4.1 Sets of matrices and sets of vectors.

0. *Assumed background.*
   - Whatever has been covered in Topics 1-3.

   Abstract. We introduce:—
   - the notion of sets (in the context of sets of numbers, sets of matrices, and sets of column/row vectors),
   - the method of specification for the construction of sets,
   - the notions of solution set of system of linear equations, null space of a matrix, and span of a collection of vectors.

1. **Motivation for the use of notion of sets in linear algebra.**

   In linear algebra, we very often encounter not individual objects, but collections of many objects which as individuals are on equal footing. And then we further want to compare such collections. In fact we have already come across with such phenomena.

   Illustration.

   Let $A = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \end{bmatrix}$, $B = \begin{bmatrix} 2 & 3 & -5 \\ -3 & 2 & 1 \end{bmatrix}$, $C = \begin{bmatrix} 4 & -3 & -1 \end{bmatrix}$.

   (a) At first sight as A, B, C are distinct as individual matrices, we should expect nothing can be said of them.

   (b) In fact, A is row-equivalent to B (and so is $[\ A \mid \mathbf{0}_2\ ]$ is row-equivalent to $[\ B \mid \mathbf{0}_2\ ]$). For this reason, the solutions of $\mathcal{LS}(A, \mathbf{0}_2)$ are the 'same' as those of $\mathcal{LS}(B, \mathbf{0}_2)$ in the sense that:—
   - For any column vector $\mathbf{t}$ with 3 entries, $\mathbf{t}$ is a solution of $\mathcal{LS}(A, \mathbf{0}_2)$ if and only if $\mathbf{t}$ is a solution of $\mathcal{LS}(B, \mathbf{0}_2)$.

   Note that this 'same-ness' is not about individual column vectors. Amongst the solutions of either systems, there are many *distinct* solutions; but 'as a whole' the collections of solutions of the respective systems are the 'same'.

   (c) Denote the first and second rows of A by $\mathbf{a}_1, \mathbf{a}_2$.

   It happens that C is a linear combination of the rows of A:

   $$C = [\ 4 \quad -3 \quad -1\ ] = 4\,[\ 1 \quad 0 \quad -1\ ] - 3\,[\ 1 \quad 0 \quad -1\ ] = 4\mathbf{a}_1 - 3\mathbf{a}_2$$

   It follows that every solution of $\mathcal{LS}(A, \mathbf{0}_2)$ is automatically a solution of $\mathcal{LS}(C, \mathbf{0}_1)$ as well. Or more formally:—
   - For any column vector $\mathbf{t}$ with 3 entries, if $\mathbf{t}$ is a solution of $\mathcal{LS}(A, \mathbf{0}_2)$ then $\mathbf{t}$ is a solution of $\mathcal{LS}(C, \mathbf{0}_1)$.

   For this reason, we will want to say that the solutions of $\mathcal{LS}(A, \mathbf{0}_3)$ are 'part of' the solutions of $\mathcal{LS}(C, \mathbf{0}_1)$.

   This relation of 'something being part of something else' is not about individual column vectors, but collections of column vectors.

2. **Purpose of set language in linear algebra.**

   To make *description of* and *comparison amongst* collections of objects more efficient, we use the language of sets.

   In almost all situations in this course where we want to relate or compare 'collections of objects', the objects under question column/row vectors of the same size (or sometimes numbers, or matrices of the same size).

3. **Real numbers, or complex numbers?**

   For some reason which is related to the nature of set language and which is beyond the scope of this course, we will need to be more precise with the use of words and symbols that have to do with numbers.

   From now on, for simplicity of presentation of theory, we will consistently work with real numbers, column/row vectors with real entries, and matrices with real entries.

   However, unless otherwise stated, all upcoming definitions, mathematical results, and their proofs can be adapted to give the corresponding versions involving complex numbers when we consistently:—
   - change the word '*real*' to the word '*complex*', and
   - change the symbol '$\mathbb{R}$' to the symbol '$\mathbb{C}$'.

4. **The notions of 'belong to', 'element of a set'.**

   Imagine we have 'collected', say, some individual objects.

   We may refer to such a collection as a **set**.

   We may assign a symbol, say, S, for such a set, and start thinking of S as an individual object on its own.

(a) Suppose u is an object amongst those 'collected into' S.
Then we agree to say that

'*the object* u **belongs to** *the set* S'

or equivalently,

'*the object* u *is an* **element of** *the set* S',

or equivalently,

'*the set* S **contains** *the object* u **as an element**',

We may present the above in short-hand as:

'$u \in S$'

(which should always be read as 'u *belongs to* S'.)
When the context is clear, some people may choose to read this as 'u *is in* S'. This is not preferrable for beginners.

(b) Suppose v is an object not amongst the objects 'collected into' S.
Then we agree to say that

'*the object* v *does not belong to the set* S'

and write

'$v \notin S$'.

5. **Special sets in linear algebra.**

Amongst various sets of objects dealt with in linear algebra, some are special, and deserve special labels.

(a) The **set of all real numbers** is denoted by $\mathbb{R}$.

(b) The **set of all complex numbers** is denoted by $\mathbb{C}$.

(c) The **set of all** $(p \times q)$**-matrices with real entries** is denoted by $M_{p,q}(\mathbb{R})$.

(d) The **set of all** $(p \times q)$**-matrices with complex entries** is denoted by $M_{p,q}(\mathbb{C})$.

We will very often work with column vectors. For this reason, the sets of all column vectors (with real or complex entries) are given special labels:

(e) The **set of all column vectors with** p **real entries** is denoted by $\mathbb{R}^p$.
(So by definition, $\mathbb{R}^p$ is the same set as $M_{p,1}(\mathbb{R})$.)

(f) The **set of all column vectors with** p **complex entries** is denoted by $\mathbb{C}^p$.
(So by definition, $\mathbb{C}^p$ is the same set as $M_{p,1}(\mathbb{C})$.)

We will occasionally work with row vectors. The sets of all row vectors (with real or complex entries) are given special labels:

(g) The **set of all row vectors with** q **real entries** is denoted by $\mathbb{R}^q_{\text{row}}$.
(So by definition, $\mathbb{R}^q_{\text{row}}$ is the same set as $M_{1,q}(\mathbb{R})$.)

(h) The **set of all row vectors with** q **complex entries** is denoted by $\mathbb{C}^q_{\text{row}}$.
(So by definition, $\mathbb{C}^q_{\text{row}}$ is the same set as $M_{1,q}(\mathbb{C})$.)

6. **Empty set.**

Sometimes we will have to consider the set to which no object belongs.

Such a set is denoted by $\emptyset$, and is called the **empty set**.

According to its definition, it happens that

$$u \notin \emptyset \quad \text{no matter what } u \text{ is.}$$

7. **Notations and terminologies for 'small' sets.**

Suppose $a, b, \cdots, c$ are 'finitely many' objects (in the sense that we can list them out exhaustively).

(It is not assumed that $a, b, \cdots, c$ are pairwise distinct: there may be repetitions.)

Then we agree that we may present the set whose elements are exactly $a, b, \cdots, c$, by writing it as

$$\left\{ \quad a, \quad b, \quad \cdots, \quad c \quad \right\}.$$

We call it **the set whose elements are the objects** $a, b, \cdots, c$.

The symbols

$$\text{‘}\left\{\right.\text{’} \qquad \text{and} \qquad \text{‘}\left.\right\}\text{’}$$

in the notation ‘$\{a, b, \cdots, c\}$’ are used for of reminding ourselves to think of this set (and read it) as an object on its own.

The symbol ‘ $\left\{\right.$ ’ signifies the beginning of the list of objects which are elements of the set concerned.

The symbol ‘ $\left.\right\}$ ’ signifies the end of this list.

8. **Conventions on the notations for ‘small’ sets.**

   There are two rules for the display of ‘small’ sets by exhaustively listing its elements.

   (a) **‘Repetition in the list’ does not count.**

   Provided an object, say, b, is an element of a ‘small’ set, say, S, the object b has to be presented at least once in the list representing S.

   However, no matter how many more times b is presented, it still counts as once.

   Illustration:

   $\{a, a, b, c, b, b\}$ is the same set as $\{a, b, c\}$.

   (Repetition of a does not count. Repetition of b does not count.)

   (b) **‘Ordering in the list’ does not matter.**

   Given any two lists, as long every object which is presented in each list is also presented in the other, the two lists will represent the same ‘small’ set, regardless of the order in which the objects is presented in each list.

   Illustration:

   $\{a, b, c\}$, $\{b, c, a\}$, $\{c, a, b\}$, $\{a, c, b\}$, $\{b, a, c\}$, $\{c, b, a\}$ all stand for the same set, namely the set whose elements are exactly the not necessarily distinct objects a, b, c.

9. **Example (1).**

   The chain of symbols

$$\text{‘} \quad \left\{ \quad \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \quad \begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 2 \\ 4 \end{bmatrix}, \quad \begin{bmatrix} 3 \\ 0 \\ 5 \end{bmatrix} \quad \right\} \quad \text{’}$$

   stands for the set with exactly four elements, namely the column vectors

$$\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \quad \begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 2 \\ 4 \end{bmatrix}, \quad \begin{bmatrix} 3 \\ 0 \\ 5 \end{bmatrix}.$$

   When we denote this set by S, we may write

$$\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \in S, \qquad \begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix} \in S, \qquad \begin{bmatrix} 0 \\ 2 \\ 4 \end{bmatrix} \in S, \qquad \begin{bmatrix} 3 \\ 0 \\ 5 \end{bmatrix} \in S.$$

   It happens that

$$\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \notin S, \quad \text{and} \quad \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} \notin S.$$

Note that the sets below are all S itself in disguise:

$$\left\{ \begin{bmatrix}1\\2\\3\end{bmatrix}, \begin{bmatrix}1\\2\\3\end{bmatrix}, \begin{bmatrix}2\\3\\1\end{bmatrix}, \begin{bmatrix}0\\2\\4\end{bmatrix}, \begin{bmatrix}3\\0\\5\end{bmatrix} \right\},$$

$$\left\{ \begin{bmatrix}1\\2\\3\end{bmatrix}, \begin{bmatrix}2\\3\\1\end{bmatrix}, \begin{bmatrix}0\\2\\4\end{bmatrix}, \begin{bmatrix}3\\0\\5\end{bmatrix}, \begin{bmatrix}1\\2\\3\end{bmatrix} \right\},$$

$$\left\{ \begin{bmatrix}1\\2\\3\end{bmatrix}, \begin{bmatrix}3\\0\\5\end{bmatrix}, \begin{bmatrix}0\\2\\4\end{bmatrix}, \begin{bmatrix}2\\3\\1\end{bmatrix} \right\},$$

$$\left\{ \begin{bmatrix}2\\3\\1\end{bmatrix}, \begin{bmatrix}1\\2\\3\end{bmatrix}, \begin{bmatrix}3\\0\\5\end{bmatrix}, \begin{bmatrix}0\\2\\4\end{bmatrix}, \begin{bmatrix}3\\0\\5\end{bmatrix}, \begin{bmatrix}2\\3\\1\end{bmatrix} \right\}.$$

10. **Construction of sets through the method of specification.**

    In many a situation, it is undesirable or impossible to describe a set, say, S, by listing its elements exhaustively, because the set under question has 'too many' elements.

    Fortunately, very often all elements of S are themselves elements of some better understood set, say, T.

    So to present the set S, we may choose to write down an appropriate '**selection criterion**' (that mathematicians call **predicate**) which instructs ourselves exactly which objects that belong to T are to be 'collected' to give the elements of S, and which not.

    Such a method for describing the set S is called the **method of specification**.

    Below we illustrate the idea in this method in two concrete examples of sets of column vectors:

    (1) Solution set of a system of linear equations,

    (2) Span of a collection of column vectors (of the same size).

    Using the same idea, we can construct other sets of interest when they are needed.

    But the method is so general that it can be apply to construct sets of other types of objects out of any given set.

11. **Construction of solution set of a system of linear equations, and null space of a matrix, through the method of specification.**

    Suppose A is an $(m \times n)$-matrix with real entries, and $\mathbf{b}$ is a column vector with m real entries.

    We may regard the system of linear equations

    $$A\mathbf{x} = \mathbf{b}$$

    with unknown column vector $\mathbf{x}$ with n entries as a 'selection criterion'.

    (a) We may apply this 'selection criterion' on column vectors with n real entries:—

    - Those objects, which, upon substitution into the '$\mathbf{x}$' in the selection criterion '$A\mathbf{x} = \mathbf{b}$', yield a vector equality, will be collected.
    - Those objects, which, upon substitution into the '$\mathbf{x}$' in the selection criterion '$A\mathbf{x} = \mathbf{b}$', does not yield a vector equality, will be discarded.

    (b) According to the definition of the phrase 'solution of a system', what are collected through the application of such a 'selection criterion'

    $$A\mathbf{x} = \mathbf{b}$$

    are exactly

    those and only those column vectors with n real entries which are solutions of $\mathcal{LS}(A, \mathbf{b})$.

    The resultant set is called the **(real) solution set of the system** $\mathcal{LS}(A, \mathbf{b})$, and is denoted by $\mathcal{S}_{\mathbb{R}}(A, \mathbf{b})$.

    i. *Simplification of notations.*
    Where there is no ambiguity about the entries of the column vectors being real numbers, we may dispense with the symbol '$\mathbb{R}$', and simply write $\mathcal{S}_{\mathbb{R}}(A, \mathbf{b})$ as $\mathcal{S}(A, \mathbf{b})$.

    ii. By definition, $\mathbf{t}$ belongs to $\mathcal{S}(A, \mathbf{b})$ if and only if

    $$\mathbf{t} \in \mathbb{R}^n \quad \text{and} \quad A\mathbf{t} = \mathbf{b}.$$

    (c) We display the set $\mathcal{S}(A, \mathbf{b})$ as this chain of symbols:—

    $$\left\{ \mathbf{x} \in \mathbb{R}^n \;\middle|\; A\mathbf{x} = \mathbf{b} \right\}.$$

    Every symbol in this chain has some role to play:—

- To the right of the 'vertical stroke', the selection criterion '$A\mathbf{x} = \mathbf{b}$' is displayed.
- To the left of the 'vertical stroke', the set of objects, namely, $\mathbb{R}^n$, on which the 'selection criterion' is applied, is indicated.
- The 'vertical stroke' has no specific meaning; its purpose is to separate the two chain of symbols '$\mathbf{x} \in \mathbb{R}^n$', '$A\mathbf{x} = \mathbf{b}$'. We may use the colon ':' instead of the 'vertical stroke'.
- The symbol '$\Big\{$' signifies the beginning of the chain of symbols that describes the set $\mathcal{S}(A, \mathbf{b})$.

  The symbol '$\Big\}$' signifies of the chain of symbols that describes the set $\mathcal{S}(A, \mathbf{b})$.
- Note the presence of the symbol '$\mathbf{x}$' in '$\mathbf{x} \in \mathbb{R}^n$' to the left of the 'vertical stroke'.
  It needs to be consistent with the choice of the symbol '$\mathbf{x}$' in the selection criterion '$A\mathbf{x} = \mathbf{b}$'.
  As long as '$\mathbf{x}$' is replaced by the same symbol in a consistent manner, the meaning of the whole chain of symbols remains unchanged. For instance,

  $$\left\{ \mathbf{x} \in \mathbb{R}^n \;\middle|\; A\mathbf{x} = \mathbf{b} \right\}, \qquad \left\{ \mathbf{y} \in \mathbb{R}^n \;\middle|\; A\mathbf{y} = \mathbf{b} \right\}, \qquad \left\{ \mathbf{z} \in \mathbb{R}^n \;\middle|\; A\mathbf{z} = \mathbf{b} \right\}$$

  refer to the same set, namely, $\mathcal{S}(A, \mathbf{b})$.

(d) (For the moment, suppose $\mathbf{b} = \mathbf{0}_m$.)
The set

$$\left\{ \mathbf{x} \in \mathbb{R}^n \;\middle|\; A\mathbf{x} = \mathbf{0}_m \right\}$$

is called the **(real) null space of** A, and is denoted by $\mathcal{N}_{\mathbb{R}}(A)$.
Simplification of notations.
Where there is no ambiguity about the entries of the column vectors being real numbers, we may dispense with the symbol '$\mathbb{R}$', and simply write $\mathcal{N}_{\mathbb{R}}(A)$ as $\mathcal{N}(A)$.

12. **Construction of span of a collection of column vectors of the same size, through the method of specification.**

Suppose $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_q$ are column vectors with p real entries.

We may regard, as a 'selection criterion', the sentence with 'indeterminate' $\mathbf{x}$

'$\mathbf{x}$ *is a linear combination of* $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_q$ *with respect to real scalars*',

or more formally,

'*There exist some* $\alpha_1, \alpha_2, \cdots, \alpha_q \in \mathbb{R}$ *such that* $\mathbf{x} = \alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \cdots + \alpha_q\mathbf{u}_q$',

(a) We may apply this 'selection criterion' on column vectors with p real entries:—
  - Those objects, which, upon substitution into the '$\mathbf{x}$' in the selection criterion yield a true statement, will be collected.
  - Those objects, which, upon substitution into the '$\mathbf{x}$' in the selection criterion yield a false statement, will be discarded.

(b) According to the definition of the phrase 'linear combination', what are collected through the application of such a 'selection criterion' are exactly

those and only those column vectors with p real entries
which are linear combinations of $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_q$ with respect to real scalars.

The resultant set is called the **(real) span of the column vectors** $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_q$, and is denoted by $\text{Span}_{\mathbb{R}}(\{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_q\})$.

  i. *Simplification of notations.*
  Where there is no ambiguity about the entries of the column vectors being real numbers, we may dispense with the symbol '$\mathbb{R}$', and simply write $\text{Span}_{\mathbb{R}}(\{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_q\})$ as $\text{Span}(\{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_q\})$.

  ii. By definition, $\mathbf{y}$ belongs to $\text{Span}(\{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_q\})$ if and only if

  $\mathbf{y} \in \mathbb{R}^p$ and $\mathbf{y}$ is a linear combination of $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_q$ with respect to some real scalars.

  iii. According to the definition of $\text{Span}(\{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_q\})$, the ordering in the presentation of the list $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_q$ does not matter.
  This is consistent with both the definition of linear combinations, and the convention on the notations for 'small' sets regarding the use of the symbols '$\{$', '$\}$'.

(c) We display the set $\text{Span}(\{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_q\})$ as this chain of symbols:—

$$\left\{ \mathbf{x} \in \mathbb{R}^p \;\middle|\; \begin{array}{c} \mathbf{x} \text{ is a linear combination of } \mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_q \\ \text{with respect to some real scalars.} \end{array} \right\}.$$

Or as:—

$$\left\{ \mathbf{x} \in \mathbb{R}^p \;\middle|\; \begin{array}{c} \text{There exist some } \alpha_1, \alpha_2, \cdots, \alpha_q \in \mathbb{R} \\ \text{such that } \mathbf{x} = \alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \cdots + \alpha_q\mathbf{u}_q. \end{array} \right\}.$$

(d) *'Short-hand' notation.*

It is also acceptable to present the set $\text{Span}(\{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_q\})$ as

$$\left\{ \alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \cdots + \alpha_q\mathbf{u}_q \;\middle|\; \alpha_1, \alpha_2, \cdots, \alpha_q \in \mathbb{R} \right\}.$$

- Such a 'short-hand' seems to be easier to read.
- It is also visually appealing: it reminds us that $\text{Span}(\{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_q\})$ is resultant from the collection of all column vectors that 'can be obtained' when the symbols $\alpha_1, \alpha_2, \cdots, \alpha_q$ in the expression

$$\alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \cdots + \alpha_q\mathbf{u}_q$$

are 'varied' amongst all possible real numbers.
- However, when you are giving an argument or doing a computation which involve such a set, you should bear in mind what it really is according to how it is constructed.

13. We put on record the definitions for the various objects that we have introduced in the above illustrations on the method of specification.

Definition. (Solution set of a system of linear equations, and null space.)

Let A be an $(m \times n)$-matrix with real entries, and $\mathbf{b}$ be a column vector with m real entries.

(1) *We define the* **(real) solution set** *of $\mathcal{LS}(A, \mathbf{b})$ to be the set $\{\mathbf{x} \in \mathbb{R}^n \mid A\mathbf{x} = \mathbf{b}\}$. We denote it by $\mathcal{S}(A, \mathbf{b})$.*

(2) *We define the* **(real) null space** *of A to be the set $\{\mathbf{x} \in \mathbb{R}^n \mid A\mathbf{x} = \mathbf{0}_m\}$. We denote it by $\mathcal{N}(A)$.*

Definition. (Span of a collection of column vectors.)

Suppose $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_q$ are column vectors with p real entries.

Then we define the **(real) span** *of $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_q$ to be the set*

$$\left\{ \mathbf{x} \in \mathbb{R}^p \;\middle|\; \begin{array}{c} \text{There exist some } \alpha_1, \alpha_2, \cdots, \alpha_q \in \mathbb{R} \\ \text{such that } \mathbf{x} = \alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \cdots + \alpha_q\mathbf{u}_q. \end{array} \right\}.$$

We denote it by $\text{Span}(\{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_q\})$.