R_2$ to make the pivot 1.

$$\left[\begin{array}{cccc|cccc} 1 & -1 & 2 & 0 & 1 & 0 & 0 & 0 \\ 0 & \boxed{1} & -2 & -2 & -1 & -1 & 0 & 0 \\ 0 & 3 & -5 & -2 & -2 & 0 & 1 & 0 \\ 0 & -1 & 2 & 2 & -1 & 0 & 0 & 1 \end{array}\right].$$

Use $R_1 \leftarrow R_1 + R_2$, $R_3 \leftarrow R_3 - 3R_2$, $R_4 \leftarrow R_4 + R_2$ to zero-out all non-pivot entries in the pivotal column to 0:

$$\left[\begin{array}{cccc|cccc} \boxed{1} & 0 & 0 & -2 & 0 & -1 & 0 & 0 \\ 0 & \boxed{1} & -2 & -2 & -1 & -1 & 0 & 0 \\ 0 & 0 & \boxed{1} & 4 & 1 & 3 & 1 & 0 \\ 0 & 0 & 0 & 0 & -2 & -1 & 0 & 1 \end{array}\right].$$

Since a zero row has appeared in the A portion, A is not invertible. And $\operatorname{rank}(A) = 3$, which is less than the order of A. The second portion of the augmented matrix has no meaning now. However, it records the elementary row operations which were carried out in the reduction process. Verify that this matrix is equal to

$$E_1[4,2]\,E_{-3}[3,2]\,E_1[1,2]\,E_{-1}[2]\,E_{-1}[4,1]\,E_{-2}[3,1]\,E_1[2,1]$$

and that the first portion is equal to this matrix times A.

For B, we proceed similarly. The augmented matrix $[B|I]$ with the first pivot looks like:

$$\left[\begin{array}{cccc|cccc} \boxed{1} & -1 & 2 & 0 & 1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 2 & 0 & 1 & 0 & 0 \\ 2 & 1 & -1 & -2 & 0 & 0 & 1 & 0 \\ 0 & -2 & 0 & 2 & 0 & 0 & 0 & 1 \end{array}\right].$$

The sequence of elementary row operations $R_2 \leftarrow R_2 + R_1$, $R_3 \leftarrow R_3 - 2R_1$ yields

$$\left[\begin{array}{cccc|cccc} \boxed{1} & -1 & 2 & 0 & 1 & 0 & 0 & 0 \\ 0 & -1 & 2 & 2 & 1 & 1 & 0 & 0 \\ 0 & 3 & -5 & -2 & -2 & 0 & 1 & 0 \\ 0 & -2 & 0 & 2 & 0 & 0 & 0 & 1 \end{array}\right].$$

Next, the pivot is -1 in $(2,2)$ position. Use $R_2 \leftarrow -R_2$ to get the pivot as 1.

$$\left[\begin{array}{cccc|cccc} \boxed{1} & -1 & 2 & 0 & 1 & 0 & 0 & 0 \\ 0 & \boxed{1} & -2 & -2 & -1 & -1 & 0 & 0 \\ 0 & 3 & -5 & -2 & -2 & 0 & 1 & 0 \\ 0 & -2 & 0 & 2 & 0 & 0 & 0 & 1 \end{array}\right].$$

And then $R_1 \leftarrow R_1 + R_2$, $R_3 \leftarrow R_3 - 3R_2$, $R_4 \leftarrow R_4 + 2R_2$ gives

$$\left[\begin{array}{cccc|cccc} \boxed{1} & 0 & 0 & -2 & 0 & -1 & 0 & 0 \\ 0 & \boxed{1} & -2 & -2 & -1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 4 & 1 & 3 & 1 & 0 \\ 0 & 0 & -4 & -2 & -2 & -2 & 0 & 1 \end{array}\right].$$

Next pivot is 1 in $(3,3)$ position. Now, $R_2 \leftarrow R_2 + 2R_3$, $R_4 \leftarrow R_4 + 4R_3$ produces

$$\left[\begin{array}{cccc|cccc} \boxed{1} & 0 & 0 & -2 & 0 & -1 & 0 & 0 \\ 0 & \boxed{1} & 0 & 6 & 1 & 5 & 2 & 0 \\ 0 & 0 & \boxed{1} & 4 & 1 & 3 & 1 & 0 \\ 0 & 0 & 0 & 14 & 2 & 10 & 4 & 1 \end{array}\right].$$

Next pivot is 14 in $(4,4)$ position. Use $R_4 \leftarrow {}^1/_4 R_4$ to get the pivot as 1:

$$\left[\begin{array}{cccc|cccc} \boxed{1} & 0 & 0 & -2 & 0 & -1 & 0 & 0 \\ 0 & \boxed{1} & 0 & 6 & 1 & 5 & 2 & 0 \\ 0 & 0 & \boxed{1} & 4 & 1 & 3 & 1 & 0 \\ 0 & 0 & 0 & \boxed{1} & 1/7 & 5/7 & 2/7 & 1/14 \end{array}\right].$$

Use $R_1 \leftarrow R_1 + 2R_4$, $R_2 \leftarrow R_2 - 6R_4$, $R_3 \leftarrow R_3 - 4R_4$ to zero-out the entries in the pivotal column:

$$\left[\begin{array}{cccc|cccc} \boxed{1} & 0 & 0 & 0 & 2/7 & 3/7 & 4/7 & 1/7 \\ 0 & \boxed{1} & 0 & 0 & 1/7 & 5/7 & 2/7 & -3/7 \\ 0 & 0 & \boxed{1} & 0 & 3/7 & 1/7 & -1/7 & -2/7 \\ 0 & 0 & 0 & \boxed{1} & 1/7 & 5/7 & 2/7 & 1/14 \end{array}\right].$$

Thus $B^{-1} = \frac{1}{7}\begin{bmatrix} 2 & 3 & 4 & 1 \\ 1 & 5 & 2 & -3 \\ 3 & 1 & -1 & -2 \\ 1 & 5 & 2 & \frac{1}{2} \end{bmatrix}$. Verify that $B^{-1}B = BB^{-1} = I$. ∎

4.5 Linear equations

We can now use our knowledge about matrices to settle some issues regarding solvability of linear equations. A system of linear equations, also called a **linear system** with m equations in n unknowns looks like:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots a_{mn}x_n &= b_m \end{aligned}$$

Using the abbreviations

$$x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \quad b = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}, \quad A = [a_{ij}],$$

the system can be written in the compact form:

$$Ax = b.$$

Here, $A \in \mathbb{F}^{m\times n}$, $x \in \mathbb{F}^{n\times 1}$ and $b \in \mathbb{F}^{m\times 1}$ so that m is the *number of equations* and n is the *number of unknowns* in the system. Notice that for linear systems, we deviate from our symbolism and write b as a column vector and x_i are unknown scalars. The system $Ax = b$ is **solvable**, also said *to have a solution*, iff there exists a vector $u \in \mathbb{F}^{n\times 1}$ such that $Au = b$.

Thus, the system $Ax = b$ is solvable iff b is a linear combination of columns of A. Also, $Ax = b$ has a unique solution iff b is a linear combination of columns of A and the columns of A are linearly independent. These issues are better tackled with the help of the corresponding **homogeneous system**

$$Ax = 0.$$

The homogeneous system always has a solution, namely, $x = 0$. It has infinitely many solutions iff it has a nonzero solution. For, if u is a solution, so is αu for any scalar α.

To study the non-homogeneous system, we use the augmented matrix $[A|b] \in \mathbb{F}^{m\times(n+1)}$ which has its first n columns as those of A in the same order, and the $(n+1)$th column is b.

(4.16) *Theorem*

Let $A \in \mathbb{F}^{m\times n}$ and let $b \in \mathbb{F}^{m\times 1}$. Then the following statements are true:

(1) *If $[A' \mid b']$ is obtained from $[A \mid b]$ by applying a finite sequence of elementary row operations, then each solution of $Ax = b$ is a solution of $A'x = b'$, and vice versa.*

(2) **(Consistency)** *$Ax = b$ has a solution iff* $\operatorname{rank}([A \mid b]) = \operatorname{rank}(A)$.

(3) *If u is a (particular) solution of $Ax = b$, then each solution of $Ax = b$ is given by $u + y$, where y is a solution of the homogeneous system $Ax = 0$.*

(4) *If $r = \operatorname{rank}([A \mid b]) = \operatorname{rank}(A) < n$, then there are $n - r$ unknowns which can take arbitrary values; and other r unknowns can be determined from the values of these $n - r$ unknowns.*

(5) *If $m < n$, then the homogeneous system has infinitely many solutions.*

(6) *$Ax = b$ has a unique solution iff* $\text{rank}([A \,|\, b]) = \text{rank}(A) = n$.

(7) *If $m = n$, then $Ax = b$ has a unique solution iff* $\det(A) \neq 0$.

(8) **(Cramer's Rule)** *If $m = n$ and* $\det(A) \neq 0$, *then the solution of $Ax = b$ is given by*
$x_j = \det(\,A_j(b)\,)/\det(A)$ *for each* $j \in \{1, \ldots, n\}$.

Proof. (1) If $[A' \,|\, b']$ has been obtained from $[A \,|\, b]$ by a finite sequence of elementary row operations, then $A' = EA$ and $b' = Eb$, where E is the product of corresponding elementary matrices. The matrix E is invertible. Now, $A'x = b'$ iff $EAx = Eb$ iff $Ax = E^{-1}Eb = b$.

(2) Due to (1), we assume that $[A \,|\, b]$ is in RREF. Suppose $Ax = b$ has a solution. If there is a zero row in A, then the corresponding entry in b is also 0. Therefore, there is no pivot in b. Hence $\text{rank}([A \,|\, b]) = \text{rank}(A)$.

Conversely, suppose that $\text{rank}([A \,|\, b]) = \text{rank}(A) = r$. Then there is no pivot in b. That is, b is a non-pivotal column in $[A \,|\, b]$. Thus, b is a linear combination of pivotal columns, which are some columns of A. Therefore, $Ax = b$ has a solution.

(3) Let u be a solution of $Ax = b$. Then $Au = b$. Now, z is a solution of $Ax = b$ iff $Az = b$ iff $Az = Au$ iff $A(z - u) = 0$ iff $z - u$ is a solution of $Ax = 0$. That is, each solution z of $Ax = b$ is expressed in the form $z = u + y$ for a solution y of the homogeneous system $Ax = 0$.

(4) Let $\text{rank}([A \,|\, b]) = \text{rank}(A) = r < n$. By (2), there exists a solution. Due to (3), we consider solving the corresponding homogeneous system. Due to (1), assume that A is in RREF. There are r number of pivots in A and $m - r$ number of zero rows. Omit all the zero rows; it does not affect the solutions. Write the system as linear equations. Rewrite the equations by keeping the unknowns corresponding to pivots on the left hand side, and taking every other term to the right hand side. The unknowns corresponding to pivots are now expressed in terms of the other $n - r$ unknowns. For obtaining a solution, we may arbitrarily assign any values to these $n - r$ unknowns, and the unknowns corresponding to the pivots get evaluated by the equations.

(5) Let $m < n$. Then $r = \text{rank}(A) \leq m < n$. Consider the homogeneous system $Ax = 0$. By (4), there are $n - r \geq 1$ number of unknowns which can take arbitrary values, and other r unknowns are determined accordingly. Each such assignment of values to the $n - r$ unknowns gives rise to a distinct solution resulting in infinite number of solutions of $Ax = 0$.

(6) It follows from (3) and (4).

(7) If $A \in \mathbb{F}^{n\times n}$, then it is invertible iff $\operatorname{rank}(A) = n$ iff $\det(A) \neq 0$. Then use (6). In this case, the unique solution is given by $x = A^{-1}b$.

(8) Recall that $A_j(b)$ is the matrix obtained from A by replacing the jth column of A with the vector b. Since $\det(A) \neq 0$, by (6), $Ax = b$ has a unique solution, say $y \in \mathbb{F}^{n\times 1}$. Write the identity $Ay = b$ in the form:

$$y_1 \begin{bmatrix} a_{11} \\ \vdots \\ a_{n1} \end{bmatrix} + \cdots + y_j \begin{bmatrix} a_{1j} \\ \vdots \\ a_{nj} \end{bmatrix} + \cdots + y_n \begin{bmatrix} a_{1n} \\ \vdots \\ a_{nn} \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}.$$

This gives

$$y_1 \begin{bmatrix} a_{11} \\ \vdots \\ a_{n1} \end{bmatrix} + \cdots + \begin{bmatrix} (y_j a_{1j} - b_1) \\ \vdots \\ (y_j a_{nj} - b_n) \end{bmatrix} + \cdots + y_n \begin{bmatrix} a_{1n} \\ \cdots \\ a_{nn} \end{bmatrix} = 0.$$

In this sum, the jth vector is a linear combination of other vectors, where $-y_j$s are the coefficients. Therefore,

$$\begin{vmatrix} a_{11} & \cdots & (y_j a_{1j} - b_1) & \cdots & a_{1n} \\ & & \vdots & & \\ a_{n1} & \cdots & (y_j a_{nj} - b_n) & \cdots & a_{nn} \end{vmatrix} = 0.$$

From Property (6) of the determinant, it follows that

$$y_j \begin{vmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ & & \vdots & & \\ a_{n1} & \cdots & a_{nj} & \cdots & a_{nn} \end{vmatrix} - \begin{vmatrix} a_{11} & \cdots & b_1 & \cdots & a_{1n} \\ & & \vdots & & \\ a_{n1} & \cdots & b_n & \cdots & a_{nn} \end{vmatrix} = 0.$$

Therefore, $y_j = \det(\, A_j(b)\,)/\det(A)$. ∎

A system of linear equations $Ax = b$ is said to be **consistent** iff $\operatorname{rank}([A|b]) = \operatorname{rank}(A)$. Due to (4.16-1) only consistent systems have solutions. And, (4.16-2) asserts that all solutions of the non-homogeneous system can be obtained by adding a particular solution to solutions of the corresponding homogeneous system.

4.6 Gauss-Jordan elimination

To determine whether a system of linear equations is consistent or not, it is enough to convert the augmented matrix $[A|b]$ to its RREF and then check whether an entry in the b portion of the augmented matrix has become a pivot or not. In fact, the pivot check shows that corresponding to the zero rows in the portion of A in the RREF of $[A|b]$, all the entries in b must be zero. Thus an entry in the b portion has become a pivot guarantees that the system is inconsistent, else the system is consistent.

(4.17) *Example*

Is the following system of linear equations consistent?

$$\begin{aligned} 5x_1 + 2x_2 - 3x_3 + x_4 &= 7 \\ x_1 - 3x_2 + 2x_3 - 2x_4 &= 11 \\ 3x_1 + 8x_2 - 7x_3 + 5x_4 &= 8 \end{aligned}$$

We take the augmented matrix and reduce it to its RREF by elementary row operations.

$$\left[\begin{array}{rrrr|r} 5 & 2 & -3 & 1 & 7 \\ 1 & -3 & 2 & -2 & 11 \\ 3 & 8 & -7 & 5 & 8 \end{array}\right] \xrightarrow{O1} \left[\begin{array}{rrrr|r} \boxed{1} & 2/5 & -3/5 & 1/5 & 7/5 \\ 0 & -17/5 & 13/5 & -11/5 & 48/5 \\ 0 & 34/5 & -26/5 & 22/5 & -19/5 \end{array}\right]$$

$$\xrightarrow{O2} \left[\begin{array}{rrrr|r} \boxed{1} & 0 & -5/17 & -1/17 & 43/17 \\ 0 & \boxed{1} & -13/17 & 11/17 & -48/17 \\ 0 & 0 & 0 & 0 & \boxed{77/5} \end{array}\right]$$

Here, $O1 = R_1 \leftarrow 1/5R_1,\ R_2 \leftarrow R_2 - R_1,\ R_3 \leftarrow R_3 - 3R_1$ and
$O2 = R_2 \leftarrow -5/17R_2,\ R_1 \leftarrow R_1 - 2/5R_2,\ R_3 \leftarrow R_3 - 34/5R_2$.
Since an entry in the b portion has become a pivot, the system is inconsistent. In fact, you can verify that the third row in A is simply first row minus twice the second row, whereas the third entry in b is not the first entry minus twice the second entry. Therefore, the system is inconsistent. ∎

(4.18) *Example*

Give conditions under which the system $x+y+5z = 3,\ x+2y+az = 5,\ x+2y+4z = b$ is consistent.

We take the augmented matrix and reduce it to its RREF by elementary row operations.

$$\left[\begin{array}{rrr|r} 1 & 1 & 5 & 3 \\ 1 & 2 & a & 5 \\ 1 & 2 & 4 & b \end{array}\right] \longrightarrow \left[\begin{array}{rrr|r} \boxed{1} & 1 & 5 & 3 \\ 0 & 1 & a-5 & 2 \\ 0 & 1 & -1 & b-3 \end{array}\right] \longrightarrow \left[\begin{array}{rrr|r} \boxed{1} & 0 & 10-a & 1 \\ 0 & \boxed{1} & a-5 & 2 \\ 0 & 0 & 4-a & b-5 \end{array}\right]$$

We see that $\text{rank}[A|b] = \text{rank}(A)$ if $4 - a \neq 0$ or if $b - 5 = 0$. That is, if $a \neq 4$, the system is consistent; and if $b = 5$, then also the system is consistent. ∎

Gauss-Jordan elimination is an application of converting the augmented matrix to its RREF for solving linear systems.

(4.19) *Example*

We change the last equation in the previous example to make it consistent. The system now looks like:

$$\begin{aligned} 5x_1 + 2x_2 - 3x_3 + x_4 &= 7 \\ x_1 - 3x_2 + 2x_3 - 2x_4 &= 11 \\ 3x_1 + 8x_2 - 7x_3 + 5x_4 &= -15 \end{aligned}$$

The reduction to echelon form will change that entry as follows: take the augmented matrix and reduce it to its echelon form by elementary row operations.

$$\left[\begin{array}{rrrr|r} 5 & 2 & -3 & 1 & 7 \\ 1 & -3 & 2 & -2 & 11 \\ 3 & 8 & -7 & 5 & -15 \end{array}\right] \xrightarrow{O1} \left[\begin{array}{rrrr|r} \boxed{1} & 2/5 & -3/5 & 1/5 & 7/5 \\ 0 & -17/5 & 13/5 & -11/5 & 48/5 \\ 0 & 34/5 & -26/5 & 22/5 & -96/5 \end{array}\right]$$

$$\xrightarrow{O2} \left[\begin{array}{rrrr|r} \boxed{1} & 0 & -5/17 & -1/17 & 43/17 \\ 0 & \boxed{1} & -13/17 & 11/17 & -48/17 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

with $O_1 = R_1 \leftarrow 1/5R_1,\ R_2 \leftarrow R_2 - R_1,\ R_3 \leftarrow R_3 - 3R_1$ and
$O_2 = R_2 \leftarrow -5/17R_2,\ R_1 \leftarrow E_1 - 2/5R_2,\ R_3 \leftarrow R_3 - 34/5R_2$.

This expresses the fact that the third equation is redundant. Now, solving the new system in RREF is easier. Writing as linear equations, we have

$$\begin{aligned} \boxed{1}\,x_1 \qquad - \tfrac{5}{17}x_3 - \tfrac{1}{17}x_4 &= \tfrac{43}{17} \\ \boxed{1}\,x_2 - \tfrac{13}{17}x_3 + \tfrac{11}{17}x_4 &= -\tfrac{48}{17} \end{aligned}$$

The unknowns corresponding to the pivots are called the **basic variables** and the other unknowns are called the **free variable**. By assigning the free variables to any arbitrary values, the basic variables can be evaluated. So, we assign a free variable x_i an arbitrary number, say α_i, and express the basic variables in terms of the free variables to get a solution of the equations.

In the above reduced system, the basic variables are x_1 and x_2; and the unknowns x_3, x_4 are free variables. We assign x_3 to α_3 and x_4 to α_4. The solution is written as follows:

$$x_1 = \tfrac{43}{17} + \tfrac{5}{17}\alpha_3 + \tfrac{1}{17}\alpha_4, \quad x_2 = -\tfrac{48}{17} + \tfrac{13}{17}\alpha_3 - \tfrac{11}{17}\alpha_4, \quad x_3 = \alpha_3, \quad x_4 = \alpha_4.$$

Notice that any solution of the system is in the form $u + v$, where

$$u = \begin{bmatrix} \frac{43}{17} \\ -\frac{48}{17} \\ 0 \\ 0 \end{bmatrix}, \quad v = \begin{bmatrix} \frac{5}{17}\alpha_3 \\ \frac{1}{17}\alpha_4 \\ \alpha_3 \\ \alpha_4 \end{bmatrix};$$

u is a particular solution of the system, and v is a solution of the corresponding homogeneous system. ▯

4.7 Exercises for Chapter 4

1. Convert the following matrices into RREF and determine their ranks.

   (a) $\begin{bmatrix} 5 & 2 & -3 & 1 & 7 \\ 1 & -3 & 2 & -2 & 11 \\ 3 & 8 & -7 & 5 & 8 \end{bmatrix}$ (b) $\begin{bmatrix} 5 & 2 & -3 & 1 & 30 \\ 1 & -3 & 2 & -2 & 11 \\ 3 & 8 & -7 & 5 & 8 \end{bmatrix}$

2. Determine linear independence of $\{(1, 2, 2, 1),\ (1, 3, 2, 1),\ (4, 1, 2, 2),\ (5, 2, 4, 3)\}$ in $\mathbb{C}^{1\times 4}$. *Ans*: Linearly dependent.

3. Compute A^{-1} using RREF and also using determinant, where

   $A = \begin{bmatrix} 4 & -7 & -5 \\ -2 & 4 & 3 \\ 3 & -5 & -4 \end{bmatrix}$. *Ans*: $\begin{bmatrix} 1 & 3 & 1 \\ -1 & 1 & 2 \\ 2 & 1 & -2 \end{bmatrix}$.

4. Solve the following system by Gauss-Jordan elimination:

$$\begin{aligned} x_1 + x_2 + x_3 + x_4 - 3x_5 &= 6 \\ 2x_1 + 3x_2 + x_3 + 4x_4 - 9x_5 &= 17 \\ x_1 + x_2 + x_3 + 2x_4 - 5x_5 &= 8 \\ 2x_1 + 2x_2 + 2x_3 + 3x_4 - 8x_5 &= 14 \end{aligned}$$

5. Check if the system is consistent. If so, determine the solution set.

   (a) $x_1 - x_2 + 2x_3 - 3x_4 = 7,\ 4x_1 + 3x_3 + x_4 = 9,\ 2x_1 - 5x_2 + x_3 = -2,$ $3x_1 - 2x_2 - 2x_3 + 10x_4 = -12.$

   (b) $x_1 - x_2 + 2x_3 - 3x_4 = 7, 4x_1 + 3x_3 + x_4 = 9, 2x_1 - 5x_2 + x_3 = -2,$ $3x_1 - 2x_2 - 2x_3 + 10x_4 = -14.$

6. Using Gauss-Jordan elimination determine the values of $k \in \mathbb{R}$ so that the system of linear equations $x + y - z = 1,\ 2x + 3y + kz = 3,\ x + ky + 3z = 2$ has (a) no solution, (b) infinitely many solutions, (c) exactly one solution.
7. Let A be an $n \times n$ matrix with integer entries and $\det(A^2) = 1$. Show that all entries of A^{-1} are also integers.
8. Let $A \in \mathbb{F}^{m\times n}$ have columns $A_1, \ldots, A_n$. Let $b \in \mathbb{F}^m$. Show the following:

   (a) The equation $Ax = 0$ has a non-zero solution iff $A_1, \ldots, A_n$ are linearly dependent.

   (b) The equation $Ax = b$ has at least one solution iff $b \in \operatorname{span}\{A_1, \ldots, A_n\}$.

   (c) Let u be a solution of $Ax = b$. Then, u is the only solution of $Ax = b$ iff $A_1, \ldots, A_n$ are linearly independent.

   (d) The equation $Ax = b$ has a unique solution iff $\operatorname{rank} A = \operatorname{rank}[A|b] =$ number of unknowns.

9. Let $A \in \mathbb{F}^{m\times n}$ have rank r. Give reasons for the following:

   (a) $\operatorname{rank}(A) \le \min\{m, n\}$.

   (b) If $n > m$, then there exist $x, y \in \mathbb{F}^{n\times 1}$ such that $x \ne y$ and $Ax = Ay$.

   (c) If $n < m$, then there exists $y \in \mathbb{F}^{m\times 1}$ such that for no $x \in \mathbb{F}^{n\times 1}$, $Ax = y$.

   (d) If $n = m$, then the following statements are equivalent:

      i. $Au = Av$ implies $u = v$ for all $u, v \in \mathbb{F}^{n\times 1}$.

      ii. Corresponding to each $y \in \mathbb{F}^{n\times 1}$, there exists $x \in \mathbb{F}^{m\times 1}$ such that $y = Ax$.

5

Matrix Eigenvalue Problem

5.1 Eigenvalues and eigenvectors

In this chapter, unless otherwise specified, we assume that any matrix is a square matrix with complex entries.

Let $A \in \mathbb{C}^{n\times n}$. A complex number λ is called an **eigenvalue** of A iff there exists a non-zero vector $v \in \mathbb{C}^{n\times 1}$ such that $Av = \lambda v$. Such a vector v is called an **eigenvector of** A **for** (or, associated with, or, corresponding to) the eigenvalue λ.

(5.1) ***Example***

Consider the matrix $A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$. It has an eigenvector $\begin{bmatrix} 1 & 0 & 0 \end{bmatrix}^T$ associated with the eigenvalue 1.

Is $\begin{bmatrix} 2 & 0 & 0 \end{bmatrix}^T$ also an eigenvector associated with the same eigenvalue 1? □

In fact, corresponding to an eigenvalue, there are infinitely many eigenvectors.

(5.2) ***Theorem***

Let $A \in \mathbb{C}^{n\times n}$. Let $v \in \mathbb{C}^{n\times 1}$, $v \neq 0$. Then, v is an eigenvector of A for the eigenvalue $\lambda \in \mathbb{C}$ iff v is a nonzero solution of the homogeneous system $(A - \lambda I)x = 0$ iff $\det(A - \lambda I) = 0$.

Proof. The complex number λ is an eigenvalue of A iff we have a nonzero vector $v \in \mathbb{C}^{n\times 1}$ such that $Av = \lambda v$ iff $(A - \lambda I)v = 0$ and $v \neq 0$ iff $A - \lambda I$ is not invertible iff $\det(A - \lambda I) = 0$. ∎

5.2 Characteristic polynomial

The polynomial $\det(A - tI)$ is called the **characteristic polynomial** of the matrix A. Thus any complex number λ that satisfies the characteristic polynomial of a matrix A, is an eigenvalue of A.

Since the characteristic polynomial of a matrix A of order n is a polynomial of degree n in t, it has exactly n, not necessarily distinct, complex zeros. And these are the eigenvalues of A. Notice that, here, we are using the fundamental theorem of algebra which says that each polynomial of degree n with complex coefficients can be factored into exactly n linear factors.

(5.3) *Example*

Find the eigenvalues and corresponding eigenvectors of the matrix

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

The characteristic polynomial is

$$\det(A - tI) = \begin{vmatrix} 1-t & 0 & 0 \\ 1 & 1-t & 0 \\ 1 & 1 & 1-t \end{vmatrix} = (1-t)^3.$$

Thus, 1 is the only eigenvalue of A.

To get an eigenvector, we solve $A \begin{bmatrix} a & b & c \end{bmatrix}^T = 1 \begin{bmatrix} a & b & c \end{bmatrix}^T$ or that

$$a = a, \; a + b = b, \; a + b + c = c.$$

It gives $a = b = 0$ and $c \in \mathbb{F}$ can be arbitrary. Since an eigenvector is nonzero, all the eigenvectors are given by $\begin{bmatrix} 0 & 0 & c \end{bmatrix}^T$, for any $c \neq 0$. □

The eigenvalue λ being a zero of the characteristic polynomial has certain multiplicity. That is, the maximum k such that $(t - \lambda)^k$ divides the characteristic polynomial is called the **algebraic multiplicity** of the eigenvalue λ.

In (5.3), the algebraic multiplicity of the eigenvalue 1 is 3.

(5.4) *Example*

For $A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$, the characteristic polynomial is $t^2 + 1$. It has eigenvalues as i and

$-i$. The corresponding eigenvectors are obtained by solving

$$A\begin{bmatrix}a\ b\end{bmatrix}^T = i\begin{bmatrix}a\ b\end{bmatrix}^T \text{ and } A\begin{bmatrix}a\ b\end{bmatrix}^T = -i\begin{bmatrix}a\ b\end{bmatrix}^T.$$

For $\lambda = i$, we have $b = ia, -a = ib$. Thus, $\begin{bmatrix}a\ ia\end{bmatrix}^T$ is an eigenvector for $a \neq 0$.
For the eigenvalue $-i$, the eigenvectors are $\begin{bmatrix}a & -ia\end{bmatrix}$ for $a \neq 0$.
Here, algebraic multiplicity of each eigenvalue is 1. □

If a matrix of order n has only real entries, then its characteristic polynomial has only real coefficients. Then complex zeros of the characteristic polynomial occur in conjugate pairs. That is, if $\alpha + i\beta$ is an eigenvalue of a matrix with real entries, where $\beta \neq 0$, then $\alpha - i\beta$ is also an eigenvalue of this matrix.

We say that $A, B \in \mathbb{C}^{n\times n}$ are **similar** iff there exists an invertible matrix $P \in \mathbb{C}^{n\times n}$ such that $B = P^{-1}AP$. Some easy consequences of our definition are listed in the following theorem.

(5.5) ***Theorem***

(1) *A matrix and its transpose have the same eigenvalues.*

(2) *Similar matrices have the same eigenvalues.*

(3) *The diagonal entries of any triangular matrix are precisely its eigenvalues.*

Proof. (1) $\det(A^T - tI) = \det((A - tI)^T) = \det(A - tI)$.

(2) $\det(P^{-1}AP - tI) = \det(P^{-1}(A - tI)P) = \det(P^{-1})\det(A - tI)\det(P) = \det(A - tI)$.

(3) If A is triangular, then $\det(A - tI) = (a_{11} - t)\cdots(a_{nn} - t)$. ∎

(5.6) ***Theorem***

$\det(A)$ *equals the product and* $\operatorname{tr}(A)$ *equals the sum of all eigenvalues of* A.

Proof. Let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of A, not necessarily distinct. Now,

$$\det(A - tI) = (\lambda_1 - t)\cdots(\lambda_n - t).$$

Put $t = 0$. It gives $\det(A) = \lambda_1 \cdots \lambda_n$.
Expand $\det(A - tI)$ and equate the coefficients of t^{n-1} to get

Coeff of t^{n-1} in $\det(A - tI)$ = Coeff of t^{n-1} in $(a_{11} - t)\cdot A_{11}$
$= \cdots =$ Coeff of t^{n-1} in $(a_{11} - t)\cdot(a_{22} - t)\cdots(a_{nn} - t) = (-1)^{n-1}\operatorname{tr}(A)$.

But Coeff of t^{n-1} in $(\lambda_1 - t)\cdots(\lambda_n - t)$ is $(-1)^{n-1}\cdot\sum\lambda_i$. ∎

(5.7) ***Theorem*** (Caley-Hamilton)

Any square matrix satisfies its characteristic polynomial.

Proof. Let $A \in \mathbb{C}^{n\times n}$. Let $p(t) = c_0+c_1t+\cdots+c_nt^n$ be the characteristic polynomial of A. We show that $p(A) = 0$, the zero matrix. (3.6-15) with the matrix $A - tI$ says that

$$p(t)\,I = \det(A - tI)\,I = \left[\operatorname{adj}(A - tI)\right](A - tI).$$

The entries in $\operatorname{adj}(A - tI)$ are polynomials in t of degree at most $n - 1$. Write

$$\operatorname{adj}(A - tI) := B_0 + tB_1 + \cdots + t^{n-1}B_{n-1},$$

where $B_0, \ldots, B_{n-1} \in \mathbb{C}^{n\times n}$. Then

$$c_0I + c_1It + \cdots + c_nIt^n = p(t)I = (B_0 + tB_1 + \cdots t^{n-1}B_{n-1})(A - t\,I).$$

Comparing the coefficients of t^k, we obtain

$$c_0I = B_0A,\ c_1I = B_1A - B_0, \ldots, c_{n-1}I = B_{n-1}A - B_{n-2},\ c_nI = -B_{n-1}.$$

Then, substituting these values in $p(A)$, we have

$$\begin{aligned} p(A) &= c_0I + c_1A + \cdots + c_nA^n = c_0I + c_1IA + \cdots + c_nIA^n \\ &= B_0A + (B_1A - B_0)A + \cdots + (B_{n-1}A - B_{n-2})A^{n-1} - B_{n-1}A^n = 0. \end{aligned}$$ ∎

Suppose a matrix $A \in \mathbb{C}^{n\times n}$ has the characteristic polynomial

$$a_0 + a_1t + \cdots + a_{n-1}t^{n-1} + (-1)^nt^n.$$

By Cayley-Hamilton theorem, $a_0I + a_1A + \cdots + (-1)^nA^n = 0$. Then

$$A^n = (-1)^{n-1}\big(a_0I + a_1A + \cdots + a_{n-1}A^{n-1}\big).$$

Thus, computation of $A^n, A^{n+1}, \ldots$ can be reduced to computing $A,\ A^2,\ \ldots,\ A^{n-1}$.

A similar approach shows that the inverse of a matrix can be expressed as a polynomial in the matrix. If A is invertible, then $\det(A) \neq 0$; so that 0 is not an eigenvalue of A. That is, $a_0 \neq 0$. Then

$$a_0I + A\big(a_1I + \cdots + a_{n-1}A^{n-2} + (-1)^nA^{n-1}\big) = 0.$$

Multiplying A^{-1} and simplifying, we obtain

$$A^{-1} = -\frac{1}{a_0}\big(a_1I + a_2A + \cdots + a_{n-1}A^{n-2} + (-1)^nA^{n-1}\big).$$

5.3 Inner product and norm of vectors

The **inner product** of two vectors $u = (a_1, \ldots, a_n)$ and $v = (b_1, \ldots, b_n)$ in $\mathbb{F}^n$ is defined as

$$\langle u, v\rangle = a_1\overline{b}_1 + \cdots + a_n\overline{b}_n.$$

In particular, if $\mathbb{F} = \mathbb{R}$, then $u, v \in \mathbb{R}^n$ and $\overline{b}_i = b_i$ so that

$$\langle u, v\rangle = a_1 b_1 + \cdots + a_n b_n.$$

For instance, if $u = (1, 2, 3) \in \mathbb{R}^3$ and $v = (2, 1, 3) \in \mathbb{R}^3$, then their inner product is

$$\langle u, v\rangle = 1 \times 2 + 2 \times 1 + 3 \times 3 = 13.$$

If $x = (1 + i, 2 - i, 1) \in \mathbb{C}^3$ and $y = (1 - i, 1 + i, 1) \in \mathbb{C}^3$, then their inner product is

$$\langle x, y\rangle = (1 + i)(1 + i) + (2 - i)(1 - i) + 1 \times 1 = 2 - i.$$

Notice that the inner product of two vectors in $\mathbb{F}^n$ is a scalar.

When we consider row or column vectors, their inner product can be given via matrix multiplication.

Let $u, v \in \mathbb{F}^{1\times n}$. Then $\langle u, v\rangle = uv^*$.
Reason: Suppose $u = \begin{bmatrix} a_1 & \cdots & a_n \end{bmatrix}$ and $v = \begin{bmatrix} b_1 & \cdots & b_n \end{bmatrix}$. Then

$$uv^* = \begin{bmatrix} a_1 & \cdots & a_n \end{bmatrix} \begin{bmatrix} \overline{b}_1 \\ \vdots \\ \overline{b}_n \end{bmatrix} = a_1\overline{b}_1 + \cdots + a_n\overline{b}_n = \langle u, v\rangle.$$

In particular, if $u, v \in \mathbb{R}^{1\times n}$ then $\langle u, v\rangle = uv^T$.

Similarly, if $u, v \in \mathbb{F}^{n\times 1}$ then $\langle u, v\rangle = v^* u$.

Verification: Suppose $u = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}$ and $v = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}$. Then

$$v^* u = \begin{bmatrix} \overline{b}_1 & \cdots & \overline{b}_n \end{bmatrix} \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} = \overline{b}_1 a_1 + \cdots + \overline{b}_n a_n = \langle u, v\rangle.$$

In particular, when $u, v \in \mathbb{R}^{n\times 1}$, $\langle u, v\rangle = v^T u$.

The inner product satisfies the following properties:

For $x, y, z \in \mathbb{F}^n$ and $\alpha, \beta \in \mathbb{F}$,

1. $\langle x, x\rangle \geq 0$.
2. $\langle x, x\rangle = 0$ iff $x = 0$.
3. $\langle x, y\rangle = \overline{\langle y, x\rangle}$.
4. $\langle x + y, z\rangle = \langle x, z\rangle + \langle y, z\rangle$.
5. $\langle z, x + y\rangle = \langle z, x\rangle + \langle z, y\rangle$.
6. $\langle \alpha x, y\rangle = \alpha\langle x, y\rangle$.
7. $\langle x, \beta y\rangle = \overline{\beta}\langle x, y\rangle$.

The inner product gives rise to the length of a vector as in the familiar case of $\mathbb{R}^{1\times 3}$. We now call the generalized version of length as the *norm*. If $u \in \mathbb{F}^n$, we define its **norm**, denoted by $\|u\|$ as the nonnegative square root of $\langle u, u\rangle$. That is, $\|u\| = \sqrt{\langle u, u\rangle}$.

Thus, if $u = (a_1, \ldots, a_n) \in \mathbb{F}^n$, then $\|u\| = \sqrt{|a_1|^2 + \cdots + |a_n|^2}$.

In particular, when $u = (a, \ldots, a_n) \in \mathbb{R}^n$, we have $\|u\| = \sqrt{a_1^2 + \cdots + a_n^2}$.

Using matrix product, we may write the norm as follows:

If $u \in \mathbb{R}^{1\times n}$, then $\|u\| = \sqrt{uu^T}$. If $u \in \mathbb{R}^{n\times 1}$, then $\|u\| = \sqrt{u^T u}$.

If $u \in \mathbb{C}^{1\times n}$, then $\|u\| = \sqrt{uu^*}$. If $u \in \mathbb{C}^{n\times 1}$, then $\|u\| = \sqrt{u^* u}$.

The norm satisfies the following properties:

For $x, y \in \mathbb{F}^n$ and $\alpha \in \mathbb{F}$,

1. $\|x\| \geq 0$.
2. $\|x\| = 0$ iff $x = 0$.
3. $\|\alpha x\| = |\alpha|\, \|x\|$.
4. $|\langle x, y\rangle| \leq \|x\|\, \|y\|$. (*Cauchy-Schwartz inequality*)
5. $\|x + y\| \leq \|x\| + \|y\|$. (*Triangle inequality*)

A proof of Cauchy-Schwartz inequality goes as follows:

If $y = 0$, then the inequality clearly holds. Else, $\langle y, y\rangle \neq 0$. Write $\alpha = \frac{\langle x,y\rangle}{\langle y,y\rangle}$. Then $\overline{\alpha} = \frac{\langle y,x\rangle}{\langle y,y\rangle}$ and $\overline{\alpha}\langle x, y\rangle = |\alpha|^2\|y\|^2$. Then

$$0 \leq \langle x - \alpha y, x - \alpha y\rangle = \langle x, x\rangle - \overline{\alpha}\langle x, y\rangle + \alpha\big(\overline{\alpha}\langle y, y\rangle - \langle y, x\rangle\big)$$
$$= \|x\|^2 - \overline{\alpha}\langle x, y\rangle = \|x\|^2 - |\alpha|^2\|y\|^2 = \|x\|^2 - \frac{|\langle x, y\rangle|^2}{\|y\|^4}\|y\|^2.$$

The triangle inequality can be proved using Cauchy-Schwartz, as in the following:

$$\|x + y\|^2 = \langle x + y, x + y\rangle = \|x\|^2 + \|y\|^2 + \langle x, y\rangle + \langle y, x\rangle \leq \|x\|^2 + \|y\|^2 + 2\|x\|\, \|y\|.$$

Let $x, y \in \mathbb{F}^n$. We say that the vectors x and y are **orthogonal**, and we write this as $x \perp y$, when $\langle x, y\rangle = 0$. That is,

$$x \perp y \quad \text{iff} \quad \langle x, y\rangle = 0.$$

It follows that if $x \perp y$, then $\|x\|^2 + \|y\|^2 = \|x+y\|^2$. This is referred to as **Pythagoras law**. The converse of Pythagoras law holds when $\mathbb{F} = \mathbb{R}$, but fails in general for $\mathbb{F} = \mathbb{C}$.

Adjoints of matrices behave in a very predictable way with the inner product.

(5.8) *Theorem*

Let $A \in \mathbb{F}^{m\times n}$, $x \in \mathbb{F}^{n\times 1}$, and let $y \in \mathbb{F}^{m\times 1}$. Then

$$\langle Ax, y\rangle = \langle x, A^*y\rangle \quad \textit{and} \quad \langle A^*y, x\rangle = \langle y, Ax\rangle.$$

Proof. Recall that in $\mathbb{F}^{r\times 1}$, $\langle u, v\rangle = v^*u$. Further, $Ax \in \mathbb{F}^{m\times 1}$ and $A^*y \in \mathbb{F}^{n\times 1}$. We are using the same notation for both the inner products in $\mathbb{F}^{m\times 1}$ and in $\mathbb{F}^{n\times 1}$. We then have

$$\langle Ax, y\rangle = y^*Ax = (A^*y)^*x = \langle x, A^*y\rangle.$$

The second equality follows from the first. ∎

Often the definition of an adjoint is taken using the identity: $\langle Ax, y\rangle = \langle x, A^*y\rangle$.

We show that unitary or orthogonal matrices preserve inner product and also the norm.

(5.9) *Theorem*

Let $A \in \mathbb{C}^{n\times n}$ be a unitary or an orthogonal matrix.

(1) *For each pair of vectors x, y, $\langle Ax, Ay\rangle = \langle x, y\rangle$. In particular, $\|Ax\| = \|x\|$ for any x.*

(2) *The columns of A are orthogonal and each is of norm* 1.

(3) *The rows of A are orthogonal, and each is of norm* 1.

Proof. (1) $\langle Ax, Ay\rangle = \langle x, A^*Ay\rangle = \langle x, y\rangle$. Take $x = y$ for the second equality.

(2) Since $A^*A = I$, the ith row of A^* multiplied with the jth column of A gives δ_{ij}. However, this product is simply the inner product of the jth column of A with the ith column of A.

(3) It follows from (2). Also, considering $AA^* = I$, we get this result. ∎

5.4 Gram-Schmidt orthogonalization

Linear independence of a finite list of vectors can be determined using the inner product.

Let $v_1, \ldots, v_n \in \mathbb{F}^n$. We say that these vectors are **orthogonal** iff $\langle v_i, v_j \rangle = 0$ for all pairs of indices i, j with $i \neq j$.

Orthogonality is stronger than linear independence, as the following theorem shows.

(5.10) *Theorem*

Any orthogonal list of nonzero vectors in $\mathbb{F}^n$ *is linearly independent.*

Proof. Let $v_1, \ldots, v_n \in \mathbb{F}^n$ be nonzero vectors. For scalars $a_1, \ldots, a_n$, let

$$a_1 v_1 + \cdots + a_n v_n = 0.$$

Take inner product of both the sides with v_1. Since $\langle v_i, v_1 \rangle = 0$ for each $i \neq 1$, we obtain $\langle a_1 v_1, v_1 \rangle = 0$. But $\langle v_1, v_1 \rangle \neq 0$. Therefore, $a_1 = 0$. Similarly, it follows that each $a_i = 0$. ∎

It will be convenient to use the following terminology. We denote the set of all linear combinations of vectors $v_1, \ldots, v_m$ by $\text{span}(v_1, \ldots, v_m)$; and read it as the **span of** the vectors $v_1, \ldots, v_m$.

Our procedure, called **Gram-Schmidt orthogonalization**, constructs orthogonal vectors $v_1, \ldots, v_k$ from the given vectors $u_1, \ldots, u_m$ so that

$$\text{span}(v_1, \ldots, v_k) = \text{span}(u_1, \ldots, u_m), \quad k \leq m.$$

It is described in (5.11) below. First, let us see how we proceed.

Given two linearly independent vectors u_1, u_2 on the plane how do we construct two orthogonal vectors?

Keep $v_1 = u_1$. Take out the projection of u_2 on u_1 to get v_2. Now, $v_2 \perp v_1$.

What is the projection of u_2 on u_1?

Its length is $\langle u_2, u_1 \rangle$. Its direction is that of u_1, i.e., $u_1/\|u_1\|$.

Thus $v_1 = u_1, \ v_2 = u_2 - \dfrac{\langle u_2, v_1 \rangle}{\langle v_1, v_1 \rangle} v_1$.

We may continue this process of taking out projections in n dimensions.

(5.11) *Theorem*

Let $u_1, u_2, \ldots, u_m \in \mathbb{F}^n$. Define

$$\begin{aligned} v_1 &= u_1 \\ v_2 &= u_2 - \frac{\langle u_2, v_1\rangle}{\langle v_1, v_1\rangle} v_1 \\ &\vdots \\ v_m &= u_m - \frac{\langle u_m, v_1\rangle}{\langle v_1, v_1\rangle} v_1 - \cdots - \frac{\langle u_m, v_{m-1}\rangle}{\langle v_{m-1}, v_{m-1}\rangle} v_{m-1} \end{aligned}$$

In the above process, if $v_i = 0$, then both u_i and v_i are ignored for the rest of the steps. After ignoring such u_is and v_is suppose we obtain the vectors as $v_{j1}, \ldots, v_{jk}$. Then $v_{j1}, \ldots, v_{jk}$ are orthogonal and $\operatorname{span}(v_{j1}, \ldots, v_{jk}) = \operatorname{span}\{u_1, u_2, \ldots, u_m\}$. *Further, if $v_i = 0$ for $i > 1$, then $u_i \in \operatorname{span}\{u_1, \ldots, u_{i-1}\}$.*

Proof Outline: We verify algebraically our geometric intuition:

$$v_1 = u_1, \quad v_2 = u_2 - \frac{\langle u_2, v_1\rangle}{\langle v_1, v_1\rangle} v_1.$$

Hence $\langle v_2, v_1\rangle = \left\langle u_2 - \frac{\langle u_2, v_1\rangle}{\langle v_1, v_1\rangle} v_1, v_1\right\rangle = \langle u_2, v_1\rangle - \frac{\langle u_2, v_1\rangle}{\langle v_1, v_1\rangle}\langle v_1, v_1\rangle = 0.$

If $v_2 = 0$, then u_2 is a scalar multiple of u_1. If $v_2 \neq 0$, then u_1, u_2 are linearly independent.

Similar to our verification of $v_2 \perp v_1$, we can prove that v_{i+1} is orthogonal to $v_1, \ldots, v_i$ for each $i \geq 1$, by using induction.

We need to prove that both the sets spans the same set. Notice that

If $x_1, \ldots, x_r \in \operatorname{span}(y_1, \ldots, y_s)$, then $\operatorname{span}(x_1, \ldots, x_r) \subseteq \operatorname{span}(y_1, \ldots, y_s)$.

For, if $v = \alpha_1 x_1 + \cdots + \alpha_r x_r$ and $x_i = a_{i1} v_1 + \cdots a_{is} v_s$, then substituting for each x_i in the previous expression and combining terms, we get

$$v = \sum_{i=1}^{s} (\alpha_1 a_{1i} + \cdots + \alpha_r a_{ri}) v_i \in \operatorname{span}(v_1, \ldots, v_s).$$

If u_i is a linear combination of $u_1, \ldots, u_{i-1}$, then $\operatorname{span}(u_1, \ldots, u_{i-1}) = \operatorname{span}(u_1, \ldots, u_i)$.

Now observe inductively that $v_1, \ldots, v_i \in \operatorname{span}(u_1, \ldots, u_i)$.

From the algorithm, it can also be observed, using induction, that $u_1, \ldots, u_i \in \operatorname{span}(v_1, \ldots, v_i)$.

Therefore, we get $\operatorname{span}(u_1, \ldots, u_i) = \operatorname{span}(v_1, \ldots, v_i)$ for each $i \geq 1$.

(5.12) *Example*

Consider the vectors $u_1 = (1,0,0),\ u_2 = (1,1,0)$ and $u_3 = (1,1,1)$. Apply Gram-Schmidt Orthogonalization.

$$\begin{aligned}
v_1 &= (1,0,0).\\
v_2 &= u_2 - \frac{\langle u_2, v_1\rangle}{\langle v_1, v_1\rangle}\, v_1 = (1,1,0) - \frac{(1,1,0)\cdot(1,0,0)}{(1,0,0)\cdot(1,0,0)}(1,0,0)\\
&= (1,1,0) - 1\,(1,0,0) = (0,1,0).\\
v_3 &= u_3 - \frac{\langle u_3, v_1\rangle}{\langle v_1, v_1\rangle}\, v_1 - \frac{\langle u_3, v_2\rangle}{\langle v_2, v_2\rangle}\, v_2\\
&= (1,1,1) - (1,1,1)\cdot(1,0,0)(1,0,0) - (1,1,1)\cdot(0,1,0)(0,1,0)\\
&= (1,1,1) - (1,0,0) - (0,1,0) = (0,0,1).
\end{aligned}$$

The set $\{(1,0,0),(0,1,0),(0,0,1)\}$ is orthogonal; and span of the new vectors is the same as span of the old ones, which is $\mathbb{R}^3$. ∎

(5.13) *Example*

The vectors $u_1 = (1,1,0), u_2 = (0,1,1), u_3 = (1,0,1)$ form a basis for $\mathbb{F}^3$. Apply Gram-Schmidt Orthogonalization.

$$\begin{aligned}
v_1 &= (1,1,0).\\
v_2 &= u_2 - \frac{\langle u_2, v_1\rangle}{\langle v_1, v_1\rangle}\, v_1 = (0,1,1) - \frac{(0,1,1)\cdot(1,1,0)}{(1,1,0)\cdot(1,1,0)}(1,1,0)\\
&= (0,1,1) - \frac{1}{2}(1,1,0) = \Big(-\frac{1}{2},\frac{1}{2},1\Big).\\
v_3 &= u_3 - \frac{\langle u_3, v_1\rangle}{\langle v_1, v_1\rangle}\, v_1 - \frac{\langle u_3, v_2\rangle}{\langle v_2, v_2\rangle}\, v_2\\
&= (1,0,1) - (1,0,1)\cdot(1,1,0)(1,1,0) - (1,0,1)\cdot\Big(-\frac{1}{2},\frac{1}{2},1\Big)\Big(-\frac{1}{2},\frac{1}{2},1\Big)\\
&= (1,0,1) - \frac{1}{2}(1,1,0) - \frac{1}{3}\Big(-\frac{1}{2},\frac{1}{2},1\Big) = \Big(-\frac{2}{3},\frac{2}{3},-\frac{2}{3}\Big).
\end{aligned}$$

The set $\Big\{(1,1,0),\ \Big(-\frac{1}{2},\frac{1}{2},1\Big),\ \Big(-\frac{2}{3},\frac{2}{3},-\frac{2}{3}\Big)\Big\}$ is orthogonal. ∎

(5.14) *Example*

Use Gram-Schmidt orthogonalization on the vectors $u_1 = (1,1,0,1), u_2 = (0,1,1,-1)$ and $u_3 = (1,3,2,-1)$.

$$\begin{aligned}
v_1 &= (1,1,0,1).\\
v_2 &= u_2 - \frac{\langle u_2, v_1\rangle}{\langle v_1, v_1\rangle} v_1 = (0,1,1,-1) - \frac{\langle (0,1,1,-1),(1,1,0,1)\rangle}{\langle (1,1,0,1),(1,1,0,1)\rangle}(1,1,0,1) = (0,1,1,-1).\\
v_3 &= u_3 - \frac{\langle u_3, v_1\rangle}{\langle v_1, v_1\rangle} v_1 - \frac{\langle u_3, v_2\rangle}{\langle v_2, v_2\rangle} v_2\\
&= (1,3,2,-1) - \frac{\langle (1,3,2,-1),(1,1,0,1)\rangle}{\langle (1,1,0,1),(1,1,0,1)\rangle}(1,1,0,1)\\
&\quad - \frac{\langle (1,3,2,-1),(0,1,1,-1)\rangle}{\langle (0,1,1,-1),(0,1,1,-1)\rangle}(0,1,1,-1)\\
&= (1,3,2,-1) - (1,1,0,1) - 2(0,1,1,-1) = (0,0,0,0).
\end{aligned}$$

Notice that since u_1, u_2 are already orthogonal, Gram-Schmidt process returned $v_2 = u_2$. Next, the process also revealed the fact that $u_3 = u_1 + 2u_2$. □

(5.15) *Example*

Use Gram-Schmidt orthogonalization on the vectors $u_1 = (1,2,2,1)$, $u_2 = (2,1,0,-1)$, $u_3 = (4,5,4,1)$ and $u_4 = (5,4,2,-1)$.

$$\begin{aligned}
v_1 &= (1,2,2,1).\\
v_2 &= (2,1,0,-1) - \frac{\langle (2,1,0,-1),(1,2,2,1)\rangle}{\langle (1,2,2,1),(1,2,2,1)\rangle}(1,2,2,-1) = \left(\frac{17}{10}, \frac{2}{5}, -\frac{3}{5}, -\frac{13}{10}\right).\\
v_3 &= (4,5,4,1) - \frac{\langle (4,5,4,1),(1,2,2,1)\rangle}{\langle (1,2,2,1),(1,2,2,1)\rangle}(1,2,2,1)\\
&\quad - \frac{\langle (4,5,4,1),(\frac{3}{2},0,-1,\frac{1}{2})\rangle}{\langle \left(\frac{17}{10}, \frac{2}{5}, -\frac{3}{5}, -\frac{13}{10}\right), \left(\frac{17}{10}, \frac{2}{5}, -\frac{3}{5}, -\frac{13}{10}\right)\rangle}\left(\frac{17}{10}, \frac{2}{5}, -\frac{3}{5}, -\frac{13}{10}\right) = (0,0,0,0).
\end{aligned}$$

So, we ignore v_3, u_3 and note that u_3 is a linear combination of u_1, u_2; and hence, a linear combination of v_1, v_2. Next, we compute

$$v_4 = u_4 - \frac{\langle u_4, v_1\rangle}{\langle v_1, v_1\rangle} v_1 - \frac{\langle u_4, v_2\rangle}{\langle v_2, v_2\rangle} v_2 = 0.$$

Now, u_4 is a linear combination of u_1, u_2, and thus, it is a linear combination of v_1, v_2. In fact, $u_3 = 2u_1 + u_2$ and $u_4 = u_1 + 2u_2$. So, $v_1 \perp v_2$ and $\text{span}(u_1, u_2, u_3, u_4) = \text{span}(v_1, v_2)$. □

5.5 Hermitian and unitary matrices

Recall that a hermitian matrix is one for which the adjoint coincides with itself, and a unitary matrix is one for which its adjoint coincides with its inverse. Real hermitian matrices are called real symmetric matrices and real unitary matrices are called orthogonal matrices. If A is an orthogonal matrix, then each column of it is orthogonal to any other column. For example, the rotation in the plane given by

$$\begin{bmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{bmatrix}$$

for any real number θ is an orthogonal matrix. It can be shown that any orthogonal 2×2 matrix is a rotation combined with a reflection on a straight line. Similarly, any orthogonal matrix in the three dimensional euclidean space is a rotation combined with a reflection on a plane.

(5.16) ***Theorem***

Let $A \in \mathbb{C}^{n\times n}$. Let λ be any eigenvalue of A.

(1) *If A is hermitian or real symmetric, then $\lambda \in \mathbb{R}$.*

(2) *If A is skew-hermitian or skew-symmetric, then λ is purely imaginary or zero.*

(3) *If A is unitary or orthogonal, then $|\lambda| = 1$.*

Proof. Let $\lambda \in \mathbb{C}$ be an eigenvalue of A with an eigenvector $v \in \mathbb{C}^{n\times 1}$. Now, $Av = \lambda v$ and $v \neq 0$. Pre-multiplying with v^*, we have $v^*Av = \lambda v^*v \in \mathbb{C}$. Taking adjoint, we obtain: $v^*A^*v = \overline{\lambda}v^*v$.

(1) Let A be hermitian, i.e., $A^* = A$. Then $\overline{\lambda}v^*v = v^*A^*v = v^*Av = \lambda v^*v$.
Since $v^*v \neq 0$, $\overline{\lambda} = \lambda$. That is, λ is real.

(2) When A is skew-hermitian, $A^* = -A$. Then $\overline{\lambda}v^*v = v^*A^*v = -v^*Av = -\lambda v^*v$.
Since $v^*v \neq 0$, $\overline{\lambda} = -\lambda$. That is, $2\mathrm{Re}(\lambda) = 0$. So, λ is purely imaginary or zero.

(3) Let A be unitary, i.e., $A^*A = I$. Now, $Av = \lambda v$. Taking adjoint, we have $v^*A^* = \overline{\lambda}v^*$. Then $v^*v = v^*Iv = v^*A^*Av = \overline{\lambda}\lambda v^*v = |\lambda|^2 v^*v$. Since $v^*v \neq 0$, $|\lambda| = 1$. ∎

Not only each eigenvalue of a real symmetric matrix is real, but also a corresponding real eigenvector can be chosen. It follows from the general fact that if a real matrix has a real eigenvalue, then there exists a corresponding real eigenvector. To see this, let $A \in \mathbb{R}^{n\times n}$ have a real eigenvalue λ with corresponding eigenvector $v = x + iy$, where $x, y \in \mathbb{R}^{n\times 1}$. Comparing the real and imaginary parts in $A(x+iy) = \lambda(x+iy)$, we have $Ax = \lambda x$ and $Ay = \lambda y$. Since $x + iy \neq 0$, at least one of x or y is nonzero. Such a nonzero vector is a real eigenvector corresponding to the eigenvalue λ of A.

5.6 Diagonalization

(5.17) *Theorem*

Eigenvectors corresponding to distinct eigenvalues of a square matrix are linearly independent.

Proof. Let $\lambda_1, \ldots, \lambda_m$ be all the distinct eigenvalues of $A \in \mathbb{C}^{n\times n}$. Let $v_1, \ldots, v_m$ be corresponding eigenvectors. We use induction on $i \in \{1, \ldots, m\}$.

For $i = 1$, since $v_1 \neq 0$, $\{v_1\}$ is linearly independent.

Induction Hypothesis: for $i = k$ suppose $\{v_1, \ldots, v_k\}$ is linearly independent. We use the characterization of linear independence as proved in (3.4).

The induction hypothesis implies that if we equate any linear combination of $v_1, \ldots, v_k$ to 0, then the coefficients in the linear combination must all be 0. Now, for $i = k+1$, we want to show that $v_1, \ldots, v_k, v_{k+1}$ are linearly independent. So, we start equating an arbitrary linear combination of these vectors to 0. Our aim is to derive that each scalar coefficient in such a linear combination must be 0. Towards this, assume that

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_k v_k + \alpha_{k+1} v_{k+1} = 0. \tag{5.6.1}$$

Then, $A(\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_k v_k + \alpha_{k+1} v_{k+1}) = 0$. Since $Av_j = \lambda_j v_j$, we have

$$\alpha_1 \lambda_1 v_1 + \alpha_2 \lambda_2 v_2 + \cdots + \alpha_k \lambda_k v_k + \alpha_{k+1} \lambda_{k+1} v_{k+1} = 0. \tag{5.6.2}$$

Multiply (5.6.1) with λ_{k+1}. Subtract from (5.6.2) to get:

$$\alpha_1(\lambda_1 - \lambda_{k+1})v_1 + \cdots + \alpha_k(\lambda_k - \lambda_{k+1})v_k = 0.$$

By the Induction Hypothesis, $\alpha_j(\lambda_j - \lambda_{k+1}) = 0$ for each $j = 1, \ldots, k$. Since $\lambda_1, \ldots, \lambda_{k+1}$ are distinct, we conclude that $\alpha_1 = \cdots = \alpha_k = 0$. Then, from (5.6.1), it follows that $\alpha_{k+1} v_{k+1} = 0$. As $v_{k+1} \neq 0$, we have $\alpha_{k+1} = 0$. ∎

Suppose an $n \times n$ matrix A has n linearly independent eigenvectors $v_1, \ldots, v_n$. Let $\lambda_1, \ldots, \lambda_n$ be the corresponding eigenvalues. We find that

$$Av_1 = \lambda_1 v_1, \; \ldots, Av_n = \lambda_n v_n.$$

Construct the matrix $P \in \mathbb{C}^{n\times n}$ by taking its columns as the eigenvectors $v_1, \ldots, v_n$. That is, let

$$P = \begin{bmatrix} v_1 \; v_2 \; \cdots \; v_{n-1} \; v_n \end{bmatrix}.$$

Also, construct the diagonal matrix $D = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$. That is,

$$D = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}.$$

Then the above product of A with the v_is can be written as a single equation $AP = PD$. Now, $\operatorname{rank}(P) = n$. So, P is an invertible matrix. Then

$$P^{-1}AP = D.$$

Let $A \in \mathbb{C}^{n\times n}$. We call A to be **diagonalizable** iff there exists an invertible matrix P such that $P^{-1}AP$ is a diagonal matrix. (That is, A is similar to a diagonal matrix.) We also say that A is **diagonalizable by the matrix** P iff $P^{-1}AP = D$.

(5.18) *Theorem*

An $n \times n$ matrix is diagonalizable iff it has n linearly independent eigenvectors.

Proof. In fact, we have already proved that if an $n \times n$ matrix A has n linearly independent eigenvectors, then A is diagonalizable.

For the converse, suppose that $A \in \mathbb{C}^{n\times n}$ is diagonalizable. So, let $P = \begin{bmatrix} v_1, \; \cdots \; v_n \end{bmatrix}$ be an invertible matrix and let $D = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ be such that $P^{-1}AP = D$. Then $AP = PD$. Then $Av_i = \lambda_i v_i$ for $1 \le i \le n$. That is, each v_i is an eigenvector of A. Moreover, P is invertible implies that $v_1, \ldots, v_n$ are linearly independent. ∎

(5.19) *Example*

Consider the matrix $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$.

Since it is upper triangular, its eigenvalues are the diagonal entries.

That is, 1 is the only eigenvalue of A with algebraic multiplicity 2. To find the eigenvectors, we solve

$$(A - 1\, I) \begin{bmatrix} a \\ b \end{bmatrix} = 0.$$

The equation can be rewritten as $a + b = a, \; b = b$. Solving the equations, we have $b = 0$ and a arbitrary. There is only one linearly independent eigenvector, namely, $\begin{bmatrix} a \; 0 \end{bmatrix}^T$ for a nonzero scalar a. Therefore, A is not diagonalizable. □

The following result is now obvious.

(5.20) *Theorem*

If $A \in \mathbb{C}^{n\times n}$ has n distinct eigenvalues $\lambda_1, \ldots, \lambda_n$, then it is similar to $\mathrm{diag}(\lambda_1, \ldots, \lambda_n)$.

We state, without proof, another sufficient condition for diagonalizability.

(5.21) *Theorem* (Spectral Theorem)

A square matrix is normal iff it is diagonalized by a unitary matrix.
Each real symmetric matrix is diagonalized by an orthogonal matrix.

Since each hermitian matrix is a normal matrix, it follows that each hermitian matrix is diagonalizable by a unitary matrix.

To **diagonalize** a matrix A means that we determine an invertible matrix P and a diagonal matrix D such that $P^{-1}AP = D$. Notice that only square matrices can possibly be diagonalized.

In general, diagonalization starts with determining eigenvalues and corresponding eigenvectors of A. We then construct the diagonal matrix D by taking the eigenvalues $\lambda_1, \ldots, \lambda_n$ of A. Next, we construct P by putting the corresponding eigenvectors $v_1, \ldots, v_n$ as columns of P in that order. Then $P^{-1}AP = D$. This work succeeds provided that the list of eigenvectors $v_1, \ldots, v_n$ in $\mathbb{C}^{n\times 1}$ are linearly independent.

Once we know that a matrix A is diagonalizable, we can give a procedure to diagonalize it. All we have to do is determine the eigenvalues and corresponding eigenvectors so that the eigenvectors are linearly independent and their number is equal to the order of A. Then, put the eigenvectors as columns to construct the matrix P. Then $P^{-1}AP$ is a diagonal matrix.

(5.22) *Example*

Consider the matrix $A = \begin{bmatrix} 1 & -1 & -1 \\ -1 & 1 & -1 \\ -1 & -1 & 1 \end{bmatrix}$.

It is real symmetric having eigenvalues -1, 2 and 2. To find the associated eigenvectors, we must solve the linear systems of the form $Ax = \lambda x$.

For the eigenvalue -1, the system $Ax = -x$ gives

$$x_1 - x_2 - x_3 = -x_1, \; -x_1 + x_2 - x_3 = -x_2, \; -x_1 - x_2 + x_3 = -x_3.$$

It has a solution: $x_1 = x_2 = x_3$. One eigenvector is $\begin{bmatrix} 1 & 1 & 1 \end{bmatrix}^T$.

For the eigenvalue 2, we have the equations as

$$x_1 - x_2 - x_3 = 2x_1, \; -x_1 + x_2 - x_3 = 2x_2, \; -x_1 - x_2 + x_3 = 2x_3.$$

It leads to $x_1 + x_2 + x_3 = 0$. We can have two linearly independent eigenvectors such as $\begin{bmatrix} -1 & 1 & 0 \end{bmatrix}^T$ and $\begin{bmatrix} -1 & -1 & 2 \end{bmatrix}^T$.

The three eigenvectors are orthogonal to each other. To orthonormalize, we divide each by its norm. We end up at the following orthonormal eigenvectors:

$$\begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}, \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix}, \begin{bmatrix} -1/\sqrt{6} \\ -1/\sqrt{6} \\ 2/\sqrt{6} \end{bmatrix}.$$

They are orthogonal vectors in $\mathbb{R}^{3\times 1}$, each of norm 1. Taking

$$P = \begin{bmatrix} 1/\sqrt{3} & -1/\sqrt{2} & -1/\sqrt{6} \\ 1/\sqrt{3} & 1/\sqrt{2} & -1/\sqrt{6} \\ 1/\sqrt{3} & 0 & 2/\sqrt{6} \end{bmatrix},$$

we have $P^{-1} = P^T,\ P^{-1}AP = P^TAP = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$ ∎

5.7 Exercises for Chapter 5

1. Find the eigenvalues and the associated eigenvectors for the matrices given below.

   (a) $\begin{bmatrix} 3 & 0 \\ 8 & -1 \end{bmatrix}$ (b) $\begin{bmatrix} 3 & 2 \\ -1 & 0 \end{bmatrix}$ (c) $\begin{bmatrix} -2 & -1 \\ 5 & 2 \end{bmatrix}$ (d) $\begin{bmatrix} -2 & 0 & 3 \\ -2 & 3 & 0 \\ 0 & 0 & 5 \end{bmatrix}$.

   Ans: (a) $\lambda_1 = 3,\ v_1 = [1\ 2]^T;\ \lambda_2 = -1,\ v_2 = [0\ 1]^T$.
   (b) $\lambda_1 = 1,\ v_1 = [1\ -1]^T;\ \lambda_2 = 2,\ v_2 = [2\ -1]^T$.
   (c) $\lambda_1 = i,\ v_1 = [1\ 2+i]^T;\ \lambda_2 = -i,\ v_2 = [1\ i-2]^T$.
   (d) $\lambda_1 = -2,\ v_1 = [5\ 2\ 0]^T;\ \lambda_2 = 3,\ v_2 = [0\ 1\ 0]^T;\ \lambda_3 = 5,\ v_3 = [3\ -3\ 7]^T$.

2. Let A be an $n \times n$ matrix and α be a scalar such that each row (or each column) sums to α. Show that α is an eigenvalue of A.

3. Let $A \in \mathbb{C}^{n\times n}$ be invertible. Show that $\lambda \in \mathbb{C}$ is an eigenvalue of A if and only if λ^{-1} is an eigenvalue of A^{-1}.

4. The vectors $u_1 = (1,2,2), u_2 = (-1,0,2), u_3 = (0,0,1)$ are linearly independent in $\mathbb{F}^3$. Apply Gram-Schmidt Orthogonalization.
   Ans: $v_1 = u_1, v_2 = (-4/3, -2/3, 4/3), v_3 = (2/9, -2/9, 1/9)$.

5. Let $A \in \mathbb{R}^{3\times 3}$ have the first two columns as $(1/\sqrt{3}, 1/\sqrt{3}, 1/\sqrt{3},)^T$ and $(1/\sqrt{2}, 0, -1/\sqrt{2})^T$. Determine the third column of A so that A is an orthogonal matrix. *Ans*: $\pm(1/\sqrt{6})(1, -2, 1)^T$.
6. Show that eigenvectors corresponding to distinct eigenvalues of a unitary (or orthogonal) matrix are orthogonal to each other.
7. Give an example of an $n \times n$ matrix that cannot be diagonalized.
   Ans: $A = [a_{ij}] \in \mathbb{C}^{n\times n}$ with $a_{12} = 1$ and all other entries as 0.
8. Find the matrix $A \in \mathbb{R}^{3\times 3}$ that satisfies the given condition. Diagonalize it if possible.
   (a) $A(a, b, c)^T = (a + b + c,\ a + b - c,\ a - b + c)^T$ for all $a, b, c \in \mathbb{R}$.
   (b) $Ae_1 = 0, \quad Ae_2 = e_1, \quad Ae_3 = e_2.$
   (c) $Ae_1 = e_2, \quad Ae_2 = e_3, \quad Ae_3 = 0.$
   (d) $Ae_1 = e_3, \quad Ae_2 = e_2, \quad Ae_3 = e_1.$
9. Show that the following matrices are diagonalizable.
   (a) $\begin{bmatrix} 3/2 & -1/2 & 0 \\ -1/2 & 3/2 & 0 \\ 1/2 & -1/2 & 1 \end{bmatrix}$ (b) $\begin{bmatrix} 3 & -1/2 & -3/2 \\ 1 & 3/2 & 3/2 \\ -1 & -1/2 & 5/2 \end{bmatrix}$.
10. Which of the following matrices is/are diagonalizable? If one is diagonalizable, then diagonalize it.
    (a) $\begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{bmatrix}$ (b) $\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$ (c) $\begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$ (d) $\begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$.

Solutions to Exercises

Series of Numbers §1.9

1. Show the following:

(a) $\lim\limits_{n\to\infty} \dfrac{\ln n}{n} = 0.$ (b) $\lim\limits_{n\to\infty} n^{1/n} = 1.$ (c) $\lim\limits_{n\to\infty} x^n = 0$ for $|x| < 1.$

(d) $\lim\limits_{n\to\infty} \dfrac{n^p}{x^n} = 0$ for $x > 1.$ (e) $\lim\limits_{n\to\infty} \dfrac{x^n}{n!} = 0$ (f) $\lim\limits_{n\to\infty} \left(1 + \dfrac{x}{n}\right)^n = e^x$

(a) $\ln x$ is defined on $[1, \infty)$. Using L' Hospital's rule, $\lim\limits_{x\to\infty} \dfrac{\ln x}{x} = \lim\limits_{x\to\infty} \dfrac{1}{x} = 0.$ Therefore, $\lim\limits_{n\to\infty} \dfrac{\ln n}{n} = \lim\limits_{x\to\infty} \dfrac{\ln x}{x} = 0.$

(b) $\lim n^{1/n} = \lim\limits_{x\to\infty} x^{1/x} = \lim\limits_{x\to\infty} e^{\frac{1}{x}\ln x} = e^{\lim_{x\to\infty} \frac{\ln x}{x}} = e^0 = 1.$

Here, we have used continuity of e^x.

(c) Write $|x| = \dfrac{1}{1+r}$ for some $r > 0$. By the Binomial theorem, $(1+r)^n \geq 1 + nr > nr$. So, $0 < |x|^n = 1/(1+r)^n < 1/(nr)$. As $|x|^n = |x^n|$, we have $-1/(nr) < x^n < 1/(nr)$. By Sandwich theorem, $\lim x^n = 0$.

(d) Let $x > 1$. We know that $\lim\limits_{t\to\infty} \dfrac{t^p}{x^t} = 0$ for $p \in \mathbb{N}$. Therefore, $\lim \dfrac{n^p}{x^n} = 0.$

If $m < p < m + 1$ for an $m \in \mathbb{N}$, then $n^p < n^{m+1}$. Use Sandwich theorem to get the limit. If $p < 1$, then similarly, use $n^p < n$.

Analogously, show that the limits in (e) and (f) hold.

2. Prove the following:

   (a) It is not possible that a series converges to a real number ℓ and also diverges to $-\infty$.
   (b) It is not possible that a series diverges to ∞ and also to $-\infty$.

(a) Suupose $\sum a_j$ converges to ℓ and also diverges to $-\infty$. Then we have natural numbers k, m such that for every $n \geq k$, $\ell - 1 < \sum_{j=1}^{n} a_j < \ell + 1$. And also for all $n \geq m$, $\sum_{j=1}^{n} a_j < \ell - 2$. Choose $M = \max\{k, m\}$. Then both inequalities hold for $n = M$. But this is not possible.

(b) Suppose $\sum a_j$ diverges to both ∞ and to $-\infty$. Then we have natural numbers k, m such that for each $n \geq k$, $\sum_{j=1}^{n} a_j > 1$ and for each $n \geq m$, $\sum_{j=1}^{n} a_j < -1$. Choose $M = \max\{k, m\}$. Then both the inequalities hold for $n = M$. But this is impossible.

3. Prove the following:

    (a) If both the series $\sum a_n$ and $\sum b_n$ converge, then the series $\sum(a_n + b_n)$, $\sum(a_n - b_n)$ and $\sum ka_n$ converge; where k is any real number.
    (b) If $\sum a_n$ converges and $\sum b_n$ diverges to $\pm\infty$, then $\sum(a_n + b_n)$ diverges to $\pm\infty$, and $\sum(a_n - b_n)$ diverges to $\mp\infty$.
    (c) If $\sum a_n$ diverges to $\pm\infty$, and $k > 0$, then $\sum ka_n$ diverges to $\pm\infty$.
    (d) If $\sum a_n$ diverges to $\pm\infty$, and $k < 0$, then $\sum ka_n$ diverges to $\mp\infty$.

(a) Suppose $\sum a_n$ converges to ℓ and $\sum b_n$ converges to s. Let $\epsilon > 0$. Then we have natural numbers k, m such that for all $n \geq k$, $|\sum_{j=1}^{n} a_j - \ell| < \epsilon/2$; and for all $n \geq m$, $|\sum_{j=1}^{n} b_j - s| < \epsilon/2$. Choose $M = \max\{k, m\}$. Then for all $n \geq M$, both the inequalities hold. So, we obtain

$$\Big|\sum_{j=1}^{n}(a_j + b_j) - (\ell + s)\Big| \leq \Big|\sum_{j=1}^{n} a_j - \ell\Big| + \Big|\sum_{j=1}^{n} b_j - s\Big| < \epsilon/2 + \epsilon/2 = \epsilon.$$

Similarly, the other two are proved.

(b) Suppose $\sum a_n$ converges to ℓ and $\sum b_n$ diverges to ∞. Let $r > 0$. Then, we have natural numbers k, m such that for all $n \geq k$, $\ell - 1 < \sum_{j=1}^{n} a_j < \ell + 1$; and for all $n \geq m$, $\sum_{j=1}^{n} b_j > r + |\ell| + 1$. Choose $M = \max\{k, m\}$. Then all the three inequalities hold for $n \geq M$. But then for all $n \geq M$,

$$\ell - 1 < \sum_{j=1}^{n} a_j, \quad r + |\ell| + 1 < \sum_{j=1}^{n} b_j.$$

That is, for all $n \geq M$, $r \leq \ell - 1 + r + |\ell| + 1 < \sum_{j=1}^{n}(a_j + b_j)$.
Similarly, other cases are proved.

(c) Suppose $\sum a_n$ diverges to $\pm\infty$, and $k > 0$. Let $r \in \mathbb{R}$. We have $m \in \mathbb{N}$ such that for all $n \geq m$, $\sum_{j=1}^{n} a_j > r/k$. Then for all such n, $\sum_{j=1}^{n}(ka_j) > r$.
Similarly other cases are proved.

(d) Suppose $\sum a_n$ diverges to ∞, and $k < 0$. Let $r \in \mathbb{R}$. We have $m \in \mathbb{N}$ such that for all $n \geq m$, $\sum_{j=1}^{n} a_j > r/k$. Then for all such n, $\sum_{j=1}^{n}(ka_j) < r$, since $k < 0$.
Similarly other cases are proved.

4. Give examples for the following:

   (a) $\sum a_n$ and $\sum b_n$ both diverge, but $\sum(a_n + b_n)$ converges to a nonzero number.
   (b) $\sum a_n$ and $\sum b_n$ both diverge, and $\sum(a_n + b_n)$ diverges to ∞.
   (c) $\sum a_n$ and $\sum b_n$ both diverge, and $\sum(a_n + b_n)$ diverges to $-\infty$.

   (a) $1 + 1 + 1 + \cdots$ diverges; $2 + (-1) + (-1) + \cdots$ also diverges.
   But $(1 + 2)) + (1 + (-1)) + \cdots = 3 + 0 + \cdots$ converges to 3.

   (b) $1 + 2 + 3 + 4 + \cdots$ diverges; $-1 - 1 - 1 - 1 - \cdots$ also diverges.
   And $(1 - 1) + (2 - 1) + (3 - 1) + \cdots = 0 + 1 + 2 + 3 + \cdots$ diverges to ∞.

   (c) $-1 - 2 - 3 - 4 - \cdots$ diverges; $1 + 1 + 1 + 1 + \cdots$ also diverges.
   And $(-1+1)+(-2+1)+(-3+1)+\cdots = 0-1-2-3-\cdots$ diverges to $-\infty$.

5. Show that the sequence 1, 1.1, 1.1011, 1.10110111, ... converges.

   Use either Cauchy sequences or monotonically increasing bounded sequences.

6. Determine whether the following series converge:

   (a) $\displaystyle\sum_{n=1}^{\infty} \frac{-n}{3n+1}$ (b) $\displaystyle\sum_{n=1}^{\infty} \frac{\ln n}{n^{3/2}}$ (c) $\displaystyle\sum_{n=1}^{\infty} \frac{1 + n\ln n}{1+n^2}$

   (a) It diverges because $\displaystyle\lim_{n\to\infty} \frac{-n}{3n+1} = -\frac{1}{3} \neq 0$.
   (b) Take $a_n = \dfrac{\ln n}{n^{3/2}}$ and $b_n = \dfrac{1}{n^{5/4}}$. Then

   $$\lim_{n\to\infty} \frac{a_n}{b_n} = \lim_{n\to\infty} \frac{\ln n}{n^{1/4}} = \lim_{n\to\infty} \frac{1/n}{(1/4)n^{-3/4}} = \lim_{n\to\infty} \frac{4}{n^{1/4}} = 0.$$

   Since $\sum b_n$ converges, by the Limit comparison test, $\sum a_n$ converges.
   (c) Take $a_n = \dfrac{1 + n\ln n}{1+n^2}$ and $b_n = \dfrac{1}{n}$. Then

   $$\lim_{n\to\infty} \frac{a_n}{b_n} = \lim_{n\to\infty} \frac{n + n^2 \ln n}{1+n^2} = \infty.$$

   As $\sum b_n$ diverges to ∞, by the Limit comparison test, $\sum a_n$ diverges to ∞.

7. Test for convergence the series $\dfrac{1}{3} + \left(\dfrac{2}{3}\right)^2 + \left(\dfrac{3}{7}\right)^3 + \cdots + \left(\dfrac{n}{2n+1}\right)^n + \cdots$.

   Using Cauchy root test, $\displaystyle\lim_{n\to\infty} (a_n)^{1/n} = \lim_{n\to\infty} \frac{n}{2n+1} = \frac{1}{2} < 1$.
   Therefore, the series converges.

8. Is the integral $\displaystyle\int_{-\infty}^{\infty} \frac{1}{1+x^2}\, dx$ convergent?

   $$\int_a^b \frac{1}{1+x^2}\, dx = \tan^{-1} b - \tan^{-1} a.$$

So,

$$\int_{-\infty}^{0} \frac{1}{1+x^2}\,dx = \lim_{a\to-\infty}\int_{a}^{0} \frac{1}{1+x^2}\,dx = \lim_{a\to-\infty}(-\tan^{-1} a) = -\Big(-\frac{\pi}{2}\Big) = \frac{\pi}{2}.$$

$$\int_{0}^{\infty} \frac{1}{1+x^2}\,dx = \lim_{b\to\infty}\int_{0}^{b} \frac{1}{1+x^2}\,dx = \lim_{b\to\infty}(\tan^{-1} b) = \frac{\pi}{2}.$$

Therefore,

$$\int_{-\infty}^{\infty} \frac{1}{1+x^2}\,dx = \int_{-\infty}^{0} \frac{1}{1+x^2}\,dx + \int_{0}^{\infty} \frac{1}{1+x^2}\,dx$$

is convergent and its value is $\pi/2 + \pi/2 = \pi$.

9. Is the area under the curve $y = (\ln x)/x^2$ for $1 \le x < \infty$ finite?

The question is whether $\displaystyle\int_1^{\infty} \frac{\ln x}{x^2}\,dx$ converges?

Let $b > 1$. Integrating by parts,

$$\int_1^{b} \frac{\ln x}{x^2}\,dx = \Big[\ln x\Big(-\frac{1}{x}\Big)\Big]_1^{b} - \int_1^{b}\Big(\frac{-1}{x}\Big)\frac{1}{x}\,dx = -\frac{\ln b}{b} - \frac{1}{b} + 1.$$

$$\lim_{b\to\infty}\int_1^{b} \frac{\ln x}{x^2}\,dx = \lim_{b\to\infty}\Big[-\frac{\ln b}{b} - \frac{1}{b} + 1\Big] = 1.$$

Therefore, the improper integral $\displaystyle\int_1^{\infty} \frac{\ln x}{x^2}\,dx$ converges to 1. That is, the required area is finite and it is equal to 1.

10. Evaluate (a) $\displaystyle\int_0^3 \frac{dx}{(x-1)^{2/3}}$ (b) $\displaystyle\int_0^3 \frac{dx}{x-1}$

(a) The integrand is not defined at $x = 1$. We consider it as an improper integral.

$$\int_0^3 \frac{dx}{(x-1)^{2/3}} = \lim_{b\to1-}\int_0^{b} \frac{dx}{(x-1)^{2/3}} + \lim_{a\to1+}\int_a^{3} \frac{dx}{(x-1)^{2/3}}.$$

$$\lim_{b\to1-}\int_0^{b} \frac{dx}{(x-1)^{2/3}} = \lim_{b\to1-} 3(x-1)^{1/3}\Big|_0^{b} = \lim_{b\to1-}(3(b-1)^{1/3} - 3(-1)^{1/3}) = 3.$$

$$\lim_{a\to1+}\int_a^{3} \frac{dx}{(x-1)^{2/3}} = \lim_{a\to1+} 3(x-1)^{1/3}\Big|_a^{3} = \lim_{a\to1+}(3(3-1)^{1/3} - 3(a-1)^{1/3}) = 3(2)^{1/3}.$$

Hence $\displaystyle\int_0^3 \frac{dx}{(x-1)^{2/3}} = 3(1+2^{1/3})$.

Had we not noticed that the integrand has discontinuity in the interior, we would have ended up at a wrong computation such as

$$\int_0^3 \frac{dx}{(x-1)^{2/3}} = 3(x-1)^{1/3}\Big|_0^3 = 3(2^{1/3} - (-1)^{1/3}),$$

even though the answer happens to be correct here. See the next problem.
(b) Overlooking the point $x = 1$, where the integrand is not defined, we may compute

$$\int_0^3 \frac{dx}{x-1} = \ln|x-1|\Big]_0^3 = \ln 2 - \ln 1 = \ln 2.$$

However, it is an improper integral and its value, if exists, must be computed as follows:

$$\int_0^3 \frac{dx}{x-1} = \lim_{b\to 1-}\int_0^b \frac{dx}{x-1} + \lim_{a\to 1+}\int_a^3 \frac{dx}{x-1}.$$

The integral converges provided both the limits are finite. However,

$$\lim_{b\to 1-}\int_0^b \frac{dx}{x-1} = \lim_{b\to 1-}\big(\ln|b-1| - \ln|-1|\big) = \lim_{b\to 1-}\ln(1-b) = -\infty.$$

Therefore, $\displaystyle\int_0^3 \frac{dx}{x-1}$ does not converge.

11. Show that $\displaystyle\int_1^\infty \frac{\sin x}{x^p}\,dx$ converges for all $p > 0$.

For $p > 1$ and $x \ge 1$, $\left|\dfrac{\sin x}{x^p}\right| \le \dfrac{1}{x^p}$. Since $\displaystyle\int_1^\infty \frac{dx}{x^p}$ converges, $\displaystyle\int_1^\infty \left|\frac{\sin x}{x^p}\right| dx$ converges. Therefore, $\displaystyle\int_1^\infty \frac{\sin x}{x^p}\,dx$ converges.

For $0 < p \le 1$, use integration by parts: $\displaystyle\int_1^b \frac{\sin x}{x^p}\,dx = -\frac{\cos b}{b^p} + \frac{\cos 1}{1^p} - p\int_1^b \frac{\cos x}{x^{p+1}}\,dx.$
Taking the limit as $b \to \infty$, we see that the first term goes to 0; the second term is already a real number, the third term, an improper integral converges as in the case for $p > 1$ above. Therefore, the given improper integral also converges in this case.

12. Show that $\displaystyle\int_0^\infty \frac{\sin x}{x^p}\,dx$ converges for $0 < p \le 1$.

For $p = 1$, the integral $\displaystyle\int_0^1 \frac{\sin x}{x}\,dx$ is not an improper integral. Since $\dfrac{\sin x}{x}$ with its value at 0 as 1 is continuous on $[0, 1]$, this integral exists.

For $0 < p < 1$ and $0 < x \le 1$, since $\dfrac{\sin x}{x^p} \le \dfrac{1}{x^p}$ and $\displaystyle\int_0^1 \frac{dx}{x^p}$ converges due to last problem; the improper integral $\int_0^1 \frac{\sin x}{x^p}\,dx$ converges.

Next, the improper integral $\int_1^\infty \frac{\sin x}{x^p}\,dx$ converges due to last problem.

Hence $\int_0^\infty \frac{\sin x}{x^p}\,dx = \int_0^1 \frac{\sin x}{x^p}\,dx + \int_1^\infty \frac{\sin x}{x^p}\,dx$ converges.

13. Show that the series $\sum_{n=2}^{\infty} \frac{1}{n(\ln n)^\alpha}$ converges for $\alpha > 1$ and diverges to ∞ for $\alpha \le 1$.
(a) $\alpha = 0$. The series is clearly divergent.
(a) $\alpha > 0$. The function $f(x) = \frac{1}{x(\ln x)^\alpha}$ is continuous, positive, and decreasing on $[2, \infty)$. By the integral test, it converges when $\int_2^\infty \frac{1}{x(\ln x)^\alpha}\,dx$ converges. Evaluating the integral, we have

$$\int_2^\infty \frac{1}{x(\ln x)^\alpha}\,dx = \int_{\ln 2}^\infty \frac{1}{t^\alpha}\,dt.$$

We conclude that the series converges for $\alpha > 1$ and diverges to ∞ for $\alpha \le 1$.
(b) $\alpha < 0$. Then $\frac{1}{n(\ln n)^\alpha} \ge \frac{1}{n}$ for $n > 3$. Comparing with the harmonic series, it follows that the series is divergent.

14. Does the series $\sum_{n=1}^{\infty} \frac{4^n (n!)^2}{(2n)!}$ converge?
The tests either give no information or are difficult to apply. However,

$$\frac{a_{n+1}}{a_n} = \frac{4^{n+1}((n+1)!)^2}{(2(n+1))!}\,\frac{(2n)!}{4^n (n!)^2} = \frac{2(n+1)}{2n+1} > 1.$$

Since $a_1 = 2$, we see that each $a_n > 2$. That is, $\lim a_n \ge 2 \ne 0$. Therefore, the series diverges. Since it is a series of positive terms, it diverges to ∞.

15. Does the series $1 - \frac{1}{4} - \frac{1}{16} + \frac{1}{9} + \frac{1}{25} + \frac{1}{49} - \cdots$ converge?
Here, the series has been made up from the terms $1/n^2$ by taking first one term, next two negative terms of squares of next even numbers, then three positive terms which are squares of next three odd numbers, and so on. This is a rearrangement of the series

$$1 - \frac{1}{2^2} + \frac{1}{3^2} - \frac{1}{4^2} + \frac{1}{5^2} - \frac{1}{6^2} + \cdots$$

which is absolutely convergent (since $\sum(1/n^2)$ is convergent). Therefore, the given series is convergent and its sum is the same as that of the alternating series $\sum(-1)^{n+1}(1/n^2)$.

16. Let (a_n) be a sequence of positive terms. Show that if $\sum_{n=1}^{\infty} \frac{a_n}{1+a_n}$ converges, then $\sum_{n=1}^{\infty} a_n$ converges.
Since $\displaystyle\sum_{n=1}^{\infty} \frac{a_n}{1+a_n}$ converges, $\dfrac{a_n}{1+a_n} \to 0$. Then $\dfrac{1}{1+a_n} = 1 - \dfrac{a_n}{1+a_n} \to 1$.
Now, $\dfrac{a_n}{1+a_n} \Big/ a_n = \dfrac{1}{1+a_n} \to 1$.
Since $\displaystyle\sum \frac{a_n}{1+a_n}$ converges, so does $\sum a_n$ by limit comparison test.

17. Let (a_n) be a sequence of positive non-increasing terms. Show that if $\sum_{n=1}^{\infty} a_n$ converges, then the sequence (na_n) converges to 0.
For any $m \in \mathbb{N}$,

$$2ma_{2m} \le 2(a_{m+1}+\cdots+a_{2m}), \quad (2m+1)a_{2m+1} \le 2(a_{m+1}+\cdots+a_{2m+1})-a_{2m+1}.$$

Since the series $\sum a_n$ converges, a_n converges to 0. Using Cauchy criterion, as $m \to \infty$,

$$2(a_{m+1}+\cdots+a_{2m}) \to 0, \quad 2(a_{m+1}+\cdots+a_{2m+1}) - a_{2m+1} \to 0.$$

Therefore, both $2ma_{2m} \to 0$ and $(2m+1)a_{2m+1} \to 0$. That is, $na_n \to 0$.

Series Representation of Functions §2.10

1. Determine the interval of convergence for each of the following power series:
(a) $\displaystyle\sum_{n=1}^{\infty} \frac{x^n}{n}$ (b) $\displaystyle\sum_{n=1}^{\infty} \frac{x^n}{n^2}$ (c) $\displaystyle\sum_{n=0}^{\infty} (-1)^n \frac{x^{n+1}}{n+1}$
(a) Its radius of convergence is $\displaystyle\lim_{n\to\infty} \frac{|a_n|}{|a_{n+1}|} = \lim_{n\to\infty} \frac{n+1}{n} = 1$.
The power series is around $x = 0$, i.e., it is in the form $\sum a_n(x-a)^n$, where $a = 0$. Thus, the power series converges at every point in the interval $(-1, 1)$.
To check at the end points:
For $x = -1$, the series $-1 + \dfrac{1}{2} - \dfrac{1}{3} + \cdots$ converges.
For $x = 1$, the series is $1 + \dfrac{1}{2} + \dfrac{1}{3} + \cdots$ diverges.
Therefore, its interval of convergence is $(-1, 1) \cup \{-1\} = [-1, 1)$.
(b) Its radius of convergence is $\displaystyle\lim_{n\to\infty} \frac{1/n^2}{1/(n+1)^2} = 1$.
At $x = \pm 1$, the series $\sum(1/n^2)$ converges.

Hence the interval of convergence is $[-1, 1]$.

(c) Here, we consider the series in the form $x \sum_{n=0}^{\infty} (-1)^n \dfrac{x^n}{n+1}$.

For the series $\sum_{n=0}^{\infty} (-1)^n \dfrac{x^n}{n+1}$, $\quad a_n = (-1)^n/(n+1)$.

Thus $\lim |a_n/a_{n+1}| = \lim (n+2)/(n+1) = 1$. Hence $R = 1$. That is, the series is convergent for all $x \in (-1, 1)$.

We know that the series converges at $x = 1$ and diverges at $x = -1$.

Therefore, the interval of convergence of the original power series is $(-1, 1]$.

2. Determine the interval of convergence of the series $\dfrac{2x}{1} - \dfrac{(2x)^2}{2} + \dfrac{(2x)^3}{3} - \cdots$.

Using Ratio test, we see that

$$\lim_{n\to\infty} \left| \frac{(2x)^{n+1}}{n+1} \, \frac{n}{(2x)^n} \right| = \lim_{n\to\infty} \left| \frac{n}{n+1} \right| |2x| = |2x|.$$

Thus the series converges for $|2x| < 1$, i.e., for $|x| < 1/2$.

Also, we find that when $x = 1/2$, the series converges and when $x = -1/2$, the series diverges.

Hence the interval of convergence of the series is $(-1/2, 1/2]$.

3. Determine power series expansion of the functions

(a) $\ln(1 + x)$ (b) $\dfrac{\ln(1+x)}{1-x}$

(a) For $-1 < x < 1$, $\dfrac{1}{1+x} = 1 - x + x^2 - x^3 + x^4 - \cdots$.

Integrating term by term and evaluating at $x = 0$, we obtain

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \cdots \quad \text{for } -1 < x < 1.$$

(b) Using the results in (a) and the geometric series for $1/(1 - x)$, we have

$$\frac{\ln(1+x)}{1-x} = \sum_{n=1}^{\infty} \frac{(-1)^{n-1} x^n}{n} \cdot \sum_{n=0}^{\infty} x^n \quad \text{for } -1 < x < 1$$

For obtaining the product of the two power series, we need to write the first in the form $\sum a_n x^n$. (Notice that for the second series, each $b_n = 1$.) Here, the first series is

$$\ln(1+x) = \sum_{n=0}^{\infty} a_n x^n, \quad \text{where } a_0 = 0 \text{ and } a_n = \frac{(-1)^{n-1}}{n} \text{ for } n \ge 1.$$

Thus the product above is $\dfrac{\ln(1+x)}{1-x} = \sum_{n=0}^{\infty} c_n x^n$, where

$$c_n = a_0 b_n + a_1 b_{n-1} + \cdots + a_n b_0 = a_0 + a_1 + \cdots a_n = 1 - \frac{1}{2} + \frac{1}{3} - \cdots + \frac{(-1)^{n-1}}{n}.$$

4. The function $\dfrac{1}{1-x}$ has interval of convergence $(-1, 1)$. However, prove that it has power series representation around any $c \neq 1$.

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

This power series converges for all x with $|x - c| < |1 - c|$, i.e., for $x \in (c - |1 - c|, c + |1 - c|)$.
We also see that the function $\dfrac{1}{1-x}$ is well defined for each $x \neq 1$.

5. Find the sum of the alternating harmonic series $\sum_{n=0}^{\infty} \dfrac{(-1)^n}{n}$.

Consider the power series representation of $\dfrac{1}{1+x}$. Integrating term by term.

$$\frac{1}{1+x} = \sum_{n=0}^{\infty} (-1)^n x^n \Rightarrow \ln(1+x) = \sum_{n=0}^{\infty} (-1)^n \frac{x^{n+1}}{n+1}.$$

Notice that the interval of convergence of the first power series is $(-1, 1)$. But the interval of convergence of the second power series is $(-1, 1]$. Thus, evaluating the second series at $x = 1$, we have

$$\ln 2 = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \cdots$$

6. Give an approximation scheme for $\displaystyle\int_0^a \frac{\sin x}{x}\, dx$ where $a > 0$.
Using the Maclaurin series for $\sin x$, we have

$$\frac{\sin x}{x} = 1 - \frac{x^2}{3!} + \frac{x^4}{5!} - \frac{x^6}{7} + \cdots$$

Integrating term by term, we get

$$\int_0^a \frac{\sin x}{x}\, dx = a - \frac{a^3}{3! \cdot 3} + \frac{a^5}{5! \cdot 5} - \frac{a^7}{7! \cdot 7} + \cdots$$

Approximations to the integral may be obtained by truncating the series suitably.

7. Show that $1+\frac{1}{2}\frac{1}{3}+\frac{1\cdot 3}{2\cdot 4}\cdot\frac{1}{5}+\frac{1\cdot 3\cdot 5}{2\cdot 4\cdot 6}\cdot\frac{1}{7}+\cdots=\frac{\pi}{2}$.

   In the binomial series $(1+x)^m = 1+mx+\frac{m(m-1)}{1\cdot 2}x^2+\frac{m(m-1)(m-2)}{1\cdot 2\cdot 3}x^3+\cdots$ for $|x|<1$, substitute $x=-t^2$ and $m=-1/2$ to obtain

$$\frac{1}{\sqrt{1-t^2}} = 1+\frac{1}{2}t^2+\frac{1\cdot 3}{2\cdot 4}t^4+\frac{1\cdot 3\cdot 5}{2\cdot 4\cdot 6}t^6+\cdots$$

   Integrating this power series from 0 to x for any $x\in(-1,1)$, we have

$$\sin^{-1}x=\int_0^x\frac{dt}{\sqrt{1-t^2}} = x+\frac{1}{2}\cdot\frac{1}{3}x^3+\frac{1\cdot 3}{2\cdot 4}\cdot\frac{1}{5}x^5+\frac{1\cdot 3\cdot 5}{2\cdot 4\cdot 6}\cdot\frac{1}{7}x^7+\cdots$$

   This series also converges for $x=1$. It may be seen as follows:

   Here, leaving the first term, $u_n=\frac{1\cdot 3\cdot 5\cdots(2n-1)}{2\cdot 4\cdots 6\cdots(2n)}\cdot\frac{1}{2n+1}$, $\quad u_{n+1}=u_n\cdot\frac{2n+1}{2n+2}\cdot\frac{2n+1}{2n+3}$.

   It follows that for $n\ge 2$, $u_{n+1}<4\big(nu_n-(n+1)u_{n+1}\big)$.

   Then $s_m=\sum_{n=2}^{m}u_n=\sum_{n=1}^{m-1}u_{n+1}<4\Big(\sum_{n=1}^{m-1}(nu_n-(n+1)u_{n+1}\Big)=4(u_1-mu_m)<4u_1$.

   Hence s_m is an increasing sequence of positive terms having an upper bound as $4u_1$.

   That is, the series $1+\sum_{n=1}^{\infty}u_n$ converges.

   Therefore, for $x=1$, the series converges to $\sin^{-1}1=\frac{\pi}{2}$.

8. Find the Fourier series of $f(x)$ given by: $f(x)=0$ for $-\pi\le x<0$; and $f(x)=1$ for $0\le x\le\pi$. Say also how the Fourier series represents $f(x)$. Hence give a series expansion of $\pi/4$.

$$a_n=\frac{1}{\pi}\int_0^{\pi}\cos nx\,dx=\begin{cases}1 & \text{for } n=0\\ 0 & \text{for } n\ne 0\end{cases}$$

$$b_n=\frac{1}{\pi}\int_0^{\pi}\sin nx\,dx=\frac{1}{\pi}\left[\frac{1-\cos n\pi}{n}\right]=\frac{1-(-1)^n}{n\pi}=\begin{cases}\frac{2}{n\pi} & \text{for } n \text{ odd}\\ 0 & \text{for } n \text{ even}\end{cases}$$

   Hence the Fourier series for $f(x)$ is $\frac{1}{2}+\frac{2}{\pi}\sum_{n=0}^{\infty}\frac{\sin(2n+1)x}{2n+1}$.

   By the convergence theorem for Fourier series, we know that this Fourier series converges to $f(x)$ for any $x\ne 0$. At $x=0$, the Fourier series converges to $1/2$.

Taking $x = \pi/2$, we have

$$1 = \frac{1}{2} + \frac{2}{\pi}\sum_{n=0}^{\infty}\frac{\sin(n+1/2)\pi}{2n+1} = \frac{1}{2} + \frac{2}{\pi}\sum_{n=0}^{\infty}\frac{(-1)^n}{2n+1}.$$

Therefore, $\dfrac{\pi}{4} = \displaystyle\sum_{n=0}^{\infty}\frac{(-1)^n}{2n+1}$.

9. Considering the fourier series for $|x|$, deduce that $\dfrac{\pi^2}{8} = \displaystyle\sum_{n=0}^{\infty}\frac{1}{(2n+1)^2}$.

Cinsider $f(x) = |x|$ in the interval $[-\pi, \pi]$; extended to $\mathbb{R}$ with period 2π. Now, it is an even function. Thus each b_n is 0. Next, $a_0 = (2/\pi)\int_0^\pi x\,dx = \pi$. And for $n > 0$,

$$a_n = \frac{2}{\pi}\int_0^\pi x\cos nx\,dx = \frac{2}{\pi}\left[\frac{(-1)^n - 1}{n^2}\right].$$

That is, $a_{2n} = 0,\ a_{2n+1} = \dfrac{-4}{\pi(2n+1)^2}$ for $n = 1, 2, 3 \ldots$

By the convergence theorem for Fourier series, we have

$$|x| = \frac{\pi}{2} - \frac{4}{\pi}\sum_{n=0}^{\infty}\frac{\cos(2n+1)x}{(2n+1)^2} \quad \text{for } x \in [-\pi, \pi].$$

Taking $x = 0$, we have $\dfrac{\pi^2}{8} = \displaystyle\sum_{n=0}^{\infty}\frac{1}{(2n+1)^2}$.

10. Considering the fourier series for x, deduce that $\dfrac{\pi}{4} = \displaystyle\sum_{n=0}^{\infty}\frac{(-1)^n}{2n+1}$.

Consider $f(x) = x$ for $x \in [-\pi, \pi]$. It is an odd function. Hence in its Fourier series, each $a_n = 0$. For $n \geq 1$,

$$b_n = \frac{2}{\pi}\int_0^\pi x\sin nx\,dx = \frac{2}{\pi}\left[-x\frac{\cos nx}{n}\right]_0^\pi + \frac{2}{\pi}\int_0^\pi \frac{\cos nx}{n}\,dx = \frac{2(-1)^{n+1}}{n}.$$

Thus the Fourier series for $f(x) = x$ in $[-\pi, \pi]$ is $2\displaystyle\sum_{n=1}^{\infty}\frac{(-1)^{n+1}}{n}\sin nx$.

Taking $x = \pi/2$, we have $\dfrac{\pi}{4} = \displaystyle\sum_{n=1}^{\infty}\frac{(-1)^{n+1}}{n}\sin\frac{n\pi}{2} = \sum_{n=0}^{\infty}\frac{(-1)^n}{2n+1}$.

11. Considering the fourier series for $f(x)$ given by: $f(x) = -1$, for $-\pi \leq x < 0$ and $f(x) = 1$ for $0 \leq x \leq \pi$. Deduce that $\dfrac{\pi}{4} = \displaystyle\sum_{n=0}^{\infty}\frac{(-1)^n}{2n+1}$.

Here, $f(x)$ is an odd function. Thus in its Fourier series, each a_n is 0.
For $n \geq 1$,

$$b_n = \frac{2}{\pi}\int_0^{\pi} \sin nx\, dx = \frac{2}{n\pi}(1 - \cos n\pi) = \frac{2}{n\pi}(1 - (-1)^n).$$

Due to the convergence theorem, $f(x) = \dfrac{4}{\pi}\displaystyle\sum_{n=0}^{\infty} \frac{\sin(2n+1)x}{2n+1}$ for $x \neq 0$.
Taking $x = \pi/2$, we obtain the desired expression for $\pi/4$.

12. Considering $f(x) = x^2$, show that for each $x \in [0, \pi]$,

$$\frac{\pi^2}{6} + 2\sum_{n=1}^{\infty} \frac{(-1)^n \cos nx}{n^2} = \sum_{n=1}^{\infty} \frac{n^2\pi^2(-1)^{n+1} + 2(-1)^n - 2}{n^3\pi} \sin nx.$$

We determine sine and cosine series expansions of $f(x) = x^2$ for $0 \leq x \leq \pi$.
The odd and even expansions of $f(x)$ are

$$f_{odd}(x) = \begin{cases} -x^2 & \text{for } -\pi \leq x < 0 \\ x^2 & \text{for } \ 0 \leq x < \pi, \end{cases} \qquad f_{even}(x) = x^2 \quad \text{for } -\pi \leq x \leq \pi.$$

We see that, as earlier, $f_{even}(x)$ has the Fourier expansion $\dfrac{\pi^2}{3} + 4\displaystyle\sum_{n=1}^{\infty} \frac{(-1)^n \cos nx}{n^2}$ for $x \in [0, \pi]$. Due to the convergence theorem of Fourier series, this series sums to x^2 in $[0, \pi]$.

For the sine series expansion, we determine the Fourier series of $f_{odd}(x)$.
Here, each a_n is 0. And for $n \geq 1$,

$$b_n = \frac{2}{\pi}\int_0^{\pi} x^2 \sin nx\, dx = 2\pi\frac{(-1)^{n+1}}{n} + \frac{4}{\pi}\left[\frac{(-1)^n - 1}{n^3}\right].$$

Due to the convergence theorem of Fourier series, $x^2 = \sum_{n=1}^{\infty} b_n \sin nx$ for $x \in [0, \pi]$.
Equating both the sine and the cosine series for $f(x) = x^2$ in $[0, \pi]$, we obtain the required result.

13. Represent the function $f(x) = 1 - |x|$ for $-1 \leq x \leq 1$ as a cosine series.

It is an even function. Thus its Fourier series is $\dfrac{a_0}{2} + \displaystyle\sum_{n=1}^{\infty} a_n \cos n\pi x$, where

$a_0 = 2\int_0^1 (1 - x)\, dx = 1.$
For $n \geq 1$,

$$a_n = \int_{-1}^{1} (1 - |x|) \cos n\pi x \, dx = 2 \int_0^1 (1 - x) \cos n\pi x \, dx = \begin{cases} 0 & \text{for } n \text{ even} \\ 4/(n^2\pi^2) & \text{for } n \text{ odd.} \end{cases}$$

Thereofore, $1 - |x| = \dfrac{1}{2} + \dfrac{4}{\pi^2} \displaystyle\sum_{n=0}^{\infty} \dfrac{\cos(2n+1)\pi x}{(2n+1)^2}$ for $-1 \le x \le 1$.

Basic Matrix Operations §3.6

1. Show that given any $n \in \mathbb{N}$ there exist matrices $A, B \in \mathbb{R}^{n\times n}$ such that $AB \neq BA$.

   Let $A = [e_2 \; e_1 \; e_3 \; e_4 \; \cdots \; e_n]$ and $B = [v \; u \; u \; \cdots \; u]$, where $e_1, \ldots, e_n$ are standard basis vectors of $\mathbb{R}^{n\times 1}$ and $u = (1, 1, 1, \ldots, 1)^T$, $v = (0, 0, 0, \ldots, 0)^T$.

2. Let $A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}$. Compute A^n.

   We show that $A^n = \begin{bmatrix} 1 & n & n(n-1) \\ 0 & 1 & 2n \\ 0 & 0 & 1 \end{bmatrix}$ for $n \in \mathbb{N}$ by induction.

   The basis case $n = 1$ is obvious. Suppose A^n is as given. Now,

$$A^{n+1} = AA^n = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & n & n(n-1) \\ 0 & 1 & 2n \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & n+1 & (n+1)n \\ 0 & 1 & 2(n+1) \\ 0 & 0 & 1 \end{bmatrix}.$$

   Notice that taking $n = 0$ in the matrix A^n, we see that $A^0 = I$.

3. Let $A \in \mathbb{F}^{m\times n}$; $B \in \mathbb{F}^{n\times k}$. Let $A_1, \ldots, A_m$ be the rows of A and let $B_1, \ldots, B_k$ be the columns of B. Show that

   (a) $A_1 B, \ldots, A_M B$ are the rows of AB. (b) $AB_1, \ldots, AB_k$ are the columns of AB.

   (a) The jth entry in $A_i B$ is $A_i \cdot B_j$, which is the (i, j)th entry in AB.

   (b) The ith entry in AB_j is $A_i \cdot B_j$, which is the (i, j)th entry in AB.

4. Let $A \in \mathbb{F}^{n\times n}$; I be the identity matrix of order n.

   Find the inverse of the $2n \times 2n$ matrix $\begin{bmatrix} I & A \\ 0 & I \end{bmatrix}$.

   Check by muiltiplying: $\begin{bmatrix} I & A \\ 0 & I \end{bmatrix}^{-1} = \begin{bmatrix} I & -A \\ 0 & I \end{bmatrix}$.

5. If A is a hermitian (symmetric) invertible matrix, then show that A^{-1} is hermitian (symmetric).
   $A^* = A$. Then $(A^{-1})^* = (A^*)^{-1} = A^{-1}$. So, A^{-1} is hermitian.
   Similarly, for symmetric take transpose instead of conjugate transpose.
6. If A is a lower (upper) triangular invertible matrix, then A^{-1} is lower (upper) triangular.
   Suppose $A^{-1} = [u_1 \ \cdots u_n]$. Then $AA^{-1} = I$ implies $Au_i = e_i$. Now, A is lower triangular with nonzero entries on the diagonal. Writing $A = [a_{ij}]$, and $u_k = [y_1, \ldots, y_n]^t$, we have
   $a_{11}y_1 = 0,\ a_{12}y_1 + a_{22}y_2 = 0, \ldots$. This gives $y_1 = 0,\ y_2 = 0,\ \ldots y_{i-1} = 0$. Thus A^{-1} is lower triangular.
   Aliter: Suppose A is a lower triangular matrix of order n. Let D be the diagonal matrix whose diagonal entries are exactly the diagonal entries of A in the correct order. Since A is invertible, D is also invertible. Then write $A = D(I + N)$. Here, N is a lower triangular matrix with all diagonal entries as 0. Then $N^n = 0$. Verify that $A^{-1} = (I - N + N^2 - \cdots + (-1)^{n-1}N^{n-1})D^{-1}$. Also, verify that this is a lower triangular matrix.
7. Show that each orthogonal 2×2 matrix is either a reflection or a rotation.
   If $A = [a_{ij}]$ is an orthogonal matrix of order 2, then $A^TA = I$ implies
   $$a_{11}^2 + a_{21}^2 = 1 = a_{12}^2 + a_{22}^2,\ a_{11}a_{12} + a_{21}a_{22} = 0.$$
   Thus, there exist α, β such that
   $a_{11} = \cos\alpha,\ a_{21} = \sin\alpha,\ a_{12} = \cos\beta,\ a_{22} = \sin\beta$ and $\cos(\alpha - \beta) = 0$.
   It then follows that A is in one of the following forms:
   $$O_1 := \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}, \quad O_2 := \begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix}.$$
   Let $\overrightarrow{(a, b)}$ be the vector in the plane that starts at the origin and ends at the point (a, b). Writing the point (a, b) as a column vector $[a\ b]^T$, we see that the matrix product $O_1[a\ b]^T$ is the end-point of the vector obtained by rotating the vector $\overrightarrow{(a, b)}$ by an angle θ. Similarly, $O_2[a\ b]^T$ gives a point obtained by reflecting (a, b) along a straight line that makes an angle $\theta/2$ with the x-axis. Thus, O_1 is said to be a *rotation by an angle* θ and O_2 is called a *reflection by an angle* $\theta/2$ along the x-axis.
8. Let $u, v, w \in \mathbb{F}^{n\times 1}$. Show that $\{u + v, v + w, w + u\}$ is linearly independent iff $\{u, v, w\}$ is linearly independent.
   $\alpha u + \beta v + \gamma w = 0 \Rightarrow \frac{\alpha+\beta-\gamma}{2}(u + v) + \frac{\beta+\gamma-\alpha}{2}(v + w) + \frac{\alpha+\gamma-\beta}{2}(w + u) = 0$

$\Rightarrow \alpha + \beta - \gamma = 0,\ \beta + \gamma - \alpha = 0,\ \alpha + \gamma - \beta = 0 \Rightarrow \alpha = \beta = \gamma = 0.$
So, $\{u, v, w\}$ is linearly independent.
Conversely, $a(u+v)+b(v+w)+c(w+u) = 0 \Rightarrow (a+c)u+(a+b)v+(b+c)w = 0$
$\Rightarrow a + c = 0,\ a + b = 0,\ b + c = 0 \Rightarrow a = 0,\ b = 0, c = 0.$
Hence $\{u + v, v + w, w + u\}$ is linearly independent.

9. Find linearly independent vectors from $U = \{(a, b, c) : 2a + 3b - 4c = 0\}$ so that the set of linear combinations of which is exactly U.
   $U = \{(a, b, c) : 2a + 3b - 4c = 0\} = \{(a, b, \frac{2a+3b}{4}) : a, b \in \mathbb{R}\}.$
   The vectors $(1, 0, 1/2)$ and $(0, 1, 3/4)$ are in U.
   $\left(a, b, \frac{2a+3b}{4}\right) = a(1, 0, 1/2) + b(0, 1, 3/4)$. So, these two vectors span U.
   Now, $a(1, 0, 1/2) + b(0, 1, 3/4) = (0, 0, 0) \Rightarrow a = 0, b = 0,\ \frac{2a+3b}{4} = 0.$
   So, the vectors are linearly independent.

10. Determine linearly independent vectors so that the set of linear combinations of which is $U = \{(a, b, c, d, e) \in \mathbb{R}^5 : a = c = e, b + d = 0\}$.
    $U = \{(a, b, a, -b, a) : a, b \in \mathbb{R}\} = \{a(1, 0, 1, 0, 1) + b(0, 1, 0, -1, 0) : a, b \in \mathbb{R}.$
    If $a(1, 0, 1, 0, 1) + b(0, 1, 0, -1, 0) = 0$, then $a = b = 0$.
    So, the vectors are $(1, 0, 1, 0, 1)$ and $(0, 1, 0, -1, 0)$.

Row Reduced Echelon Form §4.7

1. Convert the following matrices into RREF and determine their ranks.
   (a) $\begin{bmatrix} 5 & 2 & -3 & 1 & 7 \\ 1 & -3 & 2 & -2 & 11 \\ 3 & 8 & -7 & 5 & 8 \end{bmatrix}$ (b) $\begin{bmatrix} 5 & 2 & -3 & 1 & 30 \\ 1 & -3 & 2 & -2 & 11 \\ 3 & 8 & -7 & 5 & 8 \end{bmatrix}$

   (a) RREF of the matrix is $\begin{bmatrix} 1 & 0 & -5/17 & -1/17 & 0 \\ 0 & 1 & -13/17 & 11/17 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$. So its ranks is 3.
   (b) RREF of the matrix is $\begin{bmatrix} 1 & 0 & -5/17 & -1/17 & 112/17 \\ 0 & 1 & -13/17 & 11/17 & -25/17 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$. So its rank is 2.

2. Determine linear independence of $\{(1, 2, 2, 1),\ (1, 3, 2, 1),\ (4, 1, 2, 2),\ (5, 2, 4, 3)\}$ in $\mathbb{C}^{1\times 4}$.

The RREF of the matrix whose rows are the given vectors is $\begin{bmatrix} 1 & 0 & 0 & 1/3 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1/3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$.

Since a zero row has appeared in the RREF, the vectors are linearly dependent. Moreover, there had been no row exchanges in this reduction, and the fourth vector has been reduced to the zero row. Thus the fourth vector is a linear combination of the three previous ones.

Alternative: Take the matrix, where the given vectors are taken as column vectors. Reduce it to RREF. You obtain: $\begin{bmatrix} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$. Thus the fourth vector is a linear combination of the earlier ones, whose coefficients are $2, -1, 1$, respectively. You can verify that $(5, 2, 4, 3) = 2(1, 2, 2, 1) - 1(1, 3, 2, 1) + 1(4, 1, 2, 2)$. So, the set is linearly dependent.

3. Compute A^{-1} using RREF and also using determinant, where $A = \begin{bmatrix} 4 & -7 & -5 \\ -2 & 4 & 3 \\ 3 & -5 & -4 \end{bmatrix}$.

   Compute and see that $A^{-1} = \begin{bmatrix} 1 & 3 & 1 \\ -1 & 1 & 2 \\ 2 & 1 & -2 \end{bmatrix}$.

4. Solve the following system by Gauss-Jordan elimination:

$$\begin{aligned} x_1 + x_2 + x_3 + x_4 - 3x_5 &= 6 \\ 2x_1 + 3x_2 + x_3 + 4x_4 - 9x_5 &= 17 \\ x_1 + x_2 + x_3 + 2x_4 - 5x_5 &= 8 \\ 2x_1 + 2x_2 + 2x_3 + 3x_4 - 8x_5 &= 14 \end{aligned}$$

   We reduce the augmented matrix to its RREF:

$$\left[\begin{array}{ccccc|c} \boxed{1} & 1 & 1 & 1 & -3 & 6 \\ 2 & 3 & 1 & 4 & -9 & 17 \\ 1 & 1 & 1 & 2 & -5 & 8 \\ 2 & 2 & 2 & 3 & -8 & 14 \end{array}\right] \to \left[\begin{array}{ccccc|c} \boxed{1} & 1 & 1 & 1 & -3 & 6 \\ 0 & 1 & -1 & 2 & -3 & 5 \\ 0 & 0 & 0 & 1 & -2 & 2 \\ 0 & 0 & 0 & 1 & -2 & 2 \end{array}\right]$$

$$\to \left[\begin{array}{ccccc|c} \boxed{1} & 0 & 2 & -1 & 0 & 1 \\ 0 & \boxed{1} & -1 & 2 & -3 & 5 \\ 0 & 0 & 0 & 1 & -2 & 2 \\ 0 & 0 & 0 & 1 & -2 & 2 \end{array}\right] \to \left[\begin{array}{ccccc|c} \boxed{1} & 0 & 2 & 0 & -2 & 3 \\ 0 & \boxed{1} & -1 & 0 & 1 & 1 \\ 0 & 0 & 0 & \boxed{1} & -2 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right].$$

Find out the row operations used in each step. Since no pivot is on the b portion, the system is consistent. To solve this system, we consider only the pivot rows, ignoring the bottom zero rows. The basis variables are x_1, x_2, x_4 and the free variables are x_3, x_5. Write $x_3 = \alpha$ and $x_5 = \beta$. Then

$$x_1 = 3 - 2\alpha + 2\beta, \quad x_2 = 1 + \alpha - \beta, \quad x_3 = \alpha, \quad x_4 = 2 + 2\beta, \quad x_5 = \beta.$$

5. Check if the system is consistent. If so, determine the solution set.

   (a) $x_1 - x_2 + 2x_3 - 3x_4 = 7,\ 4x_1 + 3x_3 + x_4 = 9,\ 2x_1 - 5x_2 + x_3 = -2,$ $3x_1 - 2x_2 - 2x_3 + 10x_4 = -12.$

   (b) $x_1 - x_2 + 2x_3 - 3x_4 = 7, 4x_1 + 3x_3 + x_4 = 9, 2x_1 - 5x_2 + x_3 = -2,$ $3x_1 - 2x_2 - 2x_3 + 10x_4 = -14.$

   (a) RREF of $[A|b]$ is $\begin{bmatrix} 1 & 0 & 0 & 2 & 0 \\ 0 & 1 & 0 & 1/3 & 0 \\ 0 & 0 & 1 & -7/3 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$. Thus inconsistent.

   (b) RREF of $[A|b]$ is $\begin{bmatrix} 1 & 0 & 0 & 2 & -10/9 \\ 0 & 1 & 0 & 1/3 & 23/27 \\ 0 & 0 & 1 & -7/3 & 121/27 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$.

   Thus solution is $x_1 = -\frac{10}{9} + 2\alpha,\ x_2 = \frac{23}{27} + \frac{\alpha}{3},\ x_3 = \frac{121}{27} + \frac{7\alpha}{3},\ x_4 = \alpha.$

6. Using Gauss-Jordan elimination determine the values of $k \in \mathbb{R}$ so that the system of linear equations

$$x + y - z = 1,\ 2x + 3y + kz = 3,\ x + ky + 3z = 2$$

   has (a) no solution, (b) infinitely many solutions, (c) exactly one solution.

   Gauss-Jordan elimination on $[A|b]$ yields $\begin{bmatrix} \boxed{1} & 0 & -k-3 & 0 \\ 0 & \boxed{1} & k+2 & 1 \\ 0 & 0 & (k+3)(2-k) & 2-k \end{bmatrix}$.

   (a) The system has no solution when $(k+3)(2-k) = 0$ but $2 - k \neq 0$, that is, when $k = -3$.

   (b) It has infinitely many solutions when $(k+3)(2-k) = 0 = 2 - k$, that is, when $k = 2$.

   (c) It has exactly one solution when $(k+3)(2-k) \neq 0$, that is, when $k \neq -3, k \neq 2$.

7. Let A be an $n \times n$ matrix with integer entries and $\det(A^2) = 1$. Show that all entries of A^{-1} are also integers.

$\det(A^2) = [\det(A)]^2 = 1 \Rightarrow \det(A) = \pm 1$. So, A is invertible. Since A has integer entries, $\text{adj}(A)$ has also integer entries. Now, $A^{-1} = [\det(A)]^{-1}\text{adj}(A)$ has inteeger entries.

8. Let $A \in \mathbb{F}^{m\times n}$ have columns $A_1, \ldots, A_n$. Let $b \in \mathbb{F}^m$. Show the following:
   (a) The equation $Ax = 0$ has a non-zero solution iff $A_1, \ldots, A_n$ are linearly dependent.
   (b) The equation $Ax = b$ has at least one solution iff $b \in \text{span}\{A_1, \ldots, A_n\}$.
   (c) Let u be a solution of $Ax = b$. Then, u is the only solution of $Ax = b$ iff $A_1, \ldots, A_n$ are linearly independent.
   (d) The equation $Ax = b$ has a unique solution iff $\text{rank}A = \text{rank}[A|b] =$ number of unknowns.

   (a) We have scalars $\alpha_1, \ldots, \alpha_n$ not all 0 such that $\sum \alpha_i A_i = 0$. But each $A_i = Ae_i$. So, $A(\sum \alpha_i e_i) = 0$. Here, take $x = \sum \alpha_i e_i$. See that $x \neq 0$.

   (b) If b is a linear combination of the columns of A, then that linear combination provides a solution. Conversely, a solution provides a linear combination of columns of A which is equal to b.

   (c) We have $Au = b$. Assume that $A_1, \ldots, A_n$ are linearly independent. If $Av = b$, then $A(u-v) = 0$. Let $u - v = (\alpha_1, \ldots, \alpha_n)^T$. Then $A(u-v) = 0$ can be rewritten as $\alpha_1 A_1 + \cdots \alpha_n A_n = 0$. Since $A_1, \ldots, A_n$ are linearly independent, each α_i is 0. That is, $u - v = 0$. Conversely, if $A_1, \ldots, A_n$ are linearly dependent, then scalars $\beta_1, \ldots, \beta_n$ not all zero exist such that $\beta_1 A_1 + \cdots \beta_n A_n = 0$. That is, $Av = 0$ with $v = (\beta_1, \ldots, \beta_n)^T$. Then, u and $u + v$ are two solutions of $Ax = b$.

   (d) Let $A \in \mathbb{F}^{m\times n}$.
   If the system $Ax = b$ has a unique solution, then it is a consistent system and $\text{rank}(A) = n$.
   That is, $\text{rank}(A) = \text{rank}[A|b]$ and $\text{rank}(A) = n =$ number of unknowns.

9. Let $A \in \mathbb{F}^{m\times n}$ have rank r. Give reasons for the following:
   (a) $\text{rank}(A) \leq \min\{m, n\}$.
   (b) If $n > m$, then there exist $x, y \in \mathbb{F}^{n\times 1}$ such that $x \neq y$ and $Ax = Ay$.
   (c) If $n < m$, then there exists $y \in \mathbb{F}^{m\times 1}$ such that for no $x \in \mathbb{F}^{n\times 1}$, $Ax = y$.
   (d) If $n = m$, then the following statements are equivalent:
      i. $Au = Av$ implies $u = v$ for all $u, v \in \mathbb{F}^{n\times 1}$.
      ii. Corresponding to each $y \in \mathbb{F}^{n\times 1}$, there exists $x \in \mathbb{F}^{m\times 1}$ such that $y = Ax$.

   (a) $\text{rank}(A)$ is the number of pivots in the RREF. So, it is less than or equal to the number of rows, and also less than or equal to the number of columns.

(b) Suppose $n > m$. Then the RREF has at most m pivots. And, there are $n - m \geq 1$ number of non-pivotal columns. These non-pivotal columns are linear combinations of pivotal columns in A. So, there exist scalars $\alpha_1, \ldots, \alpha_n$ not all zero such that $\alpha_1 C_1 + \cdots \alpha_n C_n = 0$ where C_i is the ith column of A. Then $(\alpha_1, \ldots, \alpha_n)$ is a nonzero solution to $Ax = 0$. Now, $A0 = 0$ and $Au = 0$, where $u = (\alpha_1, \ldots, \alpha_n) \neq 0$.

(c) Suppose $n < m$. Let EA be the RREF of A. Consider the equation $Ax = E^{-1}e_{n+1}$. This has the same solutions as the system $EAx = e_{n+1}$. But $[EA|e_{n+1}]$ has a pivot in the right most column, which has no solution.

(d) Suppose $n = m$.

Assume (i). Then $Ax = 0$ has a unique solution. Then number of basic variables is n. So RREF of A is I. That is, A is invertible. Then $Ax = y$ has a solution for each y, namely, $x = A^{-1}y$. This proves (ii).

Conversely, assume (ii). That is, for each y, $Ax = y$ has a solution. In particular, $Ax = e_i$ has a solution for each i. Thus, A is invertible. Then $Ax = Ay$ implies $x = y$.

Matrix Eigenvalue Problem §5.7

1. Find the eigenvalues and the associated eigenvectors for the matrices given below.

   (a) $\begin{bmatrix} 3 & 0 \\ 8 & -1 \end{bmatrix}$ (b) $\begin{bmatrix} 3 & 2 \\ -1 & 0 \end{bmatrix}$ (c) $\begin{bmatrix} -2 & -1 \\ 5 & 2 \end{bmatrix}$ (d) $\begin{bmatrix} -2 & 0 & 3 \\ -2 & 3 & 0 \\ 0 & 0 & 5 \end{bmatrix}$.

   (d) Call the matrix A. Its characteristic polynomial is $-(2+t)(3-t)(5-t)$. So, the eigenvalues are $\lambda = -2,\ 3,\ 5$.

   For $\lambda = -2$, $A(a,b,c)^T = -2(a,b,c)^T \Rightarrow -2a + 3c = -2a,\ -2a + 3b = -2b,\ 5c = -2c$.

   One of the solutions for $(a,b,c)^T$ is $(5,2,0)^T$. It is an eigenvector for $\lambda = -2$.

   For $\lambda = 3$, $A(a,b,c)^T = 3(a,b,c)^T \Rightarrow -2a + 3c = 3a,\ -2a + 3b = 3b,\ 5c = 3c$.

   One of the solutions for $(a,b,c)^T$ is $(0,1,0)^T$. It is an eigenvector for $\lambda = 3$.

   For $\lambda = 5$, $A(a,b,c)^T = 5(a,b,c)^T \Rightarrow -2a + 3c = 5a,\ -2a + 3b = 5b,\ 5c = 5c$.

   One of the solutions for $(a,b,c)^T$ is $(3,-3,7)^T$. It is an eigenvector for $\lambda = 5$.

   Similarly, solve others.

2. Let A be an $n \times n$ matrix and α be a scalar such that each row (or each column) sums to α. Show that α is an eigenvalue of A.

If each row sums to α, then $A(1,1,\ldots,1)^T = \alpha(1,1,\ldots,1)^T$. Thus α is an eignevlaue with an eigenvector as $(1,1,\ldots,1)^T$.
If each column sums to α, then each row sums to α in A^T. Thus A^T has an eigenvalue as α. However, A^T and A have the same eigenvalues. Thus α is also an eigenvalue of A.

3. Let $A \in \mathbb{C}^{n\times n}$ be invertible. Show that $\lambda \in \mathbb{C}$ is an eigenvalue of A if and only if λ^{-1} is an eigenvalue of A^{-1}.

   Since A is invertible, its determinant is nonzero. As $\det(A)$ is the product of eigenvalues of A, no eigenvalue of A is 0.
   Also, for any nonzero λ, $Av = \lambda v$ iff $\lambda^{-1}A^{-1}Av = \lambda^{-1}A^{-1}\lambda v$ iff $\lambda^{-1}v = A^{-1}v$.
   This shows that λ is an eigenvalue of A iff λ^{-1} is an eigenvlaue of A^{-1}.

4. The vectors $u_1 = (1,2,2), u_2 = (-1,0,2), u_3 = (0,0,1)$ are linearly independent in $\mathbb{F}^3$. Apply Gram-Schmidt Orthogonalization.
   $v_1 = (1,2,2)$.
   $$v_2 = u_2 - \frac{\langle u_2, v_1\rangle}{\langle v_1, v_1\rangle} v_1 = (-1,0,2) - \frac{(-1,0,2)\cdot(1,2,2)}{(1,2,2,)\cdot(1,2,2)}(1,2,2) = \big(-4/3, -2/3, 4/3\big).$$
   $$v_3 = u_3 - \frac{\langle u_3, v_1\rangle}{\langle v_1, v_1\rangle} v_1 - \frac{\langle u_3, v_2\rangle}{\langle v_2, v_2\rangle} v_2$$
   $$= (0,0,1) - \frac{(0,0,1)\cdot(1,2,2)}{(1,2,2,)\cdot(1,2,2)}(1,2,2) - \frac{(0,0,1)\cdot(-4/3,-2/3,4/3)}{(-4/3,-2/3,4/3)\cdot(-4/3,-2/3,4/3)}(-4/3,-2/3,4/3)$$
   $$= \big(2/9, -2/9, 1/9\big).$$
   The set $\big\{(1,2,2),\ \big(-4/3,-2/3,4/3\big),\ \big(2/9,-2/9,1/9\big)\big\}$ is orthogonal.

5. Let $A \in \mathbb{R}^{3\times 3}$ have the first two columns as $(1/\sqrt{3}, 1/\sqrt{3}, 1/\sqrt{3})^T$ and $(1/\sqrt{2}, 0, -1/\sqrt{2})^T$. Determine the third column of A so that A is an orthogonal matrix.
   Notice that the first two columns of A have norm 1, and are orthogonal to each other. You can start with the third as $(0,0,1)^T$ and use Gram-Schmidt process. And then normalize the third vector.
   Aliter: Let the third column be $(a,b,c)^T$. Then the first two are orthogonal to the third implies $a+b+c=0,\ a-c=0$. This gives $(a,b,c)^T = (a,-2a,a)^T$. Now, the third column has norm 1 implies that $1 = a^2 + 4a^2 + a^2 = 6a^2 \Rightarrow a = \pm 1/\sqrt{6}$. Thus the third column of A is $\pm(1/\sqrt{6}, -2/\sqrt{6}, 1/\sqrt{6})^T$.

6. Show that eigenvectors corresponding to distinct eigenvalues of a unitary (or orthogonal) matrix are orthogonal to each other.

   Let α and β be distinct eigenvalues of a unitary matrix A with corresponding eigenvectors x and y. That is, we have: $A^*A = AA^* = I,\ Ax = \alpha x,\ Ay = \alpha y,$ $x \neq 0,\ y \neq 0$ and $\alpha \neq \beta$. We need to show that $x \perp y$. Now,
   $(Ax)^*(Ay) = (\alpha x)^*(\beta y) \Rightarrow x^*A^*Ay = \overline{\alpha}\beta x^*y \Rightarrow \big(\overline{\alpha}\beta - 1\big)x^*y = 0.$
   Since A is unitary, any eigenvalue of A has absolute value 1.

So, $|\alpha|^2 = 1 \Rightarrow \alpha\,\overline{\alpha} = 1 \Rightarrow \overline{\alpha} = 1/\alpha$.
Then $(\overline{\alpha}\beta - 1)x^*y = 0 \Rightarrow (\beta/\alpha - 1)x^*y = 0 \Rightarrow (\beta - \alpha)x^*y = 0$.
Since $\alpha \neq \beta$, we get $x^*y = 0$. That is, $x \perp y$.

7. Give an example of an $n \times n$ matrix that cannot be diagonalized.

   Take $A = [a_{ij}] \in \mathbb{C}^{n\times n}$ with $a_{12} = 1$ and all other entries as 0. Its eigenvalue is 0 with algebraic multiplicity as n. If A is diagonalizable, then A is similar to the zero matrix. But the only matrix similar to the zero matrix is the zero matrix!

8. Find the matrix $A \in \mathbb{R}^{3\times 3}$ that satisfies the given condition. Diagonalize it if possible.

   (a) $A(a, b, c)^T = (a + b + c,\ a + b - c,\ a - b + c)^T$ for all $a, b, c \in \mathbb{R}$.
   (b) $Ae_1 = 0, \quad Ae_2 = e_1, \quad Ae_3 = e_2$.
   (c) $Ae_1 = e_2, \quad Ae_2 = e_3, \quad Ae_3 = 0$.
   (d) $Ae_1 = e_3, \quad Ae_2 = e_2, \quad Ae_3 = e_1$.

   (a) $A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & -1 \\ 1 & -1 & 1 \end{bmatrix}$. Its characteristic polynomial is $-(t + 1)(t - 2)^2$.
   So, eigenvalues are -1 and 2. Solving $A(a, b, c)^T = \lambda(a, b, c)^T$ for $\lambda = -1, 2$, we have
   $\lambda = -1\ :\ a + b + c = -a,\ a + b - c = -b,\ a - b + c = -c \Rightarrow a = -c, b = c$.
   Thus a corresponding eigenvector is $(-1, 1, 1)^T$.
   $\lambda = 2\ :\ a + b + c = 2a,\ a + b - c = 2b,\ a - b + c = 2c \Rightarrow a = b + c$.
   Thus two linearly independent corresponding eigenvectors are $(1, 1, 0)^T$ and $(1, 0, 1)^T$.
   Take the matrix $P = \begin{bmatrix} -1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$. Then verify that $P^{-1}AP = diag(-1, 2, 2)$.

   (b) The eigenvalue 0 has algebraic multiplicity 3. If it is diagonalizable, then it is similar to 0. But the only matrix similar to 0, is 0. So, A is not diagonalizable.

   (c) Similar to (b).

   (d) Proceed as in (a) to get $P = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ -1 & 1 & 0 \end{bmatrix}$ and verify $P^{-1}AP = diag(-1, 1, 1)$.

9. Which of the following matrices is/are diagonalizable? If one is diagonalizable, then diagonalize it.

$$\text{(a)} \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{bmatrix} \quad \text{(b)} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{(c)} \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \quad \text{(d)} \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}.$$

(a) It is a real symmetric matrix; so diagonalizable. Its eigenvalues are $-2, -1, 2$. Also since the 3×3 matrix has three distinct eigenvalues, it is diagonalizable.
Proceed like 6(a).

(b) 1 is an eigenvalue with algebraic multipliity 3. If it is diagonalizable, then it is similar to I. But the only matrix similar to I is I. Hence, it is not diagonalizable.

(c) Its eigenvalues are 2, $(1 \pm \sqrt{3}\, i)/2$. Since three distinct eigenvalues; it is diagonalizable. Here, P will be a complex matrix. Proceed as in 6(a).

(d) $(1, 0, -1)^T$ and $(1, -1, 0)^T$ are two linearly independent eigenvectors associated with the eigenvalue -1.
$(1, 1, 1)^T$ is an eigenvector for the eigenvalue 2.
Hence taking $P = \begin{bmatrix} 1 & 1 & 1 \\ 0 & -1 & 1 \\ -1 & 0 & 1 \end{bmatrix}$, we have $P^{-1}AP = diag(-1, -1, 2)$.

Index