Linear Algebra

Adam Kelly (`ak2316@cam.ac.uk`)

November 8, 2021

This article constitutes my notes for the 'Linear Algebra' course, held in Michaelmas 2021 at Cambridge. These notes are *not a transcription of the lectures*, and differ significantly in quite a few areas. Still, all lectured material should be covered. As a warning, this course has lots of results which involve just checking that a definition holds – such proofs have been intentionally omitted.

Contents

§1 Vector Spaces

§1.1 Vector Spaces and Subspaces

Linear algebra is, somewhat obviously, primarily about studying objects that are *linear* in nature. The objects we really care about are *vector spaces*, settings in which we can add elements and multiply by scalars. We are also going to consider *linear maps*, functions on vector spaces which preserve that linear structure – but more on that later.

Throughout the following discussion (and this course), $\mathbb{F}$ is going to denote an arbitrary field[1]

Definition 1.1 ($\mathbb{F}$-Vector Space)

An $\mathbb{F}$-**vector space** is an abelian group $(V, +)$ together with a function $\mathbb{F} \times V \to V$, written $(\lambda, v) \mapsto \lambda v$ such that the following axioms hold:

(i) *Distributivity in V.* $\lambda(v_1 + v_2) = \lambda v_1 + \lambda v_2$,

(ii) *Distributivity in $\mathbb{F}$.* $(\lambda_1 + \lambda_2)v = \lambda_1 v + \lambda_2 v$,

(iii) *Associativity.* $\lambda(\mu v) = (\lambda \mu)v$,

(iv) *Identity.* $1v = v$.

We usually call elements of V **vectors** and elements of $\mathbb{F}$ **scalars**. The identity element in V is usually called the zero vector, and is written 0_V (or just 0 if the context is clear).

If $\mathbb{F}$ is $\mathbb{R}$ or $\mathbb{C}$, we use the terms 'real vector space' and 'complex vector space', since they're so common.

Example 1.2 (Examples of Vector Spaces)

(i) The set of triples
$$\{(x, y, z) \mid x, y, z \in \mathbb{R}\}$$
forms a real vector space called $\mathbb{R}^3$, because you can add any two triples component wise.

(ii) The set
$$\mathbb{Q}[\sqrt{2}] = \{a + b\sqrt{2} \mid a, b \in \mathbb{Q}\}$$
is a $\mathbb{Q}$-vector space, where we add elements and scale by rational numbers in the obvious way.

(iii) The set $\mathcal{C}[0, 1]$ of all continuous functions $f : [0, 1] \to \mathbb{R}$ forms a real vector space.

As with many new objects, it's helpful to be able to discuss its substructure. In the case of a vector space V, there's a pretty natural notion for what it means for a subset $U \subseteq V$ to still act like a vector space.

---

[1] A field $\mathbb{F}$ is a set $\mathbb{F}$ equipped with two operations $+$ ('addition') and $\cdot$ ('multiplication'). We require $\mathbb{F}$ with addition to form an abelian group, and multiplication must be associative and have an identity element 1. We also require every element except 0 to have an inverse with respect to multiplication, and multiplication must be distributive over addition.

Informally, you can think of a field as something you can do arithmetic in.

Definition 1.3 (Subspace)

Let V be a $\mathbb{F}$-vector space. A subset $U \subseteq V$ is a **subspace** of V if U is also an $\mathbb{F}$-vector space. If U is a subspace of V, we will write $U \leq V$.

Example 1.4 (Examples of Subspaces)

(i) The set of vectors $\{(x, y, z) \mid x, y, z \in \mathbb{R}, x + y + z = 0\}$ is a subspace of $\mathbb{R}^3$.

(ii) The set of polynomials with terms of even degree $\{a_0 + a_2 x^2 + a_4 x^4 + \cdots + a_{2k} x^{2k} \mid \alpha_{2i} \in \mathbb{R}, k \in \mathbb{N}\}$ is a subspace of $\mathbb{R}[X]$, the vector space of polynomials with coefficients in $\mathbb{R}$.

As you would expect, checking that something is a subspace is usually easier than checking all of the axioms for a vector space. In particular, to check that U is a subspace of an $\mathbb{F}$-vector space V, you can just check that the following hold:

- *Zero vector*[2]. $0_V \in U$,

- *Closure under addition.* $u_1, u_2 \in U$ to imply $u_1 + u_2 \in U$,

- *Closure under scaling.* $\lambda \in \mathbb{F}$ and $u \in U$ to imply $\lambda u \in U$.

There are various ways in which we can manipulate subspaces, for example we can take the intersection of two subspaces, and we will get back another subspace.

Proposition 1.5 (Intersecting Subspaces)

Let $U, W \leq V$. Then $U \cap W \leq V$.

Proof. Since U and V are both subspaces of V, we have $0_V \in U \cap V$, and also since they are both closed under addition and scaling, $u_1, u_2 \in U \cap W$ implies that $u_1 + u_2 \in U \cap W$, and $\lambda \in \mathbb{F}$ implies $\lambda u \in U \cap W$. Thus $U \cap W$ is a subspace of V. $\qquad \square$

However we can't manipulate subspaces however we want and expect magic. For example, the union of two subspaces is generally *not* a subspace, as it is typically not closed under addition. In fact, the union is only ever a subspace if one of the subspaces is contained in the other.[3]

We can however try to 'complete' the union so that it becomes a subspace.

Definition 1.6 (Sum of Subspaces)

Let V be a vector space over $\mathbb{F}$, and let $U, W \leq V$. We define the **sum** of U and W to be the set
$$U + W = \{u + w \mid u \in U, w \in W\}.$$

This definition immediately forces $U + W \leq V$, and indeed it is the minimal such space (in that any subspace of V containing both U and W must also contain $U + W$).

---

[2]You may wonder why we need to check this when we already check that we are closed under scaling. To see why, notice that we still have to ensure U is non-empty!

[3]There are some more exercises of this flavour on the example sheet.

§1.2 Quotient Spaces

Since a vector space V forms an abelian group $(V, +)$, we are able to take the quotient by any subspace $U \leq V$.

> **Definition 1.7** (Quotient Space)
>
> Let V an $\mathbb{F}$-vector space, and let $U \leq V$. The **quotient space** V/U is the abelian group V/U equipped with the scalar multiplication $F \times V/U \to V/U$ written $(\lambda, v + U) \mapsto \lambda v + U$.

With this definition, we need to check that this scalar multiplication operation is well defined. Indeed, if $v_1 + U = v_2 + U$ then

$$v_1 - v_2 \in U$$
$$\implies \lambda(v_1 - v_2) \in U$$
$$\implies \lambda v_1 + U = \lambda v_2 + U \in V/U,$$

so our operation is indeed well defined.

As you would expect, taking a quotient gives you back a vector space.

> **Proposition 1.8** (Quotient Spaces are Vector Spaces)
>
> V/U is an $\mathbb{F}$-vector space.

> *Proof Sketch.* Check definitions (most properties are inherited from V being a vector space). $\qquad\square$

§1.3 Basis and Dimension

You are likely informally familiar with the idea of *dimension*, a measure how much freedom exists in a system. Dimensionality is a rather natural concept with respect to vector spaces, but we will need to move through some technicalities to establish the results we want.

To discuss the amount of freedom, we first need a way to quantify what it means for a set of vectors to be independent from one another. This is the idea of *linear independence*.

> **Definition 1.9** (Linear Independence)
>
> We say that $\{v_1, \ldots, v_n\} \in V$ are **linearly independent** if
>
> $$\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n = 0$$
>
> implies that $\lambda_1 = \cdots = \lambda_n = 0$.

Remark. For an infinite subset $S \subseteq V$, we say it's linearly independent if every finite subset is linearly independent.

If a set of vectors is *not* linearly independent, then there's some vector in that set that can be written as a linear combination of the others – so it's not independent of them!

The next idea we need to pin down is being able to see if our set of vectors can 'generate' the rest of our vector space.

Definition 1.10 (Span)

Let V be a vector space over $\mathbb{F}$, and let $S \subset V$. We define the **span** of S, $\langle S \rangle$ to be the set of finite combinations of elements of S.

If $\langle S \rangle = V$, then we say S is **spans** or **generates** V.

Remark. By convention, we also take $\langle \emptyset \rangle = \{0\}$. An equivalent definition is that $\langle S \rangle$ is the smallest subspace of V that contains S.

Example 1.11 (Quadratic Polynomials)

Let V be the vector space of quadratic polynomials over $\mathbb{R}$,

$$V = \{ax^2 + bx + c \mid a, b, c \in \mathbb{R}\}.$$

Then the subset $S \subseteq V$ with $S = \{1, x, x^2\}$ spans V.

Putting these two concepts together gives us the idea of *bases*, which are sets of linearly independent vectors that span a vector space.

Definition 1.12 (Basis)

A subset S of a vector space V is a **basis** if S is a set of linearly independent vectors that span V.

Example 1.13 (Basis for $\mathbb{R}^n$)

The **canonical basis** of $\mathbb{R}^n$ is the set of vectors

$$S = \left\{ \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \ldots, \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} \right\}.$$

Importantly, this is not the *only* basis of $\mathbb{R}^n$, just one that is quite convenient most of the time.

Remark. Note that in the definition of a basis there is no requirement for the set of basis vectors $S \subseteq V$ to be finite – only that any element in V must be representable using finitely many elements of S.

We'd intuitively want to say that the *dimension* of a vector space is the number of elements in its basis. However, we first need to check that this is a well defined notion. We can at this point distinguish between finite and infinite dimensional vector spaces at this point though[4].

---

[4]Can you see why this is well defined already?

> **Definition 1.14** (Finite & Infinite Dimension)
>
> We say a vector space V is **finite dimensional** if it has a finite basis, and we say it is **infinite dimensional** otherwise.

The next result about bases we will prove is that they induce *unique* representations of elements in the vector space.

> **Lemma 1.15** (Unique Representations with a Basis)
>
> Let V be a vector space over $\mathbb{F}$. Then $S \subseteq V$ is a basis of V if and only if any vector $v \in V$ can be written uniquely as a linear combination of elements $v_1, \ldots, v_n \in S$.

> *Proof.* Suppose that S was a basis for V. Then if $v \in V$ can't be written as such a linear combination, then S wouldn't not span V, contradicting it being a basis. Also, if v can be written as such a linear combination non-uniquely, then taking
>
> $$v = \lambda_1 v_1 + \cdots + \lambda_n v_n = \mu_1 v_1 + \cdots + \mu_n v_n.$$
>
> where $\lambda_i \neq \mu_i$ for at least one value of i, we'd have $0 = v - v = (\lambda_1 - \mu_1)v_1 + \cdots + (\lambda_n - \mu_n)v_n$, and at least one of these coefficients must be non-zero, contradicting S being linearly independent.
>
> Alternatively, if any element in V *can* be written uniquely, then if $v_1, \ldots, v_n \in S$ with $\lambda_1 v_1 + \cdots \lambda_n v_n = 0$ implies that $\lambda_1 = \cdots = \lambda_n = 0$, giving that S must be linearly independent. Since S is also spanning by definition, we see that it therefore must be a basis of V. $\qquad\square$

With that out of the way, we can prove some results about finite dimensional vector spaces which will help us get towards our definition of dimension.

> **Lemma 1.16** (Spanning Sets Contain a Basis)
>
> Let V be a finite dimensional vector space, and let $S = \{v_1, \ldots, v_n\}$ be a set of vectors that spans V. Then there is some subset of S that is a basis of V.

> *Proof.* If $\{v_1, \ldots, v_n\}$ is linearly independent, then we are done. If it's not, then (up to reordering) we have $v_n \in \langle \{v_1, \ldots, v_{n-1}\} \rangle$. But then $\langle \{v_1, \ldots, v_n\} \rangle = \langle \{v_1, \ldots, v_{n-1}\} \rangle$, so we can not include v_n in our subset. Not including elements in this way repeatedly, since there is finitely many elements in S, we must eventually get a linearly independent set that still spans V. $\qquad\square$

> **Theorem 1.17** (Steinitz Exchange Lemma)
>
> Let V be a finite dimensional vector space. Then if $\{v_1, \ldots, v_m\}$ is a set of linearly independent vectors, and $\{w_1, \ldots, w_n\}$ spans V, then
>
>   (i) $m \leq n$
>
>   (ii) up to reordering, $\{v_1, \ldots, v_m, w_{m+1}, \ldots, v_n\}$ spans V.

Proof. We will prove this by induction. Suppose we have replaced $\ell \geq 0$ of the w_i, and that

$$\langle \{v_1, \ldots, v_\ell, w_{\ell+1}, \ldots, w_n\} \rangle = V.$$

If $m = \ell$, we are done, so assume that $\ell < m$. Then since this set is spanning, we can write $v_{\ell+1} \in V$ as

$$v_{\ell+1} = \alpha_1 v_1 + \cdots + \alpha_\ell v_\ell + \beta_{\ell+1} w_{\ell+1} + \cdots + \beta_n w_n.$$

Since having $\beta_i = 0$ for all $\ell + 1 \leq i \leq n$ would violate linear independence, we can suppose without loss of generality that $\beta_{\ell+1} \neq 0$. We also note that this implies that $\ell + 1 \leq n$, as otherwise this would not be possible.

Then $w_{\ell+1} \in \langle \{v_1, \ldots, v_{\ell+1}, w_{\ell+2}, \ldots, w_n\} \rangle$, and this set spans V.

Repeating this process, we will be done after m steps, and we have also shown (at the final step) that $m \leq n$. $\qquad\square$

Corollary 1.18 (Dimension)

Let V be a finite dimensional vector space over $\mathbb{F}$. Then any two bases of V have the same number of elements, called the **dimension** of V, denoted $\dim V$ or $\dim_{\mathbb{F}} V$.

Proof. Immediate by Steinitz exchange lemma. $\qquad\square$

So using Steinitz exchange lemma we have finally been able to pin down exactly what is meant by the dimension of a vector space – it's the size of it's basis. This should match up to the intuitive idea of 'freedom' that you had at the start of this section. Freedom in a vector space comes from varying coefficients, and in a basis we can both freely vary coefficients and also reach any element in a vector space uniquely, so the number of independent parameters really is the size of the basis.

Steinitz also gives us a few useful results for free.

Corollary 1.19

Let V be a vector space over $\mathbb{F}$ with finite dimension $n = \dim V$. Then

  (i) Any independent set of vectors has at most n elements, with equality if and only if it's a basis.

  (ii) Any spanning set has at least n elements, with equality if and only if it's a basis.

Proof. Immediate by Steinitz exchange lemma. $\qquad\square$

Working with basis and dimension can make the study of vector spaces much easier. For example, Steinitz allows us nicely extend bases. That is,

<div style="text-align:center; color:blue;">

If we have a vector space V, and a subspace $U \leq V$,
then we can pick a nice basis $\{u_1, \ldots, u_\ell\}$ of U, and
extend it to a basis $\{u_1, \ldots, u_\ell, u_{\ell+1}, \ldots, u_n\}$ of V.

</div>

Working with ideas like this can make many results easier to prove, as we will see in the

following propositions.

> **Proposition 1.20** (Dimension of the Sum of Subspaces)
>
> Let U, W be subspaces of a vector space V. if U and W are finite dimensional, then so is $U + W$, and $\dim U + W = \dim U + \dim W - \dim U \cap W$.

Proof. Pick a basis $\{v_1, \ldots, v_a\}$ of $U \cap W$, and extend by Steinitz exchange lemma to a basis $\{v_1, \ldots, v_a, u_1, \ldots, u_b\}$ of U, and to a basis $\{v_1, \ldots, v_a, w_1, \ldots, w_c\}$ of W.

It suffices to prove that $\{v_1, \ldots, v_a, u_1, \ldots, u_b, w_1, \ldots, w_c\}$ is a basis of $U + W$.

Clearly this set of vectors spans $U + W$, so we just need to check that they are linearly independent. Suppose that

$$\sum_{i=1}^{a} \alpha_i v_i + \sum_{i=1}^{b} \beta_i u_i + \sum_{i=1}^{c} \gamma_i w_i = 0.$$

Rewriting,

$$\sum_{i=1}^{a} \alpha_i v_i + \sum_{i=1}^{b} \beta_i u_i = -\sum_{i=1}^{c} \gamma_i w_i, \tag{$\dagger$}$$

where the LHS is in U and the RHS is in W. This implies that $\sum_{i=1}^{c} \gamma_i w_i \in U \cap W$, and can be written as $\sum_{i=1}^{c} \gamma_i w_i = \sum_{i=1}^{a} \mu_i v_i$, for some μ_i, and then substituting this back into ($\dagger$),

$$\sum_{i=1}^{a} (\alpha_i + \mu_i) v_i + \sum_{i=1}^{b} \beta_i u_i = 0$$

which forces $\beta_i = 0$. A similar argument also gives $\gamma_i = 0$, which then finally forces $\alpha_i = 0$, since $\{v_1, \ldots, v_a\}$ is a basis. $\qquad\square$

> **Proposition 1.21** (Dimension of the Quotient Space)
>
> If V is a finite dimensional vector space over $\mathbb{F}$ and $U \subseteq V$, then U and V/U are also finite dimensional, and $\dim V = \dim U + \dim V/U$.

Proof. Let $\{u_1, \ldots, u_\ell\}$ be a basis of U, and extend it via Steinitz exchange lemma to a basis $\{u_1, \ldots, u_\ell, w_{\ell+1}, \ldots, w_n\}$ of V.

It's easy to see that $\{w_{\ell+1} + U, \ldots, w_n + U\}$ is a basis of V/U, as it clearly spans and linear independence is inherited from it being a basis of V. The result then follows. $\qquad\square$

§1.4 Direct Sums

Previously, we were able to look at the substructure of a vector space by looking at subspaces. Given two subspaces, we were then able to construct their sum, which is the set of all linear combinations of elements in each subspace.

When studying a vector space using its subspaces in this way (considering their sum), it can be useful to impose an *additional* constraint about the way in which the subspaces

interact. In particular, it can be useful to impose a uniqueness constraint on the linear combinations that are created.

> **Definition 1.22** (Direct Sum)
>
> Let V be a vector space over $\mathbb{F}$, and let $U, W \leq V$. We say that V is the **direct sum** of U and W, written $V = U \oplus W$ if and only if every element $v \in V$ can be decomposed as
> $$v = u + w$$
> with $u \in U$ and $w \in W$, with this decomposition being unique.

Of course, we can generalize the notion of a direct sum naturally to the case of multiple subspaces, in the way that you would expect.

Remark (Warning)**.** We say that W is *a* direct complement of U in V. There is *no uniqueness* of such a complement!

> **Lemma 1.23**
>
> Let V be a vector space and let $U, W \leq V$. Then the following are equivalent.
>
> (i) $V = U \oplus W$
>
> (ii) $V = U + W$ and $U \cap W = \{0\}$
>
> (iii) For any basis B_1 of U and B_2 of W, the union $B = B_1 \cup B_2$ is a basis of V.

> *Proof. (ii) implies (i).* Let $V = U + W$ with $U \cap W = \{0\}$. Then for all $v \in V$, we can write $v = u + w$ with $u \in U$ and $w \in W$. To see that this is unique, suppose that
> $$v = u + w = u' + w'.$$
> Then $(u - u') = -(w - w')$, and thus they are both in $U \cap W$, but the only element of this is 0, and thus $u = u'$ and $w = w'$, giving us uniqueness.
>
> *(i) implies (iii).* Let B_1 be a basis of U and B_2 be a basis of W. Then $B = B_1 \cup B_2$ clearly spans V, so we need to check that it's linearly independent. Indeed, if
> $$\sum_{u \in B_1} \lambda_u u + \sum_{w \in B_2} \lambda_w w = 0,$$
> then $V = U \oplus W$ implies that since this is the sum of an element of U and an element of W, then each is zero, and linear independence follows from B_1, B_2 being sets of linearly independent vectors.
>
> *(iii) implies (ii).* Clearly we have $V = U + W$, so we just need to check that $U \cap W = \{0\}$. Let $v \in U \cap W$. Then we can write
> $$v = \sum_{u \in B_1} \lambda_u u = \sum_{w \in B_2} \lambda_w w \implies \sum_{u \in B_1} \lambda_u u - \sum_{w \in B_2} \lambda_w W = 0,$$
> and since $B_1 \cup B_2$ is a basis for V, we must have $\lambda_u, \lambda_w = 0$ for all $u, w \in B_1, B_2$, implying that $v = 0$. $\square$

This result extends as you would expect for the case of direct products using multiple

subspaces.

§2 Linear Maps

§2.1 Linear Maps and Isomorphisms

We are now going to turn our attention to *linear maps*, which are maps that preserve the underlying structure of the vector space that they act on.

Definition 2.1 (Linear Map)

Let V, W are vector spaces over $\mathbb{F}$. A map $\alpha : V \to W$ is a **linear map** if for all $\lambda_1, \lambda_2 \in \mathbb{F}$ and $v_1, v_2 \in V$ we have

$$\alpha(\lambda_1 v_1 + \lambda_2 v_2) = \lambda_1 \alpha(v_1) + \lambda_2 \alpha(v_2).$$

Example 2.2 (Examples of Linear Maps)

(i) Consider the vector space $\mathcal{C}[0,1]$ of continuous functions $f : [0,1] \to \mathbb{R}$, Then for any fixed point $x \in [0,1]$, the map $\alpha : C[0,1] \to \mathbb{R}$ with $\alpha(f) = f(x)$ is a linear map.

(ii) The map $\mathbb{R}^3 \to \mathbb{R}$ given by $(a, b, c) \mapsto 4a + 2b + c$ is a linear map.

(iii) The map $\alpha : Q[\sqrt{2}] \to Q[\sqrt{2}]$ of 'multiplying by $\sqrt{2}$' is a linear map, with $a + b\sqrt{2} \mapsto 2b + a\sqrt{2}$.

It's easy to see that the identity map is a linear map, and also that linearity is preserved by composing linear maps.

One nice thing about linear maps is that they interact well with basis, because of linearity. Indeed, specifying a linear map on the basis is sufficient to define that map on all of a vector space.

Lemma 2.3 (Extension by Linearity)

Let V and W be vector spaces, and let B be a basis for V. Then if $\alpha_0 : B \to W$ is *any* map, then there exists a *unique* linear map $\alpha : V \to W$ extending α_0, so that for all $v \in B$ we have $\alpha(v) = \alpha_0(v)$.

Proof. For $v \in V$, let $v = \sum_{i=1}^{n} \lambda_i v_i$. Then necessarily by linearity, we define

$$\alpha(v) = \alpha \left(\sum_{i=1}^{n} \lambda_i v_i \right) = \sum_{i=1}^{n} \lambda_i \alpha_0(v_i).$$

By construction this is a linear map and it extends α_0. $\qquad\square$

As with many other areas of pure mathematics, we have a special case for the structure preserving bijections on a given object.

> **Definition 2.4** (Isomorphism)
>
> If V and W are vector spaces over $\mathbb{F}$, the linear map $\alpha : V \to W$ is an **isomorphism** if it also bijective. We say two vector spaces V and W are **isomorphic**, written $V \cong W$, if there is an isomorphism between them.

Remark. if $\alpha : V \to W$ is an isomorphism, then the inverse map α^{-1} exists and importantly it's *also* linear.

> **Lemma 2.5** (Isomorphism is an Equivalence Relation)
>
> $\cong$ is an equivalence relation on the class of all vector spaces over $\mathbb{F}$.

> *Proof.* For reflexivity, we note that the identity map $id : V \to V$ is an isomorphism. For symmetry, we note that if $\alpha : V \to W$ is an isomorphism, then $\alpha^{-1} : W \to V$ is also an isomorphism. Lastly, we get transitivity because if $\alpha : U \to V$ and $\beta : V \to W$ are isomorphisms, then $\beta \circ \alpha : U \to W$ is also an isomorphism. $\qquad\square$

From one perspective, isomorphism allows us to know when two objects are really just different representations of the same thing. With this in mind, the following result tells us something interesting (yet unsurprising) about the nature of finite dimensional vector spaces.

> **Theorem 2.6** (Isomorphism for Finite Dimensional Vector Spaces)
>
> If V is an n dimensional vector space over $\mathbb{F}$, then $V \cong \mathbb{F}^n$, the vector space
>
> $$\mathbb{F}^n = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \mid x_1, \ldots, x_n \in \mathbb{F} \right\},$$
>
> where addition is component-wise.

> *Proof.* Let $B = \{v_1, \ldots, v_n\}$ be a basis for V. Then we claim that the map $\alpha : V \to \mathbb{F}^n$ defined by
>
> $$v = \sum_{i=1}^{n} \lambda_i v_i \longmapsto \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}$$
>
> is an isomorphism. Indeed, it's easy to see that we have linearity inherited from $\mathbb{F}$, and also the map is a bijection, so indeed we do have $V \cong \mathbb{F}^n$. $\qquad\square$

In particular, this result tells us that when we choose a basis for a finite dimensional vector space V, it's really the same as picking some isomorphism from V to $\mathbb{F}^n$.

This theorem also directly implies (with isomorphism being an equivalence relation) a criterion for two finite dimensional vector spaces to be isomorphic.

> **Theorem 2.7** (Isomorphism Criterion)

Let V and W be finite dimensional vector spaces over $\mathbb{F}$. Then they are isomorphic if and only if they have the same dimension.

§2.2 Images and Kernels

Stepping back from isomorphisms, we need to define some frequently used concepts relating to linear maps.

> **Definition 2.8**
>
> Let V, W be vector spaces over F, and let $\alpha : V \to W$ be a linear map. We define the **kernel** of α as
> $$\ker \alpha = \{v \in V \mid \alpha(v) = 0\},$$
> and the **image** of α as
> $$\operatorname{img} \alpha = \{w \in W \mid w = \alpha(v), v \in V\}.$$

The structure of a linear map carries over to the image and the kernel, and naturally implies that they are both subspaces of their parent vector spaces.

> **Lemma 2.9** (Image and Kernel are Subspaces)
>
> Let $\alpha : V \to W$ be a linear map. Then $\ker \alpha$ and $\operatorname{img} \alpha$ are subspaces of V and W respectively.

> *Proof Sketch.* Check definitions. $\square$

You may recognize the definitions of the kernel and image from a course in group theory, particularly in relation to the first isomorphism theorem[5]. Indeed, remembering that vector spaces with vector addition are groups, with we can use the kernel and image of a linear map to establish an isomorphism between vector spaces.

> **Theorem 2.10** (First Isomorphism Theorem For Vector Spaces)
>
> Let V and W be $\mathbb{F}$ vector spaces, and let $\alpha : V \to W$ be a linear map. Then $V/\ker \alpha \cong \operatorname{img} \alpha$.

> *Proof Sketch.* Consider the isomorphism $\overline{\alpha} : V/\ker \alpha \to \operatorname{img} \alpha$ defined by $\overline{\alpha}(v + \ker \alpha) = \alpha(v)$, and check definitions. $\square$

Now while the kernel and image being subspaces are rather straightforward, slightly more interesting is the *dimension* of these subspaces. This is described by the *rank-nullity theorem*.

> **Definition 2.11** (Rank and Nullity)
>
> Let α be a linear map. Then the **rank** of α, $r(\alpha) = \dim \operatorname{img} \alpha$, is the dimension of the image. The **nullity** of α, $n(\alpha) = \dim \ker \alpha$, is the dimension of the Kernel.

---

[5]Recall that if G and H are groups, and $\phi : G \to H$ is a homomorphism, then $G/\ker \phi \cong \operatorname{img} \phi$.

Theorem 2.12 (Rank-Nullity Theorem)

Let U and V be vector spaces such that U is finite dimensional. Let $\alpha : U \to V$ be a linear map. Then
$$\dim U = r(\alpha) + n(\alpha).$$

Proof. We know that $U/\ker \alpha \cong \operatorname{img} \alpha$. Thus $\dim(U/\ker \alpha) = \dim(\operatorname{img} \alpha)$, that is $\dim U - \dim \ker \alpha = \dim \operatorname{img} \alpha$, or $\dim U = r(\alpha) + n(\alpha)$. $\qquad\square$

This result has an important consequence which is incredibly useful for quickly checking if a map is an isomorphism.

Lemma 2.13 (Characterisation of Isomorphism)

Let V and W be vector spaces over $\mathbb{F}$ with *equal finite dimension*. Then the following are equivalent:

  (i) α is injective,

  (ii) α is surjective,

  (iii) α is an isomorphism.

Proof Sketch. Follows directly from the rank-nullity theorem. $\qquad\square$

§2.3 Matrices

When dealing with vector spaces, it's sometimes convenient to work with vector spaces in the abstract, and other times it's convenient to pick some nice basis which you can easily work with. When dealing with linear maps $\alpha : V \to W$, it it can also be useful to work with bases. The way we do this is by introducing *matrices*.

It's important to say that matrices are not the essence of linear algebra, and in may cases viewing the subject in this way will just give you the impression that linear algebra is an opaque subject. Instead, you should think of matrices as a nice representations of linear maps for which its easy to perform computations with. With that said, let's talk about linear maps.

In this section, we are going to discuss linear maps between two vector spaces V and W.

Definition 2.14 (Space of Linear Maps)

Let V and W be vector spaces over $\mathbb{F}$. We define

$$\mathcal{L}(v, w) = \{\alpha : V \to W \mid \alpha \text{ is a linear map}\}.$$

Proposition 2.15 (Space of Linear Maps is a Vector Space)

$\mathcal{L}(V, W)$ is a vector space over $\mathbb{F}$. Moreover, if V and W are finite dimensional, then

$$\dim \mathcal{L}(V, W) = \dim V \cdot \dim W.$$

> *Proof Sketch.* To see that $\mathcal{L}(V, W)$ is a vector space, one just needs to check definitions. We will leave the proof of the dimension until later. □

Now, given some linear map α, let's think about how we can represent this map with respect to a basis. Suppose that we had a finite dimensional vector space V with basis $\{v_1, \ldots, v_m\}$, and a finite dimensional vector space W with basis $\{w_1, \ldots, w_n\}$, and suppose that we had some linear map $\alpha : V \to W$.

Given some $v \in V$, if we wanted to work out $\alpha(v)$, we can write express v in the basis above, with

$$\alpha(v) = \alpha\left(\sum_{i=1}^{m} \lambda_i v_i\right) = \sum_{i=1}^{m} \lambda_i \alpha(v_i).$$

So, to compute what $\alpha(v)$ is for any $v \in V$, it suffices to know what $\alpha(v_1), \ldots, \alpha(v_m)$ is. To specify what say $\alpha(v_1)$ is, we can use the above basis of W and write

$$\alpha(v_1) = a_{11}w_1 + a_{21}w_2 + \cdots + a_{n1}w_n,$$

for some coefficients a_{i1}. Repeating this for $v_2, \ldots, v_n$ then totally specifies the linear map α. These collection of coefficients can be put into a *matrix*, which will then be a representation of α when coupled with the bases $\{v_i\}$ and $\{w_i\}$.

> **Definition 2.16** (Matrix)
>
> An $n \times m$ **matrix** over $\mathbb{F}$ is an array with n rows and m columns, with entries in $\mathbb{F}$:
>
> $$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{pmatrix}.$$
>
> We denote the set of $n \times m$ matrices over $\mathbb{F}$ by $\mathcal{M}_{n,m}(\mathbb{F})$.

> **Proposition 2.17** (Vector Space of Matrices)
>
> $\mathcal{M}_{n,m}(\mathbb{F})$ is an $\mathbb{F}$-vector space.

> *Proof Sketch.* Taking addition entry-wise we can just check definitions. □

> **Proposition 2.18** (Dimension of $\mathcal{M}_{n,m}(\mathbb{F})$)
>
> $\dim \mathcal{M}_{m,n}(\mathbb{F}) = m \times n$.

> *Proof.* We define the *elementary matrix* $E_{i,j}$ such that every entry is 0 except for the i, jth entry, which is 1. Then the set $\{E_{i,j} \mid 1 \le i \le m, 1 \le j \le n\}$ is clearly a basis, and has $n \times m$ elements. □

Following on from the discussion above, we can now write down how linear maps can be represented using matrices.

Definition 2.19 (Matrix of α in the B, C Basis)

Let $\alpha : V \to W$ be a linear map, and let $B = \{v_1, \ldots, v_m\}$ be a basis of V and $C = \{w_1, \ldots, w_n\}$ be a basis of W. Then we define the **matrix of** α to be

$$[\alpha]_{B,C} = \begin{pmatrix} | & | & & | \\ [\alpha(v_1)]_C & [\alpha(v_2)]_C & \cdots & [\alpha(v_m)]_C \\ | & | & & | \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{pmatrix},$$

where a_{ij} is the coefficient of w_i in $\alpha(v_j)$.

Remark (Notation)**.** We use $[v]_C$ to denote the column vector representing the vector v in the basis C.

Proposition 2.20 (Matrix Representation of Linear Maps)

If V and W are vector spaces over $\mathbb{F}$ such that $\dim V = n$ and $\dim W = m$, then $\mathcal{L}(V, W) \cong \mathcal{M}_{m,n}(\mathbb{F})$.

Proof. Fix a basis B, C of V and W respectively. Then consider the linear map $\theta : \mathcal{L}(V, W) \to \mathcal{M}_{m,n}(\mathbb{F})$ given by $\alpha \mapsto [\alpha]_{B,C}$.

This map is clearly linear, and it's also clearly a bijection thus it's an isomorphism and our result follows. $\square$

Remark. This implies that $\dim \mathcal{L}(L, W) = m \times n$.

What's nice about matrices is that they make working with linear maps *easy*. It's generally much easier to manipulate arbitrary matrices than it is to arbitrary linear maps (of course, this varies from case to case), and we will see this in the following set of computational results.

You should be familiar from previous courses with the notion of matrix multiplication.

Definition 2.21 (Matrix Multiplication)

Let $A = (a_{ij})$ be an $n \times p$ matrix, and $B = (b_{ij})$ be an $p \times m$ matrix. Then their **product** $C = A \cdot B$ is an $n \times m$ matrix $C = (c_{ij})$, with

$$c_{ij} = \sum_{k=1}^{p} a_{ik} b_{kj}.$$

We can use matrix multiplication in lots of nice ways.

Lemma 2.22 (Evaluating Linear Maps)

Let $\alpha : V \to W$ be a linear map, and let B, C be bases for V and W respectively. Then if $v \in V$, we have
$$[\alpha(v)]_C = [\alpha_{B,C}] \cdot [v]_B.$$

Proof. Let $v \in V$, and let $v = \sum_{j=1}^{n} \lambda_j v_j$. Then

$$\alpha(v) = \sum_{j=1}^{n} \lambda_j \alpha(v_j) = \sum_{j=1}^{n} \lambda_j \sum_{i=1}^{m} a_{ij} w_i = \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij} \lambda_j \right) w_i,$$

as required. $\qquad\square$

Lemma 2.23 (Composing Linear Maps)

Let $\alpha : V \to W$ and $\beta : U \to V$ be linear maps. Then if A, B, C are bases of U, V and W respectively, we have

$$[\alpha \circ \beta]_{A,C} = [\alpha]_{B,C} \cdot [\beta]_{A,B}.$$

Proof Sketch. Use the definition of matrix multiplication to check that this map corresponds to the correct thing. $\qquad\square$

§2.4 Change of Basis and Equivalent Matrices

Matrices represent a linear map with respect to a given basis, so a natural question is how do we rewrite a matrix in a different basis? That is, given two vector spaces V, W and a linear map $\alpha : V \to W$, we want to take the matrix $[\alpha]_{B,C}$ and express it with respect to the basis B', C'.

The way we do this is by introducing a *change of basis* matrix.

Definition 2.24 (Change of Basis Matrix)

The **change of basis** matrix from B' to B is the matrix $P = [I]_{B',B}$, where I is the identity map.

Remark. P is an invertible square matrix, and P^{-1} is the change of basis matrix from B to B'.

Given a vector represented in a basis B, the change of basis matrix allows that vector to be expressed in a different basis B'.

Lemma 2.25 (Change of Basis for a Vector)

Let B and B' be bases for a vector space V, and let $v \in V$. Then if P is the change of basis matrix from B' to B we have

$$[v]_B = P[v]_{B'}.$$

Proof. We know that $P = [I]_{B',B}$, and so $[I(v)]_B = [I]_{B',B}[v]'_B = P[v]_{B'}$. $\qquad\square$

Using this idea of change of basis matrices, we can express the matrix of a linear map with respect to another basis.

> **Proposition 2.26** (Change of Basis Formula)
>
> Let $A = [\alpha]_{B,C}$ and $A' = [\alpha]_{B',C'}$, and let P be the change of basis from B' to B and Q be the change of basis matrix from C' to C. Then
>
> $$A' = Q^{-1}AP.$$

> *Proof Sketch.* Check definitions. □

§2.5 Equivalent Matrices

A natural notion of equivalence comes from saying two matrices are equivalent if there is some choice of basis for which they represent the same linear map. Using the change of basis formula, we can pin down this notion exactly.

> **Definition 2.27** (Equivalent Matrices)
>
> We say that two matrices $A, A' \in \mathcal{M}_{m,n}(\mathcal{F})$ are **equivalent** if $A' = Q^{-1}AP$ for some invertible matrices $Q \in \mathcal{M}_{m,m}(\mathbb{F})$ and $P \in \mathcal{M}_{n,n}$.

It's easy to check that this definition gives us an equivalence relation, and indeed it turns out that the equivalence classes of matrices are remarkably straightforward.

> **Proposition 2.28** (Natural Basis for Matrices)
>
> Let V, W be vector spaces over $\mathbb{F}$, with $\dim V = n$ and $\dim W = m$. Let $\alpha : V \to W$ be a linear map, Then there exists a basis B of V and a basis C of W such that
>
> $$[\alpha]_{B,C} = \left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array} \right),$$
>
> where $r = r(\alpha)$ and I_r is the $r \times r$ identity matrix.

> *Proof.* We first choose a basis $\{v_{r+1}, \dots, v_n\}$ of $\ker \alpha$, and extend it to a basis $B = \{v_1, \dots, v_n\}$ of V. We will first show that $\{\alpha(v_1), \dots, \alpha(v_r)\}$ is a basis of $\operatorname{img} \alpha$.
>
> To see that it's spanning, given $v \in V$ we can write $\alpha(v) = \alpha \left(\sum_{i=1}^{n} \lambda_i v_i \right) = \sum_{i=1}^{n} \lambda_i \alpha(v_i) = \sum_{i=1}^{r} \lambda_i \alpha(v_i)$, since $\alpha(v_i) = 0$ for $i > r$.
>
> To see that we have linear independence, suppose that $\sum_{i=1}^{r} \lambda_i \alpha(v_i) = 0$. Then we have $\alpha \left(\sum_{i=1}^{r} \lambda_i v_i \right) = 0$ implying that $\sum_{i=1}^{r} \lambda_i v_i \in \ker \alpha$, and since the kernel is spanned by $\{v_{r+1}, \dots, v_n\}$ and $\{v_1, \dots, v_n\}$ is a basis, we must have $\lambda_i = 0$, as required.
>
> We can now extend this basis to a basis $C = \{\alpha(v_1), \dots, \alpha(v_r), w_{r+1}, \dots, w_m\}$ of W. Then we can express α as a matrix in the basis B and C as
>
> $$[\alpha]_{B,C} = \begin{pmatrix} | & & | & | & & | \\ [\alpha(v_1)]_C & \cdots & [\alpha(v_r)]_C & [\alpha(v_{r+1})]_C & \cdots & [\alpha(v_n)]_C \\ | & & | & | & & | \end{pmatrix}$$

$$= \begin{pmatrix} | & & | & | & & | \\ [\alpha(v_1)]_C & \cdots & [\alpha(v_r)]_C & [0]_C & \cdots & [0]_C \\ | & & | & | & & | \end{pmatrix},$$

and it can easily be seen that this matrix is of the form desired.   □

This result gives us a clear criterion for the equivalence of matrices, and also gives us a lovely result which is useful for choosing a basis which works nicely with the matrix representation of a linear map.

Corollary 2.29 (Matrix Equivalence Criterion)

Two $m \times n$ matrices are equivalent if the linear maps they represent have the same rank.

The equivalence criterion above is great, but doesn't give us a very easy way to see what the rank of the underlying linear map is, without doing a large amount of work. What we really want is a way to get this rank directly from the matrix. We can do this by considering the *column* and *row rank* of a matrix (which as we will see are really the same notion).

Definition 2.30 (Column and Row Rank)

Let $A \in \mathcal{M}_{m,n}(\mathbb{F})$. The **column rank** of A, $r(A)$, is the dimension of the subspace of $\mathbb{F}^m$ spanned by the column vectors of A.

Similarity, the **row rank** is the dimension of the subspace spanned by the row vectors of A (that is, it's the column rank of A^T).

Remark. If α is a linear map, represented by a matrix A with respect to some basis, then $r(A) = r(\alpha)$, so column rank is the same as the rank.

Proposition 2.31 (Equivalence from Row Rank)

Two matrices A, A' are equivalent if and only if $r(A) = r(A')$.

Proof. If they are equivalent, then they correspond to the same linear map α, expressed in two different bases. So the column rank is $r(A) = r(A') = r(\alpha)$, and we are done.

If $r(A) = r(A') = r$, then A and A' are both equivalent to

$$\left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array} \right),$$

and by transitivity of equivalent, we have A and A' equivalent.   □

The distinction between column rank and row rank is really non-existent, as the following result shows. This also means that we can refer to the column and row rank as just the 'rank' of the matrix.

Theorem 2.32 (Column Rank Equals Row Rank)

$$r(A) = r(A^T)$$

Proof. Let $r = r(A)$. Then we know that we can write

$$Q^{-1}AP = \left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array} \right),$$

then taking the transpose, we have $(Q^{-1}AP)^T = P^T A^T (Q^{-1})^T = P^T A^T (Q^T)^{-1}$ so $P^T A^T (Q^T)^{-1}$ is

$$\left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array} \right)^T,$$

which is an $n \times m$ matrix. But then $r = r(A) = r(A^T)$. $\qquad\qquad\square$

Let's now consider the special case of **endomorhisms**, that is, linear maps $\alpha : V \to V$.

If B and B' are a basis for V, and P is the change of basis matrix from B' to B, then given α we have $[\alpha]_{B',B'} = P^{-1}[\alpha]_{B,B}P$, which is a special case of the change of basis formula. This idea of writing a matrix using the same basis for the domain and range gives us *similarity*, another notion of matrices being the same.

> **Definition 2.33** (Similar Matrices)
>
> Let A, A' be two $n \times n$ square matrices. We say that A and A' are **similar** or **conjugate** if and only if $A' = P^{-1}AP$ for some invertible $n \times n$ matrix P.

This idea of similarity will be a central concept when we deal with the diagonalization of such matrices later on.

§3 Dual Space

§3.1 Defining the Dual Space

We are now going to turn our attention to an interesting set of linear maps on a vector space.

> **Definition 3.1** (Dual Space)
>
> Let V be a vector space over a field $\mathbb{F}$. The **dual space** V^* is the set of linear maps from V to $\mathbb{F}$,
>
> $$V^* = \mathcal{L}(V, \mathbb{F}).$$

So how should we understand this space? As usual, things are easier for finite dimensional vector spaces, so we are going to start there. We know from the previous section that *any* linear map is determined by its action on basis vectors. So if we have some finite dimensional vector space V with basis $\{v_1, \ldots, v_n\}$ and a linear map $\alpha : V \to \mathbb{F}$, we just need to know $\alpha(v_1), \ldots, \alpha(v_n) \in \mathbb{F}$.

This tells us exactly what the nice way to think about the dual space is – through a basis that just picks out basis vectors so that we can deal with them individually. We

call this the *dual basis*.[6]

> **Definition 3.2** (Dual Basis)
>
> Let V be a finite dimensional vector space over $\mathbb{F}$ with a basis $B = \{v_1, \ldots, v_n\}$. Then we define the **dual basis** for V^* to be $B^* = \{\varepsilon_1, \ldots, \varepsilon_n\}$ where $\varepsilon_j(v_i) = \delta_{ij}$.
>
> In particular, if $v \in V$ can be written as $v = \sum_{i=1}^n \lambda_i v_i$, then $\varepsilon_j(v) = \lambda_j$.

Since this is a basis of V^*, we can immediately see that $\dim V^* = \dim V$. Sometimes it's also convenient to think of V^* as the space of row vectors of length m over $\mathbb{F}$[7]

§3.2 The Annihilator

When we have some subset U of a vector space V, we frequently care about the set of linear maps that are zero on everything in U. This construction ends up gives us a subspace of the dual space V^*.

> **Definition 3.3** (Annihilator)
>
> If V is a vector space over $\mathbb{F}$, and $U \subseteq V$, then the **annihilator** of U is the set U^0 of linear maps $V \to \mathbb{F}$ that are zero on U,
>
> $$U^0 = \{\alpha \in V^* \mid \alpha(u) = 0 \text{ for all } u \in U\}.$$

> **Lemma 3.4** (Annihilator is a Subspace)
>
> The annihilator is always a subspace of the dual, that is, $U^0 \leq V^*$.

> *Proof Sketch.* Check definitions. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

By thinking about the annihilator, we also get a result that looks a little like rank-nullity.

> **Lemma 3.5** (Rank-Nullity for the Annihilator)
>
> If V is a finite dimensional vector space and $U \leq V$, then $\dim V = \dim U + \dim U^0$.

> *Proof Sketch.* Let $\{v_1, \ldots, v_k\}$ be a basis of U, and extend it to a basis $B = \{v_1, \ldots, v_n\}$ of V. Then the dual basis is $B' = \{\varepsilon_1, \ldots, \varepsilon_n\}$. We then just need to check that $\{\varepsilon_{k+1}, \ldots, \varepsilon_n\}$ is a basis of U^0. $\qquad\qquad$ $\square$

§3.3 The Dual Map

We will now see that thinking about the dual will give us a basis-free interpretation of what corresponds with the transpose of a matrix.

---

[6]We just write it as a definition here, but we should really check that it is indeed a basis of the dual, but I think it's relatively clear so this is left to the reader.

[7]There's also a hidden inner product structure here, but we won't dwell on that.

Definition 3.6 (Dual Map)

Let V and W be vector spaces, and suppose that $\alpha : V \to W$ is a linear map. Then we define the **dual map** of α to be linear map $\alpha^* : W^* \to V*$ given by $\varepsilon \mapsto \varepsilon \circ \alpha$.

To see what this is all about, consider the following example.

Example 3.7 (Using the Dual Map)

Let V and W be $\mathbb{R}$ vector spaces, and let $B = \{v_1, v_2, v_3\}$ be a basis for V and $C = \{w_1, w_2\}$ be a basis for W.

We define the linear map $\alpha : V \to W$ by

$$\alpha(v_1) = w_1 + w_2$$
$$\alpha(v_2) = 2w_1 - w_2$$
$$\alpha(v_3) = 6w_1 + 3w_2.$$

Now we can see what the dual map $\alpha^* : W^* \to V^*$ looks like. Let $B^* = \{v_1^*, v_2^*, v_3^*\}$ be the corresponding dual basis of V, and let $C^* = \{w_1^*, w_2^*\}$ be the corresponding dual basis of W. Then for any $v = av_1 + bv_2 + cv_3$ we have

$$[\alpha^*(w_1^*)](v) = w_1^*(\alpha(av_1 + bv_2 + cv_3))$$
$$= w_1^*((a + 2b + 6c)w_1 + (a - b + 3c)w_2)$$
$$= a + 2b + 6c,$$

and in this way, we can really write

$$\alpha^*(w_1^*) = v_1^* + 2v_2^* + 6v_3^*, \quad \text{and} \quad \alpha^*(w_2^*) = v_1^* - v_2^* + 3v_3^*,$$

which totally specifies α^*.

A slightly more interesting way of viewing this example is through the lens of matrices.

Example 3.8 (Above with Matrices)

Continuing on from the example above, we can see that in the B, C basis we have

$$[\alpha]_{B,C} = \begin{pmatrix} 1 & 2 & 6 \\ 1 & -1 & 3 \end{pmatrix}.$$

Then α^* in the C^*, B^* basis is given by

$$[\alpha^*]_{C^*,B^*} = \begin{pmatrix} 1 & 1 \\ 2 & -1 \\ 6 & 3 \end{pmatrix}.$$

Of course, in finite dimensions this works in the general case.

Theorem 3.9 (Transpose Interpretation of the Dual Map)

Let V and W be finite dimensional vector spaces over $\mathbb{F}$. Let B be a basis for V

and C be a basis for W, and let B^*, C^* be the corresponding dual spaces. Then if $\alpha : V \to W$ is a linear map, we have

$$[\alpha^*]_{C^*,B^*} = [\alpha]_{B,C}^T.$$

Proof. Let $B = \{v_1, \ldots, v_n\}$ be a basis for V and $C = \{w_1, \ldots, w_m\}$ be a basis for W, and let B^*, C^* be the corresponding dual spaces. Then

$$([\alpha]_{B,C})_{i,j} = ([\alpha(v_i)]_C)_j = ([w_j^* \circ \alpha]_{B^*})_i = ([\alpha^*]_{C^*,B^*})_{j,i},$$

as required. □

This result about viewing the matrix of α^* in terms of the matrix of α hints at a more general perspective where it can be easier to understand one of α and it's dual α^* than the other, and indeed many properties of α^* can be derived from α.

For example, the result below is *immensely important*[8] and is an example of the 'duality approach'.

Proposition 3.10 (Properties of the Dual Map)

Let V and W be vector spaces over $\mathbb{F}$, and let $\alpha : V \to W$ be a linear map. Then

  (i) $\ker \alpha^* = (\operatorname{img} \alpha)^0$, so α^* is injective if and only if α is surjective.

  (ii) $\operatorname{img} \alpha^* \leq (\ker \alpha)^0$, with equality if V and W are finite dimensional (in this case, α^* is surjective if and only if α is injective).

Proof. (i) Let $\varepsilon \in W^*$. Then $\varepsilon \in \ker \alpha^*$ if and only if $\alpha^*(\varepsilon) = 0$, that is, if $\varepsilon(\alpha(v)) = 0$ for all $v \in V$. But this occurs if and only if $\varepsilon \in (\operatorname{img} \alpha)^0$.

(ii) We will first show that $\operatorname{img} \alpha^* \leq (\ker \alpha)^0$. Indeed, let $\varepsilon \in \operatorname{img} \alpha^*$. Then $\varepsilon = \alpha^*(\phi)$ for some $\phi \in W^*$. But then for all $u \in \ker \alpha$ we have $\varepsilon(u) = (\alpha^*(\phi))(u) = \phi(\alpha(u)) = 0$. So then $\varepsilon \in (\ker \alpha)^0$, and thus $\operatorname{img} \alpha^* \leq (\ker \alpha)^0$ as required.

In finite dimensions, we can compute the dimensions of the two spaces. We have

$$\dim \operatorname{img} \alpha^* = r(\alpha^*) = r([\alpha^*]_{C^*,B^*}) = r([\alpha]_{B,C}^T) = r(\alpha).$$

So $\dim \operatorname{img} \alpha^* = r(\alpha)$, and by rank-nullity, this is $\dim V - \dim \ker \alpha = \dim (\ker \alpha)^0$. So $\operatorname{img} \alpha^* \leq (\ker \alpha)^0$ and also $\dim \operatorname{img} \alpha^* = \dim(\ker \alpha)^0$, giving us equality. □

§3.4 Double Dual

In general, there is no *obvious* relation between V and V^*, in that all of the relations we discussed in the previous section rely on choosing a basis – and choosing bases is an inherently arbitrary process. However, there *is* a canonical (that is, basis free) relation between V and $(V^*)^*$.
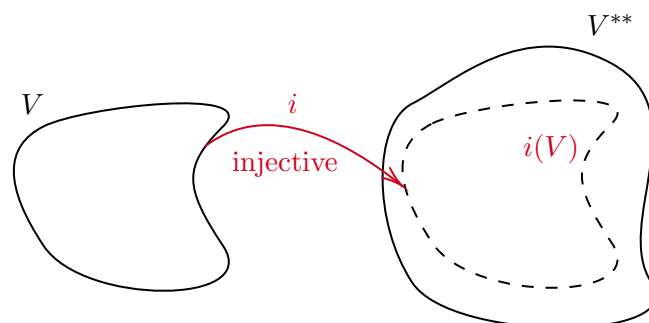
Definition 3.11 (Double Dual)

---

[8]But we won't see in this course why that is.

Let V be a vector space over $\mathbb{F}$, and $V^* = \mathcal{L}(V, \mathbb{F})$ be the dual of V. Then we define the **double dual** or **bidual** of V to be

$$V^{**} = \mathcal{L}(V^*, \mathbb{F}) = (V^*)^*.$$

Indeed, there is a *canonical* embedding from V to V^{**}, and there are non-trivial examples[9] of infinite dimensional spaces where $V \cong V^{**}$.

This canonical embedding is described by 'evaluating at v'.

Theorem 3.12 (Canonical Embedding of V into V^{**})

There is a natural injective linear map from V to V^{**}, called the **evaluation map**.

Proof. Let V be an $\mathbb{F}$ vector space, and let $v \in V$. We define the evaluation at v map

$$\hat{v} : V^* \to \mathbb{F} \quad \text{by} \quad \hat{v}(\varepsilon) = \varepsilon(v).$$

Then $\hat{v}(\varepsilon_1 + \lambda \varepsilon_2) = (\varepsilon_1 + \lambda \varepsilon_2)(v) = \varepsilon_1(v) + \lambda \varepsilon_2(v)$, so $\hat{v} \in V^{**}$.

The proof that this is injective is intentionally omitted. $\qquad\qquad\qquad\square$

Theorem 3.13 (Canonical Embedding in Finite Dimensions)

If V is finite dimensional, then the evaluation map is an isomorphism.

Proof. We know that for $v \in V$ we have the evaluation map $\hat{v} \in V^{**}$. It's easy to see that the map $v \mapsto \hat{v}$ is linear in v, so it is left to check that this map is a bijection.

Let $e \in V \backslash 0$. We can extend $\{e\}$ to a basis $B = \{e, e_2, \ldots, e_n\}$ of V. Let B^* be the dual basis of V. Then

$$\hat{e}(e^*) = e^*(e) = 1,$$

and in particular, $\hat{e} \neq 0$, so the kernel of the map $v \mapsto \hat{v}$ contains only 0, and thus it is injective[a].

To see that this map is an isomorphism, we just need to compute dimensions. We know that $\dim V = \dim V^* = \dim(V^*)^* = \dim V^{**}$, and this along with injectivity implies that the map $v \mapsto \hat{v}$ is an isomorphism. $\qquad\qquad\qquad\square$

---

[a]This step of the proof has an analogue in infinite dimensions for a certain kind of vector space

---

[9]Such as $L^P(\mathbb{R}^d) = \{f : \mathbb{R}^d \to \mathbb{R} \mid \int_{R^d} |f(x)|^p \, dx < \infty\}$.

> (Banach space), but requires the use of non-trivial results (the Hahn-Banach theorem).

This canonical embedding allows us to 'identify' V and V^{**}, which in turn can be used to prove various facts about say the annihilator.

Lemma 3.14 (Double Annihilator)

Let V be a finite dimensional vector space over $\mathbb{F}$, and $U \leq V$. Then $\hat{U} = U^{00}$, so after identification of V and V^{**}, we have $U = U^{00}$.

Proof. Let us first show that $U \leq U^{00}$. Indeed, let $u \in U$. Then for all $\varepsilon \in U^0$, $\varepsilon(u) = 0$. But then $\varepsilon(u) = \hat{u}(\varepsilon) = 0$, so $\hat{u} \in U^{00}$, and $\hat{U} \leq U^{00}$.

We can then compute dimensions, and $\dim U^{00} = \dim V - \dim U^0 = \dim U$, and since $v \mapsto \hat{v}$ is an isomorphism, $\dim \hat{U} = \dim U$. Thus $\dim \hat{U} = \dim U = \dim U^{00}$, and $\hat{U} = U^{00}$. $\qquad\square$

Remark. With the identification of V^{**} and V, for $T \leq V^*$ we can define $T^0 = \{v \in V \mid \alpha(v) = 0, \ \forall \alpha \in T\}$.

These results give us some unsurprising properties which are useful for doing computations.

Lemma 3.15 (Sums and Intersections of Annihilators)

Let V be a finite dimensional vector space over $\mathbb{F}$, and let $U_1, U_2 \leq V$. Then

  (i) $(U_1 + U_2)^0 = U_1^0 \cap U_2^0$;

  (ii) $(U_1 \cap U_2) = U_1^0 + U_2^0$.

Proof. (i) Let $\theta \in V^*$, then $\theta \in (U_1 + U_2)^0$ if and only if $\theta(u_1 + u_2) = 0$ for all $u_1 \in U_1$ and $u_2 \in U_2$. By linearity this is exactly when $\theta(u) = 0$ for $u \in U_1 \cup U_2$, which occurs if and only if $\theta \in U_1^0 \cap U_2^0$.

(ii) Take the annihilator of the result in (i), and use the fact that $U^{00} = U$. $\qquad\square$

---

§4 Bilinear Forms

§4.1 Introducing Bilinear Forms

In the previous section we considered linear maps from a vector space to the scalar field that vector space is over. In this section we will look at a related but interesting set of maps which have *two* arguments and are linear in both. Such maps are *bilinear forms*.

Definition 4.1 (Bilinear Form)

Let U and V be vector spaces over $\mathbb{F}$. Then $\phi : U \times V \to \mathbb{F}$ is a **bilinear form** if it is linear in both components.

Example 4.2 (Examples of Bilinear Forms)

The following are all bilinear forms.

   (i) A canonical example – the scalar product on $U = V = \mathbb{R}^n$.

   (ii) The map $V \times V^* \to \mathbb{F}$ given by $(v, \alpha) \mapsto \alpha(v)$.

   (iii) For $U = V = \mathcal{C}([0, 1], \mathbb{R})$ the map $\phi(f, g) = \int_0^1 f(t)g(t) \, \mathrm{d}t$.

There is a standard representation of bilinear forms using matrices.

Definition 4.3 (Matrix of a Bilinear Form)

Let $B = \{u_1, \ldots, u_n\}$ be a basis of U and $C = \{v_1, \ldots, v_m\}$ be a basis of V, and let $\phi : U \times V \to \mathbb{F}$ be a bilinear form.

We define the **matrix of** ϕ with respect to B and C by

$$([\phi]_{B,C})_{i,j} = \phi(u_i, v_j).$$

Lemma 4.4 (Evaluating a Bilinear Form)

Let $\phi : U \times V \to \mathbb{F}$ be a bilinear form, and let $u \in U$ and $v \in V$. Then if B is a basis of U and C is a basis of V, we have

$$\phi(u, v) = [u]_B^T [\phi]_{B,C} [v]_C.$$

Proof Sketch. Check definitions. $\square$

Remark. $[\phi]_{B,C}$ is the unique matrix such that this equation holds.

It's sometimes useful to study a bilinear form in just one of its arguments. This can be done by considering the induced maps.

Definition 4.5 (Induced Map of a Bilinear Form)

Let $\phi : U \times V \to \mathbb{F}$ be a bilinear form. Then ϕ determines two linear maps $\phi_L : U \to V^*$ and $\phi_R : V \to U^*$, with

$$(\phi_L(u))(v) = \phi(u, v) \quad \text{and} \quad (\phi_R(v))(u) = \phi(u, v).$$

These induced maps area easily obtained from the matrix of the bilinear form.

Lemma 4.6 (Matrices of the Induced Maps)

Let $\phi : U \times V \to \mathbb{F}$ be a bilinear map. Then if $B = \{u_1, \ldots, u_n\}$ is a basis of U and $C = \{v_1, \ldots, v_m\}$ is a basis of U, then if $A = [\phi]_{B,C}$ we have

$$[\phi_R]_{C,B^*} = A \quad \text{and} \quad [\phi_L]_{B,C^*} = A^T.$$

Proof Sketch. Check definitions. $\square$

Just as with the linear maps that we studied previously, it's possible look at the kernel of a bilinear map by considering the left and right component.

Definition 4.7 (Left/Right Kernel)

Let $\phi : U \times V \to \mathbb{F}$ be a bilinear form. Then the **left kernel** of ϕ is $\ker \phi_L$, and the **right kernel** of ϕ is $\ker \phi_R$.

Definition 4.8 (Degeneracy)

We say that a bilinear form ϕ is **non-degenerate** if $\ker \phi_L = \{0\}$ and $\ker \phi_R = \{0\}$. Otherwise, we say that ϕ is **degenerate**.

Lemma 4.9 (Matrix Degeneracy Criterion)

Let U and V be finite dimensional vector spaces, and let B be a basis of U and C be a basis of V. Suppose $\phi : U \times V \to \mathbb{F}$ is a bilinear form with matrix $[\phi]_{B,C}$. Then ϕ is non-degenerate if and only if A is invertible.

Proof. We know that ϕ is non-degenerate if and only if $\ker \phi_L = \{0\}$ and $\ker \phi_R = \{0\}$, that is, if and only if $n(A^T) = 0$ and $n(A) = 0$. But then by rank-nullity, we have this if and only if $r(A^T) = \dim U$ and $r(A) = \dim V$, that is, if and only if A is invertible. $\qquad\square$

Corollary 4.10

If $\phi : U \times V \to \mathbb{F}$ is a non-degenerate bilinear form, then $\dim U = \dim V$.

Corollary 4.11 (Chooing a Bilinear Form)

When U and V are finite dimensional vector spaces over $\mathbb{F}$, then choing a non-degenerate bilinear form $\phi : U \times V \to \mathbb{F}$ is equivalent to choosing an isomorphism $\phi_L : U \to V^*$.

As with matrices for normal linear maps $\alpha : V \to W$, we can change the basis that the matrix of a bilinear form is represented with.

Proposition 4.12 (Change of Basis for Bilinear Forms)

Let U and V be vector spaces over $\mathbb{F}$ and let B, B' be bases for U and C, C' be bases for V. Suppose that P is the change of basis matrix from B' to B and Q is the change of basis matrix from C' to C. Then if $\phi : U \times V \to \mathbb{F}$ is a bilinear form, then

$$[\phi]_{B',C'} = P^T [\phi]_{B,C} Q.$$

Proof. We have

$$\phi(u, v) = [u]_B^T [\phi]_{B,C} [v]_C = (P[u]_{B'})^T [\phi]_{B,C} (Q[v]_{C'}) = [u]_{B'}^T (P^T [\phi]_{B,C} Q)[v]_{C'},$$

and thus $[\phi]_{B',C'} = P^T [\phi]_{B,C} Q$. $\qquad\square$

This change of basis formula also allows us to define the rank of a bilinear map and have that definition be well defined.

> **Definition 4.13** (Rank of a Bilinear Map)
>
> The rank of a bilinear map ϕ, $r(\phi)$, is the rank of any matrix representation of ϕ.

To see that this is well defined, note that $r(P^T A Q) = r(A)$ for any invertible P, Q.

§4.2 Determinant and Trace

We will now look at the determinant and trace, two quantities you may be familiar with from the study of matrices. In this section, we will consider how these can be interpreted in a basis-free way.

> **Definition 4.14** (Trace)
>
> Let $A \in \mathcal{M}_{n,n}(\mathbb{F})$ be a square matrix. We define the trace of A by
> $$\operatorname{tr} A = \sum_{i=1}^{n} A_{ii}.$$

In particular, we have

$$\operatorname{tr} \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} = a_{11} + \cdots + a_{nn}.$$

Remark. The trace is a linear form $\mathcal{M}_{n,n}(\mathbb{F}) \to \mathbb{F}$ given by $A \mapsto \operatorname{tr} A$.

> **Lemma 4.15** (Trace of order)

§5 Lecture 14

§5.1 Cramer's Rule

We have an algorithm for computing the unique solution to $Ax = b$ without computing A^{-1}.

> **Proposition 5.1** (Cramer's Rule)
>
> Let $A \in \mathcal{M}_{n,n}(\mathbb{F})$ be an invertible matrix, and let $b \in \mathbb{F}^n$. Then the unique solution to $Ax = b$ is given by
> $$x_i = \frac{1}{\det A} \det(A_{\hat{i}b}),$$
> where $A_{\hat{i}b}$ is teh matrix obtained by replacing the ith column of A by b.

Proof. Since A is invertible, there exists a unique $x \in \mathbb{F}^n$ with $Ax = b$. Then computing we have

$$\det(A_{\hat{i}b}) = \det(A^{(1)}, \dots, A^{(i-1)}, b, A^{(i+1)}, A^{(n)})$$

$$= \det(A^{(1)}, \ldots, A^{(i-1)}, Ax, A^{(i+1)}, A^{(n)})$$
$$= \det(A^{(1)}, \ldots, A^{(i-1)}, \sum_{j=1}^{n} x_j A^{(j)}, A^{(i+1)}, A^{(n)})$$
$$= x_i \det(A^{(1)}, \ldots, A^{(i-1)}, A^{(i)}, A^{(i+1)}, A^{(n)})$$
$$= x_i \det A.$$

$\square$

§6 Diagonalization, Eigenvectors and Eigenvalues

§6.1 Antidiagonal Matrices

We are now going to take the first steps towards the diagonalisation of endomorphisms. Given a finite dimensional vector space V over $\mathbb{F}$ and some endomorphism $\alpha : V \to V$, we are going to be guided by the following general problem:

Can we find a basis B of V such that in this basis $[\alpha]_{B,B}$ has a 'nice' form?

With this in mind, we will first care about when this 'nice' form is a diagonal matrix, or more generally a triangular matrix.

Definition 6.1 (Diagonalizable)

Let V be a vector space and let $\alpha : V \to V$ be a linear map. We say that α is **diagonalizable** if there is a basis B of V such that

$$[\alpha]_{B,B} = \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix}$$

where $\lambda_i \in \mathbb{F}$.

Definition 6.2 (Triangulable)

Let V be a vector space and let $\alpha : V \to V$ be a linear map. We say that α is **triangulable** if there is a basis B of V such that

$$[\alpha]_{B,B} = \begin{pmatrix} \lambda_1 & \cdots & * \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix}$$

where $\lambda_i \in \mathbb{F}$.

We will see that when thinking about such matrices, we will frequently turn to the concept of eigenvectors and eigenvalues.

Definition 6.3 (Eigenvectors and Eigenvalues)

If V is a vector space over $\mathbb{F}$ and $\alpha : V \to V$ is a linear map, we say that $\lambda \in \mathbb{F}$ is a **eigenvalue** if there exists some $v \in V \backslash \{0\}$ such that $\alpha(v) = \lambda v$. We say that v is a

corresponding **eigenvector**.

Definition 6.4 (Eigenspace)

If λ is an eigenvalue of a linear map $\alpha : V \to V$, we define the corresponding **eigenspace** to be the subspace $V_\lambda = \{v \in V \mid \alpha(v) = \lambda v\}$.

When computing eigenvalues we frequently employ the following lemma.

Lemma 6.5 (Computing Eigenvalues)

Let V be a vector space over $\mathbb{F}$, and let $\alpha : V \to V$ be a linear map. Then $\lambda \in \mathbb{F}$ is an eigenvalue if and only if $\det(\alpha - \lambda\,\mathrm{id}) = 0$.

Proof. There exists $v \in V\backslash\{0\}$ with $\alpha(v) = \lambda v$ if and only if $\ker(\alpha - \lambda\,\mathrm{id}) \neq \{0\}$. This occurs if and only if $\alpha - \lambda\,\mathrm{id}$ is not injective, and since α is an endomorphism, this occurs if and only if $\alpha - \lambda\,\mathrm{id}$ is not invertible, that is, if $\det(\alpha - \lambda\,\mathrm{id}) = 0$. $\square$

Applying this lemma gives us a polynomial in λ which one must find the roots of to obtain the eigenvalues of the linear map. We call this polynomial the *characteristic polynomial*.

Definition 6.6 (Characteristic Polynomial)

Let $\alpha : V \to V$ be an endomorphism. Then the **characteristic polynomial**[a] of α is $\chi_\alpha(\lambda) = \det(\alpha - \lambda\,\mathrm{id})$.

---

[a]It should be checked that this definition is basis free, but this is straightforward.

The heart of the matter is the following criterion of triangulability.

Theorem 6.7 (Triangulability Criterion)

Let V be a vector space over $\mathbb{F}$. A linear map $\alpha : V \to V$ is triangulable if and only if

$$\chi_\alpha(t) = c\prod_{i=1}^{n}(t - \lambda_i),$$

for some $c, \lambda_i \in \mathbb{F}$.

Proof. Suppose α was triangulable. Then for some basis B of V we have

$$[\alpha]_{B,B} = \begin{pmatrix} \lambda_1 & \cdots & * \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix},$$

and so

$$\chi_\alpha(t) = \det\begin{pmatrix} \lambda_1 - t & \cdots & * \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n - t \end{pmatrix} = \prod_{i=1}^{n}(\lambda_i - 1).$$

Now suppose that this criterion holds. We will prove by induction on $n = \dim V$ that it implies triangulability. For $n = 1$, we are immediately done, so suppose that $n > 1$. Then let λ be a root of $\chi_\alpha(t)$ (which exists by assumption). Then λ is an eigenvalue of α. Let $U = V_\lambda$ be the corresponding eigenspace, and let $\{v_1, \ldots, v_k\}$ be a basis of U. We can complete this to a basis $B = \{v_1, \ldots, v_n\}$ of V. So then

$$[\alpha]_{B,B} = \left(\begin{array}{c|c} \lambda I & * \\ \hline 0 & C \end{array} \right).$$

Now α induces an endomorphism $\overline{\alpha} : V/U \to V/U$, and C is the matirx of $\overline{\alpha}$ with respect to the basis $\{v_{k+1} + U, \ldots, v_n + U\}$, and by assumption we know we can write this in triangular form by choosing a basis $\{v'_{k+1}, \ldots, v'_n\}$, and taking the basis $B' = \{v_1, \ldots, v_k, v'_{k+1}, \ldots, v'_n\}$ gives us $[\alpha]_{B'}$ in triangular form. $\qquad\square$

Remark. By the fundamental theorem of algebra, if $\mathbb{F} = \mathbb{C}$, then every matrix is triangulable.