Introduction to linear algebra

PART 1

[1]

Section Overview

- **Key points: Vectors**
  - definition of a vector in $\mathbb{R}^2$, $\mathbb{R}^3$ and $\mathbb{R}^n$
  - vector addition and scalar multiplication
  - linear combinations of vectors
  - solving systems of linear equations
  - row echelon form of a matrix
  - elementary row operations
  - homogeneous systems
  - spanning sets
  - linear independence of a set of vectors
- **Associated sections of the book:**
  - Poole Section 1.1 (p1 - 13). Omit binary vectors and modular arithmetic (p13-16).
  - Poole Section 2.1 (p58 - 63).
  - Poole Section 2.2 (p64 - 79). Omit: rank of a matrix (p71-72); linear systems over $\mathbb{Z}_p$ (p77-79)
  - Poole Section 2.3 (p88 - 97).

1.1: The Geometry and Algebra of Vectors

The set of all ordered pairs of real numbers is denoted by $\mathbb{R}^2$,

$$\mathbb{R}^2 = \{(x,y) : \; x, y \in \mathbb{R}\}\,.$$

A vector is a directed line segment corresponding to a displacement from one point A to another point B. The set of all points in the plane ie $\mathbb{R}^2$ can be identified with the set of vectors with tails at the origin O.

If A is the point (a_1, a_2) then the vector $\mathbf{a} = \overrightarrow{OA}$ has coordinates $[a_1, a_2]$. We write $\mathbf{a} = [a_1, a_2]$. Often we use column vectors instead of row vectors,

$$\begin{bmatrix} 3 \\ 1 \end{bmatrix} \text{ rather than } [3, 1].$$

The vector $\mathbf{0} = [0, 0]$ is called the zero vector.

Two vectors are equal if and only if they have the same length and the same direction. Algebraically this means that $[a_1, a_2] = [b_1, b_2]$ if and only if $a_1 = b_1$ and $a_2 = b_2$.

---

[1] Please send comments, corrections and clarifications to peter.stewart@glasgow.ac.uk and gregory.stevenson@glasgow.ac.uk.

Vector addition

In general if $\mathbf{u} = [u_1, u_2]$ and $\mathbf{v} = [v_1, v_2]$ then we define

$$\mathbf{u} + \mathbf{v} = [u_1 + v_1, u_2 + v_2].$$

Example Consider the vectors $\mathbf{u} = [2, 1]$ and $\mathbf{v} = [2, 3]$. Sketch $\mathbf{u}$ and $\mathbf{v}$. Construct and sketch $\mathbf{u} + \mathbf{v}$.

Scalar multiplication

Given a vector $\mathbf{v} = [v_1, v_2]$ and $c \in \mathbb{R}$, the scalar multiple $c\mathbf{v}$ is

$$c\mathbf{v} = [cv_1, cv_2].$$

Example Consider the vector $\mathbf{v} = [1, 2]$. Sketch $\mathbf{v}$. Construct and sketch $2\mathbf{v}$, $\frac{1}{2}\mathbf{v}$ and $-\mathbf{v}$.

We write $-\mathbf{v}$ for $(-1)\mathbf{v}$, and then use this to define subtraction by

$$\mathbf{u} - \mathbf{v} = \mathbf{u} + (-\mathbf{v}) = [u_1 - v_1, u_2 - v_2].$$

Example If $A = (-1, 2)$ and $B = (3, 4)$ compute $\overrightarrow{AB}$.

Key notational point

In this course all vectors will be written in bold font, as $\mathbf{v}$ and scalars in usual font c. When writing by hand we

always underline our vectors (in place of the bold font).

Thus $\mathbf{0}$ is the zero vector, while 0 is the real number zero, and the equation $0\mathbf{0} = \mathbf{0}$ holds. [2]

Vectors in $\mathbb{R}^3$

$\mathbb{R}^3$ is the set of all ordered triples of real numbers

$$\mathbb{R}^3 = \{(x, y, z) : \ x, y, z \in \mathbb{R}\}\,.$$

Vectors in $\mathbb{R}^3$ can be added and multiplied by scalars in an analogous way to $\mathbb{R}^2$.

Vectors in $\mathbb{R}^n$

For $n \in \mathbb{N}$, $\mathbb{R}^n$ is the set of all n-tuples of real numbers,[3]

$$\mathbb{R}^n = \{(x_1, x_2, \cdots, x_n) : \ x_1, \cdots, x_n \in \mathbb{R}\}\,.$$

A vector $\mathbf{v}$ in $\mathbb{R}^n$ can be written as a row or column vector

$$[v_1, v_2, \cdots, v_n] \text{ or } \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}.$$

We define vector addition and scalar multiplication in $\mathbb{R}^n$ componentwise. If

$$\mathbf{u} = [u_1, u_2, \cdots u_n]\,, \ \mathbf{v} = [v_1, v_2, \cdots v_n] \in \mathbb{R}^n,$$

then the i^{th} component of $\mathbf{u} + \mathbf{v}$ is $u_i + v_i$ and for $c \in \mathbb{R}$, the i^{th} component of $c\mathbf{u}$ is cu_i.

The fundamental properties of $\mathbb{R}^n$ are given by: [4]

---

[2] Why do we do this? There's many reasons: it helps us be clear what our objects are: are they vectors, or scalars. This should prevent you from performing illegal operations. You can add two vectors, but you can't add a vector and a scalar. Mathematics is heavily 'typed' subject: every object (or symbol) has a type, and you can only perform operations between objects of the required type. For example addition allows you to add two vectors, or two scalars, but not a scalar and a vector. What this means for you is that you should **always** be clear what any given object in a mathematical argument is. In future courses lecturers' policies on this will differ, but in this course it is a hard rule so as to help us get used to this important distinction.

[3] This is our first example of a generalisation. We're familiar with 2 dimensional and 3 dimensional vectors, and have a geometrical intuition for how these behave. We use our experience with $\mathbb{R}^2$ and $\mathbb{R}^3$ to generalise to $\mathbb{R}^n$.

[4] This 'theorem' looks ragingly obvious, so why do we point it out explicitly? Well it is precisely these facts which are fundamental to the study of vectors, and so they are used to define the abstract concept of a *vector space* in Section 6.1. Think about this when you're revising.

Theorem 1.1 Algebraic Properties of Addition Let $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w} \in \mathbb{R}^n$ and $c, d \in \mathbb{R}$. Then

(a) $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$ (commutativity of vector addition)

(b) $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$ (commutativity of vector addition)

(c) $\mathbf{u} + \mathbf{0} = \mathbf{u}$

(d) $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$

(e) $c(\mathbf{u} + \mathbf{v}) = c\mathbf{u} + c\mathbf{v}$ (distributivity of vector addition)

(f) $(c + d)\mathbf{u} = c\mathbf{u} + d\mathbf{u}$ (distributivity of scalar addition)

(g) $c(d\mathbf{u}) = (cd)\mathbf{u}$

(h) $1\mathbf{u} = \mathbf{u}$

Proof: Exercise. [5]

Linear combinations of vectors

Definition A vector $\mathbf{v}$ is a *linear combination* of vectors $\mathbf{v}_1, \cdots, \mathbf{v}_k$ if and only if there are scalars $c_1, c_2, \cdots, c_k$ such that

$$\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k.$$

The scalars $c_1, c_2, \cdots, c_k$ are called the *coefficients of the linear combination*. We will see linear combinations again in the setting of abstract vector spaces.

Example Write $\begin{bmatrix} -1 \\ 2 \\ -11 \end{bmatrix}$ as a linear combination of $\begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}$ and $\begin{bmatrix} 3 \\ 0 \\ 7 \end{bmatrix}$.

---

[5] With this sort of theorem it would not be a good idea to learn the proof of by heart. Instead you should work out what is going on: to prove each of these statements you work out what the i-th component of the left hand side and the right hand side are. These two components are real numbers, and various properties of real numbers are used to see they're equal. For example in (a), the i^{th} component of the left hand side is $u_i + v_i$, while the i^{th}-component of the right hand side is $v_i + u_i$. These are equal as addition is commutative in $\mathbb{R}$.

Example Let $\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ be the standard coordinate vectors in $\mathbb{R}^2$. Compute the coordinates of $\begin{bmatrix} 1 \\ 3 \end{bmatrix}$ with respect to $\mathbf{e}_1$ and $\mathbf{e}_2$.

Example Let $\mathbf{u} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$ then compute the coordinates of $\begin{bmatrix} 1 \\ 3 \end{bmatrix}$ with respect to $\mathbf{u}$ and $\mathbf{v}$.

Definition A *linear equation* in the n variables $x_1, x_2, \cdots, x_n$ is an equation that can be written in the form

$$a_1 x_1 + a_2 x_2 + \cdots a_n x_n = b.$$

A *solution* of this linear equation is a vector $[s_1, s_2, \cdots, s_n]$ so that

$$a_1 s_1 + a_2 s_2 + \cdots a_n s_n = b.$$

A *system of linear equations* is a finite set of linear equations and a *solution* to a system of linear equations is a vector that is simultaneously a solution to all equations in the system.

Example Consider the system $2x + y = 1$, $x - y = 2$.

Example Consider the system $2x + y = 1$, $4x + 2y = 2$.

Example Consider the system $2x + y = 1$, $2x + y = 0$.

A system of equations is *consistent* if it has at least one solution. Otherwise it is *inconsistent*.

2.1, 2.2: Solving systems of Linear Equations

A linear system[6]

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1, \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m, \end{aligned}$$

has *augmented matrix*

$$[A|b] = \left[\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right].$$

These augmented matrices can be used to solve systems of linear equations using Gaussian elimination. The procedure is:

a) Write the augmented matrix;

b) Use Elementary Row Operations (see below) to reduce the matrix to row echelon form.

c) If the system is consistent, use back substitution to solve the equivalent system that corresponds to the row-reduced matrix.

[6] i.e. system of linear equations (in the example m equations)

Make sure you can do this reliably.

Elementary row operations are operations that can be performed on a system of linear equations to transform it into an equivalent system.

Definition The following *elementary row operations* can be performed on a matrix:

1. Interchange two rows
2. Multiply a row by a non-zero constant
3. Add a multiple of a row to another row

Definition A matrix is in *row echelon* form if and only if

1. any all-zero rows are at the bottom
2. in each non-zero row, the first non-zero entry (the leading entry) is to the left of any leading entries below it

Row reduction is the process for reducing any matrix to row-echelon form, using elementary row operations (EROs).

Definition A matrix is in *reduced row echelon form* if and only if

1. it is in row echelon form
2. the leading entry in each nonzero row is 1
3. each column containing a leading 1 has 0s everywhere else

Definition A system of linear equations is *homogeneous* if and only if the constant term in each equation is zero.

Example Solve the system

$$\begin{aligned} 2x_2 + 3x_3 &= 8, \\ 2x_1 + 3x_2 + x_3 &= 5, \\ x_1 - x_2 - 2x_3 &= -5. \end{aligned}$$

2.3 Spanning sets

In 3-dimensions we are used to expanding out vectors in terms of the canonical vectors **i**, **j**, **k**. But as you'll know, for some problems it can be better to work with different co-ordinates. We now start to work out how this should work in our more general setting.[7]

Spanning sets If $S = \{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k\}$ is a set of vectors in $\mathbb{R}^n$, then the set of all linear combinations of $\mathbf{v}_1$, $\mathbf{v}_2, \cdots, \mathbf{v}_k$ is called the *span* of $\mathbf{v}_1$, $\mathbf{v}_2, \cdots, \mathbf{v}_k$ and is denoted $\text{span}(\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k)$ or $\text{span}(S)$. Symbolically,

$$\text{span}(S) = \{\sum_{i=1}^{k} \mathbf{v}_i c_i : c_1, c_2, \cdots, c_k \in \mathbb{R}\}.$$

Example Is $\begin{bmatrix} 8 \\ -5 \\ -15 \end{bmatrix}$ in the span of $\left\{ \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} -2 \\ 3 \\ 5 \end{bmatrix} \right\}$? Equivalently, is $\begin{bmatrix} 8 \\ -5 \\ -15 \end{bmatrix}$ a linear combination of $\begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} -2 \\ 3 \\ 5 \end{bmatrix}$?

Theorem 2.4 A system of linear equations with augmented matrix $[A|\mathbf{b}]$ is consistent if and only if $\mathbf{b}$ is a linear combination of the columns of A

Proof: Omitted.

Example Consider the system $2x + y = 1$, $x - y = 2$.

---

[7] Experience suggests that the first time students work with spanning sets and linearly independent sets (see below) it takes a while for these concepts to sink in. Look carefully at the definitions and work though plenty of examples. Make sure that you can state these definitions precisely from memory: it is not enough to have some intuitive idea of what these concepts mean roughly. We must know exactly what they means so we can prove theorems.

Example Consider the system $2x + y = 1$, $4x + 2y = 2$.

Example Consider the system $2x + y = 1$, $2x + y = 0$.

Definition Let $S = \{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k\}$ be a set of vectors in $\mathbb{R}^n$. S is a *spanning set* for $\mathbb{R}^n$ if and only if $\text{span(S)} = \mathrm{\mathbb{R}^n}$.

That is, S is a spanning set for $\mathbb{R}^n$ if and only if every vector in $\mathbb{R}^n$ can be written as a linear combination of elements of S.[8]. This combination might not be unique.[9]

Example Show that $\mathbb{R}^2 = \text{span}\left(\begin{bmatrix}2\\1\end{bmatrix}, \begin{bmatrix}1\\-1\end{bmatrix}\right)$.

---

[8] So $\{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$ is a spanning set for $\mathbb{R}^3$, while $\{\mathbf{i}, \mathbf{j}\}$ is not

[9] This is the reason behind the next definition of linear independence.

Example In $\mathbb{R}^2$ let $\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. Show that $\mathbb{R}^2 = \text{span}(\mathbf{e}_1, \mathbf{e}_2)$.

Example In $\mathbb{R}^3$ let $\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, $\mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$ and $\mathbf{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$. Show that $\mathbb{R}^3 = \text{span}(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$.

Example In $\mathbb{R}^n$ let $\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$, $\mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}$, $\ldots \mathbf{e}_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$. Show that $\mathbb{R}^n = \text{span}(\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$.

Example Find a geometric description of the span of $\begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$ and $\begin{bmatrix} 3 \\ 4 \\ 2 \end{bmatrix}$.

Linear independence

Definition A set $\{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k\}$ of vectors in $\mathbb{R}^n$ is *linearly independent* if and only if the only solution to the equation

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k = \mathbf{0}, \qquad c_i \in \mathbb{R}$$

is $c_1 = c_2 = \cdots = c_k = 0$. A set of vectors is *linearly dependent* if it is not linearly independent.

Example Are the vectors $\mathbf{u} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ and $\mathbf{w} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ linearly independent?

Example Are the vectors $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{w}$ and $\mathbf{x} = \begin{bmatrix} 5 \\ 7 \end{bmatrix}$ linearly independent?

For any vectors $\mathbf{v}_1, \cdots, \mathbf{v}_k$,

$$0\mathbf{v}_1 + \cdots + 0\mathbf{v}_k = \mathbf{0}.$$

So a set of vectors S is linearly independent if the only way to express $\mathbf{0}$ as a linear combination of the elements of S is with all coefficients 0.[10]

Theorem 2.5 Vectors $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k$ in $\mathbb{R}^n$ are linearly dependent if and only if at least one of them can be expressed as a linear combination of the others.

Proof: Omitted.

Any set of vectors containing the zero vector is linearly dependent since

$$1\mathbf{0} + 0\mathbf{v}_2 \cdots + 0\mathbf{v}_m = \mathbf{0}.$$

Two vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ are linearly dependent if and only if they are scalar multiples of each other.

Example Determine whether the set

$$\left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \right\}$$

is linearly independent.

---

[10] This tells you how to prove that a given set $\{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k\}$ is linearly independent. Set up the equation

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k = \mathbf{0}, \qquad c_i \in \mathbb{R}$$

and find out whether there are any non-zero solutions. For explicit vectors $\{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k\}$ in $\mathbb{R}^n$ you can use Gaussian elimination to do this. (See theorem 2.6 below which sets this out concretely)

Example Determine whether the set

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

is linearly independent.

Theorem 2.6 Let $\mathbf{v}_1, \cdots, \mathbf{v}_m$ be (column) vectors in $\mathbb{R}^n$ and let

$$A = [\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_m]$$

be the $n \times m$ matrix with columns $\mathbf{v}_1, \cdots, \mathbf{v}_m$. Then $\mathbf{v}_1, \cdots, \mathbf{v}_m$ are linearly dependent if and only if the homogeneous linear system with augmented matrix $[A|\mathbf{0}]$ has a non-trivial solution.

Proof: Omitted.

Theorem 2.8 If $m > n$ then any set of m vectors in $\mathbb{R}^n$ is linearly dependent.

Proof: Omitted.

Example Show that the set

$$\left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 5 \\ 7 \end{bmatrix}, \begin{bmatrix} 212 \\ 16 \end{bmatrix} \right\}$$

is linearly dependent in $\mathbb{R}^2$.