MTH102

Linear Algebra

2025–2026–I Semester

03 October 2025

Table of contents

Introduction

Chapter 1. Sets

> **Definition**
>
> A **set** is a well-defined collection of objects. The objects of a set are called its **elements** or **members**.

> **Examples**
>
> - The set of vowels a, e, i, o, u.
> - The set consisting of $2, 4, 6, 8, 10$.
> - The set consisting of $3, 15, 35, 63, 99$.

Here are further examples.

> **Examples**
>
> - The set of positive integers divisible by 3.
> - The set of prime numbers less than 100.
> - The set of positive integers which can be expressed as the product of two distinct primes.
> - The set of positive integers which can be expressed as the product of two or more distinct primes.
> - The set of positive integers which are smaller than 100 and share no common factor with 100.

> **Remark**
>
> Note that the first few sets are described by writing down all its elements, whereas in the latter examples, the sets are described by the rules or properties which determine whether a particular object is an element or not.

> **Notation**
>
> A set is usually denoted by an uppercase letter, for example,
> $$A, B, C, X, Y, Z,$$
> whereas the lowercase letters are used to denote its elements. For instance, one may use
> - a to denote an element of A,
> - b to denote an element of B,
> - c to denote an element of C,
> - x to denote an element of X,
> - y to denote an element of Y,
> - z to denote an element of Z.

Several special sets are denoted by certain standard symbols. Here are some such examples.

$\mathbb{N}$	the set of positive integers
$\mathbb{Z}$	the set of integers

If x is an element of a set X, then one says that x **belongs to** X, and writes
$$x \in X.$$
Note that 2 belongs to $\mathbb{N}$, but $\frac{1}{2}$ does not belong to $\mathbb{N}$. One writes
$$2 \in \mathbb{N}, \frac{1}{2} \notin \mathbb{N}.$$

Notation

There are two ways to write down a set. For example, consider the set of vowels, which is written as
$$V = \{\text{a, e, i, o, u}\}.$$
Note that the elements are seperated by commas and enclosed in braces $\{\}$.

Another way to write down a set is by stating the *rules* or *properties* which determine whether a particular object is an element or not. For example, the set of even positive integers is written as
$$E = \{x \in \mathbb{Z} : x \text{ is even}, x > 0\}.$$
This reads
"E is the set of elements x in $\mathbb{Z}$ such that x is even and $x > 0$".

Example

Note that the set $A = \{3, 15, 35, 63, 99\}$ is equal to
$$\{x \in \mathbb{Z} : x \text{ is equal to } n^2 - 1 \text{ for some } n \in \{2, 4, 6, 8, 10\}\}.$$
Since $1, 2$ are the roots of the polynomial $x^2 - 3x + 2$, it follows that
$$\{x \in \mathbb{N} : x^2 - 3x + 2 = 0\} = \{1, 2\}.$$
Also note that
$$\mathbb{N} = \{1, 2, 3, ...\},$$
$$\mathbb{Z} = \{0, 1, -1, 2, -2, 3, -3, ...\}.$$

One uses the symbol $:=$ to indicate that the symbol on the left is being defined by the symbol on the right.
$$\mathbb{Q} := \left\{\frac{m}{n} : m, n \in \mathbb{Z} \text{ and } n \neq 0\right\}.$$
It is called the *set of rational numbers*.

If the number of elements of a set is finite, then it is called a **finite set**. If a set is not finite, it is called an **infinite set**. A set containing exactly one element is called a **singleton set**.

Example

The sets
$$\{\text{a, e, i, o, u}\}, \{3, 15, 35, 63, 99\}, \{x \in \mathbb{N} : x^2 - 3x + 2 = 0\}$$
are finite. The sets $\mathbb{N}, \mathbb{Z}, \mathbb{Q}$ are infinite.

Exercise 1.1. Determine the elements of the following sets.
- $\{x \in \mathbb{N} : x^2 - 1 = 0\}$.
- $\{x \in \mathbb{Z} : x^2 - 1 = 0\}$.

Solution. Note that 1 is the only element of $\mathbb{N}$ satisfying the equation $x^2 - 1 = 0$. This gives

$$\boxed{\{x \in \mathbb{N} : x^2 - 1 = 0\} = \{1\}}.$$

Since $1, -1$ are precisely all the elements of $\mathbb{Z}$ satisfying $x^2 - 1 = 0$, it follows that

$$\boxed{\{x \in \mathbb{Z} : x^2 - 1 = 0\} = \{1, -1\}}.$$

$\square$

§1.1 Basic terminologies

If A, B are sets, and every element of A also belongs to B, then we say that A is a **subset** of B, and write

$$A \subseteq B.$$

If A is not a subset of B, one writes

$$A \nsubseteq B.$$

> **Example**
>
> - Note that every set is a subset of itself.
> - Note that $\mathbb{N}$ is a subset of $\mathbb{Z}$, and $\mathbb{Z}$ is a subset of $\mathbb{Q}$. One writes
> $$\mathbb{N} \subseteq \mathbb{Z}, \quad \mathbb{Z} \subseteq \mathbb{Q}.$$
> - Also note that
> $$\mathbb{Z} \nsubseteq \mathbb{N}, \mathbb{Q} \nsubseteq \mathbb{Z}, \mathbb{Q} \nsubseteq \mathbb{N}.$$

If A is a subset of B and A is not equal to B, then A is said to be a *proper subset of B*, and one writes

$$A \subsetneq B.$$

Note that

$$\mathbb{N} \subsetneq \mathbb{Z}, \mathbb{Z} \subsetneq \mathbb{Q}.$$

Exercise 1.2. What does it mean to say that A is not a subset of B?

Solution. Note that A is a subset of B is equivalent to the statement that each element of A lies in B. Hence, A is not a subset of B is equivalent to saying that some element of A does not belong to B. $\square$

Exercise 1.3. Show that if A is a subset of B and B is a subset of C, then A is a subset of C.

Solution. Let x be an element of A. Since A is a subset of B, it follows that x belongs to B. Using that B is a subset of C, we obtain that x belongs to C. Hence, the element x lies in C. This shows that any element x of A belongs to C. This proves that A is a subset of C. $\square$

Exercise 1.4. Show that no element n of $\mathbb{N}$ satisfies $n^4 - 5n^2 + 6 = 0$.

> 🔥 **Tip**
>
> Try to show that no natural number n satisfies any of the equations
> $$n^2 = 2, n^2 = 3.$$

Solution. Note that the polynomial $x^4 - 5x^2 + 6$ factorizes as

$$x^4 - 5x^2 + 6 = (x^2 - 2)(x^2 - 3).$$

Let n be a natural number. If $n = 1$, then $n^2 \neq 2$ and $n^2 \neq 3$ hold. If $n \geq 2$, then $n^2 \geq 4$, and hence, $n^2 \neq 2$ and $n^2 \neq 3$ hold. This shows that no natural number n satisfies

$$(n^2 - 2)(n^2 - 3) = 0,$$

that is, the equation

$$n^4 - 5n^2 + 6 = 0.$$

$\square$

Consider the set

$$A = \{n \in \mathbb{N} : n^4 - 5n^2 + 6 = 0\}.$$

Note that the set A has no elements. Any such set is called the **empty set** or **null set**, and is denoted by the symbol

$$\varnothing.$$

Two sets, A and B, are said to be **equal**[1] if any element of A belongs to B, and any element of B also belongs to A, or equivalently,

$$A \subseteq B \text{ and } B \subseteq A$$

holds. If two sets A, B are not equal, one writes

$$A \neq B.$$

Exercise 1.5. What does it mean to say that two sets A, B are not equal?

> 🔥 **Tip**
>
> Does Exercise 1.2 help?

Solution. Note that two sets A, B are equal is equivalent to saying that A is a subset of B, and B is a subset of C. Hence, the statement $A \neq B$ is equivalent to saying that A is **not** a subset of B, **or** B is **not** a subset of A, which is equivalent to the statement that some element of A does not belong to B, **or** some element of B does not belong to A. $\square$

§1.2 Operations on sets

Exercise 1.6. Consider the sets

$$\{1, 2, 3, 4\}, \{3, 4, 5, 6\}.$$

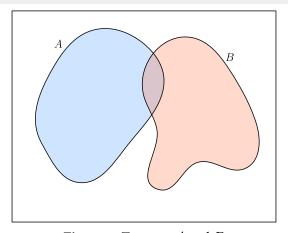Determine the elements common to these sets.

Figure 1: Two sets A and B

---

[1] *It* may appear obvious! The advantage of putting forth definitions is to set up notations and conventions, so that these are not left to interpretations!

> ### Definition
>
> If A, B are sets, their **union** is denoted by $A \cup B$, which is defined as
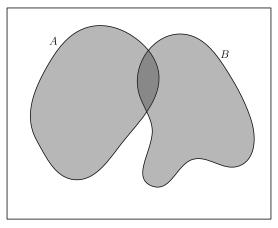> $$A \cup B := \{x : x \in A \text{ or } x \in B\}.$$
> See Figure 2.

Figure 2: The union of A and B

Exercise 1.7. Determine the union of the sets
$$\{1, 2, 3, 4\}, \{1, 3, 5, 7\}.$$

> ### ⚠ Usage of "or" in the inclusive sense
>
> The word "or" is used in the inclusive sense, allowing two or more of the conditions to be satisfied. In common terminology, this inclusive sense is denoted by "and/or".

Exercise 1.8. If A, B are sets, show that
$$A \subseteq A \cup B, B \subseteq A \cup B.$$

Solution. Let x be an element of A. Then note that x lies $A \cup B$. This shows that A is a subset of $A \cup B$. Similarly, it follows that B is a subset of $A \cup B$. ☐

Exercise 1.9. If A, B are subsets of a set C, then $A \cup B$ is a subset of C.

Solution. Let x be an element of $A \cup B$. It follows that x belongs to A or x belongs to B. If x belongs to A, then using A is a subset of C, we obtain that x lies in C. Further, if x belongs to B, then using that B is a subset of C, we obtain that x lies in C. Combining the above cases, it follows that x lies in C. Consequently, any element of $A \cup B$ is an element of C, or equivalently, $A \cup B$ is a subset of C. ☐

Exercise 1.10. If A is a set, show that
$$A \cup A = A.$$

Solution. Note that any element x of $A \cup A$ belongs to A or belongs to A, and hence it lies in A. This shows that $A \cup A$ is a subset of A. Further, for any element y of A, it belongs to $A \cup A$. Hence, A is a subset of $A \cup A$. This proves that
$$\boxed{A \cup A = A}.$$

☐

Exercise 1.11 (Commutative property). If A, B are sets, show that
$$A \cup B = B \cup A.$$

Solution. Let P, Q be sets, and let x be an element of $P \cup Q$. It follows that x belongs to P or x belongs to Q. This shows that x belongs to Q or x belongs to P. This implies that x is an element of $Q \cup P$. Hence, $P \cup Q$ is a subset of $Q \cup P$ for any two sets P, Q.

Consequently, the set $A \cup B$ is a subset of $B \cup A$, and $B \cup A$ is a subset of $A \cup B$. This shows that
$$\boxed{A \cup B = B \cup A}.$$

$\square$

> **Definition**
>
> If A, B, C are sets, their **union** is denoted by $A \cup B \cup C$, which is defined as
> $$A \cup B \cup C := \{x : x \in A \text{ or } x \in B \text{ or } x \in C\}.$$

Exercise 1.12 (Associative property). If A, B, C are sets, show that
$$A \cup B \cup C = (A \cup B) \cup C = A \cup (B \cup C).$$

Use Exercise 1.11 and Exercise 1.12, to show that for any three sets A, B, C, the following sets are equal
$$A \cup B \cup C, A \cup C \cup B, B \cup A \cup C, B \cup C \cup A, C \cup A \cup B, C \cup B \cup A.$$

> **Definition**
>
> If A, B are sets, their **intersection** is denoted by $A \cap B$, which is defined as
> $$A \cap B := \{x : x \in A \text{ and } x \in B\}.$$
>
> See Figure 3.
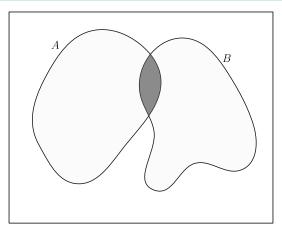
Figure 3: The intersection of A and B
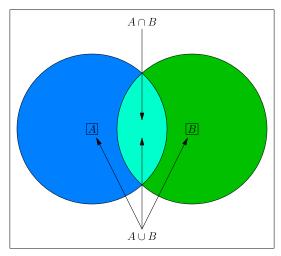
Figure 4: The union and intersection of A and B

Exercise 1.13. Determine the intersection of the sets
$$\{1, 2, 3, 4\}, \{1, 3, 5, 7\}.$$

Exercise 1.14. If A, B are sets, show that
$$A \cap B \subseteq A, A \cap B \subseteq B.$$

Solution. For any element x of $A \cap B$, it belongs to A and it belongs to B. Hence, $A \cap B$ is a subset of A, and $A \cap B$ is also a subset of B. $\square$

Exercise 1.15. If A, B, C are sets satisfying
$$C \subseteq A, C \subseteq B,$$

then show that
$$C \subseteq A \cap B$$

holds.

Exercise 1.16. Identify the integers among
$$1, 2, 3, 4, ..., 20$$

which are divisible by at least one of 2 and 5.

Solution. In the following, the integers among $1, 2, ..., 20$ divisible by 2 are marked.

$$1, \;2\;, 3, \;4\;, 5, \;6\;, 7, \;8\;, 9, \;10\;,$$

$$11, \;12\;, 13, \;14\;, 15, \;16\;, 17, \;18\;, 19, \;20\;.$$

Note that there 10 integers among $1, 2, ..., 20$, which are divisible by 2. The integers among $1, 2, ..., 20$ divisible by 5 are marked below.

$$1, 2, 3, 4, \;5\;, 6, 7, 8, 9, \;10\;,$$

$$11, 12, 13, 14, \;15\;, 16, 17, 18, 19, \;20\;.$$

Note that there are 4 integers among $1, 2, ..., 20$, which are divisible by 5. $\square$

> **i Remark**
>
> Note that some integers are counted/marked twice. This is an instance of *double counting*, which refers to the counting of certain element twice. In the above, the integers 10 and 20 are "counted twice". Thus, the number of integers among $1, 2, ..., 20$, divisible by 2 or 5, is equal to $10 + 4 - 2 = \boxed{12}$.

> **📄 Definition**
>
> Two sets A, B are said to be **disjoint** if
> $$A \cap B = \emptyset.$$

Lemma 1.17 (Inclusion-exclusion principle). *Let X be a set, and A, B be finite subsets of X. Then*
$$|A \cup B| = |A| + |B| - |A \cap B|.$$

Proof. Note that the set $A \cup B$ is equal to the union of the disjoint sets A and $B \setminus A$. This shows that
$$|A \cup B| = |A| + |B \setminus A|.$$
Also note that the set B is equal to the union of the disjoint subsets $A \cap B$ and $B \setminus A$. This gives
$$|B| = |A \cap B| + |B \setminus A|.$$
It follows that
$$\boxed{|A \cup B| = |A| + |B| - |A \cap B|}.$$

□

Exercise 1.18 (Inclusion-exclusion principle). Using Lemma 1.17 or otherwise, show that for finite subsets A, B, C of a set X,
$$|A \cup B \cup C| = |A| + |B| + |C| - |A \cap B| - |B \cap C| - |C \cap A| + |A \cap B \cap C|.$$

Solution. Note that
$$
\begin{aligned}
&|A \cup B \cup C| \\
&= |(A \cup B) \cup C| \\
&= |A \cup B| + |C| - |(A \cup B) \cap C| && \text{(by Lemma 1.17)} \\
&= |A \cup B| + |C| - |(A \cap C) \cup (B \cap C)| && \text{(by Fact 1.23)} \\
&= |A \cup B| + |C| - (|A \cap C| + |B \cap C| - |(A \cap C) \cap (B \cap C)| && \text{(by Lemma 1.17)} \\
&= |A \cup B| + |C| - |A \cap C| - |B \cap C| + |(A \cap C) \cap (B \cap C)| \\
&= |A \cup B| + |C| - |A \cap C| - |B \cap C| + |A \cap C \cap B \cap C| \\
&= |A \cup B| + |C| - |A \cap C| - |B \cap C| + |A \cap B \cap C| \\
&= |A| + |B| - |A \cap B| + |C| - |A \cap C| - |B \cap C| + |A \cap C \cap B \cap C| && \text{(by Lemma 1.17)} \\
&= |A| + |B| + |C| - |A \cap B| - |B \cap C| - |C \cap A| + |A \cap B \cap C|.
\end{aligned}
$$

□

Exercise 1.19. Determine the number of integers among
$$1, 2, 3, 4, ..., 100,$$
which are divisible by at least one of $6, 10, 15$.

Solution. Let X denote the set

$$\{1, 2, 3, ..., 100\}.$$

Let A (resp. B, C) denote the set of elements X which are divisible by 6 (resp. 10, 15). Note that

$$|A| = 16, |B| = 10, |C| = 6$$

hold. Observe that $A \cap B$ consists of the elements of X which are divisible by 6 and by 10, that is, divisible by their least common multiple, which is equal to 30. Similarly, $B \cap C$ consists of the elements of X which are divisible by the least common multiple of 10, 15, which is equal to 30. Further, it follows that $C \cap A$ also consists of the multiples of 30 lying between 1 and 100. Moreover, the set $A \cap B \cap C$ consists of the integers between 1 and 100, which are divisible by the least common multiple of 6, 10, 15, that is, the integer 30. By the inclusion-exclusion principle (Exercise 1.18), we obtain

$$
\begin{aligned}
|A \cup B \cup C| &= |A| + |B| + |C| - |A \cap B| - |B \cap C| - |C \cap A| + |A \cap B \cap C| \\
&= |A| + |B| + |C| - 3|A \cap B| + |A \cap B \cap C| \\
&= |A| + |B| + |C| - 2|A \cap B| \\
&= 16 + 10 + 6 - 2 \times 3 \\
&= 26.
\end{aligned}
$$

Note that the set $A \cup B \cup C$ consists of the integers between 1 and 100 which are divisible by at least one of 6, 10, 15. This shows that there are precisely $\boxed{26}$ such integers. $\qquad\square$

Exercise 1.20. If A is a set, show that

$$A \cap A = A.$$

> 🔥 **Tip**
>
> Does Exercise 1.10 help? If not, what about its solution?

Exercise 1.21 (Commutative property). If A, B are sets, show that

$$A \cap B = B \cap A.$$

> 🔥 **Tip**
>
> Does Exercise 1.11 help? If not, what about its solution?

> 📑 **Definition**
>
> If A, B, C are sets, their **intersection** is denoted by $A \cap B \cap C$, which is defined as
> $$A \cap B \cap C := \{x : x \in A \text{ and } x \in B \text{ and } x \in C\}.$$

Exercise 1.22 (Associative property). If A, B, C are sets, show that

$$A \cap B \cap C = (A \cap B) \cap C = A \cap (B \cap C).$$

> 🔥 **Tip**
>
> Does Exercise 1.12 help? If not, what about its solution?

Fact 1.23 (Distributive property). If A, B, C are sets, show that

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C),$$
$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C).$$

A proof of the above is provided in Chapter 4.

> **i Remark**
>
> All sets under consideration in a given context, are assumed to be contained in a 'large' fixed set, called the *universal set*. For example, while studying the positive integers (for instance, some sets consisting of the positive integers), the set $\mathbb{N}$ can be taken as the universal set. While studying the integers, the set $\mathbb{Z}$ can be taken as the universal set.

> **Definition**
>
> Let A, B be subsets of a set X. The **complement of A in B** is denoted by $B \setminus A$, and is defined by
> $$B \setminus A := \{x \in X : x \in B, x \notin A\}$$
> It is also called the **difference of B and A**.
>
> The complement of A in X, is called the *complement of A*, and is denoted by A^c, and is defined to be
> $$A^c := X \setminus A.$$

Exercise 1.24. Show that
$$A^c = \{x \in X : x \notin A\},$$
where X denotes the underlying universal set.

Solution. Since $X \setminus A$ is equal to $\{x \in X : x \notin A\}$, it follows that
$$\boxed{A^c = \{x \in X : x \notin A\}}.$$

$\square$

Exercise 1.25. Determine the complement of $\{1, 2, 3, 4\}$ in $\{1, 3, 5, 7\}$, and the complement of $\{1, 3, 5, 7\}$ in $\{1, 2, 3, 4\}$.

Exercise 1.26. If A, B are subsets of a set X, show that
$$A^c \cap B = B \setminus A.$$

Solution. Note that
$$\begin{aligned}
A^c \cap B &= (X \setminus A) \cap B \\
&= \{x \in X : x \in X \setminus A \text{ and } x \in B\} \\
&= \{x \in X : x \notin A \text{ and } x \in B\} \\
&= \{x \in X : x \in B \text{ and } x \notin A\} \\
&= B \setminus A.
\end{aligned}$$

$\square$

Exercise 1.27. If A is a subset of a set X, then show that
$$A \cup A^c = X, A \cap A^c = \emptyset, (A^c)^c = A.$$

Fact 1.28. If A, B are subsets of a set X, then show that
$$(A \cup B)^c = A^c \cap B^c, (A \cap B)^c = A^c \cup B^c.$$

A proof of the above is provided in Chapter 4.

Exercise 1.29. Let A, B be subsets of a set X. Show that the following statements are equivalent.
$$A \subseteq B,$$
$$A \cap B = A,$$
$$A \cup B = B,$$
$$B^c \subseteq A^c.$$

Fact 1.30. If A, B, C are sets, show that
1. $A \setminus (B \cup C) = (A \setminus B) \cap (A \setminus C)$,
2. $A \setminus (B \cap C) = (A \setminus B) \cup (A \setminus C)$.

A proof of the above is provided in Chapter 4.

Exercise 1.31. Let A, B be subsets of a set X. Show that the sets
$$A \setminus B, B \setminus A$$

are disjoint.

> 📑 **Definition**
>
> The **symmetric difference** of two sets A, B, denoted by $A \Delta B$, is defined as
> $$A \Delta B := (A \cup B) \setminus (A \cap B).$$

Exercise 1.32. Determine the symmetric difference of $\{1, 2, 3, 4\}$ and $\{1, 3, 5, 7\}$.

Exercise 1.33. If A, B are sets, then show that
$$A \Delta B = (A \setminus B) \cup (B \setminus A),$$
$$A \Delta B = B \Delta A,$$
$$A \cup B = (A \Delta B) \cup (A \cap B),$$
$$(A \Delta B) \cap (A \cap B) = \emptyset.$$

Solution. Note that
$$
\begin{aligned}
A \Delta B &= (A \cup B) \setminus (A \cap B) \\
&= (A \cup B) \cap (A \cap B)^c \\
&= (A \cap (A \cap B)^c) \cup (B \cap (A \cap B)^c) && \text{(by Fact 1.23)} \\
&= (A \cap (A^c \cup B^c)) \cup (B \cap (A^c \cup B^c)) && \text{(by Fact 1.28)} \\
&= ((A \cap A^c) \cup (A \cap B^c)) \cup ((B \cap A^c) \cup (B \cap B^c)) && \text{(by Fact 1.23)} \\
&= (\emptyset \cup (A \setminus B)) \cup ((B \setminus A) \cup \emptyset) && \text{(by Exercise 1.26)} \\
&= (A \setminus B) \cup (B \setminus A).
\end{aligned}
$$

Also note that

$$A\Delta B = (A \cup B) \setminus (A \cap B)$$
$$= (B \cup A) \setminus (B \cap A)$$
$$= B\Delta A.$$

Since $A \cap B$ is a subset of $A \cup B$, we obtain

$$A \cup B = ((A \cup B) \setminus (A \cap B)) \cup (A \cap B)$$
$$= (A\Delta B) \cup (A \cap B).$$

Using $A\Delta B = (A \cup B) \setminus (A \cap B)$, it follows that $A\Delta B$ contains no element of $A \cap B$. This implies that

$$(A\Delta B) \cap (A \cap B) = \emptyset.$$

$\square$

§1.3 Family of sets

> **Definition**
>
> Let $n \geq 2$ be an integer, and let $A_1, ..., A_n$ be sets. The **union of $A_1, ..., A_n$** is denoted by $A_1 \cup \cdots \cup A_n$, and is defined by
> $$A_1 \cup \cdots \cup A_n := \{x : x \text{ belongs to } A_i \text{ for some } 1 \leq i \leq n\}.$$
> The **intersection of $A_1, ..., A_n$** is denoted by $A_1 \cap \cdots \cap A_n$, and is defined by
> $$A_1 \cap \cdots \cap A_n := \{x : x \text{ belongs to } A_i \text{ for all } 1 \leq i \leq n\}.$$

§1.4 Power set

> **Example**
>
> Note that the subsets of $\{1, 2, 3\}$ are
> $$\{1, 2, 3\},$$
> $$\{1, 2\}, \{2, 3\}, \{3, 1\},$$
> $$\{1\}, \{2\}, \{3\},$$
> $$\emptyset.$$

> **Definition**
>
> Let A be a set. The set of subsets of A is denoted by $\mathcal{P}(A)$. It is called the **power set of A**.

> ### Example
>
> Note that
> $$\mathcal{P}(\{1\}) = \{\emptyset, \{1\}\},$$
> $$\mathcal{P}(\{x\}) = \{\emptyset, \{x\}\},$$
> $$\mathcal{P}(\{x, y\}) = \{\emptyset, \{x\}, \{y\}, \{x, y\}\},$$
> $$\mathcal{P}(\{x, y, z\}) = \{\emptyset, \{x\}, \{y\}, \{z\}, \{x, y\}, \{y, z\}, \{z, x\}, \{x, y, z\}\},$$
> $$\mathcal{P}(\{1, 2, 3\}) = \{\emptyset, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{2, 3\}, \{1, 3\}, \{1, 2, 3\}\},$$
> $$\mathcal{P}(\{1, 2, 3, 4\}) = \{\emptyset, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{2, 3\}, \{1, 3\}, \{1, 2, 3\},$$
> $$\{4\}, \{1, 4\}, \{2, 4\}, \{3, 4\}, \{1, 2, 4\}, \{2, 3, 4\}, \{1, 3, 4\}, \{1, 2, 3, 4\}\}.$$

> ### i Remark
>
> Convince yourself that for any integer $n \geq 0$, the power set of a set of size n has size 2^n.

Exercise 1.34. Write down the power set of each of the following sets:
$$\{1, 2\}, \{1, 2, 3\}, \{1, \{2, 3\}\}, \{1, 2, 3, 4\}.$$

§1.5 Cartesian product of sets

> ### Definition
>
> If A, B are sets, then the **cartesian product** $A \times B$ of A and B is the set of all ordered pairs of the form (a, b) with $a \in A$ and $b \in B$. That is,
> $$A \times B := \{(a, b) : a \in A, b \in B\}.$$

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

Exercise 1.35. If A, B are finite sets, then show that
$$|A \times B| = |A| \times |B|,$$
where for a set X, its number of elements is denoted by $|X|$.

§1.6 Sets and subsets of real and complex numbers

§1.6.1 Real numbers

$$\mathbb{N} = \text{the set of positive integers}$$
$$= \{1, 2, 3, ...\},$$
$$\mathbb{Z} = \text{the set of integers}$$
$$= \{0, 1, -1, 2, -2, 3, -3, ...\},$$
$$\mathbb{Q} = \text{the set of rational numbers}$$
$$= \left\{\frac{m}{n} : m \in \mathbb{Z}, n \in \mathbb{N}\right\},$$
$$\mathbb{R} = \text{the set of real numbers}.$$

> **i** **Remark**
>
> There are real numbers which are not rational. For instance, the number $\sqrt{2}$ is not rational.

Lemma 1.36. *No rational number x satisfies $x^2 = 2$.*

Proof. On the contrary, let us assume that some rational number x satisfies $x^2 = 2$. Replacing x by $-x$ if necessary, we may and do assume that x is positive. Write $x = \frac{a}{b}$ for some positive integers a, b. Let d denote the greatest common divisor of a, b. Write $a = da_0, b = db_0$ where a_0, b_0 are positive integers. Note that $\left(\frac{a_0}{b_0}\right)^2 = 2$.

Let us prove the two claims below. After establishing them, we will use them to complete the proof of the lemma.

> **Claim**
>
> None of a_0, b_0 is divisible by 3.

> **Proof of the Claim**
>
> On the contrary, let us assume that 3 divides at least one of the integers a_0, b_0.
>
> If 3 divides a_0, then 3 divides a_0^2. Using $a_0^2 = 2b_0^2$, it follows that 3 divides b_0. Further, if 3 divides b_0, then 3 divides $2b_0^2$, and hence 3 divides a_0^2. This shows that a_0 is divisible by 3.
>
> In both the cases, it follows that the greatest common divisor of a_0, b_0 is larger than 1. This shows that the greatest common divisor of da_0, db_0 is larger than d. Using $a = da_0, b = db_0$, we obtain that the greatest common divisor of a, b is larger than d, which is impossible. This contradicts the assumption that some rational number x satisfies $x^2 = 2$. This proves the claim.

> **Claim**
>
> If n is an integer not divisible by 3, then n^2 is equal to $3k + 1$ for some integer k, depending on n.

> **Proof of the Claim**
>
> Let q (resp. r) denote the quotient (resp. remainder) obtained upon dividing n by 3. This gives $n = 3q + r$. Note that
> $$n^2 = (3q + r)^2$$
> $$= 9q^2 + 6qr + r^2$$
> Since $n = 3q + r$ and 3 does not divide n, it follows that r is equal to 1 or 2. If $r = 1$, then
> $$n^2 = 3(3q^2 + 2qr) + 1.$$
> If $r = 2$, then
> $$n^2 = 3(3q^2 + 2qr) + 4 = 3(3q^2 + 2qr + 1) + 1.$$
> This proves the claim.

Using the claims above, it follows that
$$a_0^2 = 3c + 1,$$
$$b_0^2 = 3d + 1$$

for some integers c, d. This gives

$$3c + 1 = 2(3d + 1).$$

This yields $1 = 3(c - 2d)$, which is impossible.

This completes the proof of the lemma. $\qquad\square$

> **Definition**
>
> Let a, b be real numbers with $a \leq b$. Then the **closed interval** $[a, b]$ is defined by
> $$[a, b] := \{x \in \mathbb{R} : a \leq x \leq b\}.$$
> The **open interval** (a, b) is defined by
> $$(a, b) := \{x \in \mathbb{R} : a < x < b\}.$$

Exercise 1.37. Determine the cartesian product of $[1, 2]$ and $[3, 4] \cup [5, 6]$.

Exercise 1.38. Identify the set

$$\left\{ x \in \mathbb{R} \setminus \{0\} : x + \frac{1}{x} \geq 2 \right\}.$$

Solution. Let x be an element of $\mathbb{R} \setminus \{0\}$. Note that

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

hold. This shows that the condition

$$x + \frac{1}{x} \geq 2$$

is equivalent to the condition

$$\frac{(x-1)^2}{x} \geq 0. \tag{1}$$

If $x \neq 1$, then we obtain $(x - 1)^2 > 0$, and then, Equation 1 yields $\frac{1}{x} > 0$, or equivalently, $x > 0$. This gives that

$$\left\{ x \in \mathbb{R} \setminus \{0\} : x + \frac{1}{x} \geq 2 \right\} \setminus \{1\} \subseteq \{x \in \mathbb{R} : x > 0\},$$

which implies

$$\left\{ x \in \mathbb{R} \setminus \{0\} : x + \frac{1}{x} \geq 2 \right\} \subseteq \{x \in \mathbb{R} : x > 0\}.$$

Also note that if $x > 0$, then

$$\frac{(x-1)^2}{x} \geq 0,$$

or equivalently,

$$x + \frac{1}{x} \geq 2.$$

It follows that

$$\boxed{\left\{ x \in \mathbb{R} \setminus \{0\} : x + \frac{1}{x} \geq 2 \right\} = \{x \in \mathbb{R} : x > 0\}}.$$

□

Exercise 1.39. Let A, B be subsets of $\mathbb{R}$ defined by
$$A = \{x \in \mathbb{R} : x^2 \geq 0\},$$
$$B = \{x \in \mathbb{R} : x^3 \geq 0\}.$$
Determine the sets $A \cup B$, $A \cap B$, $A \setminus B$, $B \setminus A$.

Exercise 1.40. For a positive integer k, let A_k denote the set of integral multiples of k, that is,
$$A_k := \{r \in \mathbb{Z} : r = k\ell \text{ for some } \ell \in \mathbb{Z}\}.$$
Let m, n be positive integers. For each of the following statements, determine the equivalent conditions on the integers m, n.
  1. $A_m \subseteq A_n$
  2. $A_m \subsetneq A_n$
  3. $A_m \nsubseteq A_n$
  4. $A_m = A_n$
  5. $A_m \cap A_n = \emptyset$
  6. $A_m \setminus A_n \neq \emptyset$
  7. $A_m \setminus A_n \neq \emptyset$ or $A_n \setminus A_m \neq \emptyset$
  8. $A_m \setminus A_n \neq \emptyset$ and $A_n \setminus A_m \neq \emptyset$

§1.6.2 Complex numbers

Note that no real number x satisfies $x^2 = -1$. To "resolve" this, one "introduces"[2] an element i satisfying

$$\boxed{i^2 = -1.}$$

The element i can be multiplied by any real number, and the product can also be added to any real number. This leads to the numbers of the form $a + ib$, where a, b are real numbers. Such numbers are called the **complex numbers**, and the set of such numbers is denoted by $\mathbb{C}$.

> **Definition**
>
> The **set of complex numbers**, is denoted by $\mathbb{C}$, and is defined by
> $$\mathbb{C} := \{a + ib : a, b \in \mathbb{R}\}.$$

Exercise 1.41. Show that $\mathbb{R}$ is a subset of $\mathbb{C}$.

> **Definition**
>
> For a complex number $z = a + ib$ with $a, b \in \mathbb{R}$, the **real part** (resp. **imaginary part**) of z, denoted by $\mathfrak{R}(z)$ (resp. $\mathfrak{I}(z)$), is defined by
> $$\mathfrak{R}(z) = a,$$
> $$\mathfrak{I}(z) = b.$$

---

[2]To introduce is to adjoin.

> **Definition**
>
> For complex numbers $z = a + ib, w = c + id$ with $a, b, c, d \in \mathbb{R}$, the **sum** $z + w$ and the **product** $z \cdot w$ of z and w are defined by
> $$z + w := (a + c) + i(b + d),$$
> $$z \cdot w := (ac - bd) + i(ad + bc).$$

> **Definition**
>
> For a complex number $z = a + ib$ with $a, b \in \mathbb{R}$, the **conjugate** of z, denoted by $\overline{z}$, is defined by
> $$\overline{z} := a - ib.$$

> **Definition**
>
> For a complex number $z = a + ib$ with $a, b \in \mathbb{R}$, the **absolute value** of z, also called the **modulus** of z, is denoted by $|z|$, and is defined by
> $$|z| := \sqrt{a^2 + b^2}.$$

Exercise 1.42. If z, w are complex numbers, then show that
1. $\overline{z + w} = \overline{z} + \overline{w}$,
2. $\overline{z \cdot w} = \overline{z} \cdot \overline{w}$,
3. $|z| \geq 0$, and $|z| = 0$ if and only if $z = 0$,
4. $|z \cdot w| = |z|\, |w|$.

Exercise 1.43. If z is a complex number, then show that
$$z \cdot \overline{z} = |z|^2.$$

Exercise 1.44. For any real number θ, show that
$$|\cos\theta + i\sin\theta| = 1.$$

Fact 1.45. If z is a complex number satisfying $|z| = 1$, then there exists a real number θ such that
$$z = \cos\theta + i\sin\theta.$$
Consequently, if z is a nonzero complex number, then for some real number θ,
$$z = |z|\,(\cos\theta + i\sin\theta)$$
holds.

Chapter 2. Induction principles

> **Example**
>
> Show that
> $$1 + 2 + 3 + \cdots + n = \frac{n(n+1)}{2}$$
> holds for any positive integer n.

> **Solution**
>
> For any $n \in \mathbb{N}$, let $P(n)$ denote the statement that
> $$1 + 2 + 3 + \cdots + n = \frac{n(n+1)}{2}.$$
> Note that $P(1)$ holds. Let k be an element of $\mathbb{N}$, and assume that $P(k)$ holds. Note that
> $$\begin{aligned} 1 + 2 + 3 + \cdots + k + (k+1) &= \frac{k(k+1)}{2} + (k+1) \quad \text{(using } P(k)) \\ &= (k+1)\left(\frac{k}{2} + 1\right) \\ &= \frac{(k+1)(k+2)}{2}. \end{aligned}$$
> This shows that if k is a positive integer and $P(k)$ holds, then $P(k+1)$ also holds. By the principle of induction, it follows that $P(n)$ holds for any positive integer n.

> **Definition**
>
> If A is a nonempty subset of $\mathbb{N}$, then an element a_0 of A is called a **least element** of A if $a_0 \leq a$ for all $a \in A$.

Exercise 2.1. If A is a nonempty subset of $\mathbb{N}$, and a_1, a_2 are least elements of A, then show that $a_1 = a_2$.

Solution. Since a_1, a_2 are least elements of A, they belong to A. Using that a_1 is a least element of A, we obtain that $a_1 \leq a_2$. Further, using that a_2 is a least element of A, we obtain that $a_2 \leq a_1$. Combining these two inequalities, it follows that $a_1 = a_2$. $\square$

> **i Remark**
>
> By Exercise 2.1, a nonempty subset of $\mathbb{N}$ has at most one least element, to be called **the** least element, if it exists.

> **Well-ordering principle**
>
> Every nonempty subset of $\mathbb{N}$ has a least element.

> **Principle of mathematical induction**
>
> For each $n \in \mathbb{N}$, let $P(n)$ be a statement. Suppose $P(1)$ is true, and that whenever $P(n)$ holds for some $n \geq 1$, the statement $P(n+1)$ holds. Then $P(n)$ is true for all $n \geq 1$.

> **Principle of strong induction**
>
> For each $n \in \mathbb{N}$, let $P(n)$ be a statement. Suppose $P(1)$ is true, and that whenever $P(1), P(2), ..., P(n)$ hold for some $n \geq 1$, the statement $P(n+1)$ holds. Then $P(n)$ is true for all $n \geq 1$.

> **i Remark**
>
> The above three principles are equivalent. A proof of this is provided in Chapter 5 (see Theorem 5.1).

Exercise 2.2. Show that
$$\frac{1}{1 \cdot 2} + \frac{1}{2 \cdot 3} + \cdots + \frac{1}{n(n+1)} = \frac{n}{n+1}$$
for all $n \in \mathbb{N}$.

Solution. For a positive integer n, let $P(n)$ denote the statement that
$$\frac{1}{1 \cdot 2} + \frac{1}{2 \cdot 3} + \cdots + \frac{1}{n(n+1)} = \frac{n}{n+1}.$$
Note that $P(1)$ holds.

Let k be a positive integer such that $P(k)$ holds. Note that
$$\frac{1}{1 \cdot 2} + \frac{1}{2 \cdot 3} + \cdots + \frac{1}{k(k+1)} + \frac{1}{(k+1)(k+2)}$$
$$= \frac{k}{k+1} + \frac{1}{(k+1)(k+2)} \qquad \text{(using } P(k))$$
$$= \frac{k(k+2) + 1}{(k+1)(k+2)}$$
$$= \frac{(k+1)^2}{(k+1)(k+2)}$$
$$= \frac{k+1}{k+2}.$$
This shows that if k is a positive integer and $P(k)$ holds, then $P(k+1)$ also holds. By the principle of induction, it follows that $P(n)$ holds for any positive integer n. $\qquad \square$

Exercise 2.3. Show that
$$\frac{1}{1 \cdot 3} + \frac{1}{3 \cdot 5} + \cdots + \frac{1}{(2n-1)(2n+1)} = \frac{n}{2n+1}$$
for all $n \in \mathbb{N}$.

Solution. For a positive integer n, let $P(n)$ denote the statement that
$$\frac{1}{1 \cdot 3} + \frac{1}{3 \cdot 5} + \cdots + \frac{1}{(2n-1)(2n+1)} = \frac{n}{2n+1}.$$

Note that $P(1)$ holds.

Let k be a positive integer such that $P(k)$ holds. Note that

$$\frac{1}{1 \cdot 3} + \frac{1}{3 \cdot 5} + \cdots + \frac{1}{(2k-1)(2k+1)} + \frac{1}{(2(k+1)-1)(2(k+1)+1)}$$

$$= \frac{k}{2k+1} + \frac{1}{(2k+1)(2k+3)} \qquad \text{(using } P(k)\text{)}$$

$$= \frac{2k^2 + 3k + 1}{(2k+1)(2k+3)}$$

$$= \frac{(k+1)(2k+1)}{(2k+1)(2k+3)}$$

$$= \frac{k+1}{2(k+1)+1}.$$

This shows that if k is a positive integer and $P(k)$ holds, then $P(k+1)$ also holds. By the principle of induction, it follows that $P(n)$ holds for any positive integer n. $\qquad\square$

Exercise 2.4. Show that

$$\frac{1}{1 \cdot 2 \cdot 3} + \frac{1}{2 \cdot 3 \cdot 4} + \cdots + \frac{1}{n(n+1)(n+2)} = \frac{n(n+3)}{4(n+1)(n+2)}$$

for all $n \in \mathbb{N}$.

Solution. For a positive integer n, let $P(n)$ denote the statement that

$$\frac{1}{1 \cdot 2 \cdot 3} + \frac{1}{2 \cdot 3 \cdot 4} + \cdots + \frac{1}{n(n+1)(n+2)} = \frac{n(n+3)}{4(n+1)(n+2)}.$$

Note that $P(1)$ holds.

Let k be a positive integer such that $P(k)$ holds. Note that

$$\frac{1}{1\cdot 2\cdot 3}+\frac{1}{2\cdot 3\cdot 4}+\cdots+\frac{1}{k(k+1)(k+2)}+\frac{1}{(k+1)(k+2)(k+3)}$$

$$=\frac{k(k+3)}{4(k+1)(k+2)}+\frac{1}{(k+1)(k+2)(k+3)} \qquad \text{(using } P(k)\text{)}$$

$$=\frac{k(k+3)^2+4}{4(k+1)(k+2)(k+3)}$$

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

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

$$=\frac{(k+1)(k+3)^2-(k+1)(k+5)}{4(k+1)(k+2)(k+3)}$$

$$=\frac{(k+1)((k+3)^2-(k+5))}{4(k+1)(k+2)(k+3)}$$

$$=\frac{k^2+6k+9-(k+5)}{4(k+2)(k+3)}$$

$$=\frac{k^2+5k+4}{4(k+2)(k+3)}$$

$$=\frac{(k+1)((k+1)+3)}{4((k+1)+1)((k+1)+2)}.$$

This shows that if k is a positive integer and $P(k)$ holds, then $P(k+1)$ also holds. By the principle of induction, it follows that $P(n)$ holds for any positive integer n. $\qquad\square$

Exercise 2.5. Show that

$$3+11+\cdots+(8n-5)=4n^2-n$$

for all $n\in\mathbb{N}$.

Solution. For a positive integer n, let $P(n)$ denote the statement that

$$3+11+\cdots+(8n-5)=4n^2-n.$$

Note that $P(1)$ holds.

Let k be a positive integer such that $P(k)$ holds. Note that

$$3+11+\cdots+(8k-5)+(8(k+1)-5)$$

$$=(4k^2-k)+(8k+3) \qquad \text{(using } P(k)\text{)}$$

$$=4k^2+7k+3$$

$$=4k^2+8k+4-(k+1)$$

$$=4(k+1)^2-(k+1).$$

This shows that if k is a positive integer and $P(k)$ holds, then $P(k+1)$ also holds. By the principle of induction, it follows that $P(n)$ holds for any positive integer n. $\qquad\square$

Exercise 2.6. Show that

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

for all $n\in\mathbb{N}$.

Solution. For a positive integer n, let $P(n)$ denote the statement that

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

Note that $P(1)$ holds.

Let k be a positive integer such that $P(k)$ holds. Note that

$$1^2 + 3^2 + \cdots + (2k-1)^2 + (2(k+1)-1)^2$$

$$= \frac{4k^3 - k}{3} + (2k+1)^2 \qquad\qquad (\text{using } P(k))$$

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

$$= \frac{4k^3 - k + 12k^2 + 12k + 3}{3}$$

$$= \frac{4k^3 + 12k^2 + 11k + 3}{3}$$

$$= \frac{4(k+1)^3 - (k+1)}{3}.$$

This shows that if k is a positive integer and $P(k)$ holds, then $P(k+1)$ also holds. By the principle of induction, it follows that $P(n)$ holds for any positive integer n. $\qquad\square$

Exercise 2.7. Show that

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

for all $n \in \mathbb{N}$.

Solution. For a positive integer n, let $P(n)$ denote the statement that

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

Note that $P(1)$ holds.

Let k be a positive integer such that $P(k)$ holds. Note that

$$1^2 - 2^2 + 3^2 - 4^2 + \cdots + (-1)^{k+1}k^2 + (-1)^{(k+1)+1}(k+1)^2$$

$$= (-1)^{k+1}\frac{k(k+1)}{2} + (-1)^{k+2}(k+1)^2 \qquad\qquad (\text{using } P(k))$$

$$= (-1)^{k+1}\left(\frac{k(k+1)}{2} - (k+1)^2\right)$$

$$= (-1)^{k+1}\left(\frac{k(k+1) - 2(k+1)^2}{2}\right)$$

$$= (-1)^{k+1}\left(\frac{-(k+1)(k+2)}{2}\right)$$

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

This shows that if k is a positive integer and $P(k)$ holds, then $P(k+1)$ also holds. By the principle of induction, it follows that $P(n)$ holds for any positive integer n. $\qquad\square$

Exercise 2.8. Show that

$$1^2 + 2^2 + 3^2 + \cdots + n^2 = \frac{n(n+1)(2n+1)}{6},$$

$$1^3 + 2^3 + 3^3 + \cdots + n^3 = \left(\frac{n(n+1)}{2}\right)^2$$

hold for any positive integer n.

Solution. For a positive integer n, let $P(n)$ denote the statement that

$$1^2 + 2^2 + 3^2 + \cdots + n^2 = \frac{n(n+1)(2n+1)}{6},$$

$$1^3 + 2^3 + 3^3 + \cdots + n^3 = \left(\frac{n(n+1)}{2}\right)^2.$$

Note that $P(1)$ holds. Let k be a positive integer such that $P(k)$ holds. Note that

$$1^2 + 2^2 + 3^2 + \cdots + k^2 + (k+1)^2$$

$$= \frac{k(k+1)(2k+1)}{6} + (k+1)^2 \qquad (\text{using } P(k))$$

$$= \frac{k(k+1)(2k+1) + 6(k+1)^2}{6}$$

$$= \frac{(k+1)(k(2k+1) + 6(k+1))}{6}$$

$$= \frac{(k+1)(2k^2 + k + 6k + 6)}{6}$$

$$= \frac{(k+1)(2k^2 + 7k + 6)}{6}$$

$$= \frac{(k+1)(k+2)(2k+3)}{6}$$

$$= \frac{(k+1)(k+2)(2(k+1) + 1)}{6}.$$

Also note that

$$1^3 + 2^3 + 3^3 + \cdots + k^3 + (k+1)^3$$

$$= \left(\frac{k(k+1)}{2}\right)^2 + (k+1)^3 \qquad (\text{using } P(k))$$

$$= (k+1)^2 \left(\frac{k^2}{4} + (k+1)\right)$$

$$= (k+1)^2 \left(\frac{k^2}{4} + k + 1\right)$$

$$= (k+1)^2 \left(\frac{(k+2)^2}{4}\right)$$

$$= \left(\frac{(k+1)((k+1) + 1)}{2}\right)^2.$$

This shows that if k is a positive integer and $P(k)$ holds, then $P(k+1)$ also holds. By the principle of induction, it follows that $P(n)$ holds for any positive integer n.

$\square$

Exercise 2.9. Show that

$$\frac{1}{n+1} + \frac{1}{n+2} + \cdots + \frac{1}{2n} = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots + \frac{1}{2n-1} - \frac{1}{2n}$$

for all $n \in \mathbb{N}$.

Solution. For a positive integer n, let $P(n)$ denote the statement that

$$\frac{1}{n+1} + \frac{1}{n+2} + \cdots + \frac{1}{2n} = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots + \frac{1}{2n-1} - \frac{1}{2n}.$$

Note that $P(1)$ holds.

Let k be a positive integer such that $P(k)$ holds. Note that

$$\frac{1}{(k+1)+1} + \cdots + \frac{1}{2k+1} + \frac{1}{2(k+1)}$$

$$= \frac{1}{k+1} + \frac{1}{(k+1)+1} + \cdots + \frac{1}{2k} + \frac{1}{2k+1} + \frac{1}{2(k+1)} - \frac{1}{k+1}$$

$$= \frac{1}{k+1} + \frac{1}{(k+1)+1} + \cdots + \frac{1}{2k} + \frac{1}{2k+1} - \frac{1}{2(k+1)}$$

$$= \left(1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots + \frac{1}{2k-1} - \frac{1}{2k}\right) + \frac{1}{2k+1} - \frac{1}{2(k+1)}$$

$$\hspace{8cm} (\text{using } P(k))$$

$$= \left(1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots + \frac{1}{2k-1} - \frac{1}{2k}\right) + \frac{1}{2k+1} - \frac{1}{2k+2}$$

$$= 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots + \frac{1}{(2(k+1))-1} - \frac{1}{(2(k+1))}.$$

This shows that if k is a positive integer and $P(k)$ holds, then $P(k+1)$ also holds. By the principle of induction, it follows that $P(n)$ holds for any positive integer n. $\qquad\square$

> 📑 **Definition**
>
> An integer is called **divisible** by a nonzero integer m if it is equal to mk for some integer k.

Exercise 2.10. Show that $n^3 + 5n$ is divisible by 6 for all $n \in \mathbb{N}$.

Solution. For a positive integer n, let $P(n)$ denote the statement that $n^3 + 5n$ is divisible by 6.

Note that $P(1)$ holds.

Let k be a positive integer such that $P(k)$ holds. Note that

$$(k+1)^3 + 5(k+1) = (k^3 + 3k^2 + 3k + 1) + (5k + 5)$$

$$= (k^3 + 5k) + (3k^2 + 3k + 6).$$

By the induction hypothesis, the integer $k^3 + 5k$ is divisible by 6. Also note that $3k^2 + 3k + 6 = 3(k^2 + k + 2)$ is divisible by 3. Further, one of the integers $k, k+1$ is even, and hence $k^2 + k$ is even. This gives that $k^2 + k + 2$ is even, and hence $3(k^2 + k + 2)$ is divisible by 2. It follows that $3(k^2 + k + 2)$ is divisible by 6. This shows that $(k+1)^3 + 5(k+1)$ is divisible by 6. This proves that if k is a positive integer and $P(k)$ holds, then $P(k+1)$ also holds.

By the principle of induction, it follows that $P(n)$ holds for any positive integer n. $\qquad\square$

Exercise 2.11. Show that $5^{2n} - 1$ is divisible by 8 for all $n \in \mathbb{N}$.

Solution. For a positive integer n, let $P(n)$ denote the statement that $5^{2n} - 1$ is divisible by 8.

Note that $P(1)$ holds.

Let k be a positive integer such that $P(k)$ holds. Note that

$$\begin{aligned}
5^{2(k+1)} - 1 &= 5^{2k+2} - 1 \\
&= 5^2 \cdot 5^{2k} - 1 \\
&= 25 \cdot 5^{2k} - 1 \\
&= (24 + 1) \cdot 5^{2k} - 1 \\
&= 24 \cdot 5^{2k} + \left(5^{2k} - 1\right).
\end{aligned}$$

By the induction hypothesis, the integer $5^{2k} - 1$ is divisible by 8. Also note that $24 \cdot 5^{2k}$ is divisible by 8. This shows that $5^{2(k+1)} - 1$ is divisible by 8. This shows that if k is a positive integer and $P(k)$ holds, then $P(k+1)$ also holds.

By the principle of induction, it follows that $P(n)$ holds for any positive integer n. $\qquad\square$

Exercise 2.12. Show that $5^n - 4n - 1$ is divisible by 16 for all $n \in \mathbb{N}$.

Solution. For a positive integer n, let $P(n)$ denote the statement that $5^n - 4n - 1$ is divisible by 16.

Note that $P(1)$ holds.

Let k be a positive integer such that $P(k)$ holds. Note that

$$\begin{aligned}
5^{k+1} - 4(k+1) - 1 &= 5 \cdot 5^k - 4k - 4 - 1 \\
&= 5 \cdot \left(5^k - 4k - 1\right) + 5 \cdot (4k + 1) - 4k - 5 \\
&= 5 \cdot \left(5^k - 4k - 1\right) + 16k.
\end{aligned}$$

By the induction hypothesis, the integer $5^k - 4k - 1$ is divisible by 16. Using the above, it follows that $5^{k+1} - 4(k+1) - 1$ is divisible by 16. This shows that if k is a positive integer and $P(k)$ holds, then $P(k+1)$ also holds.

By the principle of induction, it follows that $P(n)$ holds for any positive integer n. $\qquad\square$

Exercise 2.13. Show that $6^n - 5n - 1$ is divisible by 25 for all $n \in \mathbb{N}$.

Exercise 2.14. Show that $n^3 + (n+1)^3 + (n+2)^3$ is divisible by 9 for all $n \in \mathbb{N}$.

Solution. For a positive integer n, let $P(n)$ denote the statement that $n^3 + (n+1)^3 + (n+2)^3$ is divisible by 9.

Note that $P(1)$ holds.

Let k be a positive integer such that $P(k)$ holds. Note that

$$\begin{aligned}
&(k+1)^3 + (k+2)^3 + (k+3)^3 \\
&= \left(k^3 + (k+1)^3 + (k+2)^3\right) - k^3 + (k+3)^3 \\
&= \left(k^3 + (k+1)^3 + (k+2)^3\right) - k^3 + \left(k^3 + 9k^2 + 27k + 27\right) \\
&= \left(k^3 + (k+1)^3 + (k+2)^3\right) + \left(9k^2 + 27k + 27\right).
\end{aligned}$$

By the induction hypothesis, the integer $k^3 + (k+1)^3 + (k+2)^3$ is divisible by 9. Also note that $9k^2 + 27k + 27$ is divisible by 9. This shows that $(k+1)^3 + (k+2)^3 + (k+3)^3$ is divisible by 9. This implies that if k is a positive integer and $P(k)$ holds, then $P(k+1)$ also holds.

By the principle of induction, it follows that $P(n)$ holds for any positive integer n. $\qquad\square$

Exercise 2.15. Show that

$$\frac{1}{\sqrt{1}} + \frac{1}{\sqrt{2}} + \cdots + \frac{1}{\sqrt{n}} > \sqrt{n}$$

for all $n \in \mathbb{N}$ satisfying $n > 1$.

Solution. For a positive integer n, let $P(n)$ denote the statement that

$$\frac{1}{\sqrt{1}} + \frac{1}{\sqrt{2}} + \cdots + \frac{1}{\sqrt{n}} > \sqrt{n}$$

holds.

Note that $P(2)$ holds, since

$$\frac{1}{\sqrt{1}} + \frac{1}{\sqrt{2}} = 1 + \frac{\sqrt{2}}{2} > \sqrt{2}.$$

Let k be a positive integer such that $k \geq 2$ and $P(k)$ holds. Note that

$$\frac{1}{\sqrt{1}} + \frac{1}{\sqrt{2}} + \cdots + \frac{1}{\sqrt{k}} + \frac{1}{\sqrt{k+1}}$$

$$> \sqrt{k} + \frac{1}{\sqrt{k+1}} \qquad \text{(using } P(k))$$

$$\geq \sqrt{k+1},$$

where the last inequality follows since

$$k - \left(\sqrt{k+1} - \frac{1}{\sqrt{k+1}}\right)^2 = k - (k+1) + 2 - \frac{1}{k+1}$$

$$= 1 - \frac{1}{k+1}$$

$$> 0.$$

This shows that if k is a positive integer satisfying $k \geq 2$ and $P(k)$ holds, then $P(k+1)$ also holds. By the principle of induction, it follows that $P(n)$ holds for any positive integer n satisfying $n \geq 2$. □

Exercise 2.16. Show that $3^n \geq n^2$ for all $n \in \mathbb{N}$.

Solution. For any $n \in \mathbb{N}$, let $P(n)$ denote the statement that $3^n \geq n^2$. Note that $P(1)$ holds. Let k be a positive integer such that $P(k)$ holds. Note that

$$(k+1)^2 = k^2 + 2k + 1$$

$$\leq 3^k + 2k + 1 \quad \text{(using } P(k))$$

$$\leq 3^k + 3^k + 3^k \quad \text{(using } k \geq 1)$$

$$= 3 \cdot 3^k$$

$$= 3^{k+1}.$$

This shows that if k is a positive integer and $P(k)$ holds, then $P(k+1)$ also holds. By the principle of induction, it follows that $P(n)$ holds for any positive integer n. □

Exercise 2.17. Show that $n! > 2^n$ for all $n \in \mathbb{N}$ satisfying $n \geq 4$.

Solution. For a positive integer n, let $P(n)$ denote the statement that $n! > 2^n$.

Note that $P(4)$ holds, since $4! = 24 > 16 = 2^4$.

Let k be a positive integer such that $k \geq 4$ and $P(k)$ holds. Note that

$$(k+1)! = (k+1)k!$$
$$\geq (k+1) \cdot 2^k \quad \text{(using } P(k)\text{)}$$
$$> 2 \cdot 2^k \qquad \text{(using } k \geq 4\text{)}$$
$$= 2^{k+1}.$$

This shows that if k is a positive integer satisfying $k \geq 4$ and $P(k)$ holds, then $P(k+1)$ also holds.

By the principle of induction, it follows that $P(n)$ holds for any positive integer n satisfying $n \geq 4$. $\qquad \square$

> ⚠️ **Warning**
>
> It is important to verify the base case, that is, verifying that $P(1)$ holds, even if it is trivial. For example, if we define $P(n)$ to be the statement that
> $$n = n + 1,$$
> then $P(n)$ implies $P(n+1)$ for any $n \in \mathbb{N}$. However, $P(1)$ does not hold. Another example is to take $Q(n)$ to be the statement that $2n + 1$ is even, then $Q(n)$ implies $Q(n+1)$ for any $n \in \mathbb{N}$. However, $Q(1)$ does not hold.

Here are a few exercises where strong induction is useful.

Exercise 2.18. Show that every positive integer greater than 1 is a prime or is a product of prime numbers.

Solution. For any integer $n \geq 2$, let $P(n)$ denote the statement that n is a prime or is a product of prime numbers.

Note that $P(2)$ holds, since 2 is a prime.

Let $k \geq 2$ be an integer such that $P(m)$ holds for all integers m satisfying $2 \leq m \leq k$. If $k+1$ is a prime, then $P(k+1)$ holds. If $k+1$ is not a prime, then there exist integers a, b such that $k+1 = ab$ and $1 < a \leq b < k+1$. This gives that $2 \leq a \leq k$ and $2 \leq b \leq k$. By the induction hypothesis, both a and b are either primes or products of primes. It follows that $k+1 = ab$ is a product of primes. This shows that if $k \geq 2$ is an integer such that $P(m)$ holds for all integers m satisfying $2 \leq m \leq k$, then $P(k+1)$ also holds.

By the principle of strong induction, it follows that $P(n)$ holds for any integer $n \geq 2$. $\qquad \square$

Exercise 2.19. Let the integers $x_1, x_2, \ldots$ be defined by
$$x_1 := 1,$$
$$x_2 := 2,$$
$$x_{n+2} := \frac{1}{2}(x_n + x_{n+1})$$
for all $n \in \mathbb{N}$. Show that $1 \leq x_n \leq 2$ for all $n \in \mathbb{N}$.

> 🔥 **Tip**
>
> For any $n \in \mathbb{N}$, let $P(n)$ denote the statement that $1 \leq x_m \leq 2$ for all $m \in \mathbb{N}$ satisfying $m \leq n$.

Exercise 2.20. The Fibonacci sequence $F_0, F_1, F_2, \ldots$ is defined by

$$F_0 := 0,$$
$$F_1 := 1,$$
$$F_2 := 1,$$
$$F_{n+2} := F_n + F_{n+1}$$

for all $n \in \mathbb{N}$. Show that

$$F_n = \frac{1}{\sqrt{5}}\left(\left(\frac{1 + \sqrt{5}}{2}\right)^n - \left(\frac{1 - \sqrt{5}}{2}\right)^n\right)$$

for all $n \in \mathbb{N} \cup \{0\}$.

Exercise 2.21. Show that

$$F_0^2 + F_1^2 + F_2^2 + \cdots + F_n^2 = F_n F_{n+1}$$

for all $n \in \mathbb{N} \cup \{0\}$.

Linear algebra

Chapter 3. Matrices

§3.1 Vectors

§3.1.1 Definitions and examples

A vector is a column of numbers. The numbers in the vector are called its **components** or **coordinates**.

> **i** **Remark**
>
> Throughout this course, most of the vectors to be considered, will have real or complex components.

Recall that for two sets A, B, their cartesian product $A \times B$ is defined as
$$A \times B := \{(a, b) \mid a \in A, b \in B\}.$$
The cartesian product of $\mathbb{R}$ with itself is denoted by $\mathbb{R}^2$. In other words,
$$\boxed{\mathbb{R}^2 := \{(x, y) \mid x \in \mathbb{R}, y \in \mathbb{R}\}}.$$
Similarly, one defines $\mathbb{R}^3$ as $\mathbb{R} \times \mathbb{R} \times \mathbb{R}$, that is,
$$\boxed{\mathbb{R}^3 := \{(x, y, z) \mid x \in \mathbb{R}, y \in \mathbb{R}, z \in \mathbb{R}\}}.$$

> **Definition**
>
> If $n \geq 2$ is an integer and $A_1, A_2, ..., A_n$ are sets, then their **cartesian product** $A_1 \times A_2 \times \cdots \times A_n$ is defined as
> $$A_1 \times A_2 \times \cdots \times A_n := \{(a_1, a_2, ..., a_n) \mid a_i \in A_i \text{ for all } i = 1, 2, ..., n\}.$$

The n-fold cartesian product of $\mathbb{R}$ with itself, that is, the cartesian product of n copies of $\mathbb{R}$, is denoted by $\mathbb{R}^n$. In other words,
$$\boxed{\mathbb{R}^n := \{(x_1, x_2, ..., x_n) \mid x_i \in \mathbb{R} \text{ for all } i = 1, 2, ..., n\}}.$$
Similarly, one defines $\mathbb{C}^n$ as the n-fold cartesian product of $\mathbb{C}$ with itself, that is,
$$\boxed{\mathbb{C}^n := \{(z_1, z_2, ..., z_n) \mid z_i \in \mathbb{C} \text{ for all } i = 1, 2, ..., n\}}.$$

> ### Example
>
> The elements
> $$\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$
> of $\mathbb{R}^2$ are vectors. The elements
> $$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 10 \\ 20 \\ -2025 \end{pmatrix}$$
> of $\mathbb{R}^3$ are vectors too. The elements
> $$\begin{pmatrix} 1+i \\ 2-i \\ -3i \end{pmatrix}, \begin{pmatrix} 0 \\ \pi \\ 1+i \end{pmatrix}$$
> of $\mathbb{C}^3$ are vectors as well.

§3.1.2 Vector operations

> ### Definition
>
> Two vectors $\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ and $\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$ in $\mathbb{R}^n$ are added component-wise. That is, the **sum** of two elements $\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ and $\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$ in $\mathbb{R}^n$ is defined as
> $$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} := \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{pmatrix}.$$
> If $c \in \mathbb{R}$ is a real number and $v = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ is a vector in $\mathbb{R}^n$, then the **scalar multiplication** of c with the vector v is defined as
> $$c \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} := \begin{pmatrix} c \cdot x_1 \\ c \cdot x_2 \\ \vdots \\ c \cdot x_n \end{pmatrix}.$$
> Note that $1 \cdot v$ is equal to v. Instead of $(-1) \cdot v$, we write $-v$. Further, $c \cdot v$ is often written as cv.

> ### ⚠ Warning
>
> In the above definition, we have used the same notation "+" for addition of two vectors and addition of two real numbers, and the same notation "·" for scalar multiplication of a vector by a real number and multiplication of two real numbers. The meaning will be clear from the context.

> **ℹ Remark**
>
> Note that the sum of two vectors in $\mathbb{R}^n$ is again a vector in $\mathbb{R}^n$. Also, the scalar multiplication of a vector in $\mathbb{R}^n$ by a real number is again a vector in $\mathbb{R}^n$.

> **ℹ Remark**
>
> Similar definitions work for vectors in $\mathbb{C}^n$.

> **✦ Example**
>
> Consider the vectors $u = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ and $v = \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix}$ in $\mathbb{R}^3$. Note that
>
> $$u + v = \begin{pmatrix} 1+4 \\ 2+5 \\ 3+6 \end{pmatrix} = \begin{pmatrix} 5 \\ 7 \\ 9 \end{pmatrix},$$
>
> and if $c = 3$, then
>
> $$c \cdot u = 3 \cdot \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 3 \cdot 1 \\ 3 \cdot 2 \\ 3 \cdot 3 \end{pmatrix} = \begin{pmatrix} 3 \\ 6 \\ 9 \end{pmatrix}.$$

§3.2 Matrix

§3.2.1 Definitions and examples

> **📑 Definition**
>
> A **matrix** is a rectangular array of numbers arranged in rows and columns. A matrix with m rows and n columns is called an $m \times n$ **matrix**. The numbers in the matrix are called its **entries** or **elements**. A matrix is called a **square matrix** if its number of rows, and its number of columns are equal.

> **📑 Definition**
>
> Two matrices are **equal** if they have the same size and their corresponding entries are equal.

> **ℹ Remark**
>
> The **size** of a matrix is always given in the form "(the number of rows) $\times$ (the number of columns)".

> 🔗 **Example**
>
> The following is a 2×3 matrix:
> $$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}.$$
>
> The following is a 3×2 matrix:
> $$B = \begin{pmatrix} 7 & 8 \\ 9 & 10 \\ 11 & 12 \end{pmatrix}.$$

> 🔗 **Example**
>
> Here is an example of two matrices which are not equal.
> $$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \neq \begin{pmatrix} 1 & 2 \\ 3 & 5 \end{pmatrix}$$

If A is a matrix, its (i, j)-th entry, that is, the entry common to the i-th row and the j-th column is denoted by a_{ij}. Thus an $m \times n$ matrix A can be represented as

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

One also writes $A = \left(a_{ij}\right)_{\substack{1 \leq i \leq m, \\ 1 \leq j \leq n}}$ to denote the matrix A, or one writes $A = \left(a_{ij}\right)$ if the size of A is clear from the context.

> ℹ️ **Remark**
>
> Note that for an $m \times n$ matrix A, one writes $A = (\text{the } (i, j)\text{-th entry of } A)_{\substack{1 \leq i \leq m, \\ 1 \leq j \leq n}}.$

If $A = \left(a_{ij}\right)$ is an $n \times n$ matrix, then the entries $a_{11}, a_{22}, ..., a_{nn}$ are called its **diagonal entries**. A square matrix is called a **diagonal matrix** if its entries, other the diagonal entries, are equal to zero.

§3.2.2 Matrix times a vector

The product

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \begin{pmatrix} 7 \\ 8 \\ 9 \end{pmatrix}$$

is defined by "combining" each row of the matrix with the vector as follows:

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \begin{pmatrix} 7 \\ 8 \\ 9 \end{pmatrix} = \begin{pmatrix} 1 \times 7 + 2 \times 8 + 3 \times 9 \\ 4 \times 7 + 5 \times 8 + 6 \times 9 \end{pmatrix},$$

which is equal to the vector $\begin{pmatrix} 50 \\ 122 \end{pmatrix}$.

As we see, for this to work, the number of columns of the matrix must be equal to the number of rows of the vector. Thus, when the product of an $m \times n$ matrix with a vector in $\mathbb{R}^n$ is defined similarly, the result is a vector in $\mathbb{R}^m$.

> **📑 Definition**
>
> If
> $$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix}$$
> is an $m \times n$ matrix with entries in $\mathbb{R}$ and $\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ is a vector in $\mathbb{R}^n$, then one defines
> $$\begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} := \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{pmatrix}.$$
> Note that the result is a vector in $\mathbb{R}^m$.

§3.2.3 Matrix times a matrix

Similar to the case of vectors, one defines addition, subtraction, and scalar multiplication of matrices entry-wise. For addition and subtraction, the two matrices must be of the same size. Indeed, if
$$A = \left(a_{ij}\right)_{\substack{1 \leq i \leq m, \\ 1 \leq j \leq n}}, B = \left(b_{ij}\right)_{\substack{1 \leq i \leq m, \\ 1 \leq j \leq n}}$$
are two $m \times n$ matrices, then their **sum** $A + B$ and **difference** $A - B$ are defined as
$$A + B := \left(a_{ij} + b_{ij}\right)_{\substack{1 \leq i \leq m, \\ 1 \leq j \leq n}},$$
and
$$A - B := \left(a_{ij} - b_{ij}\right)_{\substack{1 \leq i \leq m, \\ 1 \leq j \leq n}}$$
respectively. If c is a real number and $A = \left(a_{ij}\right)_{\substack{1 \leq i \leq m, \\ 1 \leq j \leq n}}$ is an $m \times n$ matrix, then the **scalar multiplication** of A with c is defined as
$$c \cdot A := \left(c \cdot a_{ij}\right)_{\substack{1 \leq i \leq m, \\ 1 \leq j \leq n}}.$$
Note that $1 \cdot A$ is equal to A. Instead of $-1 \cdot A$, we write $-A$. Further, $c \cdot A$ is often written as cA.

> **ℹ Remark**
>
> Note that the sum of two matrices with entries in $\mathbb{R}$ is a matrix with entries in $\mathbb{R}$. Also, the scalar multiplication of a matrix with entries in $\mathbb{R}$ by a real number is again a matrix with entries in $\mathbb{R}$.

> **ℹ Remark**
>
> Similar definitions work for matrices with entries in $\mathbb{C}$.

> ### ⌁ Example
>
> If
> $$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$$
> and
> $$B = \begin{pmatrix} 7 & 8 & 9 \\ 10 & 11 & 12 \end{pmatrix},$$
> then
> $$A + B = \begin{pmatrix} 1+7 & 2+8 & 3+9 \\ 4+10 & 5+11 & 6+12 \end{pmatrix} = \begin{pmatrix} 8 & 10 & 12 \\ 14 & 16 & 18 \end{pmatrix},$$
> and if $c = 3$, then
> $$c \cdot A = 3 \cdot \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} = \begin{pmatrix} 3 \cdot 1 & 3 \cdot 2 & 3 \cdot 3 \\ 3 \cdot 4 & 3 \cdot 5 & 3 \cdot 6 \end{pmatrix} = \begin{pmatrix} 3 & 6 & 9 \\ 12 & 15 & 18 \end{pmatrix}.$$

Exercise 3.1. If A, B, C are matrices of the same size, show that
$$A + (B + C) = (A + B) + C.$$

Exercise 3.2. If A, B are matrices of the same size, show that
$$A + B = B + A.$$

Exercise 3.3. If A is an $m \times n$ matrix, show that
$$A + 0_{m \times n} = 0_{m \times n} + A = A$$
holds, where $0_{m \times n}$ denotes the $m \times n$ zero matrix, that is, the $m \times n$ matrix whose entries are all zero.

Exercise 3.4. Show that if A is a matrix and c, d are real numbers, then
$$(c + d)A = cA + dA$$
and
$$c(dA) = (cd)A.$$

> **Definition**
>
> The **transpose** of an $m \times n$ matrix
> $$A = \left(a_{ij} \right)_{\substack{1 \le i \le m, \\ 1 \le j \le n}}$$
> is the $n \times m$ matrix
> $$A^T := \left(a_{ji} \right)_{\substack{1 \le i \le n, \\ 1 \le j \le m}}.$$
> That is, if
> $$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix},$$
> then its **transpose** is the $n \times m$ matrix
> $$A^T := \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{pmatrix}.$$

> **Example**
>
> If
> $$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix},$$
> then its transpose is the 3×2 matrix
> $$A^T = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}.$$

> **Remark**
>
> Note that the transpose of a matrix is obtained by interchanging its rows and columns.

Exercise 3.5. If A is a matrix, what is $\left(A^T \right)^T$?

Exercise 3.6. Show that if A, B are matrices of the same size, then
$$(A + B)^T = A^T + B^T.$$

Solution. Write
$$A = \left(a_{ij} \right)_{\substack{1 \le i \le m, \\ 1 \le j \le n}}, B = \left(b_{ij} \right)_{\substack{1 \le i \le m, \\ 1 \le j \le n}}.$$

Note that

$$(A+B)^T = \left(\left(a_{ij} + b_{ij} \right)_{\substack{1 \leq i \leq m, \\ 1 \leq j \leq n}} \right)^T$$

$$= \left(a_{ji} + b_{ji} \right)_{\substack{1 \leq i \leq n, \\ 1 \leq j \leq m}}$$

$$= \left(a_{ji} \right)_{\substack{1 \leq i \leq n, \\ 1 \leq j \leq m}} + \left(b_{ji} \right)_{\substack{1 \leq i \leq n, \\ 1 \leq j \leq m}}$$

$$= A^T + B^T.$$

$\square$

Exercise 3.7. Show that if A denotes a matrix, then

$$\left(A^T \right)^T = A.$$

Solution. Write

$$A = \left(a_{ij} \right)_{\substack{1 \leq i \leq m, \\ 1 \leq j \leq n}}.$$

Note that

$$\left(A^T \right)^T = \left(\left(a_{ji} \right)_{\substack{1 \leq i \leq n, \\ 1 \leq j \leq m}} \right)^T$$

$$= \left(a_{ij} \right)_{\substack{1 \leq i \leq m, \\ 1 \leq j \leq n}}$$

$$= A.$$

$\square$

> ### 📑 Definition
>
> The **product** of an $m \times n$ matrix A with an $n \times p$ matrix B is defined as follows:
> $$AB := \left(c_{ij}\right)_{\substack{1 \leq i \leq m,, \\ 1 \leq j \leq p}}$$
> where
> $$c_{ij} := a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj}$$
> for all $1 \leq i \leq m$ and $1 \leq j \leq p$. In other words, given an $m \times n$ matrix
> $$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$
> and an $n \times p$ matrix
> $$B = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1p} \\ b_{21} & b_{22} & \cdots & b_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{np} \end{pmatrix},$$
> we have
> $$AB = \begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1p} \\ c_{21} & c_{22} & \cdots & c_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ c_{m1} & c_{m2} & \cdots & c_{mp} \end{pmatrix},$$
> where
> $$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj}$$
> for all $1 \leq i \leq m$ and $1 \leq j \leq p$. Note that the result is an $m \times p$ matrix.

In other words, the (i, j)-th entry of the product AB is obtained by "combining" the i-th row of A with the j-th column of B.

> ### ℹ️ Remark
>
> Note that the product AB is defined only when the number of columns of A is equal to the number of rows of B.

> ⚡ **Example**
>
> If
> $$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$$
> and
> $$B = \begin{pmatrix} 7 & 8 \\ 9 & 10 \\ 11 & 12 \end{pmatrix},$$
> then the product AB is the 2×2 matrix
> $$AB = \begin{pmatrix} 1 \times 7 + 2 \times 9 + 3 \times 11 & 1 \times 8 + 2 \times 10 + 3 \times 12 \\ 4 \times 7 + 5 \times 9 + 6 \times 11 & 4 \times 8 + 5 \times 10 + 6 \times 12 \end{pmatrix}$$
> $$= \begin{pmatrix} 58 & 64 \\ 139 & 154 \end{pmatrix}.$$

> ❓ **Question**
>
> Suppose A, B are matrices.
> - Under what conditions are the products AB and BA defined?
> - Under what conditions is the product $B^T A^T$ defined?

Exercise 3.8. Show that if A, B, C are matrices such that the products $A(BC)$ and $(AB)C$ are defined, then
$$A(BC) = (AB)C.$$

Exercise 3.9. Provide examples to show that in general, matrix multiplication is not commutative, that is, $AB \neq BA$ for some matrices A, B.

Exercise 3.10 (**). Show that for any $n \in \mathbb{N}$ with $n \geq 2$, there are $n \times n$ matrices A, B such that
$$AB \neq BA.$$

> ❓ **Question**
>
> Can induction be used for the above exercise?

Exercise 3.11. If A is an $n \times n$ matrix, show that
$$AI_n = I_n A = A$$
holds, where I_n denotes the $n \times n$ diagonal matrix, with all diagonal entries equal to 1, that is,
$$I_n = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}.$$

> **i Remark**
>
> The matrix I_n as above, is called the **identity matrix** of order n, or the $n \times n$ **identity matrix**.

Exercise 3.12. If A is an $m \times n$ matrix, show that
$$I_m A = A,\ AI_n = A$$
where I_m (resp. I_n) denote the $m \times m$ (resp. $n \times n$) identity matrix.

Exercise 3.13. If A is an $m \times n$ matrix with entries in $\mathbb{R}$ and c is a real number, then
$$cA = (cI_n)A = A(cI_m),$$
where I_n (resp. I_m) denotes the $n \times n$ (resp. $m \times m$) identity matrix.

Exercise 3.14. Show that if A, B are matrices, then
$$(AB)^T = B^T A^T.$$

> **i Remark**
>
> Given a square matrix A, that is, a matrix with the same number of rows and columns, the product AA is often denoted by A^2. Similarly, one defines $A^3, A^4, \ldots$.

> **Definition**
>
> Let A be an $n \times n$ matrix, and k be a positive integer. The **k-th power** of A, denoted by A^k, is defined as the product of k copies of A.

Exercise 3.15. Show that if
$$A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix},$$
then
$$A^n = \begin{pmatrix} F_{n-1} & F_n \\ F_n & F_{n+1} \end{pmatrix}$$
for all $n \in \mathbb{N}$ with $n \geq 1$, where F_n denotes the n-th Fibonacci number.

Exercise 3.16. Compute
$$\begin{pmatrix} 0 & \cdots & 0 & 1 \\ 0 & \cdots & 1 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 1 & \cdots & 0 & 0 \end{pmatrix}^2.$$

Exercise 3.17. Show that if A, B, C are matrices such that the sum $B + C$, and the products $AB, AC, A(B+C)$ are defined, then
$$A(B+C) = AB + AC.$$

Exercise 3.18. Show that if A, B, C are matrices such that the sum $B + C$, and the products $AB, AC, A(B + C)$ are defined, then

$$(A + B)C = AC + BC.$$

§3.2.4 Few to many more!

> ✎ **Exercise**
>
> Suppose A is a 3×2 matrix, and
>
> $$A\begin{pmatrix} 1 \\ 4 \end{pmatrix} = \begin{pmatrix} 2 \\ 8 \\ 14 \end{pmatrix},$$
>
> $$A\begin{pmatrix} 3 \\ 7 \end{pmatrix} = \begin{pmatrix} -4 \\ 8 \\ 20 \end{pmatrix}.$$
>
> Determine
>
> $$A\begin{pmatrix} 103 \\ 407 \end{pmatrix}, A\begin{pmatrix} 97 \\ 393 \end{pmatrix}.$$

Solution. Observe that

$$\begin{pmatrix} 103 \\ 407 \end{pmatrix} = 100\begin{pmatrix} 1 \\ 4 \end{pmatrix} + \begin{pmatrix} 3 \\ 7 \end{pmatrix}.$$

This yields

$$\begin{aligned}
A\begin{pmatrix} 103 \\ 407 \end{pmatrix} &= A\left(100\begin{pmatrix} 1 \\ 4 \end{pmatrix} + \begin{pmatrix} 3 \\ 7 \end{pmatrix}\right) \\
&= 100A\begin{pmatrix} 1 \\ 4 \end{pmatrix} + A\begin{pmatrix} 3 \\ 7 \end{pmatrix} \\
&= 100\begin{pmatrix} 2 \\ 8 \\ 14 \end{pmatrix} + \begin{pmatrix} -4 \\ 8 \\ 20 \end{pmatrix} \\
&= \begin{pmatrix} 200 - 4 \\ 800 + 8 \\ 1400 + 20 \end{pmatrix} \\
&= \begin{pmatrix} 196 \\ 808 \\ 1420 \end{pmatrix}.
\end{aligned}$$

Also note that

$$\begin{pmatrix} 97 \\ 393 \end{pmatrix} = 100\begin{pmatrix} 1 \\ 4 \end{pmatrix} - \begin{pmatrix} 3 \\ 7 \end{pmatrix}.$$

Using this we obtain

$$A \begin{pmatrix} 97 \\ 393 \end{pmatrix} = A \left(100 \begin{pmatrix} 1 \\ 4 \end{pmatrix} - \begin{pmatrix} 3 \\ 7 \end{pmatrix} \right)$$

$$= 100 A \begin{pmatrix} 1 \\ 4 \end{pmatrix} - A \begin{pmatrix} 3 \\ 7 \end{pmatrix}$$

$$= 100 \begin{pmatrix} 2 \\ 8 \\ 14 \end{pmatrix} - \begin{pmatrix} -4 \\ 8 \\ 20 \end{pmatrix}$$

$$= \begin{pmatrix} 200 + 4 \\ 800 - 8 \\ 1400 - 20 \end{pmatrix}$$

$$= \begin{pmatrix} 204 \\ 792 \\ 1380 \end{pmatrix}.$$

$\square$

> **i Remark**
>
> The above exercise indicates that if we know how a matrix acts on a bunch of vectors, then we can determine how it acts on any vector that can be expressed as a *linear combination* of those vectors.
>
>   A **linear combination** of vectors $v_1, v_2, ..., v_k$ is a vector of the form
> $$c_1 v_1 + c_2 v_2 + \cdots + c_k v_k,$$
> where $c_1, c_2, ..., c_k$ are real numbers[3].
>
>   Note that any vector in $\mathbb{R}^2$ is a linear combination of the vectors
> $$\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$
>
> So, if A is a 3×2 matrix and we know how A acts on the vectors
> $$\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$
> that is, we know the products
> $$A \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad A \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$
> then we can determine how A acts on any vector in $\mathbb{R}^2$. Indeed, for any $\begin{pmatrix} x \\ y \end{pmatrix}$ in $\mathbb{R}^2$, we have
> $$\begin{pmatrix} x \\ y \end{pmatrix} = x \begin{pmatrix} 1 \\ 0 \end{pmatrix} + y \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$
> and hence we obtain
> $$A \begin{pmatrix} x \\ y \end{pmatrix} = A \left(x \begin{pmatrix} 1 \\ 0 \end{pmatrix} + y \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right)$$
> $$= x A \begin{pmatrix} 1 \\ 0 \end{pmatrix} + y A \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

---

[3]One can also allow the c_i's to be complex numbers, depending on the situation.

> 🧪 **Exercise**
>
> Show that every vector in $\mathbb{R}^2$ can be expressed as a linear combination of the vectors
> $$\begin{pmatrix} 1 \\ 4 \end{pmatrix}, \begin{pmatrix} 3 \\ 7 \end{pmatrix}.$$

Solution. Note that for any $\begin{pmatrix} x \\ y \end{pmatrix}$ in $\mathbb{R}^2$, we have

$$\begin{pmatrix} x \\ y \end{pmatrix} = x \begin{pmatrix} 1 \\ 0 \end{pmatrix} + y \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

that is, $\begin{pmatrix} x \\ y \end{pmatrix}$ is a linear combination of the vectors

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

So, it suffices[4] to show that

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

can be expressed as linear combinations of the vectors

$$\begin{pmatrix} 1 \\ 4 \end{pmatrix}, \begin{pmatrix} 3 \\ 7 \end{pmatrix}.$$

Note that

$$7 \begin{pmatrix} 1 \\ 4 \end{pmatrix} - 4 \begin{pmatrix} 3 \\ 7 \end{pmatrix} = \begin{pmatrix} -5 \\ 0 \end{pmatrix},$$

$$3 \begin{pmatrix} 1 \\ 4 \end{pmatrix} - 1 \begin{pmatrix} 3 \\ 7 \end{pmatrix} = \begin{pmatrix} 0 \\ 5 \end{pmatrix},$$

which yields

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} = -\frac{7}{5} \begin{pmatrix} 1 \\ 4 \end{pmatrix} + \frac{4}{5} \begin{pmatrix} 3 \\ 7 \end{pmatrix},$$

$$\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \frac{3}{5} \begin{pmatrix} 1 \\ 4 \end{pmatrix} - \frac{1}{5} \begin{pmatrix} 3 \\ 7 \end{pmatrix}.$$

$\square$

> ℹ️ **Remark**
>
> A careful reading of the above exercises and the remark indicates that if there are a few vectors $v_1, v_2, ..., v_k$ in $\mathbb{R}^n$ such that every vector in $\mathbb{R}^n$ can be expressed as a linear combination of those vectors, then knowing how a matrix A acts on the vectors $v_1, v_2, ..., v_k$ is sufficient to determine how A acts on any vector in $\mathbb{R}^n$.

---

[4] How does it suffice?

Example

Note that

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 4 \\ 7 \end{pmatrix},$$

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 5 \\ 8 \end{pmatrix},$$

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 6 \\ 9 \end{pmatrix}$$

hold. The vectors

$$e_1 := \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, e_2 := \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, e_3 := \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

are called the **standard basis vectors** of $\mathbb{R}^3$. The standard basis vectors of $\mathbb{R}^2$ are

$$e_1 := \begin{pmatrix} 1 \\ 0 \end{pmatrix}, e_2 := \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

In general, the vectors

$$e_1 := \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, e_2 := \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, e_3 := \begin{pmatrix} 0 \\ 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, ..., e_n := \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

lying in $\mathbb{R}^n$ are called the **standard basis vectors** of $\mathbb{R}^n$.

Example

Note that

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \\ 5 \end{pmatrix},$$

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 4 \\ 6 \end{pmatrix}$$

hold. The vectors

$$e_1 := \begin{pmatrix} 1 \\ 0 \end{pmatrix}, e_2 := \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

are called the **standard basis vectors** of $\mathbb{R}^2$.

Exercise 3.19. If A is an $m \times n$ matrix, then show that for all $i = 1, 2, ..., n$, the product Ae_i is equal to the i-th column of A, where e_i denotes the i-th standard basis vector of $\mathbb{R}^n$, that is,

$$e_i = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix}$$

with 1 in the i-th position, and 0's elsewhere.

ⓘ Remark

Exercise 3.19 shows that the columns of a matrix A can be obtained by multiplying A with the standard basis vectors. Further, if A is an $m \times n$ matrix, and we know how A acts on the standard basis vectors of $\mathbb{R}^n$, then we can determine all the entries of A, and hence we can determine how A acts on any vector v in $\mathbb{R}^n$. Note that this is achieved by determining the entries of A.

Moreover, the first remark in Section 3.2.4 indicates that if we know how A acts on the standard basis vectors of $\mathbb{R}^n$, then using that every element of $\mathbb{R}^n$ can be expressed as a linear combination of the standard basis vectors, we can determine how A acts on any vector v in $\mathbb{R}^n$.

Furthermore, if $v_1, v_2, ..., v_k$ are elements of $\mathbb{R}^n$, then knowing how a matrix A acts on $v_1, v_2, ..., v_k$ is sufficient to determine how A acts on any vector which can be expressed as a linear combination of $v_1, ..., v_k$.

§3.3 System of linear equations

🧪 Exercise

Solve the system of equations in the variables x, y:
$$2x + 3y = 8,$$
$$5x + 7y = 19.$$

Solution. Multiplying the first equation by 7, and the second equation by 3, and taking the difference of the equations obtained, we get
$$7 \cdot 2x - 3 \cdot 5x = 7 \cdot 8 - 3 \cdot 19,$$
which yields
$$x = \frac{7 \cdot 8 - 3 \cdot 19}{7 \cdot 2 - 3 \cdot 5}.$$

Similarly, multiplying the first equation by 5, and the second equation by 2, and taking the difference of the equations obtained, we get
$$5 \cdot 3y - 2 \cdot 7y = 5 \cdot 8 - 2 \cdot 19,$$
which yields
$$y = \frac{5 \cdot 8 - 2 \cdot 19}{5 \cdot 3 - 2 \cdot 7}.$$

$\square$

Alternate solution. Note that the given system of equations can be rewritten as
$$\begin{pmatrix} 2 & 3 \\ 5 & 7 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 8 \\ 19 \end{pmatrix}, \tag{2}$$
which can be thought as a concise way of putting the information

$$\begin{pmatrix} 2x + 3y \\ 5x + 7y \end{pmatrix} = \begin{pmatrix} 8 \\ 19 \end{pmatrix}. \tag{3}$$

Recall that in the previous solution,

- to solve for x, the first equation was multiplied by 7, and the second equation by 3, and next, we considered the difference of the equations obtained,
- to solve for y, the first equation was multiplied by 5, and the second equation was multiplied by 2, and then, we considered the difference of the equations obtained.

The two steps can be put together in the matrix form as

$$\begin{pmatrix} 7 & -3 \\ 5 & -2 \end{pmatrix} \begin{pmatrix} 2x + 3y \\ 5x + 7y \end{pmatrix}.$$

Using Equation 3, it yields

$$\begin{pmatrix} 7 & -3 \\ 5 & -2 \end{pmatrix} \begin{pmatrix} 2x + 3y \\ 5x + 7y \end{pmatrix} = \begin{pmatrix} 7 & -3 \\ 5 & -2 \end{pmatrix} \begin{pmatrix} 8 \\ 19 \end{pmatrix}.$$

The advantage of this process is that, first y gets eliminated in the first step and helps to find x, and in the next step, x gets eliminated and leads to finding y. Stated in matrix form, we have that

$$\begin{pmatrix} 7 & -3 \\ 5 & -2 \end{pmatrix} \begin{pmatrix} 2x + 3y \\ 5x + 7y \end{pmatrix} = \begin{pmatrix} (7 \cdot 2 - 3 \cdot 5)x \\ (5 \cdot 3 - 2 \cdot 7)y \end{pmatrix}.$$

We could have also written

$$\begin{pmatrix} 7 & -3 \\ -5 & 2 \end{pmatrix} \begin{pmatrix} 2x + 3y \\ 5x + 7y \end{pmatrix} = \begin{pmatrix} (7 \cdot 2 - 3 \cdot 5)x \\ (-5 \cdot 3 + 2 \cdot 7)y \end{pmatrix}.$$

Thus, rewriting the given sysem of equations in the matrix form as in Equation 3, and multiplying it from the left by the matrix

$$\begin{pmatrix} 7 & -3 \\ -5 & 2 \end{pmatrix},$$

we obtain

$$\begin{pmatrix} 7 & -3 \\ -5 & 2 \end{pmatrix} \begin{pmatrix} 2x + 3y \\ 5x + 7y \end{pmatrix} = \begin{pmatrix} 7 & -3 \\ -5 & 2 \end{pmatrix} \begin{pmatrix} 8 \\ 19 \end{pmatrix},$$

which gives

$$\begin{pmatrix} (7 \cdot 2 - 3 \cdot 5)x \\ (-5 \cdot 3 + 2 \cdot 7)y \end{pmatrix} = \begin{pmatrix} 7 & -3 \\ -5 & 2 \end{pmatrix} \begin{pmatrix} 8 \\ 19 \end{pmatrix},$$

that is, we have

$$(7 \cdot 2 - 3 \cdot 5) \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 7 & -3 \\ -5 & 2 \end{pmatrix} \begin{pmatrix} 8 \\ 19 \end{pmatrix},$$

and hence, the solution to the given system of equations is given by

$$\begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{7 \cdot 2 - 3 \cdot 5} \begin{pmatrix} 7 & -3 \\ -5 & 2 \end{pmatrix} \begin{pmatrix} 8 \\ 19 \end{pmatrix}.$$

□

Summarizing the above. The given system of equations can be rewritten as

$$\begin{pmatrix} 2x + 3y \\ 5x + 7y \end{pmatrix} = \begin{pmatrix} 8 \\ 19 \end{pmatrix}, \tag{4}$$

or equivalently,

$$\begin{pmatrix} 2 & 3 \\ 5 & 7 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 8 \\ 19 \end{pmatrix}, \tag{5}$$

Multiplying the above from the left by the matrix

$$\frac{1}{7 \cdot 2 - 3 \cdot 5} \begin{pmatrix} 7 & -3 \\ -5 & 2 \end{pmatrix}, \tag{6}$$

we obtain

$$\frac{1}{7 \cdot 2 - 3 \cdot 5} \begin{pmatrix} 7 & -3 \\ -5 & 2 \end{pmatrix} \begin{pmatrix} 2x + 3y \\ 5x + 7y \end{pmatrix} = \frac{1}{7 \cdot 2 - 3 \cdot 5} \begin{pmatrix} 7 & -3 \\ -5 & 2 \end{pmatrix} \begin{pmatrix} 8 \\ 19 \end{pmatrix},$$

which shows that the solution to the given system of equations is given by

$$\begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{7 \cdot 2 - 3 \cdot 5} \begin{pmatrix} 7 & -3 \\ -5 & 2 \end{pmatrix} \begin{pmatrix} 8 \\ 19 \end{pmatrix}.$$

□

> **ⓘ Remark**
>
> In the above, Equation 5 was multiplied by the matrix as in Equation 6, since multiplying the matrix
>
> $$\begin{pmatrix} 2 & 3 \\ 5 & 7 \end{pmatrix}$$
>
> as in Equation 5 **from the left** by the matrix
>
> $$\frac{1}{7 \cdot 2 - 3 \cdot 5} \begin{pmatrix} 7 & -3 \\ -5 & 2 \end{pmatrix},$$
>
> as in Equation 6 yields
>
> $$\frac{1}{7 \cdot 2 - 3 \cdot 5} \begin{pmatrix} 7 & -3 \\ -5 & 2 \end{pmatrix} \begin{pmatrix} 2 & 3 \\ 5 & 7 \end{pmatrix} = \frac{1}{7 \cdot 2 - 3 \cdot 5} \begin{pmatrix} 7 \cdot 2 - 3 \cdot 5 & 0 \\ 0 & -5 \cdot 3 + 2 \cdot 7 \end{pmatrix}$$
> $$= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$
>
> which is the 2×2 identity matrix.

§3.4 Invertible matrices

> **🗎 Definition**
>
> Let A be an $n \times n$ matrix. If there is an $n \times n$ matrix B such that
> $$AB = BA = I_n,$$
> then B is called an **inverse** of A. If A admits an inverse, then A is called **invertible**.

Fact 3.20. Show that if B, C are inverses of an $n \times n$ matrix A, then $B = C$.

A proof of the above is provided in Chapter 6.

> **ⓘ Remark**
>
> If A is an $n \times n$ matrix, and A admits an inverse, then it is called **the** inverse of A, and is denoted by A^{-1}.

Exercise 3.21. Let A, B be invertible $n \times n$ matrices. Show that AB is invertible, and that

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

Compare the above exercise with Exercise 3.14.

Solution. Note that

$$\begin{aligned}
(AB)(B^{-1}A^{-1}) &= A(BB^{-1})A^{-1} \\
&= AI_nA^{-1} \\
&= AA^{-1} \\
&= I_n,
\end{aligned}$$

and

$$\begin{aligned}
(B^{-1}A^{-1})(AB) &= B^{-1}(A^{-1}A)B \\
&= B^{-1}I_nB \\
&= B^{-1}B \\
&= I_n.
\end{aligned}$$

This shows that AB is invertible, and that

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

$\square$

Lemma 3.22. *Let*

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

be a 2×2 matrix. The following statements are equivalent.

- *A is invertible,*
- *$ad - bc$ is nonzero,*
- *every vector $\begin{pmatrix} e \\ f \end{pmatrix}$ can be expressed as a linear combination of the columns of A.*

Moreover, if A is invertible, then its inverse is equal to

$$\frac{1}{ad - bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Proof. Note that

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = \begin{pmatrix} ad - bc & 0 \\ 0 & ad - bc \end{pmatrix},$$

$$\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} ad - bc & 0 \\ 0 & ad - bc \end{pmatrix}$$

hold. Assume that if A is invertible. It follows that

$$(ad - bc)A^{-1} = \begin{pmatrix} ad - bc & 0 \\ 0 & ad - bc \end{pmatrix} A^{-1}$$

$$= \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} A^{-1}$$

$$= \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} AA^{-1}$$

$$= \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} I_2$$

$$= \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Note that if $ad - bc = 0$, then a, b, c, d are equal to 0, which implies $A = 0$, which is impossible since $AA^{-1} = I_2$. Hence, $ad - bc$ is nonzero, and

$$A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Conversely, note that if $ad - bc \neq 0$, then it follows that the matrix

$$\frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

is the inverse[5] of A.

Also note that if every vector $\begin{pmatrix} e \\ f \end{pmatrix}$ can be expressed as a linear combination of the columns of A, then in particular, the vectors $\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ can be expressed as a linear combination of the columns of A. This implies that there are real numbers x_1, y_1, x_2, y_2 such that

$$A \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix},$$

and

$$A \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

This shows that

$$A \begin{pmatrix} x_1 & x_2 \\ y_1 & y_2 \end{pmatrix} = I_2,$$

implying that A is invertible[6]. Further, if $ad - bc$ is nonzero, then as observed above, it follows that

$$\frac{d}{ad - bc} \begin{pmatrix} a \\ c \end{pmatrix} + \frac{-c}{ad - bc} \begin{pmatrix} b \\ d \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix},$$

$$\frac{-b}{ad - bc} \begin{pmatrix} a \\ c \end{pmatrix} + \frac{a}{ad - bc} \begin{pmatrix} b \\ d \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

hold, which shows that every vector $\begin{pmatrix} e \\ f \end{pmatrix}$ can be expressed as a linear combination of the columns of A. $\square$

---

[5]Why?
[6]How?

> 📑 **Definition**
>
> If $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is a 2×2 matrix, then its **determinant** is denoted by $\det A$, and is defined as
> $$\det A := ad - bc.$$

> 🅘 **Remark**
>
> The determinant of a 2×2 matrix can be used to determine whether the matrix is invertible or not. Indeed, by Lemma 3.22, the following statements are equivalent.
> - The matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is invertible.
> - The determinant $\det\left(\begin{pmatrix} a & b \\ c & d \end{pmatrix}\right)$ is nonzero.
>
> Moreover, if $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is invertible, then its inverse is given by
> $$\frac{1}{\det\left(\begin{pmatrix} a & b \\ c & d \end{pmatrix}\right)} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Exercise 3.23. If A, B are 2×2 matrices, then show that
$$\det(AB) = \det(A)\det(B).$$

Solution. Let
$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$
and
$$B = \begin{pmatrix} e & f \\ g & h \end{pmatrix}$$
be 2×2 matrices. Note that
$$AB = \begin{pmatrix} ae + bg & af + bh \\ ce + dg & cf + dh \end{pmatrix}.$$
Using the definition of determinant, we obtain
$$\begin{aligned}
\det(AB) &= (ae + bg)(cf + dh) - (af + bh)(ce + dg) \\
&= acef + adeh + bcfg + bdgh - acef - adfg - bceh - bdgh \\
&= adeh - adfg + bcfg - bceh \\
&= (ad - bc)(eh - fg) \\
&= \det(A)\det(B).
\end{aligned}$$

$\square$

Exercise 3.24. Solve the system of equations in the variables x, y:
$$\begin{aligned}
5x + 2y &= 11, \\
3x + 4y &= 8.
\end{aligned}$$

Solution. Note that the given system of equations can be rewritten as
$$\begin{pmatrix} 5 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 11 \\ 8 \end{pmatrix}, \tag{7}$$
Note that

$$\det\left(\begin{pmatrix} 5 & 2 \\ 3 & 4 \end{pmatrix}\right) = 5 \cdot 4 - 2 \cdot 3 = 14,$$

which is nonzero. Hence, the matrix $\begin{pmatrix} 5 & 2 \\ 3 & 4 \end{pmatrix}$ is invertible, and

$$\begin{pmatrix} 5 & 2 \\ 3 & 4 \end{pmatrix}^{-1} = \frac{1}{14}\begin{pmatrix} 4 & -2 \\ -3 & 5 \end{pmatrix}.$$

Multiplying Equation 7 from the left by the inverse of the matrix $\begin{pmatrix} 5 & 2 \\ 3 & 4 \end{pmatrix}$, we obtain

$$\begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{14}\begin{pmatrix} 4 & -2 \\ -3 & 5 \end{pmatrix}\begin{pmatrix} 11 \\ 8 \end{pmatrix},$$

which yields

$$\begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{14}\begin{pmatrix} 44 - 16 \\ -33 + 40 \end{pmatrix} = \begin{pmatrix} 2 \\ \frac{1}{2} \end{pmatrix}.$$

$\square$

Exercise 3.25. Solve the system of equations in the variables x, y:
$$12x - 25y = -47,$$
$$-7x + 30y = 51.$$

Solution. The given system of equations can be put in the matrix form as

$$\begin{pmatrix} 12 & -25 \\ -7 & 30 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -47 \\ 51 \end{pmatrix}, \tag{8}$$

Note that

$$\det\left(\begin{pmatrix} 12 & -25 \\ -7 & 30 \end{pmatrix}\right) = 12 \cdot 30 - (-25) \cdot (-7) = 360 - 175 = 185,$$

which is nonzero. It follows that the matrix $\begin{pmatrix} 12 & -25 \\ -7 & 30 \end{pmatrix}$ is invertible, and

$$\begin{pmatrix} 12 & -25 \\ -7 & 30 \end{pmatrix}^{-1} = \frac{1}{185}\begin{pmatrix} 30 & 25 \\ 7 & 12 \end{pmatrix}.$$

Multiplying Equation 8 from the left by the inverse of the matrix $\begin{pmatrix} 12 & -25 \\ -7 & 30 \end{pmatrix}$, we obtain

$$\begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{185}\begin{pmatrix} 30 & 25 \\ 7 & 12 \end{pmatrix}\begin{pmatrix} -47 \\ 51 \end{pmatrix},$$

which yields

$$\begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{185}\begin{pmatrix} -1410 + 1275 \\ -329 + 612 \end{pmatrix} = \begin{pmatrix} -\frac{135}{185} \\ \frac{283}{185} \end{pmatrix} = \begin{pmatrix} -\frac{27}{37} \\ \frac{283}{185} \end{pmatrix}.$$

$\square$

Exercise 3.26. Solve the system of equations in the variables x, y, z:
$$2x + 3y + z = 1,$$
$$4x + y + 2z = 2,$$
$$3x + 2y + 3z = 3.$$

Solution. The given system of equations can be rewritten as

$$A\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \tag{9}$$

where

$$A = \begin{pmatrix} 2 & 3 & 1 \\ 4 & 1 & 2 \\ 3 & 2 & 3 \end{pmatrix}.$$

Note that A^{-1} is equal to[7]

$$\frac{1}{15} \begin{pmatrix} 1 & 7 & -5 \\ 6 & -3 & 0 \\ -5 & -5 & 10 \end{pmatrix}.$$

Indeed, observe that

$$A \times \frac{1}{15} \begin{pmatrix} 1 & 7 & -5 \\ 6 & -3 & 0 \\ -5 & -5 & 10 \end{pmatrix} = \frac{1}{15} \begin{pmatrix} 2 & 3 & 1 \\ 4 & 1 & 2 \\ 3 & 2 & 3 \end{pmatrix} \begin{pmatrix} 1 & 7 & -5 \\ 6 & -3 & 0 \\ -5 & -5 & 10 \end{pmatrix}$$

$$= \frac{1}{15} \begin{pmatrix} 15 & 0 & 0 \\ 0 & 15 & 0 \\ 0 & 0 & 15 \end{pmatrix}$$

$$= I_3,$$

and

$$\frac{1}{15} \begin{pmatrix} 1 & 7 & -5 \\ 6 & -3 & 0 \\ -5 & -5 & 10 \end{pmatrix} \times A = \frac{1}{15} \begin{pmatrix} 1 & 7 & -5 \\ 6 & -3 & 0 \\ -5 & -5 & 10 \end{pmatrix} \begin{pmatrix} 2 & 3 & 1 \\ 4 & 1 & 2 \\ 3 & 2 & 3 \end{pmatrix}$$

$$= \frac{1}{15} \begin{pmatrix} 15 & 0 & 0 \\ 0 & 15 & 0 \\ 0 & 0 & 15 \end{pmatrix}$$

$$= I_3.$$

Multiplying Equation 9 from the left by A^{-1}, we obtain

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = A^{-1} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix},$$

which yields

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \frac{1}{15} \begin{pmatrix} 1 & 7 & -5 \\ 6 & -3 & 0 \\ -5 & -5 & 10 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \frac{1}{15} \begin{pmatrix} 1 + 14 - 15 \\ 6 - 6 + 0 \\ -5 - 10 + 30 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

□

Exercise 3.27. Let A, B be 2×2 matrices. Show that if AB is invertible, then both A and B are invertible.

Compare the above with Exercise 3.70, Exercise 3.71, Exercise 3.87.

Exercise 3.28. Let A be an $n \times n$ matrix. If P is an invertible $n \times n$ matrix, then show that

$$\left(PAP^{-1}\right)^k = PA^kP^{-1}$$

holds for any positive integer k.

---

[7]One may wonder how to find the inverse of an *invertible* 3×3 matrix. One way is to use the Gaussian elimination method, to be discussed in Section 3.7. Can one also have a formula for the inverse of an *invertible* 3×3 matrix, similar to the formula for the inverse of a 2×2 matrix as in Lemma 3.22? See Lemma 3.96 for the details.

Solution. We prove the statement by induction. Let P be an invertible $n \times n$ matrix. For a positive integer k, let $P(k)$ denote the statement

$$(PAP^{-1})^k = PA^kP^{-1}.$$

Note that $P(1)$ holds. Assume that $P(k)$ holds for some positive integer k. Then, using the induction hypothesis, we obtain

$$\begin{aligned}
(PAP^{-1})^{k+1} &= (PAP^{-1})^k(PAP^{-1}) \\
&= (PA^kP^{-1})(PAP^{-1}) \\
&= PA^k(P^{-1}P)AP^{-1} \\
&= PA^kI_nAP^{-1} \\
&= PA^{k+1}P^{-1}.
\end{aligned}$$

This shows that $P(k+1)$ holds. Hence, by the principle of mathematical induction, $P(k)$ holds for all positive integers k. □

§3.5 Systems of linear equations again

§3.5.1 Linear combination of the columns of a matrix

Here are further examples of systems of linear equations.

System of linear equations	Matrix form	Solution(s)	Associated matrix
$x + 2y = 1,$ $3x - 5y = -7$	$\begin{pmatrix} 1 & 2 \\ 3 & -5 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ -7 \end{pmatrix}$	Unique	Invertible
$x + 2y = 1,$ $2x + 4y = 2$	$\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$	Infinitely many[8]	Not invertible
$x + 2y = 1,$ $3x + 6y = 2$	$\begin{pmatrix} 1 & 2 \\ 3 & 6 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$	None	Not invertible

> **? Question**
>
> Which of such systems of equations can be solved?

---

[8]Why?

> **ℹ Remark**
>
> Consider the following system of linear equations.
> $$12345x + 67890y = 11111,$$
> $$54321x + 9876y = 22222.$$
> To solve it, we need to find a vector $\begin{pmatrix} x \\ y \end{pmatrix}$ in $\mathbb{R}^2$ satisfying
> $$\begin{pmatrix} 12345 & 67890 \\ 54321 & 9876 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 11111 \\ 22222 \end{pmatrix}.$$
> In other words, we need to find real numbers x, y such that
> $$\begin{pmatrix} 12345 & 67890 \\ 54321 & 9876 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 12345 & 67890 \\ 54321 & 9876 \end{pmatrix} (xe_1 + ye_2)$$
> $$= x \begin{pmatrix} 12345 & 67890 \\ 54321 & 9876 \end{pmatrix} e_1 + y \begin{pmatrix} 12345 & 67890 \\ 54321 & 9876 \end{pmatrix} e_2$$
> $$= x \begin{pmatrix} 12345 \\ 54321 \end{pmatrix} + y \begin{pmatrix} 67890 \\ 9876 \end{pmatrix}$$
> is equal to $\begin{pmatrix} 11111 \\ 22222 \end{pmatrix}$, where e_1, e_2 are the standard basis vectors of $\mathbb{R}^2$.
>
> It is left to determining whether some linear combination of the columns of the matrix $\begin{pmatrix} 12345 & 67890 \\ 54321 & 9876 \end{pmatrix}$ is equal to $\begin{pmatrix} 11111 \\ 22222 \end{pmatrix}$. If such a linear combination exists, then the system of equations has a solution, and if such a linear combination does not exist, then the system of equations does not have a solution. If there is more than one such linear combination, then the system of equations has more than one solution, etc.

> **ℹ Remark**
>
> The above discussion shows that solving a system of linear equations is equivalent to determining whether **some linear combination** of the columns of a matrix is equal to the given vector. This is the reason why the notions[9] of *column space* and *linear span of vectors* are important in the study of systems of linear equations.

> **ℹ Remark**
>
> Consider a matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$, and a vector $\begin{pmatrix} \alpha \\ \beta \end{pmatrix}$ where $a, b, c, d, \alpha, \beta$ are real numbers. If $\begin{pmatrix} \alpha \\ \beta \end{pmatrix}$ is **not** a linear combination of the columns of $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$, that is, if $\begin{pmatrix} \alpha \\ \beta \end{pmatrix}$ is **not** a linear combination of the vectors $\begin{pmatrix} a \\ c \end{pmatrix}$ and $\begin{pmatrix} b \\ d \end{pmatrix}$, then the system of equations
> $$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$$
> admits **no** solution.

Lemma 3.29. *Let A be a 2×2 matrix, and let $\begin{pmatrix} \alpha \\ \beta \end{pmatrix}$ be a vector of $\mathbb{R}^2$. Show that the following statements are equivalent.*
- *The vector $\begin{pmatrix} \alpha \\ \beta \end{pmatrix}$ is a linear combination of the columns of the matrix A.*
- *The system of equations*

---

[9]The notions of column space and linear span of vectors will be introduced formally at a later stage.

$$A\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$$

admits at least one solution.

Proof. Assume that $\begin{pmatrix} \alpha \\ \beta \end{pmatrix}$ is a linear combination of the columns of the matrix A. This shows that there are real numbers x_0, y_0 such that

$$A\begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}.$$

Hence, the system of equations

$$A\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$$

admits at least one solution.

Conversely, assume that the system of equations

$$A\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$$

admits at least one solution, say $\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}$. It follows that

$$A\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}.$$

This shows that $\begin{pmatrix} \alpha \\ \beta \end{pmatrix}$ is a linear combination of the columns of the matrix A. □

§3.5.2 Homogeneous system of linear equations

Let us first consider the solutions to the system of equations

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{10}$$

Note that $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ is a linear combination of the columns of $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$, since

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix} = 0\begin{pmatrix} a \\ c \end{pmatrix} + 0\begin{pmatrix} b \\ d \end{pmatrix}.$$

This shows that $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ is a solution to Equation 10. It is called the **trivial solution** to Equation 10.

Case 1

Let us first consider the case that the matrix A **is** invertible.

Note that under this hypothesis, Equation 10 has no solution other than $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$. Indeed, if $\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}$ is another solution to Equation 10, then

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

This shows that

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1}\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

which yields

$$\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Hence, if A is invertible, then the only solution to Equation 10 is $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$.

Observation

If Equation 10 has at least one solution other than $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$, then Equation 10 has infinitely many solutions. Indeed, if $\begin{pmatrix} x_2 \\ y_2 \end{pmatrix}$ is a solution to Equation 10 other than $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$, then for any real number t, the vector $t\begin{pmatrix} x_2 \\ y_2 \end{pmatrix}$ is also a solution to Equation 10, and hence, Equation 10 has infinitely many solutions.

Case 2

Now let us consider the case that the matrix A is **not** invertible. By Lemma 3.22, we have $ad - bc = 0$, this implies that

$$d\begin{pmatrix} a \\ c \end{pmatrix} - c\begin{pmatrix} b \\ d \end{pmatrix} = \begin{pmatrix} ad - bc \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

$$b\begin{pmatrix} a \\ c \end{pmatrix} - a\begin{pmatrix} b \\ d \end{pmatrix} = \begin{pmatrix} 0 \\ bc - ad \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

and consequently,

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} d \\ -c \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} b \\ -a \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

This shows that if A is **not** invertible, then the vectors $\begin{pmatrix} d \\ -c \end{pmatrix}$ and $\begin{pmatrix} b \\ -a \end{pmatrix}$ are solutions to Equation 10.

Subcase 2a

If the vectors $\begin{pmatrix} d \\ -c \end{pmatrix}$ and $\begin{pmatrix} b \\ -a \end{pmatrix}$ are equal to the zero vector, equivalently, if a, b, c, d are all equal to 0, then any vector $\begin{pmatrix} x \\ y \end{pmatrix}$ of $\mathbb{R}^2$ is a solution to Equation 10.

Subcase 2b

If not both of the vectors $\begin{pmatrix} d \\ -c \end{pmatrix}$ and $\begin{pmatrix} b \\ -a \end{pmatrix}$ are equal to the zero vector (equivalently, at least one of them is nonzero) and A is not invertible, then Equation 10 admits a solution other than $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$, and moreover, not all vectors of $\mathbb{R}^2$ are solutions to Equation 10 since not all of $\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ are solutions to Equation 10.

Lemma 3.30. *Consider the system of linear equations*

$$A\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \tag{11}$$

where A is a 2×2 matrix. The following statements are equivalent.
- *The matrix A is invertible.*
- *The system of equations as in Equation 11 admits a unique solution, that is, it admits no solution other than the trivial solution $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$.*
- *The trivial linear combination (that is, the linear combination using zeroes as the coefficients) of the columns of the matrix A is the only linear combination of the columns of A that is equal to $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$.*

Furthermore, the following statements are equivalent.
- *Any element of $\mathbb{R}^2$ is a solution to Equation 11.*
- *The matrix A is the zero matrix.*

Moreover, the following statements are equivalent too.
- *The system of equations as in Equation 11 admits a solution other than the trivial solution $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$, and not all elements of $\mathbb{R}^2$ are solutions to Equation 11.*
- *The matrix A is not invertible and nonzero.*

A proof of the above is provided in Chapter 6.

Exercise 3.31. Let A be a 2×2 matrix. Let V be the set of solutions to the system of equations

$$A \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

that is,

$$V = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 : A \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\}.$$

Show that for any $u, v \in V$ and $c \in \mathbb{R}$, the element $u + v$ of $\mathbb{R}^2$ lies in V, and the element cu also lies in V.

Solution. Let u, v be elements of V, and let c be a real number. Note that

$$A(u + v) = Au + Av = \begin{pmatrix} 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

and

$$A(cu) = c(Au) = c \begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

This shows that both $u + v$ and cu lie in V. $\square$

> **i** **Remark**
>
> What about the difference of two elements of V? If u, v are elements of V and c, d are real numbers, then what can be said about the element $cu + dv$ of $\mathbb{R}^2$?

Exercise 3.32. Let A be a 2×2 matrix. Let V be the set of vectors which can be expressed as a linear combination of the columns of A, that is[10],

$$V = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 : \begin{pmatrix} x \\ y \end{pmatrix} = A \begin{pmatrix} s \\ t \end{pmatrix} \text{ for some } s, t \in \mathbb{R} \right\}.$$

Show that for any $u, v \in V$ and $c \in \mathbb{R}$, the element $u + v$ of $\mathbb{R}^2$ lies in V, and the element cu also lies in V.

Solution. Let u, v be elements of V, and let c be a real number. By the definition of V, there exist real numbers s_1, t_1, s_2, t_2 such that

---

[10]Observe that

$$A \begin{pmatrix} s \\ t \end{pmatrix} = sC_1 + tC_2,$$

where C_1, C_2 denote the first and second columns of A respectively.

$$u = A\begin{pmatrix} s_1 \\ t_1 \end{pmatrix},$$

$$v = A\begin{pmatrix} s_2 \\ t_2 \end{pmatrix}.$$

Note that

$$u + v = A\begin{pmatrix} s_1 \\ t_1 \end{pmatrix} + A\begin{pmatrix} s_2 \\ t_2 \end{pmatrix} = A\left(\begin{pmatrix} s_1 \\ t_1 \end{pmatrix} + \begin{pmatrix} s_2 \\ t_2 \end{pmatrix}\right) = A\begin{pmatrix} s_1 + s_2 \\ t_1 + t_2 \end{pmatrix},$$

and

$$cu = c\left(A\begin{pmatrix} s_1 \\ t_1 \end{pmatrix}\right) = A\left(c\begin{pmatrix} s_1 \\ t_1 \end{pmatrix}\right) = A\begin{pmatrix} cs_1 \\ ct_1 \end{pmatrix}.$$

This shows that both $u + v$ and cu lie in V. $\qquad\square$

Exercise 3.33. Let A be an $m \times n$ matrix. Let V be the set of solutions to the system of equations
$$Au = 0,$$
that is,
$$V = \{v \in \mathbb{R}^n : Av = 0\}.$$
Show that for any $u, v \in V$ and $c, d \in \mathbb{R}$, the element $cu + dv$ of $\mathbb{R}^m$ lies in V.

Solution. Let u, v be elements of V, and let c, d be real numbers. Note that
$$A(cu + dv) = c(Au) + d(Av) = c \cdot 0 + d \cdot 0 = 0.$$
This shows that $cu + dv$ lies in V. $\qquad\square$

Compare the following exercise with .

Exercise 3.34. Let A be an $m \times n$ matrix. Let V be the set of vectors which can be expressed as a linear combination of the columns of A, that is,
$$V = \{v \in \mathbb{R}^m : v = Au \text{ for some } u \in \mathbb{R}^n\}.$$
Show that for any $u, v \in V$ and $c, d \in \mathbb{R}$, the element $cu + dv$ of $\mathbb{R}^m$ lies in V.

Solution. Let u, v be elements of V, and let c, d be real numbers. By the definition of V, there exist vectors $u_1, u_2 \in \mathbb{R}^n$ such that
$$u = Au_1,$$
$$v = Au_2.$$

Note that
$$cu + dv = c(Au_1) + d(Au_2) = A(cu_1) + A(du_2) = A(cu_1 + du_2).$$
This shows that $cu + dv$ lies in V. $\qquad\square$

§3.5.3 General system of linear equations

If $\begin{pmatrix} \alpha \\ \beta \end{pmatrix}$ is a linear combination of the columns of the matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$, then there exist real numbers x_0, y_0 such that

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}.$$

This shows that the system of equations

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$$

has at least one solution, namely $\begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$. If $\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}$ is another solution to the above system of equations, then

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}.$$

This shows that

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}.$$

Hence,

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\left(\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} - \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}\right) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

holds. This shows that the difference of any two solutions to the system of equations

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$$

is a solution to the system of equations

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Conversely, if $\begin{pmatrix} x_2 \\ y_2 \end{pmatrix}$ is a solution to the system of equations

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

then

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\left(\begin{pmatrix} x_0 \\ y_0 \end{pmatrix} + \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}\right) = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x_0 \\ y_0 \end{pmatrix} + \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x_2 \\ y_2 \end{pmatrix}$$
$$= \begin{pmatrix} \alpha \\ \beta \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$
$$= \begin{pmatrix} \alpha \\ \beta \end{pmatrix}.$$

This shows that if $\begin{pmatrix} x_2 \\ y_2 \end{pmatrix}$ is a solution to the system of equations

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

then $\begin{pmatrix} x_0 \\ y_0 \end{pmatrix} + \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}$ is also a solution to the system of equations

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}.$$

In summary, if $\begin{pmatrix} \alpha \\ \beta \end{pmatrix}$ is a linear combination of the columns of $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$, then the solutions to the system of equations

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$$

are in one-to-one correspondence with the solutions to the system of equations

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

> **i** Remark
>
> The above discussion shows that if $\begin{pmatrix} \alpha \\ \beta \end{pmatrix}$ is a linear combination of the columns of $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$, then the solutions to the system of equations
>
> $$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$$
>
> are in one-to-one correspondence with the solutions to the system of equations
>
> $$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$
>
> Hence, to determine the number of solutions to the system of equations
>
> $$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$$
>
> it suffices to determine the number of solutions to the system of equations
>
> $$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Lemma 3.35. *Let A be a 2×2 matrix, and let $\begin{pmatrix} \alpha \\ \beta \end{pmatrix}$ be a vector of $\mathbb{R}^2$. Assume that $\begin{pmatrix} \alpha \\ \beta \end{pmatrix}$ is a linear combination of the columns of the matrix A. Then the solutions to the system of equations*

$$A \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$$

are in one-to-one correspondence with the solutions to the system of equations

$$A \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

More specifically, for any fixed solution $\begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$ to the system of equations

$$A \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix},$$

the map

$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x \\ y \end{pmatrix} - \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$$

is a bijection from the set of solutions to the system of equations

$$A \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$$

to the set of solutions to the system of equations

$$A \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

having the following map as its inverse:

$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}.$$

Proof. Deferred. □

> **i Remark**
>
> The terms "one-to-one correspondence" and "bijection" will be introduced formally at a later stage. The proof of the above lemma is deferred until then.

§3.6 Special types of matrices

§3.6.1 Diagonal, scalar, and triangular matrices

Recall that a matrix is called a **square** matrix if it has the same number of rows and columns.

> **Definition**
>
> A square matrix A is called **diagonal** if all its non-diagonal entries are equal to zero.

> **Example**
>
> The following matrices are diagonal.
> $$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}, \begin{pmatrix} 5 & 0 \\ 0 & -7 \end{pmatrix}, \begin{pmatrix} i & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \pi \end{pmatrix}.$$
> Some non-diagonal matrices are as follows.
> $$\begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}.$$

> **Definition**
>
> A square matrix A is called **scalar** if it is a diagonal matrix and all its diagonal entries are equal. In other words, an $n \times n$ matrix A is called **scalar** if
> $$A = \lambda I_n$$
> for some scalar λ.

> **Example**
>
> The following matrices are scalar.
> $$\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}, \begin{pmatrix} -3 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & -3 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$
> Some non-scalar matrices are as follows.
> $$\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}, \begin{pmatrix} 5 & 0 & 0 \\ 0 & -7 & 0 \\ 0 & 0 & \pi \end{pmatrix}, \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}.$$

> **Definition**
>
> A square matrix A is called **triangular** if all its entries below the main diagonal are equal to zero, or if all its entries above the main diagonal are equal to zero.
>
> A triangular matrix A is called **upper-triangular** if all its entries below the main diagonal are equal to zero.
>
> A triangular matrix A is called **lower-triangular** if all its entries above the main diagonal are equal to zero.

> **Example**
>
> The following matrices are upper-triangular.
> $$\begin{pmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{pmatrix}, \begin{pmatrix} 7 & 8 \\ 0 & 9 \end{pmatrix}.$$
>
> The following matrices are lower-triangular.
> $$\begin{pmatrix} 1 & 0 & 0 \\ 2 & 3 & 0 \\ 4 & 5 & 6 \end{pmatrix}, \begin{pmatrix} 7 & 0 \\ 8 & 9 \end{pmatrix}.$$
>
> Some non-triangular matrices are as follows.
> $$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 0 \\ 0 & 3 & 4 \\ 5 & 0 & 6 \end{pmatrix}.$$

§3.6.2 Symmetric, and skew-symmetric matrices

> **Definition**
>
> A square matrix A is called **symmetric** if
> $$A^T = A.$$

> **Example**
>
> The following matrices are symmetric.
> $$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{pmatrix}, \begin{pmatrix} 7 & 8 \\ 8 & 9 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$
>
> Some non-symmetric matrices are as follows.
> $$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}.$$

> **Definition**
>
> A square matrix A is called **skew-symmetric** if
> $$A^T = -A.$$

> **Example**
>
> The following matrices are skew-symmetric.
> $$\begin{pmatrix} 0 & 2 & 3 \\ -2 & 0 & 5 \\ -3 & -5 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 8 \\ -8 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$
>
> Here are some matrices that are not skew-symmetric.
> $$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}.$$

Exercise 3.36. Let A be a square matrix. Show that if A is symmetric and skew-symmetric, then A is the zero matrix.

Solution. Assume that A is symmetric and skew-symmetric. It follows that
$$A^T = A,$$
$$A^T = -A.$$
This shows that
$$A = -A,$$
which yields
$$2A = 0,$$
and hence, A is the zero matrix. □

Fact 3.37. Any square matrix can be uniquely expressed as the sum of a symmetric matrix and a skew-symmetric matrix.

A proof of the above is provided in Chapter 6.

§3.6.3 Orthogonal matrices

> **Definition**
>
> An $n \times n$ matrix A with **real entries** is called **orthogonal** if
> $$AA^T = A^T A = I_n.$$

> **Example**
>
> The following matrices are orthogonal.
> $$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}.$$
>
> Some non-orthogonal matrices are as follows.
> $$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}.$$

> **i Remark**
>
> There are infinitely many orthogonal matrices. For instance, for any real number θ, the matrix
> $$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$
> is an orthogonal matrix since
> $$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}^T \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$
> $$= \begin{pmatrix} \cos^2\theta + \sin^2\theta & 0 \\ 0 & \cos^2\theta + \sin^2\theta \end{pmatrix}$$
> $$= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$
>
> and
> $$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}^T = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$$
> $$= \begin{pmatrix} \cos^2\theta + \sin^2\theta & 0 \\ 0 & \cos^2\theta + \sin^2\theta \end{pmatrix}$$
> $$= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Exercise 3.38. Show that for any $n \geq 2$, there are infinitely many $n \times n$ orthogonal matrices.

Exercise 3.39. Let A be an orthogonal matrix. Show that A is invertible and that
$$A^{-1} = A^T.$$

Solution. Let A be an $n \times n$ orthogonal matrix. It follows that
$$AA^T = A^T A = I_n,$$
which shows that A is invertible and that
$$A^{-1} = A^T.$$

$\square$

Exercise 3.40. Let A, B be orthogonal matrices of the same size. Show that the matrix product AB is also an orthogonal matrix.

Solution. Let A, B be $n \times n$ orthogonal matrices. Note that

$$(AB)(AB)^T = ABB^T A^T = AI_n A^T = AA^T = I_n,$$

and similarly,

$$(AB)^T(AB) = B^T A^T AB = B^T I_n B = B^T B = I_n.$$

This shows that the matrix product AB is also an orthogonal matrix. $\square$

§3.6.4 Hermitian and skew-hermitian matrices

> **📑 Definition**
>
> If A is a matrix with complex entries, then the **complex conjugate** of A, denoted by $\overline{A}$, is defined as the matrix of the same size, obtained by taking the complex conjugate of each entry of A.

> **Example**
>
> The complex conjugate of the matrix
> $$\begin{pmatrix} 1+2i & 3-4i \\ -i & 5 \end{pmatrix}$$
> is the matrix
> $$\begin{pmatrix} 1-2i & 3+4i \\ i & 5 \end{pmatrix}.$$

Exercise 3.41. Let A, B be square matrices with complex entries. Show that
$$\overline{AB} = \overline{A}\,\overline{B}.$$

Exercise 3.42. Let A be a square matrix with complex entries. Show that
$$\overline{A^T} = \left(\overline{A}\right)^T.$$

> **📑 Definition**
>
> If A is a matrix with complex entries, then the **conjugate transpose** of A, denoted by A^*, is defined as the transpose of the matrix obtained by taking the complex conjugate of each entry of A, that A^* is equal to $\overline{A}^T$.

> **Example**
>
> The conjugate transpose of the matrix
> $$\begin{pmatrix} 1+2i & 3-4i \\ -i & 5 \end{pmatrix}$$
> is the matrix
> $$\begin{pmatrix} 1-2i & i \\ 3+4i & 5 \end{pmatrix}.$$

Compare the following exercises with

Exercise 3.43. Let A, B be square matrices of the same size with complex entries. Show that

$$(A + B)^* = A^* + B^*.$$

Solution. Note that

$$(A + B)^* = \left(\overline{A + B}\right)^T = \left(\overline{A} + \overline{B}\right)^T = \overline{A}^T + \overline{B}^T = A^* + B^*.$$

$\square$

Exercise 3.44. Let A be a matrix with complex entries. Show that
$$(A^*)^* = A.$$

Solution. Note that

$$(A^*)^* = \left(\overline{A^T}\right)^* = \left(\overline{A^T}\right)^* = \overline{\overline{A^T}}^T = \left(A^T\right)^T = A.$$

$\square$

Exercise 3.45. Let A, B be matrices with complex entries. Show that
$$(AB)^* = B^* A^*.$$

Solution. Note that

$$(AB)^* = \left(\overline{AB}\right)^T = \left(\overline{A}\,\overline{B}\right)^T = \overline{B}^T\,\overline{A}^T = B^* A^*.$$

$\square$

> **Definition**
>
> A square matrix A with complex entries is called **hermitian** if
> $$A^* = A,$$
> where A^* denotes the conjugate transpose of A.

> **Example**
>
> The following matrices are hermitian.
> $$\begin{pmatrix} 1 & 2 + i \\ 2 - i & 3 \end{pmatrix}, \begin{pmatrix} 5 & i \\ -i & 7 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$
> Some non-hermitian matrices are as follows.
> $$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}.$$

Exercise 3.46. Let A be an $n \times n$ hermitian matrix. Show that all the diagonal entries of A are real numbers. More generally, show that for any $1 \leq i, j \leq n$, the (i, j)-entry and the (j, i)-entry of A are complex conjugates of each other.

Exercise 3.47. Let A, B be hermitian matrices of the same size. Show that the matrix $A + B$ is also a hermitian matrix. Is the matrix AB also a hermitian matrix?

Solution. Let A, B be $n \times n$ hermitian matrices. Note that
$$(A + B)^* = A^* + B^* = A + B.$$
This shows that the matrix $A + B$ is also a hermitian matrix.

The matrix product AB is not necessarily a hermitian matrix. For instance, let
$$A = \begin{pmatrix} 1 & i \\ -i & 1 \end{pmatrix},$$
$$B = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}.$$
Note that both A and B are hermitian matrices. Also note that
$$AB = \begin{pmatrix} 2 & 3i \\ -2i & 3 \end{pmatrix},$$
and hence, the product AB is not a hermitian matrix. ☐

> **Definition**
>
> A square matrix A with complex entries is called **skew-hermitian** if
> $$A^* = -A,$$
> where A^* denotes the conjugate transpose of A.

> **Example**
>
> The following matrices are skew-hermitian.
> $$\begin{pmatrix} 0 & 2+i \\ -2+i & 0 \end{pmatrix}, \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$
> Some non-skew-hermitian matrices are as follows.
> $$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}.$$

Exercise 3.48. State and prove an analogue of Exercise 3.46 for skew-hermitian matrices.

Exercise 3.49. Let A, B be skew-hermitian matrices of the same size. Show that the matrix $A + B$ is also a skew-hermitian matrix. Is the matrix AB also a skew-hermitian matrix?

Solution. Let A, B be $n \times n$ skew-hermitian matrices. Note that
$$(A + B)^* = A^* + B^* = -A - B = -(A + B).$$
This shows that the matrix $A + B$ is also a skew-hermitian matrix.

The matrix product AB is not necessarily a skew-hermitian matrix. For instance, let
$$A = \begin{pmatrix} 0 & 2+i \\ -2+i & 0 \end{pmatrix},$$
$$B = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$
Note that both A and B are skew-hermitian matrices. Also note that
$$AB = \begin{pmatrix} -2-i & 0 \\ 0 & -2+i \end{pmatrix},$$
which is not a skew-hermitian matrix. ☐

Exercise 3.50. Let A be a square matrix with complex entries. Show that if A is hermitian and skew-hermitian, then A is the zero matrix.

Solution. Assume that A is hermitian and skew-hermitian. It follows that

$$A^* = A,$$
$$A^* = -A.$$

This shows that

$$A = -A,$$

which yields

$$2A = 0,$$

and hence, A is the zero matrix. $\qquad\square$

Fact 3.51. Let A be a square matrix with complex entries. Show that A can be uniquely expressed as the sum of a hermitian matrix and a skew-hermitian matrix.

A proof of the above is provided in Chapter 6.

§3.6.5 Unitary and normal matrices

> **Definition**
>
> An $n \times n$ matrix A with complex entries is called **unitary** if
> $$AA^* = A^*A = I_n.$$

> **Example**
>
> The following matrices are unitary.
> $$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{i}{\sqrt{2}} \\ -\frac{i}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}.$$
>
> Some non-unitary matrices are as follows.
> $$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}.$$

Exercise 3.52. Let A be a unitary matrix. Show that A is invertible and that
$$A^{-1} = A^*.$$

Solution. Let A be a unitary matrix. It follows that
$$AA^* = A^*A = I_n,$$
which shows that A is invertible and that
$$A^{-1} = A^*.$$

$\qquad\square$

Exercise 3.53. Let A, B be unitary matrices of the same size. Show that the matrix product AB is also a unitary matrix.

Solution. Let A, B be $n \times n$ unitary matrices. Note that
$$(AB)(AB)^* = ABB^*A^* = AI_nA^* = AA^* = I_n,$$
and similarly,
$$(AB)^*(AB) = B^*A^*AB = B^*I_nB = B^*B = I_n.$$

This shows that the matrix product AB is also a unitary matrix. □

Exercise 3.54. Show that any orthogonal matrix is a unitary matrix. Is the converse true?

Solution. Let A be an orthogonal matrix. It follows that
$$A^T A = I_n,$$
which shows that A is invertible and that
$$A^{-1} = A^T.$$
Since A has real entries, we have $A^* = A^T$, and hence,
$$A^* A = A^T A = I_n.$$
Similarly, we have
$$AA^* = AA^T = I_n.$$
This shows that any orthogonal matrix is a unitary matrix.

The converse is not true in general. For instance, the matrix
$$\begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{i}{\sqrt{2}} \\ -\frac{i}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}$$
is a unitary matrix but not an orthogonal matrix. □

Exercise 3.55. Determine whether the following matrices are unitary, hermitian, both, or neither.
$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Exercise 3.56. Let z, w be complex numbers such that $|z|^2 + |w|^2 = 1$. Show that
$$\begin{pmatrix} z & -\overline{w} \\ w & \overline{z} \end{pmatrix}$$
is a unitary matrix.

Solution. Let z, w be complex numbers such that $|z|^2 + |w|^2 = 1$. Note that
$$\begin{pmatrix} z & -\overline{w} \\ w & \overline{z} \end{pmatrix}\begin{pmatrix} z & -\overline{w} \\ w & \overline{z} \end{pmatrix}^* = \begin{pmatrix} z & -\overline{w} \\ w & \overline{z} \end{pmatrix}\begin{pmatrix} \overline{z} & \overline{w} \\ -w & z \end{pmatrix}$$
$$= \begin{pmatrix} z\overline{z} + \overline{w}w & z\overline{w} - \overline{w}z \\ w\overline{z} - \overline{z}w & w\overline{w} + \overline{z}z \end{pmatrix}$$
$$= \begin{pmatrix} |z|^2 + |w|^2 & 0 \\ 0 & |z|^2 + |w|^2 \end{pmatrix}$$
$$= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$
hold. Also note that

$$\begin{pmatrix} z & -\overline{w} \\ w & \overline{z} \end{pmatrix}^* \begin{pmatrix} z & -\overline{w} \\ w & \overline{z} \end{pmatrix} = \begin{pmatrix} \overline{z} & \overline{w} \\ -w & z \end{pmatrix} \begin{pmatrix} z & -\overline{w} \\ w & \overline{z} \end{pmatrix}$$

$$= \begin{pmatrix} \overline{z}z + \overline{w}(-w) & \overline{z}(-\overline{w}) + \overline{w}z \\ -wz + zw & -w(-\overline{w}) + z\overline{z} \end{pmatrix}$$

$$= \begin{pmatrix} |z|^2 + |w|^2 & 0 \\ 0 & |z|^2 + |w|^2 \end{pmatrix}$$

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

hold. This shows that

$$\begin{pmatrix} z & -\overline{w} \\ w & \overline{z} \end{pmatrix}$$

is a unitary matrix. $\square$

Exercise 3.57. Let A be a diagonal matrix with complex entries. Show that A is a unitary matrix if and only if all the diagonal entries of A have absolute value equal to 1.

Definition

A square matrix A with complex entries is called **normal** if
$$AA^* = A^*A,$$
where A^* denotes the conjugate transpose of A.

Example

The following matrices are normal.
$$\begin{pmatrix} 1 & 2+i \\ 2-i & 3 \end{pmatrix}, \begin{pmatrix} 5 & i \\ -i & 7 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$
Some non-normal matrices are as follows.
$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}.$$

Exercise 3.58. Show that any unitary matrix is a normal matrix. Is the converse true?

Exercise 3.59. Show that any hermitian matrix is a normal matrix. Is the converse true?

Exercise 3.60. Show that any skew-hermitian matrix is a normal matrix. Is the converse true?

Here is a summary of some properties of some of the special types of matrices introduced above.

Matrix	Some property
Diagonal	Entries outside the main diagonal are zero
Scalar	Diagonal matrix with all diagonal entries equal
Triangular	Entries below or above the main diagonal are zero
Upper-triangular	Entries below the main diagonal are zero
Lower-triangular	Entries above the main diagonal are zero
Symmetric	$A^T = A$
Skew-symmetric	$A^T = -A$
Orthogonal	$AA^T = A^T A = I_n$
Hermitian	$A^* = A$
Skew-hermitian	$A^* = -A$
Unitary	$AA^* = A^* A = I_n$
Normal	$AA^* = A^* A$

§3.7 Row reduction

§3.7.1 Matrix units

> **Definition**
>
> The **matrix unit** e_{ij} is the $n \times n$ matrix whose (i, j)-entry is equal to 1 and all other entries are equal to 0.

> **Example**
>
> The matrix unit e_{23} for $n = 4$ is the matrix
> $$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

> **i Remark**
>
> For a 2×2 matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, we have
> $$A = a\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + b\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + c\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} + d\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$
> $$= ae_{11} + be_{12} + ce_{21} + de_{22}.$$
> More generally, if $A = \begin{pmatrix} a_{ij} \end{pmatrix}$ is an $n \times n$ matrix, then
> $$\begin{aligned} A = &a_{11}e_{11} + a_{12}e_{12} + \cdots + a_{1n}e_{1n} \\ &+ a_{21}e_{21} + a_{22}e_{22} + \cdots + a_{2n}e_{2n} \\ &+ \cdots \\ &+ a_{n1}e_{n1} + a_{n2}e_{n2} + \cdots + a_{nn}e_{nn} \end{aligned}$$
> holds.

> **Example**
>
> Let
> $$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}.$$
>
> Note that
> $$e_{12}A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} = \begin{pmatrix} 4 & 5 & 6 \\ 0 & 0 & 0 \end{pmatrix},$$
> $$e_{21}A = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 2 & 3 \end{pmatrix},$$
> $$e_{11}A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 0 & 0 \end{pmatrix},$$
> $$e_{22}A = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 4 & 5 & 6 \end{pmatrix}.$$

> **i   Remark**
>
> More generally, if A is a matrix with n rows, then multiplying A from the left by the matrix unit e_{ij} produces the matrix whose i-th row is equal to the j-th row of A, and all other rows are equal to zero.

Exercise 3.61. Let e_{ij}, e_{kl} be matrix units of the same size. Show that $e_{ij}e_{kl}$ is equal to the matrix unit e_{il} if $j = k$, and is equal to the zero matrix if $j \neq k$, that is,
$$e_{ij}e_{kl} = \begin{cases} e_{il} & \text{if } j = k, \\ 0 & \text{if } j \neq k. \end{cases}$$

§3.7.2 Matrix multiplication and linear combinations of rows

> **i   Remark**
>
> If A is an $m \times n$ matrix and B is an $n \times p$ matrix, then the matrix product AB is an $m \times p$ matrix. Recall that the i-th **column of AB** is a linear combination of the **columns of A** with **coefficients** from the i-th **column of B**.

> **i   Remark**
>
> Consider the matrices
> $$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix},$$
>
> $$B = \begin{pmatrix} 7 & 8 \\ 9 & 10 \\ 11 & 12 \end{pmatrix}.$$
>
> Note that
> $$AB = \begin{pmatrix} 1 \times 7 + 2 \times 9 + 3 \times 11 & 1 \times 8 + 2 \times 10 + 3 \times 12 \\ 4 \times 7 + 5 \times 9 + 6 \times 11 & 4 \times 8 + 5 \times 10 + 6 \times 12 \end{pmatrix} = \begin{pmatrix} 58 & 64 \\ 139 & 154 \end{pmatrix}.$$
>
> Note that first row of AB is a linear combination of the rows of B with coefficients from the first row of A, and the second row of AB is a linear combination of the rows of B with coefficients from the second row of A. Indeed,
> $$(1 \times 7 + 2 \times 9 + 3 \times 11, \ 1 \times 8 + 2 \times 10 + 3 \times 12)$$
> $$= 1 \times (7, \ 8) + 2 \times (9, \ 10) + 3 \times (11, \ 12),$$
> $$(4 \times 7 + 5 \times 9 + 6 \times 11, \ 4 \times 8 + 5 \times 10 + 6 \times 12)$$
> $$= 4 \times (7, \ 8) + 5 \times (9, \ 10) + 6 \times (11, \ 12).$$

> **i   Remark**
>
> If A is an $m \times n$ matrix and B is an $n \times p$ matrix, then the matrix product AB is an $m \times p$ matrix, and the i-th **row of AB** is a linear combination of the **rows of B** with **coefficients** from the i-th **row of A**. Thus, left multiplication by a matrix A (in particular, by a square matrix A) can be viewed as a transformation that transforms each row of B into a linear combination of the rows of B, and is often called a **row operation**.

Denote the set of $m \times n$ matrices with entries from $\mathbb{R}$ by $M_{m,n}(\mathbb{R})$.

Compare the following exercise with .

Exercise 3.62. Let A be an $m \times n$ matrix. Let V be the set of row vectors which can be expressed as a linear combination of the rows of A, that is,
$$V = \left\{ v \in M_{1,n}(\mathbb{R}) : v = u^T A \text{ for some } u \in \mathbb{R}^m \right\}.$$
Show that for any $u, v \in V$ and $c, d \in \mathbb{R}$, the element $cu + dv$ of $M_{1,n}(\mathbb{R})$ lies in V.

Solution. Let $u, v \in V$ and $c, d \in \mathbb{R}$. By the definition of V, there exist $x, y \in \mathbb{R}^m$ such that
$$u = x^T A,$$
$$v = y^T A.$$
Note that
$$cu + dv = cx^T A + dy^T A = (cx^T + dy^T)A = (cx + dy)^T A.$$
Since $cx + dy \in \mathbb{R}^m$, it follows that $cu + dv \in V$. $\qquad \square$

§3.7.3 Elementary row operations and elementary matrices

> **Definition**
>
> An **elementary row operation** on a matrix is one of the following operations:
>   1. Interchanging two rows.
>   2. Multiplying a row by a nonzero scalar.
>   3. Adding a scalar multiple of one row to another row.

To study matrices that perform elementary row operations, we introduce the following special types of matrices, called **elementary matrices**.

> **Definition**
>
> The **elementary matrices** are of three types, and these are obtained by performing the elementary row operations on the identity matrix.
>   1. The matrix obtained by interchanging the i-th row and the j-th row of the identity matrix, that is, it is given by
> $$I_n - e_{ii} - e_{jj} + e_{ij} + e_{ji} \text{ with } i \neq j,$$
>   2. The matrix obtained by multiplying the i-th row of the identity matrix by a nonzero scalar, that is, it is given by
> $$I_n + (\lambda - 1)e_{ii} \text{ with } \lambda \neq 0,$$
>   3. The matrix obtained by adding a scalar multiple of the j-th row of the identity matrix to its i-th row, that is, it is given by
> $$I_n + \lambda e_{ij} \text{ with } i \neq j.$$

> **Example**
>
> The elementary 2×2 matrices are as follows.
> $$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} \lambda & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & \lambda \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ \lambda & 1 \end{pmatrix}, \begin{pmatrix} 1 & \lambda \\ 0 & 1 \end{pmatrix}, \text{where } \lambda \neq 0.$$

> **Example**
>
> The elementary 3×3 matrices are as follows.
> $$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix},$$
> $$\begin{pmatrix} \lambda & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \lambda \end{pmatrix} \text{ where } \lambda \neq 0,$$
> $$\begin{pmatrix} 1 & 0 & 0 \\ \lambda & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \lambda & 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & \lambda & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & \lambda & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 & \lambda \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & \lambda \\ 0 & 0 & 1 \end{pmatrix} \text{ where } \lambda \neq 0.$$

Lemma 3.63. *Left multiplication by an elementary matrix on a matrix A performs the corresponding elementary row operation on the matrix A. More precisely, if E is an elementary matrix obtained by performing an elementary row operation on the identity matrix, then the matrix product EA is the matrix obtained by performing the same elementary row operation on the matrix A.*

A proof of the above is provided in Chapter 6.

Lemma 3.64. *Any elementary matrix is invertible, and its inverse is also an elementary matrix.*

A proof of the above is provided in Chapter 6.

> **_i_ Remark**
>
> Note that the elementary row operations on a matrix A are precisely the operations of left multiplying A by the elementary matrices. Thus, if $E_1, E_2, ..., E_k$ are elementary matrices, then the matrix
>
> $$E_k...E_2 E_1 A$$
>
> is obtained by performing a sequence of elementary row operations on the matrix A. Moreover, any matrix that can be obtained from A by performing a sequence of elementary row operations can be expressed in the form
>
> $$EA,$$
>
> where E is an invertible matrix which is a product of elementary matrices.

§3.7.4 Row reduction of a matrix

> **📑 Definition**
>
> Multiplying a matrix A by a matrix E from the left is called **row reduction** of the matrix A if E is a product of elementary matrices. In other words, row reduction of a matrix A is the process of performing a sequence of elementary row operations on the matrix A. It is also known as **Gaussian elimination**.

> ### ⌇ Example
>
> Consider the matrix
> $$A = \begin{pmatrix} 1 & 2 & -1 & -4 \\ 2 & 3 & -1 & -11 \\ -2 & 0 & -3 & 22 \end{pmatrix}.$$
>
> We perform row reduction on the matrix A as follows.
>
> 1. First, we add -2 times the first row to the second row, and add 2 times the first row to the third row, to obtain the matrix
> $$\begin{pmatrix} 1 & 2 & -1 & -4 \\ 0 & -1 & 1 & -3 \\ 0 & 4 & -5 & 14 \end{pmatrix}.$$
>
> 2. Next, we add 4 times the second row to the third row, to obtain the matrix
> $$\begin{pmatrix} 1 & 2 & -1 & -4 \\ 0 & -1 & 1 & -3 \\ 0 & 0 & -1 & 2 \end{pmatrix}.$$
>
> 3. Finally, we multiply the second row by -1, and multiply the third row by -1, to obtain the matrix
> $$\begin{pmatrix} 1 & 2 & -1 & -4 \\ 0 & 1 & -1 & 3 \\ 0 & 0 & 1 & -2 \end{pmatrix}.$$

We now illustrate how row reduction can be used to solve a system of linear equations.

Exercise 3.65. Solve the system of linear equations in five variables x_1, x_2, x_3, x_4, x_5:
$$\begin{aligned} x_1 + 2x_2 - x_3 + 4x_4 + x_5 &= 7, \\ 2x_1 + 3x_2 + x_3 + 5x_4 + 2x_5 &= 14, \\ -x_1 + 4x_2 - 2x_3 - 3x_4 + x_5 &= -10. \end{aligned} \tag{12}$$

Solution. The above system of equations can be expressed in matrix form as

$$\begin{pmatrix} 1 & 2 & -1 & 4 & 1 \\ 2 & 3 & 1 & 5 & 2 \\ -1 & 4 & -2 & -3 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 7 \\ 14 \\ -10 \end{pmatrix}.$$

We perform row reduction on the augmented matrix

$$(A \mid b) = \left(\begin{array}{ccccc|c} 1 & 2 & -1 & 4 & 1 & 7 \\ 2 & 3 & 1 & 5 & 2 & 14 \\ -1 & 4 & -2 & -3 & 1 & -10 \end{array} \right).$$

1. First, we add -2 times the first row to the second row, and add 1 times the first row to the third row, to obtain the matrix
$$\left(\begin{array}{ccccc|c} 1 & 2 & -1 & 4 & 1 & 7 \\ 0 & -1 & 3 & -3 & 0 & 0 \\ 0 & 6 & -3 & 1 & 2 & -3 \end{array} \right).$$

2. Next, we add 6 times the second row to the third row, to obtain the matrix

$$\left(\begin{array}{ccccc|c} 1 & 2 & -1 & 4 & 1 & 7 \\ 0 & -1 & 3 & -3 & 0 & 0 \\ 0 & 0 & 15 & -17 & 2 & -3 \end{array}\right).$$

3. Finally, we multiply the second row by -1, and multiply the third row by $\frac{1}{15}$, to obtain the matrix

$$\left(\begin{array}{ccccc|c} 1 & 2 & -1 & 4 & 1 & 7 \\ 0 & 1 & -3 & 3 & 0 & 0 \\ 0 & 0 & 1 & -\frac{17}{15} & \frac{2}{15} & -\frac{1}{5} \end{array}\right).$$

The solutions to the system of equations Equation 12 are precisely the solutions to the system of equations

$$\begin{pmatrix} 1 & 2 & -1 & 4 & 1 \\ 0 & 1 & -3 & 3 & 0 \\ 0 & 0 & 1 & -\frac{17}{15} & \frac{2}{15} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 7 \\ 0 \\ -\frac{1}{5} \end{pmatrix}.$$

Hence, the solutions of the given system of equations are precisely the elements $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix}$ of $\mathbb{R}^5$ satisfying

$$x_1 + 2x_2 - x_3 + 4x_4 + x_5 = 7,$$
$$x_2 - 3x_3 + 3x_4 = 0,$$
$$x_3 - \frac{17}{15}x_4 + \frac{2}{15}x_5 = -\frac{1}{5},$$

or equivalently,

$$x_1 = 7 - 2x_2 + x_3 - 4x_4 - x_5,$$
$$x_2 = 3x_3 - 3x_4,$$
$$x_3 = -\frac{1}{5} + \frac{17}{15}x_4 - \frac{2}{15}x_5,$$

which is equivalent to

$$x_3 = -\frac{1}{5} + \frac{17}{15}s - \frac{2}{15}t,$$
$$x_2 = 3\left(-\frac{1}{5} + \frac{17}{15}s - \frac{2}{15}t\right) - 3s$$
$$= -\frac{3}{5} + \frac{51}{15}s - \frac{2}{5}t - 3s$$
$$= -\frac{3}{5} + \frac{2}{5}s - \frac{2}{5}t$$
$$x_1 = 7 - 2\left(-\frac{3}{5} + \frac{2}{5}s - \frac{2}{5}t\right) + \left(-\frac{1}{5} + \frac{17}{15}s - \frac{2}{15}t\right) - 4s - t$$
$$= 7 + \frac{6}{5} - \frac{4}{5}s + \frac{4}{5}t - \frac{1}{5} + \frac{17}{15}s - \frac{2}{15}t - 4s - t$$
$$= 8 - \frac{11}{3}s - \frac{1}{3}t.$$

In other words, the set of solutions of the given system of equations is equal to

$$\left\{ \left(8 - \frac{11}{3}s - \frac{1}{3}t, -\frac{3}{5} + \frac{2}{5}s - \frac{2}{5}t, -\frac{1}{5} + \frac{17}{15}s - \frac{2}{15}t, s, t \right)^T : s, t \in \mathbb{R} \right\}.$$

$\square$

Theorem 3.66. *Let b denote the vector $\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$, and let A be an $m \times n$ matrix. Let E be a product of elementary matrices. Write $X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$. The solutions to the system of equations*

$$AX = b$$

are precisely the solutions to the system of equations

$$A'X = b',$$

where

$$b' = Eb, \quad A' = EA.$$

A proof of the above is provided in Chapter 6.

> **ⓘ Remark**
>
> Note that while solving Equation 12, we performed row reduction on the augmented matrix
>
> $$(A \mid b) = \begin{pmatrix} 1 & 2 & -1 & 4 & 1 & 7 \\ 2 & 3 & 1 & 5 & 2 & 14 \\ -1 & 4 & -2 & -3 & 1 & -10 \end{pmatrix},$$
>
> and obtained the matrix
>
> $$\begin{pmatrix} 1 & 2 & -1 & 4 & 1 & 7 \\ 0 & 1 & -3 & 3 & 0 & 0 \\ 0 & 0 & 1 & -\frac{17}{15} & \frac{2}{15} & -\frac{1}{5} \end{pmatrix}.$$
>
> Note that one can perform further row reductions. Adding the third row to the first row, and adding 3 times the third row to the second row, we obtain the matrix
>
> $$\begin{pmatrix} 1 & 2 & 0 & \frac{43}{15} & \frac{17}{15} & \frac{34}{5} \\ 0 & 1 & 0 & -\frac{2}{5} & \frac{2}{5} & -\frac{3}{5} \\ 0 & 0 & 1 & -\frac{17}{15} & \frac{2}{15} & -\frac{1}{5} \end{pmatrix}.$$
>
> Next, we add -2 times the second row to the first row, to obtain the matrix
>
> $$\begin{pmatrix} 1 & 0 & 0 & \frac{11}{3} & \frac{1}{3} & 8 \\ 0 & 1 & 0 & -\frac{2}{5} & \frac{2}{5} & -\frac{3}{5} \\ 0 & 0 & 1 & -\frac{17}{15} & \frac{2}{15} & -\frac{1}{5} \end{pmatrix}.$$
>
> This shows that the augmented matrix corresponding to the system of equations Equation 12 can be row reduced to the matrix
>
> $$\begin{pmatrix} 1 & 0 & 0 & \frac{11}{3} & \frac{1}{3} & 8 \\ 0 & 1 & 0 & -\frac{2}{5} & \frac{2}{5} & -\frac{3}{5} \\ 0 & 0 & 1 & -\frac{17}{15} & \frac{2}{15} & -\frac{1}{5} \end{pmatrix},$$
>
> which shows that the solutions of the system of equations Equation 12 are precisely the elements of the set
>
> $$\boxed{\left\{ \left(8 - \frac{11}{3}s - \frac{1}{3}t, -\frac{3}{5} + \frac{2}{5}s - \frac{2}{5}t, -\frac{1}{5} + \frac{17}{15}s - \frac{2}{15}t, s, t \right)^T : s, t \in \mathbb{R} \right\}}.$$

Exercise 3.67. Solve the system of equations in the variables x_1, x_2, x_3:

$$x_1 + 2x_2 - x_3 = 1,$$
$$2x_1 + 3x_2 + x_3 = 2,$$
$$-x_1 + 4x_2 - 2x_3 = -3,$$
$$3x_1 - x_2 + 4x_3 = 4,$$
$$5x_1 + 2x_2 + 3x_3 = 5.$$

Solution. The above system of equations can be expressed in matrix form as

$$\begin{pmatrix} 1 & 2 & -1 \\ 2 & 3 & 1 \\ -1 & 4 & -2 \\ 3 & -1 & 4 \\ 5 & 2 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ -3 \\ 4 \\ 5 \end{pmatrix}.$$

We perform row reduction on the augmented matrix

$$(A \mid b) = \left(\begin{array}{ccc|c} 1 & 2 & -1 & 1 \\ 2 & 3 & 1 & 2 \\ -1 & 4 & -2 & -3 \\ 3 & -1 & 4 & 4 \\ 5 & 2 & 3 & 5 \end{array} \right).$$

1. First, we add -2 times the first row to the second row, add 1 times the first row to the third row, add -3 times the first row to the fourth row, and add -5 times the first row to the fifth row, to obtain the matrix

$$\left(\begin{array}{ccc|c} 1 & 2 & -1 & 1 \\ 0 & -1 & 3 & 0 \\ 0 & 6 & -3 & -2 \\ 0 & -7 & 7 & 1 \\ 0 & -8 & 8 & 0 \end{array} \right).$$

2. Next, we add 6 times the second row to the third row, add -7 times the second row to the fourth row, and add -8 times the second row to the fifth row, to obtain the matrix

$$\left(\begin{array}{ccc|c} 1 & 2 & -1 & 1 \\ 0 & -1 & 3 & 0 \\ 0 & 0 & 15 & -2 \\ 0 & 0 & -14 & 1 \\ 0 & 0 & -16 & 0 \end{array} \right).$$

3. Finally, we multiply the second row by -1, multiply the third row by $\frac{1}{15}$, multiply the fourth row by $-\frac{1}{14}$, and multiply the fifth row by $-\frac{1}{16}$, to obtain the matrix

$$\left(\begin{array}{ccc|c} 1 & 2 & -1 & 1 \\ 0 & 1 & -3 & 0 \\ 0 & 0 & 1 & -\frac{2}{15} \\ 0 & 0 & 1 & -\frac{1}{14} \\ 0 & 0 & 1 & 0 \end{array} \right).$$

Note that the above matrix corresponds to the system of equations

$$\begin{pmatrix} 1 & 2 & -1 \\ 0 & 1 & -3 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ -\frac{2}{15} \\ \frac{1}{14} \\ 0 \end{pmatrix},$$

which has no solutions. Hence, the given system of equations has no solutions. $\square$

Exercise 3.68. Solve the system of equations in the variables x_1, x_2, x_3, x_4:

$$x_1 + 2x_2 - x_3 + 4x_4 = 5,$$
$$2x_1 + 3x_2 + x_3 + 5x_4 = 8,$$
$$-x_1 + 4x_2 - 2x_3 - 3x_4 = -4,$$
$$3x_1 - x_2 + 4x_3 + 2x_4 = 7.$$

Solution. The above system of equations can be expressed in matrix form as

$$\begin{pmatrix} 1 & 2 & -1 & 4 \\ 2 & 3 & 1 & 5 \\ -1 & 4 & -2 & -3 \\ 3 & -1 & 4 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 5 \\ 8 \\ -4 \\ 7 \end{pmatrix}.$$

We perform row reduction on the augmented matrix

$$(A \mid b) = \begin{pmatrix} 1 & 2 & -1 & 4 & \vline & 5 \\ 2 & 3 & 1 & 5 & \vline & 8 \\ -1 & 4 & -2 & -3 & \vline & -4 \\ 3 & -1 & 4 & 2 & \vline & 7 \end{pmatrix}.$$

1. First, we add -2 times the first row to the second row, add 1 times the first row to the third row, and add -3 times the first row to the fourth row, to obtain the matrix

$$\begin{pmatrix} 1 & 2 & -1 & 4 & \vline & 5 \\ 0 & -1 & 3 & -3 & \vline & -2 \\ 0 & 6 & -3 & 1 & \vline & 1 \\ 0 & -7 & 7 & -10 & \vline & -8 \end{pmatrix}.$$

2. Next, we add 6 times the second row to the third row, and add -7 times the second row to the fourth row, to obtain the matrix

$$\begin{pmatrix} 1 & 2 & -1 & 4 & \vline & 5 \\ 0 & -1 & 3 & -3 & \vline & -2 \\ 0 & 0 & 15 & -17 & \vline & -11 \\ 0 & 0 & -14 & 11 & \vline & 6 \end{pmatrix}.$$

3. Multiplying the second row by -1, the third row by $\frac{1}{15}$, and the fourth row by $-\frac{1}{14}$, we obtain the matrix

$$\begin{pmatrix} 1 & 2 & -1 & 4 & \vline & 5 \\ 0 & 1 & -3 & 3 & \vline & 2 \\ 0 & 0 & 1 & -\frac{17}{15} & \vline & -\frac{11}{15} \\ 0 & 0 & 1 & -\frac{11}{14} & \vline & -\frac{3}{7} \end{pmatrix}.$$

4. Adding -1 times the third row to the fourth row (that is, subtracting the third row from the fourth row), we obtain the matrix

$$\begin{pmatrix} 1 & 2 & -1 & 4 & \vline & 5 \\ 0 & 1 & -3 & 3 & \vline & 2 \\ 0 & 0 & 1 & -\frac{17}{15} & \vline & -\frac{11}{15} \\ 0 & 0 & 0 & \frac{73}{210} & \vline & \frac{32}{105} \end{pmatrix}.$$

5. Multiplying the fourth row by $\frac{210}{73}$, we obtain the matrix

$$\begin{pmatrix} 1 & 2 & -1 & 4 & \Big| & 5 \\ 0 & 1 & -3 & 3 & \Big| & 2 \\ 0 & 0 & 1 & -\frac{17}{15} & \Big| & -\frac{11}{15} \\ 0 & 0 & 0 & 1 & \Big| & \frac{64}{73} \end{pmatrix}.$$

6. Adding -4 times the fourth row to the first row, and adding -3 times the fourth row to the second row, and adding $\frac{17}{15}$ times the fourth row to the third row, we obtain the matrix

$$\begin{pmatrix} 1 & 2 & -1 & 0 & \Big| & \frac{109}{73} \\ 0 & 1 & -3 & 0 & \Big| & -\frac{46}{73} \\ 0 & 0 & 1 & 0 & \Big| & \frac{19}{73} \\ 0 & 0 & 0 & 1 & \Big| & \frac{64}{73} \end{pmatrix}$$

7. Adding 3 times the third row to the second row, and adding 1 times the third row to the first row, we obtain the matrix

$$\begin{pmatrix} 1 & 2 & 0 & 0 & \Big| & \frac{128}{73} \\ 0 & 1 & 0 & 0 & \Big| & \frac{11}{73} \\ 0 & 0 & 1 & 0 & \Big| & \frac{19}{73} \\ 0 & 0 & 0 & 1 & \Big| & \frac{64}{73} \end{pmatrix}$$

8. Adding -2 times the second row to the first row, we obtain the matrix

$$\begin{pmatrix} 1 & 0 & 0 & 0 & \Big| & \frac{106}{73} \\ 0 & 1 & 0 & 0 & \Big| & \frac{11}{73} \\ 0 & 0 & 1 & 0 & \Big| & \frac{19}{73} \\ 0 & 0 & 0 & 1 & \Big| & \frac{64}{73} \end{pmatrix}.$$

Hence, the given system of equations admits the unique solution

$$\begin{pmatrix} \frac{106}{73} \\ \frac{11}{73} \\ \frac{19}{73} \\ \frac{64}{73} \end{pmatrix}.$$

$\square$

Exercise 3.69. Solve the system of equations in the variables x_1, x_2, x_3:
$$x_1 + 2x_2 - 5x_3 = 20,$$
$$2x_1 + 5x_2 - 7x_3 = 33,$$
$$-x_1 - 2x_2 + 4x_3 = -17.$$

Solution. The above system of equations can be expressed in matrix form as

$$\begin{pmatrix} 1 & 2 & -5 \\ 2 & 5 & -7 \\ -1 & -2 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 20 \\ 33 \\ -17 \end{pmatrix}.$$

We perform row reduction on the augmented matrix

$$(A \mid b) = \begin{pmatrix} 1 & 2 & -5 & \Big| & 20 \\ 2 & 5 & -7 & \Big| & 33 \\ -1 & -2 & 4 & \Big| & -17 \end{pmatrix}.$$

1. First, we add -2 times the first row to the second row, to obtain the matrix

$$\left(\begin{array}{ccc|c} 1 & 2 & -5 & 20 \\ 0 & 1 & 3 & -7 \\ -1 & -2 & 4 & -17 \end{array}\right).$$

2. Adding 1 times the first row to the third row, we obtain the matrix

$$\left(\begin{array}{ccc|c} 1 & 2 & -5 & 20 \\ 0 & 1 & 3 & -7 \\ 0 & 0 & -1 & 3 \end{array}\right).$$

3. Next, we multiply the third row by -1, to obtain the matrix

$$\left(\begin{array}{ccc|c} 1 & 2 & -5 & 20 \\ 0 & 1 & 3 & -7 \\ 0 & 0 & 1 & -3 \end{array}\right).$$

4. Adding -3 times the third row to the second row, we obtain the matrix

$$\left(\begin{array}{ccc|c} 1 & 2 & -5 & 20 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & -3 \end{array}\right).$$

5. Adding 5 times the third row to the first row, we obtain the matrix

$$\left(\begin{array}{ccc|c} 1 & 2 & 0 & 5 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & -3 \end{array}\right).$$

6. Finally, adding -2 times the second row to the first row, we obtain the matrix

$$\left(\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & -3 \end{array}\right).$$

Hence, the given system of equations admits the unique solution

$$\boxed{\begin{pmatrix} 1 \\ 2 \\ -3 \end{pmatrix}}.$$

$\square$

Exercise 3.70. Let A, B be square matrices of the same size. Suppose that A is invertible and $AB = I$ holds. Show that B is invertible and $B = A^{-1}$.

Exercise 3.71. Let A, B be square matrices of the same size. Suppose that A is invertible and $BA = I$ holds. Show that B is invertible and $B = A^{-1}$.

Compare the above two exercises with Exercise 3.27, Exercise 3.87.

Exercise 3.72. Let A, B be matrices, suppose that the number of columns of A is equal to the number of rows of B. Suppose AB has at least two columns. Let C be the matrix obtained from AB by removing the last column of AB. Show that there exists a matrix B' such that $C = AB'$.

Exercise 3.73. Let A, B be matrices, suppose that the number of columns of A is equal to the number of rows of B. Suppose AB has at least two rows. Let C be the matrix obtained from AB by removing the last row of AB. Show that there exists a matrix A' such that $C = A'B$.

Remark 3.74. Note that while solving the above system of equations, we performed row reduction on the augmented matrix

$$(A \mid b) = \begin{pmatrix} 1 & 2 & -5 & 20 \\ 2 & 5 & -7 & 33 \\ -1 & -2 & 4 & -17 \end{pmatrix},$$

and obtained the matrix

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

This shows that multiplying the matrix

$$A = \begin{pmatrix} 1 & 2 & -5 \\ 2 & 5 & -7 \\ -1 & -2 & 4 \end{pmatrix}$$

from the left by the product of some elementary matrices yields the identity matrix I_3, which implies that the matrix A is invertible. Indeed, if $EA = I_3$ where E is the product of some elementary matrices, then using that E is invertible, it follows that $A = E^{-1}EA = E^{-1}I_3 = E^{-1}$, and hence $AE = I_3$, which shows that A is invertible, and A^{-1} is equal to E.

The elementary matrices used to row reduce the matrix A to I_3 are

$$E_1 = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, E_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}, E_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix},$$

$$E_4 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -3 \\ 0 & 0 & 1 \end{pmatrix}, E_5 = \begin{pmatrix} 1 & 0 & 5 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, E_6 = \begin{pmatrix} 1 & -2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Note that $EA = I_3$ holds for

$$E = E_6 E_5 E_4 E_3 E_2 E_1$$

$$= \begin{pmatrix} 1 & -2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 5 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -3 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & -2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 5 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -3 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & -2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 5 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -3 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -1 & 0 & -1 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & -2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 5 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 3 \\ -1 & 0 & -1 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & -2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} -4 & 0 & -5 \\ 1 & 1 & 3 \\ -1 & 0 & -1 \end{pmatrix}$$

$$= \begin{pmatrix} -6 & -2 & -11 \\ 1 & 1 & 3 \\ -1 & 0 & -1 \end{pmatrix}.$$

This yields

$$A^{-1} = E = \begin{pmatrix} -6 & -2 & -11 \\ 1 & 1 & 3 \\ -1 & 0 & -1 \end{pmatrix}.$$

Exercise 3.75. Solve the system of equations in the variables x_1, x_2:

$$2x_1 - 3x_2 = 7,$$
$$-4x_1 + 5x_2 = -11.$$

Solution. The above system of equations can be expressed in matrix form as

$$\begin{pmatrix} 2 & -3 \\ -4 & 5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 7 \\ -11 \end{pmatrix}.$$

We perform row reduction on the augmented matrix

$$(A \mid b) = \begin{pmatrix} 2 & -3 & 7 \\ -4 & 5 & -11 \end{pmatrix}.$$

1. First, we add 2 times the first row to the second row, to obtain the matrix

$$\begin{pmatrix} 2 & -3 & 7 \\ 0 & -1 & 3 \end{pmatrix}.$$

2. Next, we multiply the second row by -1, to obtain the matrix

$$\begin{pmatrix} 2 & -3 & 7 \\ 0 & 1 & -3 \end{pmatrix}.$$

3. Add 3 times the second row to the first row to obtain the matrix

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

4. Finally, we multiply the first row by $\frac{1}{2}$, to obtain the matrix

$$\left(\begin{array}{cc|c} 1 & 0 & -1 \\ 0 & 1 & -3 \end{array}\right).$$

Hence, the given system of equations admits the unique solution

$$\boxed{\begin{pmatrix} -1 \\ -3 \end{pmatrix}}.$$

$\square$

ℹ **Remark**

Let A denote the matrix

$$\begin{pmatrix} 2 & -3 \\ -4 & 5 \end{pmatrix}.$$

Note that $EA = I_2$ holds where E is the product of some elementary matrices. More specifically, $EA = I_2$ holds for

$$E = E_4 E_3 E_2 E_1,$$

where

$$E_1 = \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}, E_2 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, E_3 = \begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix}, E_4 = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & 1 \end{pmatrix}.$$

One can argue that A is invertible, and that

$$\begin{aligned} A^{-1} &= E \\ &= E_4 E_3 E_2 E_1 \\ &= \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix} \\ &= \begin{pmatrix} \frac{1}{2} & \frac{3}{2} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -2 & -1 \end{pmatrix} \\ &= \begin{pmatrix} -\frac{5}{2} & -\frac{3}{2} \\ -2 & -1 \end{pmatrix}. \end{aligned}$$

---

§3.7.5 Row echelon form

📑 **Definition**

A matrix is said to be in **row echelon form** if the following conditions hold.
1. The first nonzero entry in each nonzero row is 1 (called a **leading 1** or a **pivot**).
2. The leading 1 in each nonzero row is to the right of the leading 1 in the previous row (if any).
3. All zero rows (if any) are at the bottom of the matrix.
4. If a column contains a pivot, then all its entries above the pivot are zero.

> ### Example
>
> The following matrices are in row echelon form:
> $$\begin{pmatrix} 0 & 1 & -5 & 0 & 0 & 0 & 0 & 0 & 4 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$
>
> The following matrices are not in row echelon form:
> $$\begin{pmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}.$$

Fact 3.76. Through a sequence of elementary row operations, any matrix can be transformed into a matrix in row echelon form.

Exercise 3.77. Transform the matrix
$$\begin{pmatrix} 0 & 2 & -3 & 0 & 0 \\ 1 & -1 & 4 & 0 & 5 \\ 0 & 0 & 0 & 1 & -2 \\ 0 & 3 & -5 & 0 & 1 \\ 0 & 3 & -5 & -1 & 2 \end{pmatrix}$$
into a matrix in row echelon form using elementary row operations.

Solution. We perform row reduction on the matrix
$$A = \begin{pmatrix} 0 & 2 & -3 & 0 & 0 \\ 1 & -1 & 4 & 0 & 5 \\ 0 & 0 & 0 & 1 & -2 \\ 0 & 3 & -5 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

1. Interchanging the first row and the second row, we obtain the matrix
$$\begin{pmatrix} 1 & -1 & 4 & 0 & 5 \\ 0 & 2 & -3 & 0 & 0 \\ 0 & 0 & 0 & 1 & -2 \\ 0 & 3 & -5 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

2. Next, multiplying the second row by $\frac{1}{2}$, we obtain the matrix
$$\begin{pmatrix} 1 & -1 & 4 & 0 & 5 \\ 0 & 1 & -\frac{3}{2} & 0 & 0 \\ 0 & 0 & 0 & 1 & -2 \\ 0 & 3 & -5 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

3. Adding -3 times the second row to the fourth row, we obtain the matrix

$$\begin{pmatrix} 1 & -1 & 4 & 0 & 5 \\ 0 & 1 & -\frac{3}{2} & 0 & 0 \\ 0 & 0 & 0 & 1 & -2 \\ 0 & 0 & -\frac{1}{2} & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

4. Interchanging the third row and the fourth row, we obtain the matrix

$$\begin{pmatrix} 1 & -1 & 4 & 0 & 5 \\ 0 & 1 & -\frac{3}{2} & 0 & 0 \\ 0 & 0 & -\frac{1}{2} & 0 & 1 \\ 0 & 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

5. Multiplying the third row by -2, we obtain the matrix

$$\begin{pmatrix} 1 & -1 & 4 & 0 & 5 \\ 0 & 1 & -\frac{3}{2} & 0 & 0 \\ 0 & 0 & 1 & 0 & -2 \\ 0 & 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

6. Adding -4 times the third row to the first row, and adding $\frac{3}{2}$ times the third row to the second row, we obtain the matrix

$$\begin{pmatrix} 1 & -1 & 0 & 0 & 13 \\ 0 & 1 & 0 & 0 & -3 \\ 0 & 0 & 1 & 0 & -2 \\ 0 & 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

7. Finally, adding the second row to the first row, we obtain the matrix

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 10 \\ 0 & 1 & 0 & 0 & -3 \\ 0 & 0 & 1 & 0 & -2 \\ 0 & 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

which is in row echelon form.

□

Exercise 3.78. Solve the system of equations in the variables x_1, x_2, x_3, x_4:
$$2x_2 - 3x_3 = 0,$$
$$x_1 - x_2 + 4x_3 = 5,$$
$$x_4 = -2,$$
$$3x_2 - 5x_3 = 1,$$
$$3x_2 - 5x_3 - x_4 = 2.$$

Solution. The above system of equations can be expressed in matrix form as

$$\begin{pmatrix} 0 & 2 & -3 & 0 \\ 1 & -1 & 4 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 3 & -5 & 0 \\ 0 & 3 & -5 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 5 \\ -2 \\ 1 \\ 2 \end{pmatrix}.$$

We perform row reduction on the augmented matrix. Adding the third row to the fifth row, we obtain the matrix

$$\left(\begin{array}{cccc|c} 0 & 2 & -3 & 0 & 0 \\ 1 & -1 & 4 & 0 & 5 \\ 0 & 0 & 0 & 1 & -2 \\ 0 & 3 & -5 & 0 & 1 \\ 0 & 3 & -5 & 0 & 0 \end{array} \right).$$

Adding -1 times the fourth row to the fifth row, we obtain the matrix

$$\left(\begin{array}{cccc|c} 0 & 2 & -3 & 0 & 0 \\ 1 & -1 & 4 & 0 & 5 \\ 0 & 0 & 0 & 1 & -2 \\ 0 & 3 & -5 & 0 & 1 \\ 0 & 0 & 0 & 0 & -1 \end{array} \right).$$

Multiplying the fifth row by -1, we obtain the matrix

$$\left(\begin{array}{cccc|c} 0 & 2 & -3 & 0 & 0 \\ 1 & -1 & 4 & 0 & 5 \\ 0 & 0 & 0 & 1 & -2 \\ 0 & 3 & -5 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{array} \right).$$

Performing several more elementary row operations, as done in the previous exercise, yields the matrix

$$\left(\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 10 \\ 0 & 1 & 0 & 0 & -3 \\ 0 & 0 & 1 & 0 & -2 \\ 0 & 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 0 & 1 \end{array} \right).$$

The corresponding system of equations has no solution, since the last column of the above matrix contains a pivot. □

Fact 3.79. Let

$$M = (A \,|\, b)$$

be the matrix obtained by augmenting a matrix A with a column vector b. If M is in its **row echelon form**, then the following statements are equivalent.

1. The system of equations $AX = b$ admits a solution.
2. The last column of the augmented matrix M contains no pivot.

If one (and hence, both) of the above statements holds, then the solutions to the system of equations $AX = b$ can be obtained by assigning arbitrary values to the variables corresponding to the non-pivot columns of the augmented matrix M, and then solving for the variables corresponding to the pivot columns of M.

A **pivot column** of a matrix in row echelon form is a column that contains a pivot.

> 📑 **Definition**
>
> If $M = (A \mid b)$ denotes the augmented matrix of a system of equations $AX = b$, and $M' = (A' \mid b')$ is the matrix in row echelon form obtained by performing a sequence of elementary row operations on M, then a variable corresponding to a non-pivot column of A' is called a **free variable**.

Recall that we obtained the following matrices earlier while reducing some augmented matrices. The first and the third matrices are in row echelon form. The second matrix is **not** in row echelon form.

$$
\begin{pmatrix}
1 & 0 & 0 & \frac{11}{3} & \frac{1}{3} & 8 \\
0 & 1 & 0 & -\frac{2}{5} & \frac{2}{5} & -\frac{3}{5} \\
0 & 0 & 1 & -\frac{17}{15} & \frac{2}{15} & -\frac{1}{5}
\end{pmatrix},
\quad
\begin{pmatrix}
1 & 2 & -1 & 1 \\
0 & 1 & -3 & 0 \\
0 & 0 & 1 & -\frac{2}{15} \\
0 & 0 & 1 & -\frac{1}{14} \\
0 & 0 & 1 & 0
\end{pmatrix},
\quad
\begin{pmatrix}
1 & 0 & 0 & 1 \\
0 & 1 & 0 & 2 \\
0 & 0 & 1 & -3
\end{pmatrix}.
$$

Fact 3.80. Every homogeneous system of linear equations with more variables than equations has at least one free variable, and hence, admits infinitely many solutions. That is, if A is an $m \times n$ matrix with $m < n$, then the homogeneous system of equations $AX = 0$ admits a nonzero solution, and hence, infinitely many solutions.

A proof of the above is provided in Chapter 6.

Fact 3.81. Let A be an $n \times n$ matrix. The following statements are equivalent.
1. The matrix A can be transformed into the identity matrix I_n by performing a sequence of elementary row operations.
2. The matrix A is a product of elementary matrices.
3. The matrix A is invertible.
4. The system of equations $AX = b$ admits a unique solution for every column vector b having n entries.

A proof of the above is provided in Chapter 6.

Exercise 3.82. Let A be a square matrix. If a sequence of elementary row operations can be performed on A to obtain the identity matrix, then A is invertible, and the same sequence of elementary row operations, when performed on the identity matrix, yields the inverse of A. Indeed, if there exists a sequence of elementary matrices

$$E_1, E_2, ..., E_k$$

such that

$$E_k...E_2 E_1 A = I,$$

then show that A is invertible, and that

$$A^{-1} = E_k...E_2 E_1 I.$$

> 📑 **Definition**
>
> The **rank** of a matrix A is the number of pivots in the row echelon form of A. The rank of a matrix A is denoted by $\mathrm{rk}(A)$.

Exercise 3.83. Find the rank of the matrix

$$\begin{pmatrix} 1 & 2 & -1 & 1 \\ 0 & 1 & -3 & 0 \\ 0 & 0 & 1 & -\frac{2}{15} \\ 0 & 0 & 1 & -\frac{1}{14} \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

Exercise 3.84. If A is an $m \times n$ matrix, then show that
$$\mathrm{rk}(A) \le \min\{m, n\}.$$

Exercise 3.85. If A is an $n \times n$ matrix, then show that the following statements are equivalent.
1. The matrix A is invertible.
2. The rank of A is n.

Solution. Using Fact 3.81, it follows that following statements are equivalent.
1. The matrix A is invertible.
2. The matrix A can be transformed into the identity matrix by performing a sequence of elementary row operations.
3. The row echelon form of A has n pivots.
4. The rank of A is equal to n.

□

Exercise 3.86. If A is an invertible square matrix, then show that A^{-1} is also invertible, and that
$$\left(A^{-1}\right)^{-1} = A.$$

Solution. Suppose A is an invertible square matrix. We have
$$AA^{-1} = A^{-1}A = I.$$
This shows that A^{-1} is also invertible, and that
$$\left(A^{-1}\right)^{-1} = A.$$

□

> **Example**
>
> The matrix
> $$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$
> is **not** in row echelon form, since the second row is a zero row, but is not at the bottom of the matrix. The rank of this matrix is 2, since its row echelon form is
> $$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

Exercise 3.87. If A, B are square matrices of the same size satisfying $BA = I$, then show that the matrices A, B are invertible, and are the inverses of each other.

Compare the above with Exercise 3.70, Exercise 3.71, Exercise 3.27.

> **i** Remark
>
> Does considering the row echelon form of B help?

Solution. Let E be a product of elementary matrices such that EB is in row echelon form. Note that
$$(EB)A = E$$
holds. Since EB is in row echelon form, it has a pivot in every nonzero row. If EB has a zero row, then the corresponding row of E is also zero, which is a contradiction. This shows that EB has no zero row, and hence, has a pivot in every row. This implies that EB is the identity matrix[11]. This shows that B is invertible, and that $B^{-1} = E$. Moreover, we obtain $A = E$, which shows that A is invertible, and that
$$A^{-1} = E^{-1} = (B^{-1})^{-1} = B$$
holds. $\qquad\square$

§3.8 Systems of linear equations in three variables and determinants of 3×3 matrices

Exercise 3.88. Solve the system of equations in the variables x, y, z:
$$x + 2y + 3z = 1,$$
$$4x + 5y + 6z = 2,$$
$$7x + 8y + 10z = 3.$$

---

[11]Why?

> **i** Remark
>
> If the matrix
> $$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 10 \end{pmatrix}$$
> is invertible, then the above system of equations admits a unique solution, which can be found by multiplying A^{-1} from the left to the matrix form of the system of equations, that is, multiplying
> $$A \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$
> from the left by A^{-1}, **provided A^{-1} exists**. Note that
> $$\frac{1}{3} \begin{pmatrix} -2 & -4 & 3 \\ -2 & 11 & -6 \\ 3 & -6 & 3 \end{pmatrix} A = \frac{1}{3} \begin{pmatrix} -2 & -4 & 3 \\ -2 & 11 & -6 \\ 3 & -6 & 3 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 10 \end{pmatrix} = I_3,$$
> and
> $$A \cdot \frac{1}{3} \begin{pmatrix} -2 & -4 & 3 \\ -2 & 11 & -6 \\ 3 & -6 & 3 \end{pmatrix} = \frac{1}{3} \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 10 \end{pmatrix} \begin{pmatrix} -2 & -4 & 3 \\ -2 & 11 & -6 \\ 3 & -6 & 3 \end{pmatrix} = I_3.$$
> This shows that A is invertible, and that
> $$A^{-1} = \frac{1}{3} \begin{pmatrix} -2 & -4 & 3 \\ -2 & 11 & -6 \\ 3 & -6 & 3 \end{pmatrix}.$$

> **Definition**
>
> Let A be a 3×3 matrix. Write $A = \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix}$. The **determinant** of A is denoted by $\det(A)$, and is defined to be
> $$\det(A) := a_1(b_2 c_3 - b_3 c_2) - b_1(a_2 c_3 - a_3 c_2) + c_1(a_2 b_3 - a_3 b_2).$$
> The **adjoint** of A is denoted by $\mathrm{adj}(A)$, and is defined to be the matrix
> $$\mathrm{adj}(A) := \begin{pmatrix} b_2 c_3 - b_3 c_2 & -(b_1 c_3 - b_3 c_1) & b_1 c_2 - b_2 c_1 \\ -(c_3 a_2 - c_2 a_3) & c_3 a_1 - c_1 a_3 & -(c_2 a_1 - c_1 a_2) \\ a_2 b_3 - a_3 b_2 & -(a_1 b_3 - a_3 b_1) & a_1 b_2 - a_2 b_1 \end{pmatrix}.$$

Exercise 3.89. Show that the determinant of the identity matrix I_3 is 1, and that
$$\mathrm{adj}(I_3) = I_3.$$

Exercise 3.90. Let
$$A = \begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{pmatrix},$$
where a, b, c are scalars. Show that $\det(A) = abc$, and that

$$\text{adj}(A) = \begin{pmatrix} bc & 0 & 0 \\ 0 & ca & 0 \\ 0 & 0 & ab \end{pmatrix}.$$

Exercise 3.91. Let

$$A = \begin{pmatrix} 1 & -5 & 6 \\ -4 & 8 & -9 \\ 7 & 2 & 10 \end{pmatrix}.$$

Find $\det(A)$ and $\text{adj}(A)$.

Exercise 3.92. Let

$$A = \begin{pmatrix} 2 & 1 & -1 \\ 3 & 2 & 1 \\ 4 & 1 & 2 \end{pmatrix}.$$

Find $\det(A)$ and $\text{adj}(A)$.

📝 **Definition**

For a 2×2 matrix

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

its **adjoint** is denoted by $\text{adj}(A)$, and is defined to be

$$\text{adj}(A) := \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Exercise 3.93. Let

$$A = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}$$

where a, b are scalars. Show that $\det(A) = ab$, and that

$$\text{adj}(A) = \begin{pmatrix} b & 0 \\ 0 & a \end{pmatrix}.$$

Prove that for a 2×2 matrix A, the equalities
$$A \cdot \text{adj}(A) = \text{adj}(A) \cdot A = \det(A)I_2$$
hold (cf. Lemma 3.22). Moreover, if A, B are 2×2 matrices, then
$$\det(AB) = \det(A)\det(B)$$
holds, as shown in Exercise 3.23.

Fact 3.94. For any 3×3 matrix A, the following holds.
$$A \cdot \text{adj}(A) = \text{adj}(A) \cdot A = \det(A)I_3.$$

Fact 3.95. If A, B are 3×3 matrices, then
$$\det(AB) = \det(A)\det(B).$$

The following is an analogue of Lemma 3.22 for 3×3 matrices.

Lemma 3.96. *Let A be a 3×3 matrix. The following statements are equivalent.*

1. *The matrix A is invertible.*
2. *The determinant $\det(A)$ of A is nonzero.*
3. *Every vector*

$$\begin{pmatrix} e \\ f \\ g \end{pmatrix}$$

can be expressed as a linear combination of the columns of A. Moreover, if A is invertible, then

$$A^{-1} = \frac{1}{\det(A)} \operatorname{adj}(A)$$

holds.

A proof of the above is provided in Chapter 6.

Exercise 3.97. Determine whether the matrix

$$A = \begin{pmatrix} 1 & -5 & 6 \\ 2 & 0 & 9 \\ -1 & 2 & -1 \end{pmatrix}$$

is invertible. If A is invertible, then find its inverse.

Solution. Note that

$$\det(A) = 1 \times (-18) - (-5) \times 7 + 6 \times 4 = -18 + 35 + 24 = 41.$$

Since $\det(A)$ is nonzero, it follows that the matrix A is invertible.

Note that

$$\operatorname{adj}(A) = \begin{pmatrix} -18 & -7 & -45 \\ 7 & 5 & 3 \\ 4 & 3 & 10 \end{pmatrix}.$$

Hence, we obtain

$$A^{-1} = \frac{1}{\det(A)} \operatorname{adj}(A) = \frac{1}{41} \begin{pmatrix} -18 & -7 & -45 \\ -7 & 5 & 3 \\ 4 & 3 & 10 \end{pmatrix}.$$

$\square$

Exercise 3.98. Use Gaussian elimination to determine the row echelon form of the matrix

$$A = \begin{pmatrix} 1 & -5 & 6 \\ 2 & 0 & 9 \\ -1 & 2 & -1 \end{pmatrix}.$$

Use the row echelon form to determine whether A is invertible. If A is invertible, then find its inverse using elementary matrices, corresponding to the elementary row operations used to transform A into its row echelon form.

Solution. Adding -2 times the first row to the second row, we obtain the matrix

$$\begin{pmatrix} 1 & -5 & 6 \\ 0 & 10 & -3 \\ -1 & 2 & -1 \end{pmatrix}.$$

Adding the first row to the third row, we obtain the matrix

$$\begin{pmatrix} 1 & -5 & 6 \\ 0 & 10 & -3 \\ 0 & -3 & 5 \end{pmatrix}.$$

Next, multiplying the second row by $\frac{1}{10}$, we obtain the matrix

$$\begin{pmatrix} 1 & -5 & 6 \\ 0 & 1 & -\frac{3}{10} \\ 0 & -3 & 5 \end{pmatrix}.$$

Adding 3 times the second row to the third row, we obtain the matrix

$$\begin{pmatrix} 1 & -5 & 6 \\ 0 & 1 & -\frac{3}{10} \\ 0 & 0 & \frac{41}{10} \end{pmatrix}.$$

Multiplying the third row by $\frac{10}{41}$, we obtain the matrix

$$\begin{pmatrix} 1 & -5 & 6 \\ 0 & 1 & -\frac{3}{10} \\ 0 & 0 & 1 \end{pmatrix},$$

Adding $\frac{3}{10}$ times the third row to the second row, we obtain the matrix

$$\begin{pmatrix} 1 & -5 & 6 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Adding -6 times the third row to the first row, we obtain the matrix

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

Finally, adding 5 times the second row to the first row, we obtain the matrix

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = I_3.$$

The above is the row echelon form of A. Since the row echelon form of A is the identity matrix I_3, it follows that A is invertible.

The elementary matrices corresponding to the above elementary row operations are

$$E_1 = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, E_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}, E_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{10} & 0 \\ 0 & 0 & 1 \end{pmatrix}, E_4 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 3 & 1 \end{pmatrix},$$

$$E_5 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{10}{41} \end{pmatrix}, E_6 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & \frac{3}{10} \\ 0 & 0 & 1 \end{pmatrix}, E_7 = \begin{pmatrix} 1 & 0 & -6 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, E_8 = \begin{pmatrix} 1 & 5 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

It follows that

$$A^{-1} = E_8 E_7 E_6 E_5 E_4 E_3 E_2 E_1$$

$$= \begin{pmatrix} 1 & 5 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & -6 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & \frac{3}{10} \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{10}{41} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 3 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{10} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 5 & -6 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & \frac{3}{41} \\ 0 & 0 & \frac{10}{41} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{10} & 0 \\ 0 & \frac{3}{10} & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 5 & -\frac{45}{41} \\ 0 & 1 & \frac{3}{41} \\ 0 & 0 & \frac{10}{41} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ -\frac{2}{10} & \frac{1}{10} & 0 \\ \frac{4}{10} & \frac{3}{10} & 1 \end{pmatrix}$$

$$= \frac{1}{41} \begin{pmatrix} -18 & 7 & -45 \\ -7 & 5 & 3 \\ 4 & 3 & 10 \end{pmatrix}.$$

□

Exercise 3.99. Solve the following system of equations in the variables x, y, z.

$$x + 2y + 3z = 1,$$
$$4x - 5y + 6z = 2,$$
$$7x - y + 10z = 3.$$

Appendix

The content of the appendix is supplementary and not required for the main course. However, it may be useful for further studies.

Chapter 4. Sets

Proof of Fact 1.23. Note that if P is a subset of Q, then we claim that $A \cup P$ is a subset of $A \cup Q$. Indeed, for an element x of $A \cup P$, note that x lies in A, or it lies in P. If x lies in A, then it lies in $A \cup Q$. If x lies in P, it also lies in Q, and hence it lies in $A \cup Q$. This proves the claim. Consequently, $A \cup (B \cap C)$ is a subset of both of the sets

$$A \cup B, A \cup C,$$

and hence, we obtain

$$\boxed{A \cup (B \cap C) \subseteq (A \cup B) \cap (A \cup C)}.$$

Let y be an element of $(A \cup B) \cap (A \cup C)$. Note that y lies in $A \cup B$ and y also lies in $A \cup C$. If y lies in A, then it also lies in $A \cup (B \cap C)$. Thus, it remains to consider the case that y does not lie in A, which we assume from now on. Since y lies in $A \cup B$ and y does not lie in A, it follows that y lies in B. Similarly, it also follows that y lies in C. This shows that y lies in $B \cap C$, and hence it lies in $A \cup (B \cap C)$. This proves that

$$\boxed{(A \cup B) \cap (A \cup C) \subseteq A \cup (B \cap C)}.$$

This establishes that

$$\boxed{A \cup (B \cap C) = (A \cup B) \cap (A \cup C)}.$$

Let us now establish that

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C).$$

Let x be an element of $A \cap (B \cup C)$. Note that x lies in A and also lies in $B \cup C$. If x lies in B, then x lies in $A \cap B$, and hence x belongs to $(A \cap B) \cup (A \cap C)$. It remains to consider the case that x does not lie in B, which we assume from now on. Since x lies in $B \cup C$ and x does not lie in B, it follows that x lies in C. Using that x lies in A, we obtain that x lies in $A \cap C$, and hence, x lies in $(A \cap B) \cup (A \cap C)$. Combining these cases, we obtain

$$\boxed{A \cap (B \cup C) \subseteq (A \cap B) \cup (A \cap C)}.$$

Note that if P is a subset of Q, then $A \cap P$ is contained in $A \cap Q$. It follows that $A \cap B$ is a subset of $A \cap (B \cup C)$, and $A \cap C$ is also a subset of $A \cap (B \cup C)$. This shows that

$$\boxed{(A \cap B) \cup (A \cap C) \subseteq A \cap (B \cup C)}.$$

This proves that

$$\boxed{A \cap (B \cup C) = (A \cap B) \cup (A \cap C)}.$$

$\square$

Proof of Fact 1.28. Note that

$$(A \cup B)^c = X \setminus (A \cup B)$$
$$= \{x \in X : x \notin (A \cup B)\}$$
$$= \{x \in X : x \notin A \text{ and } x \notin B\}$$
$$= \{x \in X : x \in A^c \text{ and } x \in B^c\}$$
$$= \{x \in X : x \in A^c \cap B^c\}$$
$$= A^c \cap B^c$$

holds for any subsets A, B of X.

As a consequence, we obtain

$$(A^c \cup B^c)^c = ((A^c)^c \cap (B^c)^c),$$

which yields

$$(A^c \cup B^c)^c = A \cap B.$$

This implies

$$((A^c \cup B^c)^c)^c = (A \cap B)^c,$$

or equivalently,

$$\boxed{(A \cap B)^c = A^c \cup B^c}.$$

$\square$

Proof of Fact 1.30. Let X denote the underlying universal set. Note that
$$X \setminus (B \cup C) = (B \cup C)^c$$
$$= B^c \cap C^c$$
$$= (X \setminus B) \cap (X \setminus C).$$

This shows that
$$A \cap (X \setminus (B \cup C)) = A \cap ((X \setminus B) \cap (X \setminus C)),$$

which implies that
$$A \cap (X \setminus (B \cup C)) = (A \cap (X \setminus B)) \cap (A \cap (X \setminus C)).$$

Note that for any subset Y of X, we have
$$A \cap (X \setminus Y) = \{x \in X : x \in A \text{ and } x \in X \setminus Y\}$$
$$= \{x \in X : x \in A \text{ and } x \notin Y\}$$
$$= A \setminus Y.$$

Consequently, we obtain

$$\boxed{A \setminus (B \cup C) = (A \setminus B) \cap (A \setminus C)}.$$

Similarly, we obtain

$$X \setminus (B \cap C) = (B \cap C)^c$$
$$= B^c \cup C^c$$
$$= (X \setminus B) \cup (X \setminus C).$$

This shows that
$$A \cap (X \setminus (B \cap C)) = A \cap ((X \setminus B) \cup (X \setminus C)).$$

Using the distributive property of intersection over union (see Fact 1.23), we obtain
$$A \cap (X \setminus (B \cap C)) = (A \cap (X \setminus B)) \cup (A \cap (X \setminus C)).$$

Since for any subset Y of X, the equality

$$A \cap (X \setminus Y) = A \setminus Y$$

holds, we conclude that

$$\boxed{A \setminus (B \cap C) = (A \setminus B) \cup (A \setminus C)} \; .$$

$\square$

Exercise 4.1 (*). Do there exist three finite sets A, B, C satisfying
$$|A \Delta B| = 1, |B \Delta C| = 2, |C \Delta A| = 4?$$

Solution. Note that for any three sets A, B, C,

$$\begin{aligned}
C \setminus A &= (C \setminus A) \cap (B \cup B^c) \\
&= ((C \setminus A) \cap B) \cup ((C \setminus A) \cap B^c) \\
&\subseteq (A^c \cap B) \cup (C \cap B^c) \\
&= (B \setminus A) \cup (C \setminus B) \\
&\subseteq (A \Delta B) \cup (B \Delta C)
\end{aligned}$$

hold. Hence, for any three sets A, B, C, we also obtain

$$\begin{aligned}
A \setminus C &\subseteq (C \Delta B) \cup (B \Delta A) \\
&= (A \Delta B) \cup (B \Delta C).
\end{aligned}$$

This shows that for any three sets A, B, C, the union of $A \setminus C, C \setminus A$ is a subset of $(A \Delta B) \cup (B \Delta C)$, that is, $C \Delta A$ is a subset of $(A \Delta B) \cup (B \Delta C)$. If A, B, C are finite sets, it follows that

$$\boxed{|C \Delta A| \le |(A \Delta B) \cup (B \Delta C)| \le |A \Delta B| + |B \Delta C|} \; .$$

This shows that there are no three finite sets A, B, C satisfying the given conditions. $\square$

Exercise 4.2. If z, w are complex numbers, then show that
  1. $|z + w| \le |z| + |w|$ (triangle inequality).

> ### i Remark (**)
>
> More formally, one defines $\mathbb{C}$, as the following set
> $$\{(a, b) : a, b \in \mathbb{R}\},$$
> equipped with addition and multiplication defined by
> $$(a, b) + (c, d) := (a + c, b + d),$$
> $$(a, b) \cdot (c, d) := (ac - bd, ad + bc),$$
> for any $(a, b), (c, d)$ with $a, b, c, d \in \mathbb{R}$.

Chapter 5. Induction principles

Theorem 5.1. *The well-ordering principle, the principle of mathematical induction, and the principle of strong induction are equivalent.*

Proof.

$\square$

Exercise 5.2 (Infinitude of primes, by Saidak)**.** Let $a_1 = 2$, and $a_{n+1} = a_n(a_n + 1)$ for any positive integer n. For any $n \in \mathbb{N}$, show that a_n has at least n distinct prime divisors.

Solution. For a positive integer n, let $P(n)$ denote the statement that $a_n \geq 1$ and a_n has at least n distinct prime divisors.

Since $a_1 = 2$, it follows that $P(1)$ is true.

Let n be a positive integer such that $P(n)$ holds. This gives that $a_n \geq 1$, and hence $a_n + 1 \geq 2$. Also note that the integers $a_n, a_n + 1$ have no common prime divisor, since any of their common divisors divides their difference, which is equal to 1. Hence, no prime divisor of a_n is a prime divisor of $a_n + 1$. By the induction hypothesis, a_n has at least n distinct prime divisors. Since $a_n + 1 \geq 2$, it admits at least one prime divisor. It follows that the product $a_n(a_n + 1)$ has at least $n + 1$ distinct prime divisors. Also note that $a_n(a_n + 1) \geq 2$ holds. This shows that $P(n + 1)$ holds.

By induction, the statement $P(n)$ holds for all $n \in \mathbb{N}$. In particular, a_n has at least n distinct prime divisors. $\square$

Exercise 5.3 (*)**.** Show that there are infinitely many prime numbers of the form $4m + 3$, where m is a nonnegative integer.

Solution. Le $a_1, a_2, \ldots$ be a sequence of positive integers defined by
$$a_1 := 7,$$
$$a_{n+1} := a_n(4a_n + 3)$$
for all $n \in \mathbb{N}$. For any $n \in \mathbb{N}$, let $P(n)$ denote the statement that a_n is a positive integer, 3 does not divide a_n, and a_n has at least n distinct prime divisors of the form $4m + 3$.

Since $a_1 = 7$, it follows that $P(1)$ holds.

Let k be a positive integer such that $P(k)$ holds. Since a_k is a positive integer, it follows that $4a_k + 3$ is an odd positive integer ≥ 3. Hence it has[12] a prime divisor of the form $4m + 3$. Since 3 does not divide a_k, it follows[13] that 3 does not divide $4a_k + 3$. Also note that any common factor of the integers a_k and $4a_k + 3$ divides
$$(4a_k + 3) - 4a_k,$$
which is equal to 3. Since 3 does not divide a_k, we obtain that the integers a_k and $4a_k + 3$ have no common prime divisor. Hence, any prime divisor of $4a_k + 3$ of the form $4m + 3$ is different from the prime divisors of a_k. This shows that a_{k+1} is a positive integer, 3 does not divide a_{k+1}, and a_{k+1} has at least $k + 1$ distinct prime divisors of the form $4m + 3$. By induction, it follows that for any $n \in \mathbb{N}$, the integer a_n has n distinct prime divisors of the form $4m + 3$.

This shows that there are infinitely many prime numbers of the form $4m + 3$. $\square$

Exercise 5.4. Show that

---

[12]How does it follow? Prove it by induction that for any positive integer n, the product of any integers of the form $4m + 1$ is also of the form $4m + 1$.

[13]Why?

$$2^{2^n} - 1$$

has at least n distinct prime divisors for all $n \in \mathbb{N}$.

Solution. For a positive integer n, let $P(n)$ denote the statement that $2^{2^n} - 1$ has at least n distinct prime divisors.

Since $2^{2^1} - 1 = 3$, it follows that $P(1)$ holds.

Let k be a positive integer such that $P(k)$ holds. Note that

$$2^{2^{k+1}} - 1 = \left(2^{2^k} - 1\right)\left(2^{2^k} + 1\right).$$

Also note that the integers $2^{2^k} - 1$ and $2^{2^k} + 1$ have no common prime divisors, since these are odd integers, and any of their common divisors divides their difference, which is equal to 2. By the induction hypothesis, the integer $2^{2^k} - 1$ has at least k distinct prime divisors. Since $2^{2^k} + 1 \geq 5$, it admits at least one prime divisor. It follows that the product $\left(2^{2^k} - 1\right)\left(2^{2^k} + 1\right)$ has at least $k + 1$ distinct prime divisors. This shows that $P(k+1)$ holds.

By induction, the statement $P(n)$ holds for all $n \in \mathbb{N}$. $\qquad\square$

> **i Remark**
>
> Note that it is false that
> $$2^{2^n} + 1$$
> has at least n distinct prime divisors for all $n \in \mathbb{N}$. Indeed, the integer $2^{2^2} + 1 = 17$ is a prime.

Exercise 5.5. Let S be a finite set of size n. Show that the number of subsets of S is 2^n.

Solution. For a nonnegative integer n, let $P(n)$ denote the statement that if A is a set with n elements, then the number of subsets of A is 2^n.

Since the empty set has exactly one subset, namely itself, it follows that $P(0)$ holds.

Let k be a nonnegative integer such that $P(k)$ holds, that is, for **any set** with k elements, the number of its subsets is 2^k. We show that $P(k+1)$ holds.

Let A be a set with $k + 1$ elements. Choose an element $a \in A$, and let $B = A \setminus \{a\}$. Note that B has exactly k elements. By the induction hypothesis, the number of subsets of B is 2^k. Any subset of A either contains a or does not contain a. The number of subsets of A that do not contain a is equal to the number of subsets of B, which is 2^k.

Note that any subset of A that contains a is of the form $C \cup \{a\}$, where C is a subset of B. Further, different subsets of B give rise to different subsets of A that contain a. That is, if C_1, C_2 are different subsets of B, then $C_1 \cup \{a\} \neq C_2 \cup \{a\}$. This shows that the number of subsets of A that contain a is also equal to the number of subsets of B, which is again 2^k. It follows that the total number of subsets of A is equal to

$$2^k + 2^k = 2^{k+1}.$$

This shows that $P(k+1)$ holds.

By induction, it follows that for any nonnegative integer n, if A is a set of size n, then the number of subsets of A is 2^n. $\qquad\square$

Exercise 5.6 (*). Show that

$$\left(1 + \frac{1}{1^3}\right)\left(1 + \frac{1}{2^3}\right)\left(1 + \frac{1}{3^3}\right)\cdots\left(1 + \frac{1}{n^3}\right) < 3$$

holds for all $n \in \mathbb{N}$.

> **i** Remark (*)
>
> Does taking $P(n)$ to be the statement that
> $$\left(1+\frac{1}{1^3}\right)\left(1+\frac{1}{2^3}\right)\left(1+\frac{1}{3^3}\right)\cdots\left(1+\frac{1}{n^3}\right) < 3$$
> help?

Solution. For any $n \in \mathbb{N}$, let $P(n)$ denote the statement that
$$\left(1+\frac{1}{1^3}\right)\left(1+\frac{1}{2^3}\right)\left(1+\frac{1}{3^3}\right)\cdots\left(1+\frac{1}{n^3}\right) \le 3 - \frac{1}{n}.$$
Note that $P(1)$ holds. Let k be a positive integer such that $P(k)$ holds. Note that

$$\left(1+\frac{1}{1^3}\right)\left(1+\frac{1}{2^3}\right)\left(1+\frac{1}{3^3}\right)\cdots\left(1+\frac{1}{k^3}\right)\left(1+\frac{1}{(k+1)^3}\right)$$

$$\le \left(3 - \frac{1}{k}\right)\left(1+\frac{1}{(k+1)^3}\right)$$

$$= 3 - \frac{1}{k} + \frac{3}{(k+1)^3} - \frac{1}{k(k+1)^3}$$

$$= 3 - \frac{1}{k+1} + \frac{1}{k+1} - \frac{1}{k} + \frac{3}{(k+1)^3} - \frac{1}{k(k+1)^3}$$

$$= 3 - \frac{1}{k+1} - \frac{1}{k(k+1)} + \frac{3}{(k+1)^3} - \frac{1}{k(k+1)^3}$$

$$= 3 - \frac{1}{k+1} - \frac{(k+1)^2 - 3k + 1}{k(k+1)^3}$$

$$= 3 - \frac{1}{k+1} - \frac{k^2 - k + 2}{k(k+1)^3}$$

$$= 3 - \frac{1}{k+1} - \frac{k(k-1) + 2}{k(k+1)^3}$$

$$< 3 - \frac{1}{k+1}$$

holds. This proves that if k is a positive integer and $P(k)$ holds, then $P(k+1)$ also holds. By the principle of induction, it follows that $P(n)$ holds for any positive integer n. This implies the given inequality for all $n \in \mathbb{N}$. $\qquad\square$

Exercise 5.7 (*). Show that
$$1 + \frac{1}{2^3} + \frac{1}{3^3} + \cdots + \frac{1}{n^3} < \frac{3}{2}$$
for all $n \in \mathbb{N}$.

Solution. For any $n \in \mathbb{N}$, let $P(n)$ denote the statement that
$$1 + \frac{1}{2^3} + \frac{1}{3^3} + \cdots + \frac{1}{n^3} \le \frac{3}{2} - \frac{1}{2n^2}.$$
Note that $P(1)$ holds. Let k be a positive integer such that $P(k)$ holds. Note that

$$1 + \frac{1}{2^3} + \frac{1}{3^3} + \cdots + \frac{1}{k^3} + \frac{1}{(k+1)^3}$$

$$\leq \left(\frac{3}{2} - \frac{1}{2k^2}\right) + \frac{1}{(k+1)^3} \qquad \text{(using } P(k)\text{)}$$

$$= \frac{3}{2} - \frac{1}{2k^2} + \frac{1}{(k+1)^3}$$

$$= \frac{3}{2} - \frac{1}{2(k+1)^2} + \frac{1}{2(k+1)^2} - \frac{1}{2k^2} + \frac{1}{(k+1)^3}$$

$$= \frac{3}{2} - \frac{1}{2(k+1)^2} - \frac{(k+1)^3 - k^2(k+1) - 2k^2}{2k^2(k+1)^3}$$

$$= \frac{3}{2} - \frac{1}{2(k+1)^2} - \frac{k^3 + 3k^2 + 3k + 1 - k^3 - k^2 - 2k^2}{2k^2(k+1)^3}$$

$$= \frac{3}{2} - \frac{1}{2(k+1)^2} - \frac{3k+1}{2k^2(k+1)^3}$$

$$< \frac{3}{2} - \frac{1}{2(k+1)^2}.$$

This proves that if k is a positive integer and $P(k)$ holds, then $P(k+1)$ also holds. By the principle of induction, it follows that $P(n)$ holds for any positive integer n. This implies the given inequality for all $n \in \mathbb{N}$.

$\square$

Exercise 5.8 (*). Show that

$$\left(1 + \frac{1}{2}\right)\left(1 + \frac{1}{2^2}\right)\left(1 + \frac{1}{2^3}\right)\cdots\left(1 + \frac{1}{2^n}\right) < \frac{5}{2}$$

holds for all $n \in \mathbb{N}$.

Solution. For any $n \in \mathbb{N}$, let $P(n)$ denote the statement that

$$\left(1 + \frac{1}{2}\right)\left(1 + \frac{1}{2^2}\right)\left(1 + \frac{1}{2^3}\right)\cdots\left(1 + \frac{1}{2^n}\right) \leq \frac{5}{2}\left(1 - \frac{2}{2^{n+1} + 1}\right).$$

Note that $P(1)$ holds. Let k be a positive integer such that $P(k)$ holds. Note that

$$\left(1 + \frac{1}{2}\right)\left(1 + \frac{1}{2^2}\right)\left(1 + \frac{1}{2^3}\right)\cdots\left(1 + \frac{1}{2^k}\right)\left(1 + \frac{1}{2^{k+1}}\right)$$

$$\leq \frac{5}{2}\left(1 - \frac{2}{2^{k+1} + 1}\right)\left(1 + \frac{1}{2^{k+1}}\right)$$

$$\leq \frac{5}{2} \times \frac{2^{k+1} - 1}{2^{k+1}}$$

$$\leq \frac{5}{2}\left(1 - \frac{1}{2^{k+1}}\right)$$

$$\leq \frac{5}{2}\left(1 - \frac{1}{2^{k+2} + 1}\right).$$

This proves that if k is a positive integer and $P(k)$ holds, then $P(k+1)$ also holds. By the principle of induction, it follows that $P(n)$ holds for any positive integer n. This implies the given inequality for all $n \in \mathbb{N}$.

$\square$

Chapter 6. Matrices

Let us revisit Exercise 3.16, in the light of Exercise 3.19.

Exercise 6.1. Compute B^2 where B is the following $n \times n$ matrix

$$B = \begin{pmatrix} 0 & \cdots & 0 & 1 \\ 0 & \cdots & 1 & 0 \\ \vdots & \iddots & \vdots & \vdots \\ 1 & \cdots & 0 & 0 \end{pmatrix}.$$

Solution. Let $e_1, e_2, ..., e_n$ denote the standard basis vectors of $\mathbb{R}^n$. Note that B sends $e_1, e_2, ..., e_{n-1}, e_n$ to the vectors $e_n, e_{n-1}, ..., e_2, e_1$ respectively. That is,

$$Be_i = e_{n+1-i}$$

holds for all $1 \le i \le n$. Hence, for any integer $1 \le i \le n$, it follows that

$$B^2 e_i = B(Be_i)$$
$$= B(e_{n+1-i})$$
$$= e_{n+1-(n+1-i)}$$
$$= e_i.$$

This shows that[14]

$$B^2 = I_n.$$

$\square$

Proof of Fact 3.20. Using $AB = I_n$, we obtain

$$CAB = CI_n = C.$$

Using $CA = I_n$, we get

$$CAB = I_n B = B.$$

This shows that $B = C$. $\square$

> **i Remark**
>
> Note that if A is a matrix and P is an invertible matrix such that PAP^{-1} is a diagonal matrix, then using Exercise 3.28, the powers of A can be computed. Indeed, if
>
> $$PAP^{-1} = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix},$$
>
> then
>
> $$A^k = P^{-1}(PA^kP^{-1})P = P^{-1}(PAP^{-1})^k P = P^{-1} \begin{pmatrix} \lambda_1^k & 0 & \cdots & 0 \\ 0 & \lambda_2^k & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n^k \end{pmatrix} P$$
>
> holds for any positive integer k. Thus, Exercise 3.28 can be used to compute powers of matrices which are similar to diagonal matrices.

Exercise 6.2. Does the following hold?

---

[14]How does it follow? Does Exercise 3.19 help?

$$\begin{pmatrix} -\frac{\sqrt{5}+1}{2} & \frac{\sqrt{5}-1}{2} \\ 1 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} -\frac{\sqrt{5}+1}{2} & \frac{\sqrt{5}-1}{2} \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} \frac{-\sqrt{5}+1}{2} & 0 \\ 0 & \frac{1+\sqrt{5}}{2} \end{pmatrix}$$

> **? Question**
>
> What can be said about the Fibonacci numbers using Exercise 3.15 and Exercise 3.28?

Proof of Lemma 3.30. In the following, the first and the second column of A are denoted by C_1 and C_2 respectively. Write

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Let us first show that the first three statements are equivalent. Assume that A is invertible. If $\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}$ is a solution to Equation 11, then

$$A \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

This shows that

$$A^{-1} A \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

which yields

$$\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

This implies that the solution $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ is the only solution to Equation 11. This shows that the first statement implies the second statement. Assume that the second statement holds, that is, Equation 11 admits no solution other than $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$. If s, t are real numbers such that

$$sC_1 + tC_2 = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

then

$$A \begin{pmatrix} s \\ t \end{pmatrix} = sC_1 + tC_2 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

This shows that $\begin{pmatrix} s \\ t \end{pmatrix}$ is a solution to Equation 11, and hence, $s = t = 0$. This shows that the second statement implies the third statement. Assume that the third statement holds, that is, the trivial linear combination of the columns of A is the only linear combination of the columns of A that is equal to $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$. If A is not invertible, then by Lemma 3.22, we have $ad - bc = 0$, which implies that

$$dC_1 - cC_2 = \begin{pmatrix} ad - bc \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

$$-bC_1 + aC_2 = \begin{pmatrix} 0 \\ ad - bc \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

and consequently, a, b, c, d are equal to 0, and hence, any linear combination of the columns of A is equal to $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$, contradicting the third statement. This shows that if the third statement holds, then A is invertible, and hence, the third statement implies the first statement. This proves the equivalence of the first three statements.

For the equivalence of the next two statements, note that the following statements are equivalent.

- any element of $\mathbb{R}^2$ is a solution to Equation 11

- $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ are solutions to Equation 11
- the matrix A is the zero matrix.

$\square$

Proof of Fact 3.37. Let A be a square matrix. Note that

$$A = \frac{1}{2}(A + A^T) + \frac{1}{2}(A - A^T).$$

Note that the matrix $\frac{1}{2}(A + A^T)$ is symmetric, and that the matrix $\frac{1}{2}(A - A^T)$ is skew-symmetric. This shows that A can be expressed as the sum of a symmetric matrix and a skew-symmetric matrix.

Assume that

$$A = S_1 + K_1 = S_2 + K_2,$$

where S_1, S_2 are symmetric matrices, and K_1, K_2 are skew-symmetric matrices. It follows that

$$S_1 - S_2 = K_2 - K_1.$$

Note that the left-hand side of the above equality is a symmetric matrix, and that the right-hand side of the above equality is a skew-symmetric matrix. This shows that each of $S_1 - S_2, K_2 - K_1$ is both symmetric and skew-symmetric. Since the only matrix that is both symmetric and skew-symmetric is the zero matrix, we have $S_1 - S_2 = K_2 - K_1 = 0$. This yields

$$S_1 = S_2,$$
$$K_1 = K_2.$$

This implies that any square matrix A can be **uniquely** expressed as the sum of a symmetric matrix and a skew-symmetric matrix. $\square$

Proof of Fact 3.51. Note that

$$A = \frac{1}{2}(A + A^*) + \frac{1}{2}(A - A^*).$$

Note that the matrix $\frac{1}{2}(A + A^*)$ is hermitian, and that the matrix $\frac{1}{2}(A - A^*)$ is skew-hermitian. This shows that A can be expressed as the sum of a hermitian matrix and a skew-hermitian matrix.

Assume that

$$A = H_1 + K_1 = H_2 + K_2,$$

where H_1, H_2 are hermitian matrices, and K_1, K_2 are skew-hermitian matrices. It follows that

$$H_1 - H_2 = K_2 - K_1.$$

Note that the left-hand side of the above equality is a hermitian matrix, and that the right-hand side of the above equality is a skew-hermitian matrix. This shows that each of $H_1 - H_2, K_2 - K_1$ is both hermitian and skew-hermitian. Since the only matrix that is both hermitian and skew-hermitian is the zero matrix, we have $H_1 - H_2 = K_2 - K_1 = 0$. This yields

$$H_1 = H_2,$$
$$K_1 = K_2.$$

This implies that any square matrix A with complex entries can be **uniquely** expressed as the sum of a hermitian matrix and a skew-hermitian matrix. $\square$

Proof of Lemma 3.63. Let A be an $n \times m$ matrix. If

$$E = I_n + \lambda e_{ij} \text{ with } i \neq j,$$

then

$$EA = (I_n + \lambda e_{ij})A = A + \lambda e_{ij}A$$

holds, and hence, the k-th row of EA is equal to the k-th row of A if $k \neq i$, and is equal to the sum of the i-th row of A and the row vector obtained by multiplying the j-th row of A by λ if $k = i$.

Moreover, if

$$E = I_n - e_{ii} - e_{jj} + e_{ij} + e_{ji} \text{ with } i \neq j,$$

then

$$EA = \left(I_n - e_{ii} - e_{jj} + e_{ij} + e_{ji}\right)A = A - e_{ii}A - e_{jj}A + e_{ij}A + e_{ji}A$$

holds, and hence, the k-th row of EA is equal to the k-th row of A if $k \neq i, j$, and the i-th row of EA is equal to the j-th row of A, and the j-th row of EA is equal to the i-th row of A.

Finally, if

$$E = I_n + (\lambda - 1)e_{ii} \text{ with } \lambda \neq 0,$$

then

$$EA = (I_n + (\lambda - 1)e_{ii})A = A + (\lambda - 1)e_{ii}A$$

holds, and hence, the k-th row of EA is equal to the k-th row of A if $k \neq i$, and is equal to the vector obtained by multiplying the i-th row of A by the scalar λ if $k = i$.

This shows that left multiplication by an elementary matrix on a matrix A performs the corresponding elementary row operation on the matrix A. □

Proof of Lemma 3.64. Let E be an elementary matrix. We consider the three types of elementary matrices separately.

Suppose E is the elementary matrix obtained by interchanging the i-th row and the j-th row of the identity matrix. Since multiplying any matrix by E from the left interchanges its i-th and j-th rows, we obtain

$$E^2 = EE = I_n.$$

This shows that E is invertible and that $E^{-1} = E$.

Now consider the case that E is the elementary matrix obtained by multiplying the i-th row of the identity matrix by a nonzero scalar λ. Since multiplying any matrix A by E from the left has the same effect of multiplying the i-th row of A by the same nonzero scalar λ, it follows that

$$E\left(I_n + \left(\frac{1}{\lambda} - 1\right)e_{ii}\right) = I_n.$$

Similarly, it follows that

$$\left(I_n + \left(\frac{1}{\lambda} - 1\right)e_{ii}\right)E = I_n.$$

This shows that E is invertible and that

$$E^{-1} = I_n + \left(\frac{1}{\lambda} - 1\right)e_{ii}.$$

Finally, consider the case that E is the elementary matrix obtained by adding a scalar multiple of the j-th row of the identity matrix to its i-th row, that is,

$$E = I_n + \lambda e_{ij} \text{ with } i \neq j,$$

where λ is a scalar. Using Exercise 3.61, it follows that

$$E(I_n - \lambda e_{ij}) = (I_n + \lambda e_{ij})(I_n - \lambda e_{ij}) = I_n,$$

and

$$(I_n - \lambda e_{ij})E = (I_n - \lambda e_{ij})(I_n + \lambda e_{ij}) = I_n,$$

which shows that E is invertible and that

$$E^{-1} = I_n - \lambda e_{ij}.$$

This shows that any elementary matrix is invertible, and its inverse is also an elementary matrix. □

Proof of Theorem 3.66. Let X_0 be a solution to the system of equations $AX = b$, that is, $AX_0 = b$ holds. It follows that

$$A'X_0 = EAX_0 = Eb = b'$$

holds, which shows that X_0 is a solution to the system of equations $A'X = b'$.

Conversely, let Y_0 be a solution to the system of equations $A'X = b'$, that is, $A'Y_0 = b'$ holds. Since the elementary matrices are invertible, and E is a product of elementary matrices, it follows that E is invertible. This yields

$$AY_0 = E^{-1}A'Y_0 = E^{-1}b' = b,$$

which shows that Y_0 is a solution to the system of equations $AX = b$. $\qquad\square$

Proof of Fact 3.80. Let $M = (A \mid 0)$ denote the augmented matrix corresponding to the homogeneous system of equations $AX = 0$. Performing row reduction on M, we obtain a matrix $M' = (A' \mid 0)$ in row echelon form. Since A has more columns than rows, the matrix A' has more columns than rows. Hence, the matrix A' has at least one non-pivot column, which implies that the system of equations $A'X = 0$ has at least one free variable. Assigning an arbitrary nonzero value to this free variable, and then solving for the other variables, we obtain a nonzero solution to the system of equations $A'X = 0$. Since M' is obtained from M by performing a sequence of elementary row operations, it follows that the system of equations $AX = 0$ also admits a nonzero solution. Considering the scalar multiples of any such nonzero solution (cf. Exercise 3.33), we see that the homogeneous system of equations $AX = 0$ admits infinitely many solutions. $\qquad\square$

Proof of Fact 3.81. Note that if A can be transformed into the identity matrix by performing a sequence of elementary row operations, then there are elementary matrices $E_1, E_2, ..., E_k$ such that

$$E_k...E_2E_1A = I_n,$$

which implies that

$$A = (E_k...E_2E_1)^{-1} = E_1^{-1}E_2^{-1}...E_k^{-1},$$

and consequently, A is a product of elementary matrices.

If A is a product of elementary matrices, then using the fact that every elementary matrix is invertible, it follows that A is invertible.

Note that the vector $A^{-1}b$ is a solution to the $AX = b$ If A is invertible, then for any solution X_0 to the system of equations $AX = b$, we have $AX_0 = b$, which implies that $A^{-1}AX_0 = A^{-1}b$, which yields $X_0 = A^{-1}b$. This shows that the system of equations $AX = b$ admits a unique solution if A is invertible.

Finally, let us assume that the system of equations $AX = b$ admits a unique solution for every column vector b having n entries. In particular, the homogeneous system of equations $AX = 0$ admits only the zero solution. Performing row reduction on the augmented matrix $(A \mid 0)$, we obtain a matrix $(A' \mid 0)$ in row echelon form. Since the homogeneous system of equations $AX = 0$ admits only one solution, each column of A' must contain a pivot of A'. This implies that the matrix A' has no non-pivot column, and hence, A' has n pivots. In other words, A' is equal to the identity matrix I_n. Since $(A' \mid 0)$ is obtained from $(A \mid 0)$ by performing a sequence of elementary row operations, it follows that A can be transformed into the identity matrix by performing a sequence of elementary row operations.

This proves the equivalence of the four statements. $\qquad\square$

§6.1 Systems of linear equations in three variables and determinants of 3×3 matrices

Let us revisit the solution to Exercise 3.88.

> **?** **Question**
>
> Can one **come up** with a formula for the inverse of an *invertible* 3×3 matrix? Can one also have a notion of determinant of a 3×3 matrix, and use it to determine whether a 3×3 matrix is invertible or not? Can one have an analogue of Lemma 3.22 and Exercise 3.23 for 3×3 matrices?

Consider the system of linear equations in three variables x, y, z:

$$a_1 x + b_1 y + c_1 z = d_1, \tag{13}$$

$$a_2 x + b_2 y + c_2 z = d_2, \tag{14}$$

$$a_3 x + b_3 y + c_3 z = d_3, \tag{15}$$

where a_i, b_i, c_i, d_i are real numbers for $i = 1, 2, 3$. This system of equations can be expressed in matrix form as

$$\begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} d_1 \\ d_2 \\ d_3 \end{pmatrix}.$$

The matrix

$$A = \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix}$$

is called the **coefficient matrix** of the system of equations, and the vector

$$\begin{pmatrix} d_1 \\ d_2 \\ d_3 \end{pmatrix}$$

is called the **constant vector** of the system of equations. The system of equations is said to be **homogeneous** if

$$\begin{pmatrix} d_1 \\ d_2 \\ d_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Otherwise, the system of equations is said to be **non-homogeneous**.

> **i** **Remark**
>
> The discussion in Section 3.3 and Section 3.5 can be adapted to the case of systems of linear equations in three variables. In particular, if $\begin{pmatrix} d_1 \\ d_2 \\ d_3 \end{pmatrix}$ is a linear combination of the columns of the coefficient matrix A, then the solutions to the system of equations
>
> $$A \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} d_1 \\ d_2 \\ d_3 \end{pmatrix}$$
>
> are in one-to-one correspondence with the solutions to the system of equations
>
> $$A \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Multiplying Equation 14 by c_3 and Equation 15 by $-c_2$, and adding the results, we obtain the equation

$$(a_2 c_3 - a_3 c_2)x + (b_2 c_3 - b_3 c_2)y = d_2 c_3 - d_3 c_2. \tag{16}$$

Similarly, multiplying Equation 14 by $-b_3$ and Equation 15 by b_2, and adding the results, we obtain the equation

$$(a_2 b_3 - a_3 b_2)x + (c_2 b_3 - c_3 b_2)z = d_2 b_3 - d_3 b_2. \tag{17}$$

Multiplying Equation 13 by $b_2 c_3 - b_3 c_2$ and Equation 16 by $-b_1$, and adding the results, we obtain the equation

$$\begin{aligned}
(a_1(b_2 c_3 - b_3 c_2) &- b_1(a_2 c_3 - a_3 c_2))x + (c_1(b_2 c_3 - b_3 c_2))z \\
&= d_1(b_2 c_3 - b_3 c_2) - b_1(d_2 c_3 - d_3 c_2).
\end{aligned} \tag{18}$$

Multiplying Equation 17 by c_1 and adding to Equation 18, we obtain the equation

$$\begin{aligned}
(a_1(b_2 c_3 - b_3 c_2) &- b_1(a_2 c_3 - a_3 c_2) + c_1(a_2 b_3 - a_3 b_2))x \\
&= d_1(b_2 c_3 - b_3 c_2) - b_1(d_2 c_3 - d_3 c_2) + c_1(d_2 b_3 - d_3 b_2),
\end{aligned} \tag{19}$$

or equivalently, we obtain

$$\begin{aligned}
(a_1(b_2 c_3 - b_3 c_2) &- b_1(a_2 c_3 - a_3 c_2) + c_1(a_2 b_3 - a_3 b_2))x \\
&= d_1(b_2 c_3 - b_3 c_2) - d_2(b_1 c_3 - b_3 c_1) + d_3(b_1 c_2 - b_2 c_1).
\end{aligned} \tag{20}$$

Similarly[15], it follows that

$$\begin{aligned}
(b_2(c_3 a_1 - c_1 a_3) &- c_2(b_3 a_1 - b_1 a_3) + a_2(b_3 c_1 - b_1 c_3))y \\
&= d_2(c_3 a_1 - c_1 a_3) - d_3(c_2 a_1 - c_1 a_2) + d_1(c_2 a_3 - c_3 a_2),
\end{aligned} \tag{21}$$

and

$$\begin{aligned}
(c_3(a_1 b_2 - a_2 b_1) &- a_3(c_1 b_2 - c_2 b_1) + b_3(c_1 a_2 - c_2 a_1))z \\
&= d_3(a_1 b_2 - a_2 b_1) - d_1(a_3 b_2 - a_2 b_3) + d_2(a_3 b_1 - a_1 b_3).
\end{aligned} \tag{22}$$

This may indicate that it could be useful to consider the product of the matrix

$$\begin{pmatrix}
b_2 c_3 - b_3 c_2 & -(b_1 c_3 - b_3 c_1) & b_1 c_2 - b_2 c_1 \\
c_2 a_3 - c_3 a_2 & c_3 a_1 - c_1 a_3 & -(c_2 a_1 - c_1 a_2) \\
-(a_3 b_2 - a_2 b_3) & a_3 b_1 - a_1 b_3 & a_1 b_2 - a_2 b_1
\end{pmatrix}$$

$$= \begin{pmatrix}
b_2 c_3 - b_3 c_2 & -(b_1 c_3 - b_3 c_1) & b_1 c_2 - b_2 c_1 \\
-(c_3 a_2 - c_2 a_3) & c_3 a_1 - c_1 a_3 & -(c_2 a_1 - c_1 a_2) \\
a_2 b_3 - a_3 b_2 & -(a_1 b_3 - a_3 b_1) & a_1 b_2 - a_2 b_1
\end{pmatrix}$$

with the matrix A. Consider the product

$$M := \begin{pmatrix}
b_2 c_3 - b_3 c_2 & -(b_1 c_3 - b_3 c_1) & b_1 c_2 - b_2 c_1 \\
-(c_3 a_2 - c_2 a_3) & c_3 a_1 - c_1 a_3 & -(c_2 a_1 - c_1 a_2) \\
a_2 b_3 - a_3 b_2 & -(a_1 b_3 - a_3 b_1) & a_1 b_2 - a_2 b_1
\end{pmatrix} \begin{pmatrix}
a_1 & b_1 & c_1 \\
a_2 & b_2 & c_2 \\
a_3 & b_3 & c_3
\end{pmatrix}.$$

> **? Question**
>
> Does it follow that if one of m_{12}, m_{23}, m_{31} is identically zero, then the remaining two are also identically zero. Does it also follow that if one of m_{13}, m_{21}, m_{32} is identically zero, then the remaining two are also identically zero. Does it follow that M is a diagonal matrix? Does it also follow that M is a scalar matrix, that is, its diagonal entries are equal?

Fact 6.3. The product M is a diagonal matrix whose diagonal entries are equal to each other, and is equal to

---

[15]How does it follow? What would be a useful **similarity** to consider?

$$\det(A)I_3,$$

where

$$\det(A) := a_1(b_2c_3 - b_3c_2) - b_1(a_2c_3 - a_3c_2) + c_1(a_2b_3 - a_3b_2).$$

Moreover, it follows that

$$\begin{pmatrix} b_2c_3 - b_3c_2 & -(b_1c_3 - b_3c_1) & b_1c_2 - b_2c_1 \\ -(c_3a_2 - c_2a_3) & c_3a_1 - c_1a_3 & -(c_2a_1 - c_1a_2) \\ a_2b_3 - a_3b_2 & -(a_1b_3 - a_3b_1) & a_1b_2 - a_2b_1 \end{pmatrix} \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix}$$

$$= \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix} \begin{pmatrix} b_2c_3 - b_3c_2 & -(b_1c_3 - b_3c_1) & b_1c_2 - b_2c_1 \\ -(c_3a_2 - c_2a_3) & c_3a_1 - c_1a_3 & -(c_2a_1 - c_1a_2) \\ a_2b_3 - a_3b_2 & -(a_1b_3 - a_3b_1) & a_1b_2 - a_2b_1 \end{pmatrix}$$

$$= \det(A)I_3.$$

📑 **Definition**

Let A be a 3×3 matrix. Write $A = \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix}$. The **determinant** of A is denoted by $\det(A)$,

and is defined to be

$$\det(A) := a_1(b_2c_3 - b_3c_2) - b_1(a_2c_3 - a_3c_2) + c_1(a_2b_3 - a_3b_2).$$

The **adjoint** of A is denoted by $\mathrm{adj}(A)$, and is defined to be the matrix

$$\mathrm{adj}(A) := \begin{pmatrix} b_2c_3 - b_3c_2 & -(b_1c_3 - b_3c_1) & b_1c_2 - b_2c_1 \\ -(c_3a_2 - c_2a_3) & c_3a_1 - c_1a_3 & -(c_2a_1 - c_1a_2) \\ a_2b_3 - a_3b_2 & -(a_1b_3 - a_3b_1) & a_1b_2 - a_2b_1 \end{pmatrix}.$$

Using the fact above, the following analogue of Lemma 3.22 for 3×3 matrices can be established[16].

Lemma 6.4. *Let A be a 3×3 matrix. The following statements are equivalent.*
1. *The matrix A is invertible.*
2. *The determinant $\det(A)$ of A is nonzero.*
3. *Every vector*

$$\begin{pmatrix} e \\ f \\ g \end{pmatrix}$$

can be expressed as a linear combination of the columns of A. Moreover, if A is invertible, then

$$A^{-1} = \frac{1}{\det(A)}\,\mathrm{adj}(A)$$

holds.

Note that the above lemma is same as Lemma 3.96.

Proof of Lemma 3.96. If $\det(A)$ is nonzero, then using

$$A \cdot \mathrm{adj}(A) = \mathrm{adj}(A) \cdot A = \det(A)I_3,$$

it follows that A is invertible.

If A is invertible, then every vector

---

[16] In fact, the first few steps of the proof of Lemma 3.22 is precisely the analogue of Fact 3.94 for 2×2 matrices.

$$\begin{pmatrix} e \\ f \\ g \end{pmatrix}$$

can be expressed as a linear combination of the columns of A, since the standard basis vectors can be expressed as a linear combination of the columns of A.

Moreover, if every vector can be expressed as a linear combination of the columns of A, then in particular, the standard basis vectors can be expressed as a linear combination of the columns of A. This shows that there are real numbers $x_1, y_1, z_1, x_2, y_2, z_2, x_3, y_3, z_3$ such that

$$A \begin{pmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{pmatrix} = I_3.$$

This implies that A has nonzero determinant by Fact 3.95. $\qquad\square$