ITCS 531: Linear Algebra - Vector spaces over fields

Rob Egrot

What is linear algebra?

- Linear algebra is an abstract approach to thinking about Euclidean space.
- What is a Euclidean space?
- Examples include the plane in 2 dimensions, and the 3 dimensional grid.
- These spaces have an *origin* (the point at zero), and two, three, or some other number of dimensions.
- For each dimension we have an axis, and we can define the position of points by how far along each axis they are.
- Euclidean space is not curved. So, for example the Euclidean plane is a flat plane in space. It's not curved around the surface of a sphere, or in any other way.

What is linear algebra for?

- The axioms of linear algebra allow geometric facts to be proved with very clean arguments.
- By abstracting away from intuitions about physical space we can see the underlying mathematics more clearly.
- Conversely, by taking an abstract approach we can 'see' systems that are not obviously geometric as 'spaces in disguise'.
- We can use geometric reasoning about these 'secret spaces'.

Where is linear algebra used?

- Linear algebra is used almost everywhere mathematics is used.

- Physicists need it to understand e.g. quantum systems.
- Statisticians use it, e.g. principal component analysis.
- Pure mathematicians like to reformulate problems as linear algebra problems so they can solve them.
- Computer scientists use linear algebra too, e.g:
  - The Google page rank algorithm.
  - Machine learning, e.g. ANN, SVM.
  - 3D graphics.

What will we cover on this course?

- Since this is a short course we will only scratch the surface.
- We will introduce the basic abstract definitions and try to understand how they relate to the idea of a space.
- We will prove some fundamental results using abstract arguments.
- At the end of the course we will use these abstract results to prove some geometric facts.
- The idea is that the rigorous approach taken here will give you the background you need to go deeper.

Complex numbers

- The complex numbers $\mathbb{C}$ are obtained by adding a root for the equation $x^2 + 1 = 0$ to the real numbers $\mathbb{R}$.
- This root is a new number called i.
- It turns out that if we add i, then we get roots for every other polynomial equation too.
- So every polynomial over $\mathbb{C}$ factorizes into linear factors (the *Fundamental Theorem of Algebra*).
- We can define $\mathbb{C}$ as the set of all numbers $a + bi$ where $a, b \in \mathbb{R}$.
- We have

$$(a + bi) \pm (c + di) = (a \pm c) + (b \pm d)i$$

and

$$(a + bi) \times (c + di) = ac - bd + (ad + bc)i.$$

Complex arithmetic

Lemma 1

Let α, β and γ be complex numbers. Then:

1. $\alpha + \beta = \beta + \alpha$, *and* $\alpha\beta = \beta\alpha$ *(commutativity).*
2. $\alpha + (\beta + \gamma) = (\alpha + \beta) + \gamma$, *and* $\alpha(\beta\gamma) = (\alpha\beta)\gamma$ *(associativity).*
3. $0 + \alpha = \alpha$, *and* $1\alpha = \alpha$ *(identities).*
4. *There is a unique* $-\alpha \in \mathbb{C}$ *such that* $\alpha + (-\alpha) = 0$ *(inverse for addition).*
5. *If* $\alpha \neq 0$ *there is a unique* α^{-1} *such that* $\alpha\alpha^{-1} = 1$ *(inverse for multiplication).*
6. $\alpha(\beta + \gamma) = \alpha\beta + \alpha\gamma$ *(distributivity).*

Complex arithmetic - proof

We'll prove part 5. Part 4 is in the notes and the rest are exercises.

- Given $\alpha = a + bi$, suppose $(a + bi)(c + di) = 1$.
- Then $ac - bd + (ad + bc)i = 1$.
- So

  $$ac - bd = 1, \tag{$\dagger$}$$

  and

  $$ad = -bc. \tag{$\ddagger$}$$

- If $b = 0$ then $\alpha^{-1} = \frac{1}{a}$, so we assume $b \neq 0$.
- So we can rewrite ($\ddagger$) as $c = \frac{-ad}{b}$.
- Substituting into ($\dagger$) gives $d = \frac{-b}{a^2+b^2}$.
- Substituting this value for d into ($\ddagger$) produces $c = \frac{a}{a^2+b^2}$.
- So, we define $\alpha^{-1} = \frac{a-bi}{a^2+b^2}$.

Fields

We can define division for complex numbers:

Definition 2

Let $\alpha, \beta \in \mathbb{C}$, and suppose $\beta \neq 0$. Then $\frac{\alpha}{\beta} = \alpha\beta^{-1}$.

- A **field** is a mathematical structure generalizing the arithmetic of real numbers.
- Fields have special elements zero and one, have addition and multiplication operations, and also inverses for non-zero elements.
- E.g. $\mathbb{C}$ is a field.
- Fields behave like real numbers, but with important differences - e.g. they can be finite!
- Abstract linear algebra can be done with arbitrary fields, but we will just use $\mathbb{R}$ and $\mathbb{C}$.

Vector spaces over fields

Definition 3

For a field $\mathbb{F}$, a *vector space over* $\mathbb{F}$ is a set V with operations $+ : V \times V \to V$ and $\cdot : \mathbb{F} \times V \to V$ satisfying:

1. $u + v = v + u$ for all $u, v \in V$.
2. $u + (v + w) = (u + v) + w$ for all $u, v, w \in V$.
3. $(ab)v = a(bv)$ for all $a, b \in \mathbb{F}$ and for all $v \in V$.
4. There is a special element $0 \in V$ such that $0 + v = v$ for all $v \in V$.
5. For all $v \in V$ there is $w \in V$ such that $v + w = 0$.
6. $1v = v$ for all $v \in V$ (i.e. scalar multiplication by 1 does not change v).
7. $a(u + v) = au + av$ for all $a \in \mathbb{F}$ and for all $u, v \in V$.
8. $(a + b)v = av + bv$ for all $a, b \in \mathbb{F}$ and for all $v \in V$.

Real and complex vector spaces

- $+$ is known as **vector addition**.
- $\cdot$ is known as **scalar multiplication**.
- When $\mathbb{F} = \mathbb{R}$ we say V is a *real vector space*.
- When $\mathbb{F} = \mathbb{C}$ we say V is a *complex vector space*.
- We refer to elements of V as *vectors*, or *points*.

Examples of vector spaces

- Any field as a vector space over itself. E.g. $\mathbb{R}$ is a real vector space.
- $\mathbb{R} \times \mathbb{R}$, i.e. the Euclidean plane, is a real vector space.
- For any $n \in \mathbb{N} \setminus \{0\}$, $\mathbb{F}^n$ is a vector space over $\mathbb{F}$
  - Define $(x_1, \ldots, x_n) + (y_1, \ldots, y_n) = (x_1 + y_1, \ldots, x_n + y_n)$, and $a(x_1, \ldots x_n) = (ax_1, \ldots, ax_n)$.
- Let $\mathbb{R}[x]$ be the set of all polynomials with the variable x. So

  $$\mathbb{R}[x] = \{a_0 + a_1 x + \ldots + a_n x^n : n \in \mathbb{N} \text{ and } a_i \in \mathbb{R} \text{ for all } i\}.$$

  Then $\mathbb{R}[x]$ is a vector space over $\mathbb{R}$.

Properties of vector spaces

Proposition 4

Let V be a vector space over $\mathbb{F}$. Then:

1. *The additive identity 0 is unique.*
2. *The additive inverse of v is unique for all $v \in V$ (we call it $-v$).*
3. $0v = 0$ *for all $v \in V$.*
4. $-1v = -v$ *for all $v \in V$.*

Properties of vector spaces - proof

Proof.

1. Suppose 0 and $0'$ are both additive identities for V. Then $0 = 0 + 0' = 0'$.
2. Suppose $v + u = 0$ and $v + u' = 0$. Then $(v + u) + u' = u'$, and so $(v + u') + u = u'$, which means $u = u'$.
3. $0v = (0 + 0)v = 0v + 0v$, so $0v + (-0v) = (-0v) + 0v + 0v$, and so $0 = 0v$.
4. Exercise 1.3.

$\square$

Subspaces

Definition 5

Let V be a vector space over $\mathbb{F}$. Then a subset U of V is a *subspace* of V if it has the following properties:

1. $0 \in U$.
2. $u + v \in U$ for all $u, v \in U$ (closure under vector addition).
3. $au \in U$ for all $a \in \mathbb{F}$ and for all $u \in U$ (closure under scalar multiplication).

Another view of subspaces

Lemma 6

If V is a vector space over $\mathbb{F}$ then $U \subseteq V$ is a subspace of V if and only if it is also a vector space over $\mathbb{F}$ with the addition and scalar multiplication inherited from V.

Proof.

- If U is a vector space with the inherited operations then it must be closed under the inherited operations and contain 0.
- Conversely, if U satisfies the conditions of definition 5 then it automatically satisfies all conditions of definition 3 except (5).
- To see that (5) also holds in U note that, by proposition 4(4), given $u \in U$ we have $-u = -1u$, which is in U by definition 5(3).

$\square$

Sums of subspaces

Definition 7

Given subspaces $U_1, \ldots, U_n$ of V, the *sum* $U_1 + \ldots + U_n$ is the smallest subspace of V containing $\bigcup_{i=1}^{n} U_i$.

Lemma 8

If $U_1, \ldots, U_n$ are subspaces of V, then

$$U_1 + \ldots + U_n = \{u_1 + \ldots + u_n : u_i \in U_i \textit{ for all } i \in \{1, \ldots, n\}\}.$$

Proof.

- $\{u_1 + \ldots + u_n : u_i \in U_i \text{ for all } i \in \{1, \ldots, n\}\}$ contains $\bigcup_{i=1}^{n} U_i$ because $u_i = 0 + \ldots + 0 + u_i + 0 + \ldots + 0$ for all $u_i \in U_i$.
- It is a subspace by the definition of a vector space.
- It must be the smallest subspace containing $\bigcup_{i=1}^{n} U_i$, because any such subspace must be closed under vector addition.

$\square$

Direct sums

Definition 9
If $U_1, \ldots, U_n$ are subspaces of V, then the sum $U_1 + \ldots + U_n$ is a *direct sum* if, for all $u \in U_1 + \ldots + U_n$, there is exactly one choice of $\{u_1, \ldots, u_n\}$ such that $u_i \in U_i$ for all i and $u = u_1 + \ldots + u_n$. In this case we write $U_1 \oplus \ldots \oplus U_n$.

- So direct sum is a sum where there is no redundancy.
- Every element in a direct sum is formed in exactly one way using the subspaces that make up the sum.

Direct sums - expressing zero

Lemma 10

If $U_1, \ldots, U_n$ are subspaces of V, then $U_1 + \ldots + U_n$ is a direct sum if and only if there is exactly one choice of $\{u_1, \ldots, u_n\}$ such that $u_i \in U_i$ for all i and $0 = u_1 + \ldots + u_n$.

Proof.

- If $U = U_1 + \ldots + U_n$ is a direct sum, then by definition there is only one way to express 0 (i.e. $0 = 0 + \ldots + 0$).
- Conversely, suppose there is only one way to express 0.
- Let $u \in U$, and suppose $u = u_1 + \ldots + u_n = u_1' + \ldots + u_n'$. Then

  $$0 = u_1 + \ldots + u_n - (u_1' + \ldots + u_n') = (u_1 - u_1') + \ldots + (u_n - u_n').$$

  So $(u_i - u_i') = 0$ for all i, as there is only one way to express 0.
- Thus $u_i = u_i'$ for all i.

$\square$

Direct sums - two subspaces

Lemma 11

Let U and W be subspaces of V. Then $U + W$ is a direct sum if and only if $U \cap W = \{0\}$.

Proof.

- If there is $v \in U \cap W$ then $v = 0 + v$ and $v = v + 0$, so $U + W$ is not a direct sum.
- Conversely, suppose $U \cap W = \{0\}$ and that $v = u + w$ and $v = u' + w'$.
- Then $u - u' = w' - w$, and so $u - u'$ and $w' - w$ are both in $U \cap W$, and thus are both 0.
- This implies $u = u'$ and $w = w'$, so $U + W$ is a direct sum.

$\square$

Direct sums - Example

- Let $V = \mathbb{R}^3$, let $U_1 = \{(2x, 0, z) : x, z \in \mathbb{R}\}$, let $U_2 = \{(0, y, 0) : y \in \mathbb{R}\}$, and let $U_3 = \{(0, z, z) : z \in \mathbb{R}\}$.
- Then $\mathbb{R}^3 = U_1 + U_2 + U_3$, because given $(a, b, c) \in \mathbb{R}^3$ we have

$$(a, b, c) = (2(\frac{a}{2}), 0, 0) + (0, b - c, 0) + (0, c, c).$$

- However, $U_1 + U_2 + U_3$ is not a direct sum as

$$(0, 0, 0) = (0, 0, 1) + (0, 1, 0) + (0, -1, -1).$$

- I.e., 0 is not uniquely expressible.
- However, $U_i \cap U_j = \{0\}$ for all $i \neq j$, which indicates that lemma 11 only applies to binary sums.

Span

Definition 12

Given a vector space V over $\mathbb{F}$, and vectors $v_1, \dots v_n \in V$, we say the *span* of $(v_1, \dots, v_n)$ is the smallest subspace of V containing $\{v_1, \dots, v_n\}$.
By convention we define $\mathrm{span}() = \{0\}$. If $\mathrm{span}(v_1, \dots, v_n) = V$ we say $(v_1, \dots, v_n)$ *spans* V.

Span

Lemma 13

If V is vector space over $\mathbb{F}$, and $v_1, \ldots v_n \in V$, then

$$\mathrm{span}(v_1, \ldots v_n) = \{a_1 v_1 + \ldots + a_n v_n : a_i \in \mathbb{F} \textit{ for all } i\}.$$

Proof.

- Let $U = \{a_1 v_1 + \ldots + a_n v_n : a_i \in \mathbb{F} \text{ for all } i\}$.
- Then clearly $U \subseteq \mathrm{span}(v_1, \ldots v_n)$, as $\mathrm{span}(v_1, \ldots v_n)$ is closed under vector addition and scalar multiplication.
- Moreover, U is closed under vector addition and scalar multiplication, so U is a subspace of V.
- Since $\{v_1, \ldots, v_n\} \subseteq U$, it follows from the definition that $\mathrm{span}(v_1, \ldots v_n) \subseteq U$.
- Thus $U = \mathrm{span}(v_1, \ldots v_n)$ as required.

$\square$

Linear independence

Definition 14

Let V be vector space over $\mathbb{F}$, and let $v_1, \ldots v_n \in V$. Then $(v_1, \ldots, v_n)$ is *linearly independent* if whenever $a_1 v_1 + \ldots + a_n v_n = 0$ we have $a_1 = \ldots = a_n = 0$.
If $(v_1, \ldots, v_n)$ is not linearly independent then we say it is *linearly dependent*.

Examples

1. The vectors $(1, 0, 0)$, $(0, 1, 0)$ and $(0, 0, 1)$ are linearly independent and span $\mathbb{R}^3$ and $\mathbb{C}^3$.
2. The span of a single vector v is $\{av : a \in \mathbb{F}\}$. Single vectors are always linearly independent.
3. The vectors $(2, 3, 1)$, $(1, -1, 2)$ and $(7, 3, c)$ are linearly independent so long as $c \neq 8$.
4. Every list of vectors containing 0 is linearly dependent.