

# Part IB — Linear Algebra

Based on lectures by Dr. A. Keating (amk50@cam.ac.uk)

Notes taken by Christopher Turnbull

Michaelmas 2017

## 0 Introduction

Linear algebra is an important component of undergraduate mathematics. At the practical level, matrix theory and the related vector-space concepts provide a language and a powerful computational framework for posing and solving important problems.

Beyond this, elementary linear algebra is a valuable introduction to mathematical abstraction and logical reasoning because the theoretical development is self-contained, consistent, and accessible to most students.

# 1 Vector Spaces

## 1.1 Vector Spaces

**Definition.** An  $\mathbb{F}$ -Vector space (a vector space on  $\mathbb{F}$ ) is an abelian group  $(V, +)$  equipped with a function<sup>1</sup>  $F \times V \rightarrow V$ ,  $(\lambda, v) \mapsto \lambda v$

$$\lambda(v_1 + v_2) = \lambda v_1 + \lambda v_2$$

$$(\lambda_1 + \lambda_2)v = \lambda_1 v + \lambda_2 v$$

$$\lambda(\mu v) = \lambda\mu v$$

$$1v = v$$

$$v + \mathbf{0} = v$$

for all  $\lambda_i, \lambda, \mu \in F$ ,  $v_i \in V$

Note that we will not be underlining our vectors, as this is cumbersome here. We will however be using  $\mathbf{0}$  to denote the zero vector.

**Example.** For all  $n \in \mathbb{N}$ ,  $\mathbb{F}^n$  = space of column vectors of length  $n$ , entries in  $\mathbb{F}$ . We understand the definition as entry-wise addition, entry-wise scalar multiplication

**Example.**  $M_{m,m}(\mathbb{F})$ , the set of  $m \times m$  matrices with entries in  $\mathbb{F}$

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} + \begin{pmatrix} e & f \\ g & h \end{pmatrix} = \begin{pmatrix} a+e & b+f \\ c+g & d+h \end{pmatrix}$$

again all operations defined entry-wise

**Example.** For any set  $X$ ,  $\mathbb{R}^X = \{f : X \rightarrow \mathbb{R}\}$  Addition and scalar multiplication defined pointwise =  $f_1(x) + f_2(x)$ .

**Exercise.** Show that the above examples satisfy the axioms

**Proposition.**  $0v = \mathbf{0}$  for all  $v \in V$ .

*Proof.*  $((0+0)v = 0v \iff 0v + 0v = 0v \iff 0v = \mathbf{0})$  □

**Exercise.** Show<sup>2</sup> that  $(-1)v = -v$

**Definition.** Let  $V$  be an  $\mathbb{F}$ -vector space. A subset  $U$  of  $V$  is a subspace ( $U \leq V$ ) if:

- (i)  $\mathbf{0} \in U$
- (ii)  $u_1, u_2 \in U \Rightarrow u_1 + u_2 \in U$  “ $U$  is closed under addition...”
- (iii)  $u \in U$ , any  $\lambda \in \mathbb{F} \Rightarrow \lambda u \in U$  “...and scalar multiplication”

<sup>1</sup>scalar multiplication

<sup>2</sup>Hint: Use the previous proposition

**Exercise.** If  $U$  is a subspace of  $V$ , then  $U$  is also an  $\mathbb{F}$ -vector space.

**Example.** Let  $V = \mathbb{R}^{\mathbb{R}}$ , then  $f : \mathbb{R} \rightarrow \mathbb{R}$ . The set of all continuous functions  $C(\mathbb{R})$  are a subspace. An even smaller subspace is the set of all polynomials.

**Exercise.** Define  $U \subseteq \mathbb{R}^3$  as:

$$\left\{ \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \mid a_1 + a_2 + a_3 = t \right\}$$

for some constant  $t$ . Check that this is a subspace of  $\mathbb{R}^3$  if and only if  $t = 0$ .

**Proposition.** Let  $V$  be an  $F$ -vector space,  $U, W \leq V$ . Then  $U \cap W \leq V$ .

*Proof.* (i)  $0 \in U, 0 \in W \Rightarrow 0 \in U \cap W$

(ii) Suppose  $u, v \in U \cap W$ ,  $\lambda, \mu \in F$ .  $U$  is a subspace  $\Rightarrow \lambda u + \mu v \in W$ .  
Similarly  $\lambda u + \mu v \in U \in W$ , so it is in the intersection. □

**Example.**  $V = \mathbb{R}^3$ ,  $U = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mid x = 0 \right\}$ ,  $W = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mid y = 0 \right\}$  then  $U \cap W = U = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mid x = 0, y = 0 \right\}$  (intersect along the  $z$ -axis)

Note: union of family of subspaces is almost never a subspace itself.

**Definition.** Let  $V$  be an  $F$ -vector space,  $U, W \leq V$ . The *sum* of  $U$  and  $W$  is the set:

$$U + W = \{u + w \mid u \in U, w \in W\}$$

**Proposition.**  $U + W \leq V$

*Proof.*  $0 \in U, W \Rightarrow 0 + 0 = 0 \in U + W$

$u_1, u_2 \in U, w_1, w_2 \in W$ ,

$$(u_1 + w_1) + (u_2 + w_2) = \underbrace{(u_1 + u_2)}_{\in U} + \underbrace{(w_1 + w_2)}_{\in W}$$

Similarly for scalar multiplication (ex.) □

Note:  $U + W$  is the smallest subspace containing both  $U$  and  $W$ . (This is because all elements of the form  $u + w$  are forced to be in such a subspace by the “closed under addition” axiom)

**Definition.**  $V$  is an  $\mathbb{F}$ -vector space,  $U \leq V$ . The quotient space<sup>3</sup>  $V/U$  is the abelian group  $V/U$  equipped with scalar multiplication;

$$F \times V/U \rightarrow V/U$$

$$(\lambda, v + U) \mapsto \lambda v + U$$

---

<sup>3</sup>think of this as the collection of cosets of  $U$  in  $V$

**Proposition.** This is well-defined, and  $V/U$  is an  $F$ -vector space.

*Proof.* Well-defined: Suppose  $v_1 + U = v_2 + U \in V/U$ .  $\Rightarrow (v_2 - v_1) \in U \Rightarrow (\lambda v_2 - \lambda v_1) \in U \Rightarrow \lambda v_2 + U = \lambda v_1 + U \in V/U$

To show that it is an  $\mathbb{F}$ -vector space, we must show that the axioms hold. These follow from the axioms of  $V$ .  $\lambda(\mu(v + U)) = \lambda(\mu v + U) = \lambda(\mu v) + U = (\lambda\mu)v + U = \lambda\mu(v + U)$  (scalar multiplication on  $V/U$ ).

Ex. Other axioms follow similarly from using vector space axioms

□

## 1.2 Bases

**Definition.**  $V$  is an  $\mathbb{F}$ -vector space,  $S \subset V$ . The *span* of  $S$  is denoted by

$$\langle S \rangle = \left\{ \sum_{s \in S} \lambda_s s \mid \lambda_s \in \mathbb{F} \right\}$$

ie. the set of all finite linear combinations, all but finitely many of the  $\lambda_s$  are zero.

Remark:  $\langle S \rangle$  is the smallest subspace of  $V$  which contains<sup>4</sup> all of the elements of  $S$

Convention:  $\langle \emptyset \rangle = \{\mathbf{0}\}$ .

**Example.**  $V = \mathbb{R}^3$ ,

$$S = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 3 \\ -2 \\ -4 \end{pmatrix} \right\}$$

$$\langle S \rangle = \left\{ \begin{pmatrix} a \\ b \\ 2b \end{pmatrix} \mid a, b \in \mathbb{R} \right\}$$

ie. we have took linear combinations of the first two. We don't need the third one.

**Example.** For  $X$  a set, define  $\delta_x(y) : X \rightarrow \mathbb{F}$  as

$$\delta_x(y) = \begin{cases} 1 & \text{if } x = y \\ 0 & \text{if } x \neq y \end{cases}$$

$$\langle \delta_x \mid x \in X \rangle = \{f \in \mathbb{R}^X \mid f \text{ has finite support}\}$$

$$= \langle x \in X \mid f(x) \neq 0 \rangle$$

**Definition.**  $S$  spans  $V$  if  $\langle S \rangle = V$

**Definition.**  $V$  is *finite dimensional* over  $\mathbb{F}$  if it is spanned by a set that is finite.

<sup>4</sup>This is essentially a tautology

**Definition.** The vectors  $v_1, \dots, v_n$  are *linearly independent* over  $\mathbb{F}$  if

$$\sum_{i=1}^n \lambda_i v_i = 0 \Rightarrow \lambda_i = 0 \text{ for all } i$$

some coefficients  $\lambda_i \in \mathbb{F}$ .  $S \subset V$  is linearly independent if every finite subset of it is.

**Example.** The first example,  $u, v, w$  are not linearly independent<sup>5</sup>, but the set  $\{\delta_X \mid x \in X\}$  is linearly independent.

A lesson to be learnt from our example is that a linearly dependent spanning set contains redundant information. In a sense, a linearly independent spanning set is a minimal spanning set and hence represents the most efficient way of characterizing the subspace. This idea leads to the following definition.

**Definition.**  $\mathcal{B}$  is a *basis* of  $V$  if it is linearly independent and spans  $V$

**Example.** –  $\mathbb{F}^n$  standard basis:  $\{e_1, e_2, \dots, e_n\}$ .

- $V = \mathbb{C}$  over  $\mathbb{C}$  has natural basis  $\{1\}$ , over  $\mathbb{R}$  has natural basis  $\{1, i\}$
- $V = \mathcal{P}(\mathbb{R})$  space of all polynomials, has natural basis

$$\{1, x, x^2, x^3, \dots\}$$

**Exercise.** Check this carefully

**Lemma.**  $V$  is an  $\mathbb{F}$ -vector space. The vectors  $v_1, \dots, v_n$  form a basis of  $V$  iff each vector  $v \in V$  has a unique expression

$$v = \sum_{i=1}^n \lambda_i v_i, \text{ with } \lambda_i \in \mathbb{F}$$

*Proof.* ( $\Rightarrow$ ) Fix  $v \in V$ . The  $v_i$  span, so

$$\exists \lambda_i \in \mathbb{F} \text{ s.t. } v = \sum \lambda_i v_i$$

Suppose also  $v = \sum \mu_i v_i$  for some  $\mu_i \in \mathbb{F}$ .  $\sum (\mu_i - \lambda_i) v_i = \mathbf{0}$ .

The  $v_i$  are linearly independent so  $\mu_i - \lambda_i = 0$  for all  $i$ ,  $\lambda_i = \mu_i$

( $\Leftarrow$ ) The  $v_i$  span  $V$ , since any  $v \in V$  is a linear combination of them. IF  $\sum_{i=1}^n \lambda_i v_i = \mathbf{0}$ . Note that  $\mathbf{0} = \sum_{i=1}^n 0 v_i$ . By uniqueness (applied to  $\mathbf{0}$ ),  $\lambda_i = 0$  for all  $i$ .  $\square$

**Lemma.** If  $v_1, \dots, v_n$  span  $V$  (over  $\mathbb{F}$ ), then some subset of  $v_1, \dots, v_n$  is a basis for  $V$  (over  $\mathbb{F}$ ).

*Proof.* If  $v_1, \dots, v_n$  linearly independent, done. Otherwise for some  $l$ , there exist  $\alpha_1, \dots, \alpha_{l-1} \in \mathbb{F}$  such that

$$v_l = \alpha_1 v_1 + \dots + \alpha_{l-1} v_{l-1}$$

<sup>5</sup>If *not* linearly independent, say a set is linearly dependent.

( If  $\sum \lambda_i v_i = \mathbf{0}$ , not all  $\lambda_i = 0$ . Take  $l$  maximaml with  $\lambda_i \neq 0$ , just  $\alpha_i = -\lambda_i/\lambda_l$  ).

Now  $v_2, \dots, v_{l-1}, v_{l+1}, \dots, v_n$  still span  $V$ . Continue iteratively until get linear independence.  $\square$

**Theorem.** (Steinitz exchange lemma) Let  $V$  be a finite dimensional vector space over  $\mathbb{F}$ . Take  $v_1, \dots, v_m$  to be linearly independent  $w_1, \dots, w_n$  to span  $V$ .

Then  $m \leq n$ , and reordering the spanning set if needed,

$$v_1, \dots, v_m, w_{m+1}, \dots, w_n$$

span  $V$ .

*Proof.* (Induction) Suppose that we've replaced  $l(\geq 0)$  of the  $w_i$ . Reordering the  $w_i$  if needed,  $v_1, \dots, v_l, w_{l+1}, \dots, w_n$  span  $V$ .

If  $l = m$ , done.

If  $l < m$ , then

$$v_{l+1} = \sum_{i=1}^l \alpha_i v_i + \sum_{i>l} \beta_i w_i$$

$\alpha_i, \beta_i \in \mathbb{F}$ . As the  $v_i$  are lin. indep,  $\beta_i \neq 0$  for some  $i$ . (After reordering,  $\beta_{l+1} \neq 0$ ).

$$w_{l+1} = \frac{1}{\beta_{l+1}} \left( v_{l+1} - \sum_{i \leq l} \alpha_i v_i - \sum_{i>l+1} \beta_i w_i \right)$$

This  $v_1, \dots, v_{l+1}, w_{l+2}, \dots, w_n$  also spans  $V$ . After  $m$  steps,  $w_i$  will have replaced  $m$  of the  $w_i$  by  $v_i$ . Thus  $m \leq n$ .  $\square$

**Theorem.** If  $V$  is a finite dimensional vector space over  $\mathbb{F}$ , then any two bases for  $V$  have the same number of elements. This is what we call the *dimension* of  $V$ , denoted  $\dim_{\mathbb{F}} V$ .

*Proof.* If  $\{v_1, \dots, v_n\}$  is a basis and  $w_1, \dots, w_m$  is another basis, the  $\{v_i\}$  span and  $\{w_i\}$  is linearly indepndent' so by Steinitz  $m \leq n$ . Likewise,  $n \leq m$ .  $\square$

**Example.**  $\dim_{\mathbb{C}} \mathbb{C} = 1$ ,  $\dim_{\mathbb{R}} \mathbb{C} = 2$

**Theorem.**  $V$ , finite dim,  $v$ -space over  $\mathbb{F}$ . If  $w_1, \dots, w_l$  is a linearly indepndent set of vectors, we can extend it to a basis  $w_1, \dots, w_l, v_{l+1}, \dots, v_n$

*Proof.* Apply Steinitiz to  $w_1, \dots, w_l$  (lin indep) and any basis  $v_1, \dots, v_n$ .

Or directrly, if  $V = \langle w_1, \dots, w_l \rangle$ , stop.

Otherwise take  $v_{l+1} \in V \setminus \langle w_1, \dots, w_l \rangle$ , now  $w_1, \dots, w_l, v_{l+1}$  is linearly indep. iterate  $\square$

**Corollary.** Suppose  $V$  is a finite dimensional vector space, with dimension  $n$ .

- (i) Any linearly independent set of vectors has at most  $n$  elements with equality iff it's a basis

- (ii) Any spanning set of vectors must have at least  $n$  elements, with equality if and only if it's a basis.

Slogan "Choose the best basis for the job"

**Theorem.** Let  $U, W$  be subspaces of  $V$ . If  $U, W$  are finite dim, so is  $U + W$  and  $\dim(U + W) = \dim U + \dim W - \dim(U \cap W)$

*Proof.* Pick basis  $v_1, \dots, v_l$  of  $U \cap W$ . Extend it to basis  $v_1, \dots, v_l, u_1, \dots, u_m$  of  $U$ . Extend it to basis  $v_1, \dots, v_l, w_1, \dots, w_n$  of  $W$ .

Claim:  $v_1, \dots, v_l, u_1, \dots, u_m, w_1, \dots, w_n$  is a basis for  $U + W$ .

- (i) Span:  $u \in U$ , then  $u = \sum \alpha_i v_i + \sum \beta_i u_i$ ,  $\alpha_i, \beta_i \in \mathbb{F}$   $w \in W$ , then  $w = \sum \gamma_i v_i + \sum \delta_i w_i$ ,  $\gamma_i, \delta_i \in \mathbb{F}$

$$u + w = \sum (\alpha_i + \gamma_i) v_i + \sum (\beta_i + \delta_i) u_i$$

- (ii) lin indep:  $u = \sum \alpha_i v_i + \sum \beta_i u_i + \sum \gamma_i w_i = \mathbf{0}$

$$\Rightarrow u = \underbrace{\sum \alpha_i v_i + \sum \beta_i u_i}_{\in U} - \underbrace{\sum \gamma_i w_i}_{\in W} \in U \cap W$$

This is equal to  $\sum \delta_i v_i$  for some  $\delta_i \in \mathbb{F}$  because  $v_i$  are basis for  $U \cap W$ .

As  $v_i$  and  $w_i$  are lin indep,  $(*) \Rightarrow \gamma_i = \delta_i = 0$  for all  $i$ .

$\Rightarrow \sum \alpha_i v_i + \sum \beta_i u_i = 0 \Rightarrow \alpha_i = \beta_i = 0$  because  $v_i$  and  $u_i$  form a basis for  $U$ .

□

**Theorem.** Let  $V$  be a finite dim  $\mathbb{F}$ -vector space,  $U \leq V$ , then  $U$  and  $V/U$  are also of finite dim, and

$$\dim V = \dim U + \dim V/U$$

*Proof.*

**Exercise.** Show that  $U$  is finite dim.

Let  $u_1, \dots, u_l$  be a basis for  $U$ . Extend it to a basis for  $V$ . Say  $u_1, \dots, u_l, w_{l+1}, \dots, w_n$  of  $V$ .

**Exercise.** Check:  $w_{l+1} + U, \dots, w_n + U$  form a basis for  $V/U$ .

□

**Corollary.** If  $U$  is a proper subspace of  $V$ ,  $V$  is finite dimensional,  $\dim U < \dim V$ .

*Proof.*  $V/U \neq \{0\} \Rightarrow \dim V/U > 0 \Rightarrow \dim U < \dim V$

□

**Definition.** Let  $V$  be an  $\mathbb{F}$ -vector space,  $U, W \leq V$  Then  $V = U \oplus W$  ( $V$  is an internal direct sum of  $U$  and  $W$ ) if every element of  $V$  can be written as  $v = u + w, w \in W, u \in U$ , uniquely.

$W$  is a *direct complement* of  $U$  in  $V$



**Lemma.**  $U, W \leq V$ . The following are equivalent

- (i)  $V = U \oplus W$ , ie. every element of  $V$  can be written uniquely as  $u + w$ , for  $u \in U, w \in W$
- (ii)  $V = U + W$  and  $U \cap W = \{0\}$
- (iii)  $B_1$  any basis of  $U$ ,  $B_2$  is any basis of  $W$ , then  $B = B_1 \cup B_2$  is a basis of  $V$ .

*Proof.* (ii)  $\Rightarrow$  (i). Any  $v \in V$  is  $u + w$  for some  $u \in U$ ,  $w \in W$ .

Suppose that

$$u_1 + w_1 = u_2 + w_2$$

Then

$$\Rightarrow u_1 - u_2 = -w_1 + w_2 \in U \cap W = \{0\} \Rightarrow w_1 = w_2, u_1 = u_2$$

Thus uniqueness of expressions.

(i)  $\Rightarrow$  (iii)  $B$  spans, any  $v \in V$  is  $u + w$ , for some  $u \in U$ ,  $w \in W$ , write  $u$  in terms of  $B_1$ ,  $w$  in terms of  $B_2$ , Then  $u + w$  is a lin comb. of elements of  $B$ .

$B$  indep?

$$\begin{aligned} \sum_{v \in B} \lambda_v v = 0 &= 0_U + 0_W \\ \underbrace{\sum_{v \in B_1} \lambda_v v}_{\in U} + \sum_{v \in B_2} \lambda_v v &= 0 \end{aligned}$$

By uniqueness of expressions,

$$\sum_{v \in B_1} \lambda_v v = 0_U \quad \sum_{v \in B_2} \lambda_v v = 0_W$$

As  $B_1$  and  $B_2$  are basis, all of the  $\lambda_v$  are zero.

$$(iii) \Rightarrow (ii). \text{ If } v \in V, v = \sum_{x \in B} \lambda_x x = \underbrace{\sum_{u \in B_1} \lambda_u u}_{\in U} + \underbrace{\sum_{w \in B_2} \lambda_w w}_{\in W}$$

$$\Rightarrow v \in U + W.$$

If  $v \in U \cap W$ ,  $v = \sum_{u \in B_1} \lambda_u u$ ,  $v = \sum_{w \in B_2} \lambda_w w \Rightarrow$  All  $\lambda_u, \lambda_w$  are zero, because  $B_1 \cup B_2$  is lin. indep.

□

**Lemma.** Let  $V$  be an  $f$ -dim vector space.  $U \leq V$ . Then there exists a direct complement to  $U$  in  $V$

*Proof.* Let  $u_1, \dots, u_l$  be a basis for  $U$ . Extend it to a basis for  $V$ ,

$$u_1, \dots, u_l, w_{l+1}, \dots, w_n$$

Then  $\langle w_{l+1}, \dots, w_n \rangle$  is a direct complement of  $U$ .

□

Note! Direct compliments are not at all unique. In general, if you pick different ways of extending this you will get different direct compliments.

Pick  $V = \mathbb{R}^2$ . Pick  $U$  as the  $y$ -axis, then any one of the following green lines are direct compliments.:

**Definition.** Def  $v_1, \dots, v_l \leq V$ ,

$$\sum V_i = V_1 + \dots + V_l = \{v_1 + \dots + v_l \mid v_i \in V_i\}$$

The sum is direct if

$$v_1 + \dots + v_l = v'_1 + \dots + v'_l \Rightarrow v_i = v'_i \text{ for all } i$$

(“unique expressions”)

Notation:

$$\bigoplus_{i=1}^l V_i$$

**Exercise.**  $V_1, \dots, V_l \leq V$ . TFAE

- (i) The sum  $\sum V_i$  is direct
- (ii)  $V_i \cap \sum_{j \neq i} V_j = \{0\}$  for all  $i$
- (iii) The  $B_i$  are pairwise disjoint and their union is a basis for  $\sum V_i$

We also discuss *external* direct sums, though will not touch them much in this course. This is simply an internal direct sum  $U_1 \oplus U_2$ , except now the  $U_i$ 's are not subspaces of  $V$ , they can be any old vector space.

**Definition.** Let  $U, W$  be  $\mathbb{F}$ -vector spaces. External direct sum

$$U \oplus V = \{(u, w) \mid u \in U, w \in W\}$$

$$\begin{aligned} \text{with } (u, w) + (x, y) &= (u + x, w + y), \\ \lambda(u, w) &= (\lambda u, \lambda w) \end{aligned}$$

Note that when we talk about dimension in this course, we have not shown yet that the dimension of an *infinite* vector space is well defined<sup>6</sup>. We will come to this later.

---

<sup>6</sup>It is!

## 2 Linear Maps

### 2.1 Linear Maps

**Definition.**  $V, W$  are  $\mathbb{F}$ -vector spaces. A map  $\alpha : V \rightarrow W$  is linear if

- (i)  $\alpha(v_1 + v_2) = \alpha(v_1) + \alpha(v_2)$
- (ii)  $\alpha(\lambda v) = \lambda \alpha(v)$

Can be combined concisely as:

$$\alpha(\lambda_1 v_1 + \lambda_2 v_2) = \lambda_1 \alpha(v_1) + \lambda_2 \alpha(v_2) \quad \lambda_i \in \mathbb{F}, v_i \in V$$

**Example.** A  $n \times m$  matrix with coeff in  $\mathbb{F}$

$$\begin{aligned} \alpha : \mathbb{F}^n &\rightarrow \mathbb{F}^m \\ v &\mapsto Av \end{aligned}$$

**Example.** The set of all polynomials with real coefficients:

$$\mathcal{D} : \mathcal{P}(\mathbb{R}) \rightarrow \mathcal{P}(\mathbb{R})$$

$$f \mapsto \frac{df}{dx}$$

**Example.** The set of continuous functions over  $[0, 1]$

$$I : \mathcal{C}[0, 1] \rightarrow \mathcal{C}[0, 1]$$

$$f \mapsto I(f)$$

$$\text{where } I(f)(x) = \int_0^x f(t) \, dt$$

**Example.** Fix  $x \in [0, 1]$

$$\begin{aligned} \mathcal{C}[0, 1] &\rightarrow \mathbb{R} \\ f &\mapsto f(x) \end{aligned}$$

Notes: If  $U, V, W$  are v spaces over  $\mathbb{F}$ , then

- (i) The identity map  $\text{id} : V \rightarrow V$  is linear
- (ii) If  $U \xrightarrow{\alpha} V \xrightarrow{\beta} W$  with  $\alpha, \beta$  both linear, then  $\beta \circ \alpha$  is linear.

**Lemma.** Let  $V, W$  be  $\mathbb{F}$ -vector spaces, and let  $\mathcal{B}$  is a basis for  $V$ . If  $\alpha_0 : \mathcal{B} \rightarrow W$  is *any* map, then there exists a unique linear map<sup>7</sup>  $\alpha : V \rightarrow W$  extending  $\alpha_0$ , ie.

$$\alpha(v) = \alpha_0(v)$$

for any basis element  $v \in \mathcal{B}$ .

<sup>7</sup>ie. if I tell you *any* mapping of the basis vectors  $\alpha_0$  (it could be a non-linear mapping), then you have enough information to construct a linear map from this.

*Proof.* Let  $v \in V$ . Then  $v = \sum \lambda_i v_i$ ,  $v_i \in B$ ,  $\lambda_i \in \mathbb{F}$ , unique expression.

Now Linearity forces

$$\begin{aligned}\alpha(v) &= \alpha\left(\sum \lambda_i v_i\right) \\ &= \sum \lambda_i \alpha(v_i) \\ &= \sum \lambda_i \alpha_0(v_i)\end{aligned}$$

linear, exists. expression forced to be unique.  $\square$

Note

- (i) True for infinite dimensional vector space also
- (ii) Very often, to define a linear map, define it on a basis and ‘extend linearly’
- (iii) Let  $\alpha_1, \alpha_2 : V \rightarrow W$  be linear maps. If they agree on any basis, then they are equal.

**Definition.** (Isomorphism)

Let  $V, W$  be vector spaces over  $F$ . The map  $\alpha : V \rightarrow W$  is an *isomorphism* if it is linear and bijective. Notation:  $V \simeq W$

**Lemma.**  $\simeq$  is an equivalence notation on the set (score out set and write class) of all vector spaces over  $\mathbb{F}$ . That is,

- (i)  $i_V : V \rightarrow V$  is an iso
- (ii) If  $\alpha : V \rightarrow W$  is an iso, then the inverse map  $\alpha^{-1} : W \rightarrow V$  is also linear, hence an iso.
- (iii) If

$$U \xrightarrow{\beta} V \xrightarrow{\alpha} W$$

then

$$U \xrightarrow{\beta \circ \alpha} W$$

is also an iso

*Proof.* (i) immediate

- (ii)  $\alpha$  bijective  $\Rightarrow \alpha^{-1}$  exists. Check: linear.  $w_2 \in W, w_2 = \alpha(v_2)$ ,  $v_2 \in V$ , unique.  $\alpha^{-1}(w_1 + w_2) = \alpha^{-1}(\alpha(v_1) + \alpha(v_2)) = \alpha^{-1}(\alpha(v_1 + v_2)) = v_1 + v_2 = \alpha^{-1}(w_1) + \alpha^{-1}(w_2)$ .

Similarly,  $\lambda \in \mathbb{F}$ ,  $w \in W$ ,

$$\alpha^{-1}(\lambda w) = \lambda \alpha^{-1}(w)$$

- (iii) immediate  $\square$

**Theorem.** If  $V$  vector space over  $\mathbb{F}$  of dimension  $n$ , then  $V \simeq \mathbb{F}^n$ .

*Proof.* Choose a basis  $\mathcal{B}$  for  $V$ , say  $v_1, \dots, v_n$

$$V \rightarrow \mathbb{F}^n$$

$$\sum \lambda_i v_i \mapsto \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} \text{ is an iso}$$

□

Remark: Choosing an iso  $V \simeq \mathbb{F}^n$  is equivalent to choosing a basis for  $V$ .

**Theorem.**  $V, W$   $v$  spaces over  $\mathbb{F}$ , finite dim, are isomorphic iff they have the same dimension

*Proof.* ( $\Leftarrow$ ) Both  $V$  and  $W$  are isomorphic

$$\mathbb{F}^{\dim V} = \mathbb{F}^{\dim W}$$

( $\Rightarrow$ ) Let  $\alpha : V \rightarrow W$  iso,  $\mathcal{B}$  a basis for  $V$ .

Claim:  $\alpha(\mathcal{B})$  is a basis for  $W$ .

**Exercise.**  $\alpha(\mathcal{B})$  spans  $W$  because of surjectivity of  $\alpha$ .

**Exercise.**  $\alpha(\mathcal{B})$  lin indep: follows from injectivity of  $\alpha$ .

□

**Definition.** (Null space/ Kernel of a linear map) Let  $\alpha : V \rightarrow W$  be a linear map, the *null space* of  $\alpha$  is given by

$$N(\alpha) = \ker \alpha = \{v \in V \mid \alpha(v) = \mathbf{0}\} \leq V$$

**Definition.** (Image of a linear map) Let  $\alpha : V \rightarrow W$  be a linear map, the *image* of  $\alpha$  is defined as:

$$\text{Im}(\alpha) = \{w \in W \mid w = \alpha(v), \text{ some } v \in V\} \leq W$$

**Definition.** (Injective map)  $\alpha$  is injective if and only if  $N(\alpha) = \{\mathbf{0}\}$

**Definition.** (Surjective map)<sup>8</sup>  $\alpha$  is surjective if and only if  $\text{Im}(\alpha) = W$

**Example.** Let  $\alpha : \mathcal{C}^\infty(\mathbb{R}) \rightarrow \mathcal{C}^\infty(\mathbb{R})$  be defined by

$$\alpha(f)(t) = f''(t) + 2f'(t) - 5f$$

$$\ker \alpha \text{ is solns to } f'' + 2f' + 5f = 0$$

$$g \in \text{Im } \alpha \text{ if } \exists \text{ soln } f \text{ to } f'' + 2f' + 5f = g$$

---

<sup>8</sup>I mean, all definitions are iff statements really. Sometimes we leave it out and just use 'if'

## 2.2 The First Isomorphism Theorem

**Theorem.** (First Isomorphism Theorem) Let  $\alpha : V \rightarrow W$  be a linear map. It induces an iso :

$$V/\ker \alpha \xrightarrow{\bar{\alpha}} \text{Im}(\alpha)$$

defined by

$$\bar{\alpha}(v + \ker \alpha) = \alpha(v)$$

*Proof.* (i)  $\bar{\alpha}$  is well defined:

$$\begin{aligned} v + \ker \alpha = v' + \ker \alpha \\ \iff v - v' \in \ker \alpha \Rightarrow \alpha(v) \\ \Rightarrow \alpha(v) = \alpha(v') \end{aligned}$$

(ii)  $\bar{\alpha}$  is linear; immediate from linearity of  $\alpha$ .

(iii)  $\bar{\alpha}$  bijective?

$$\begin{aligned} \bar{\alpha}(v + \ker \alpha) &= \mathbf{0} \\ \Rightarrow \alpha(v) &= 0 \\ \Rightarrow v &\in \ker \alpha \end{aligned}$$

(iv) surjective: by defn of  $\text{Im}(\alpha)$ . □

**Definition.** (Rank and Nullity of a linear map) The *rank* of a linear map  $r(\alpha) = rk(\alpha)$  is given by  $\dim(\text{Im } \alpha)$ , and the *nullity*  $n(\alpha)$  is likewise given as  $\dim(N(\alpha))$

**Theorem.** (Rank-nullity theorem) Let  $U, V$  be vector spaces over  $\mathbb{F}$ ,  $\dim_{\mathbb{F}} U < \infty$ . Let  $\alpha : U \rightarrow V$  linear. Then:

$$\dim U = r(\alpha) + n(\alpha)$$

*Proof.*

$$U/\ker \alpha \simeq \text{Im}(\alpha) \Rightarrow \dim(U) - \dim(\ker \alpha) = \dim(\text{Im}(\alpha))$$

□

**Lemma.** Let  $V, W$  be v spaces over  $\mathbb{F}$ , of equal finite dim. Let  $\alpha : V \rightarrow W$  linear.

TFAE

- (i)  $\alpha$  injective
- (ii)  $\alpha$  surjective

(iii)  $\alpha$  isomorphism

**Definition.** The *space of linear maps* from  $V$  to  $W$  is denoted by

$$L(V, W) = \{\alpha : V \rightarrow W \text{ linear}\}$$

**Proposition.**  $L(V, W)$  is a v-space over  $\mathbb{F}$  under operators

- $(\alpha_1 + \alpha_2)(v) = \alpha_1(v) + \alpha_2(v)$  for all  $\alpha_i \in L(V, W)$
- $(\lambda\alpha)(v) = \lambda(\alpha(v))$  for all  $v \in V, \lambda \in \mathbb{F}$

If both  $V$  and  $W$  are finite dim, then so is  $L(V, W)$  and  $\dim(L(V, W)) = \dim(V) \times \dim(W)$ .

*Proof.*  $\alpha_1 + \alpha_2, \lambda\alpha$  defined above are well-defined linear maps. The v-space axioms are satisfied.

Claim about finite dim: See later □

### 2.3 Representation of Linear Maps by Matrices

**Definition.** An  $m \times n$  matrix over  $\mathbb{F}$  is an array with  $m$  rows and  $n$  columns, entries in  $\mathbb{F}$ .

$$A = (a_{ij}), \quad a_{ij} \in \mathbb{F}, \quad 1 \leq i \leq m, 1 \leq j \leq n$$

$M_{m,n}(\mathbb{F})$  is the set of all such matrices

**Proposition.**  $M_{m,n}(\mathbb{F})$  is an  $\mathbb{F}$  vector space, under operations

$$(a_{ij}) + (b_{ij}) = (a_{ij} + b_{ij})$$

$$\lambda(a_{ij}) = (\lambda a_{ij})$$

$$\text{and } \dim(M_{m,n}(\mathbb{F})) = m \times n$$

*Proof.* v-space okay, see 1.1. And dim? A standard basis for  $M_{m,n}(\mathbb{F})$  is

$$E_{ij} = \begin{pmatrix} 0 & \cdots & 0 \\ \vdots & 1 & \vdots \\ 0 & \cdots & 0 \end{pmatrix}$$

(ie a matrix of zeroes, with 1 in  $i^{\text{th}}$  row and  $j^{\text{th}}$  column)

$(a_{ij}) = \sum_{i,j} a_{ij} E_{ij}$ , from which span and LI follows

This basis has cardinality  $mn$  □

**Definition.** (Coordinate Vectors)

Let  $V, W$  be v-spaces over  $\mathbb{F}$ , of finite dim, with  $\alpha : V \rightarrow W$ , linear. Basis  $\mathcal{B}$  for  $V$ ,  $v_1, \dots, v_n$  basis  $\mathcal{C}$  for  $W$ ,  $w_1, \dots, w_n$ . If  $v \in V$ ,  $v = \sum \lambda_i v_i$ , write

$$[v]_{\mathcal{B}} = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} \in \mathbb{F}^n, \text{ called coordinate vector of } v \text{ wrt } \mathcal{B}. \text{ Similarly, } [w]_{\mathcal{C}} \in \mathbb{F}^m.$$

**Definition.** (Matrix)  $[\alpha]_{\mathcal{B},\mathcal{C}}$  matrix of  $\alpha$  wrt  $\mathcal{B}$  and  $\mathcal{C}$

$$[\alpha]_{\mathcal{B},\mathcal{C}} = ([\alpha(v_1)]_{\mathcal{C}} \mid [\alpha(v_2)]_{\mathcal{C}} \mid \cdots \mid [\alpha(v_n)]_{\mathcal{C}}) \in M_{m,n}(\mathbb{F})$$

$$= (a_{ij})$$

The notation says  $\alpha(v_j) = \sum \alpha_{ij} w_i$

**Lemma.** For any  $v \in V$ ,

$$[\alpha(v)]_{\mathcal{C}} = [\alpha]_{\mathcal{B},\mathcal{C}} \cdot [v]_{\mathcal{B}}$$

where the dot denotes matrix applied to vector

*Proof.* Fix  $v \in V$ ,  $v = \sum_{j=1}^n \lambda_j v_j$ , so  $[v]_{\mathcal{B}} = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}$

$$\begin{aligned} \alpha(v) &= \alpha\left(\sum \lambda_j v_j\right) = \sum \lambda_j \alpha(v_j) = \sum_j \lambda_j \left(\sum_i \alpha_{ij} w_i\right) \\ &= \sum_i \underbrace{\left(\sum_j \alpha_{ij} \lambda_j\right)}_{i^{\text{th}} \text{ entry of } [\alpha]_{\mathcal{B},\mathcal{C}} [v]_{\mathcal{B}}} w_i \end{aligned}$$

□

**Lemma.** Let  $\alpha, \beta$  be linear maps, with  $U \xrightarrow{\beta} V \xrightarrow{\alpha} W$  and  $\alpha \circ \beta$  linear. Let  $\mathcal{A}, \mathcal{B}, \mathcal{C}$  be basis for  $U, W, V$  reps. Then

$$[\alpha \circ \beta]_{\mathcal{A},\mathcal{C}} = \underbrace{[\alpha]_{\mathcal{B},\mathcal{C}}}_{=(a_{ij})} \circ \underbrace{[\beta]_{\mathcal{A},\mathcal{B}}}_{=(b_{ji})}$$

*Proof.*

$$\begin{aligned} (\alpha \circ \beta) \left( \overbrace{u_i}^{\text{in } \mathcal{A}} \right) &= \alpha(\beta(u_i)) = \alpha\left(\sum_j b_{ji} \overbrace{v_j}^{\text{in } \mathcal{B}}\right) \\ &= \sum_j b_{ji} \alpha(v_j) \\ &= \sum_j b_{ji} \sum_i a_{ij} \overbrace{w_i}^{\text{in } \mathcal{C}} \\ &= \sum_i \underbrace{\left(\sum_j a_{ij} b_{ji}\right)}_{(i,j)^{\text{th}} \text{ entry of } [\alpha]_{\mathcal{B},\mathcal{C}} [\beta]_{\mathcal{A},\mathcal{B}}} w_i \end{aligned}$$

□



**Proposition.** If  $V, W$  are  $v$ -spaces over  $\mathbb{F}$  with  $\dim V = n$ ,  $\dim W = m$ , then  $L(V, W) \simeq M_{m,n}(\mathbb{F})$

*Proof.* Fix bases

$$\mathcal{B} \text{ of } V : v_1, v_2, \dots, v_n$$

$$\mathcal{C} \text{ of } W : w_1, w_2, \dots, w_m$$

Claim:

$$L(v, w) \rightarrow M_{m,n}(\mathbb{F})$$

$$\alpha \mapsto [\alpha]_{\mathcal{B}, \mathcal{C}}$$

is an iso.

- $\theta$  linear  $[\lambda_1 \alpha_1 + \lambda_2 \alpha_2]_{\mathcal{B}, \mathcal{C}} = \lambda_1 [\alpha_1]_{\mathcal{B}, \mathcal{C}} + \lambda_2 [\alpha_2]_{\mathcal{B}, \mathcal{C}}$
- $\theta$  surjective: given  $A = (a_{ij})$ . Let  $\alpha : v_j \mapsto \sum_{i=1}^m a_{ij} w_i$ , and extend linearly. Then  $\alpha \in L(V, W)$ ,  $b\theta(\alpha) = A$ .
- $\theta$  injective,  $[\alpha]_{\mathcal{B}, \mathcal{C}} = 0$  matrix  $\Rightarrow \alpha$  is zero-map from  $V$  to  $W$ .

□

**Corollary.**

$$\dim(L(V, W)) = (\dim V)(\dim W)$$

**Example.**  $\alpha : V \rightarrow W, Y \leq V, Z \leq W$ . Say  $\alpha(Y) \subseteq Z$ .

Basis of  $V$  :

$$\mathcal{B} : \underbrace{v_1, \dots, v_k}_{\text{Basis for } Y, \mathcal{B}'}, v_{k+1}, \dots, v_n$$

Basis of  $W$  :

$$\mathcal{C} : \underbrace{w_1, \dots, w_l}_{\text{Basis for } Z, \mathcal{C}'}, w_{l+1}, \dots, w_m$$

Then

$$[\alpha]_{\mathcal{B}, \mathcal{C}} = \begin{pmatrix} A & \cdots & B_1 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & C_1 \end{pmatrix}$$

because for  $1 \leq j \leq k$ ,  $\alpha(v_j)$  is a lin combo of  $w_i$ , where  $1 \leq i \leq l$ .

And

$$[\alpha|_Y]_{\mathcal{B}', \mathcal{C}'} = A_1$$

Claim:  $\alpha$  induces

$$\bar{\alpha} : V/Y \rightarrow W/Z$$

$$v + Y \mapsto \alpha(v) + Z$$

Well defined?

$$\begin{aligned} v_1 + Y = v_2 + Y &\Rightarrow v_1 - v_2 \in Y \\ &\Rightarrow \alpha(v_1 - v_2) \in Z \\ &\Rightarrow \alpha(v_1) + Z = \alpha(v_2) + Z \end{aligned}$$

**Exercise.** Linear from linearity of  $\alpha$

Basis for  $Y/V$ ,

$$\mathcal{B}'' : v_{k+1} + Y, \dots, v_n + Y$$

Basis for  $W/Z$ ,

$$\mathcal{B}'' : v_{k+1} + Y, \dots, v_n + Y$$

**Exercise.**  $[\bar{\alpha}]_{\mathcal{B}'', \mathcal{C}''}$

## 2.4 Change of Basis

Let  $V$  and  $W$  be  $v$ -spaces over  $\mathbb{F}$  with the following basis

$$\begin{array}{cc} V & W \\ \mathcal{B} = \{v_1, \dots, v_n\} & \mathcal{C} = \{w_1, \dots, w_m\} \\ \mathcal{B}' = \{v'_1, \dots, v'_n\} & \mathcal{C}' = \{w'_1, \dots, w'_m\} \end{array}$$

**Definition.** The *change of basis matrix* from  $\mathcal{B}$  to  $\mathcal{B}'$  is  $P = (p_{ij})$  given by  $v'_j = \sum p_{ij} v_i$ .  
Equivalently,

$$P = \left( \begin{array}{ccc|ccc} [v'_1]_{\mathcal{B}} & & & & & \\ [v'_2]_{\mathcal{B}} & & & & & \\ & \cdots & & & & \\ & & [v'_n]_{\mathcal{B}} & & & \end{array} \right) = [\text{id}]_{\mathcal{B}', \mathcal{B}}$$

**Lemma.**  $[v]_{\mathcal{B}} = P[v]_{\mathcal{B}'}$

*Proof.*

$$P[v]_{\mathcal{B}'} = [\text{id}]_{\mathcal{B}', \mathcal{B}}[v]_{\mathcal{B}'} = [v]_{\mathcal{B}}$$

□

**Lemma.**  $P$  is an invertible  $n \times n$  matrix, and  $P^{-1}$  is the change of basis matrix from  $\mathcal{B}$  to  $\mathcal{B}'$

*Proof.*

$$[\text{id}]_{\mathcal{B}, \mathcal{B}'}[\text{id}]_{\mathcal{B}', \mathcal{B}} = [\text{id}]_{\mathcal{B}', \mathcal{B}'} = I_n$$

$$[\text{id}]_{\mathcal{B}', \mathcal{B}}[\text{id}]_{\mathcal{B}, \mathcal{B}'} = [\text{id}]_{\mathcal{B}, \mathcal{B}} = I_n$$

□

Let  $Q$  be the change of basis matrix from  $\mathcal{C}'$  to  $\mathcal{C}$ .  $Q$  also invertible  $m \times m$ .

**Proposition.** Let  $\alpha : V \rightarrow W$  linear,  $A = [\alpha]_{\mathcal{B}, \mathcal{C}}$ ,  $A' = [\alpha]_{\mathcal{B}', \mathcal{C}'}$ . Then

$$A' = Q^{-1}AP$$

*Proof.*

$$\begin{aligned} Q^{-1}AP &= [\text{id}]_{\mathcal{C}, \mathcal{C}'} [\alpha]_{\mathcal{B}, \mathcal{C}} [\text{id}]_{\mathcal{B}', \mathcal{B}} \\ &= [\text{id} \circ \alpha \circ \text{id}]_{\mathcal{B}', \mathcal{C}'} \\ &= A' \end{aligned}$$

□

**Definition.**  $A, A' \in M_{m,n}(\mathbb{F})$  are *equivalent* if  $A' = Q^{-1}AP$  for some invertible  $P \in M_{n,n}(\mathbb{F})$ ,  $Q \in M_{m,m}(\mathbb{F})$

Note: this defines an equivalence relation on  $M_{m,n}(\mathbb{F})$ , eg.  $A' = Q^{-1}AP$ ,  $A'' = (Q^{-1})^{-1}A'P' \Rightarrow A'' = (QQ^{-1})^{-1}APP'$

**Proposition.** Let  $V, W$  be  $\mathbb{F}$ -vector spaces of dim  $n$  and  $m$  resp. Let  $\alpha : V \rightarrow W$  be a linear map. Then there exists bases  $\mathcal{B}$  of  $V$  and  $\mathcal{C}$  of  $W$ , and some  $r \leq m, n$  st.

$$[\alpha]_{\mathcal{B}, \mathcal{C}} = \begin{pmatrix} I_r & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{pmatrix}$$

where  $I_r$  is the identity matrix.

Note:  $r = \text{rank}(\alpha) = r(\alpha)$

*Proof.* Fix  $r$  st.  $N(\alpha)$  has dim  $n-r$ . Fix a basis for  $N(\alpha)$ , say  $v_{r+1}, v_{r+2}, \dots, v_n$ . Extend this to a basis for  $V$ , say  $\underbrace{v_1, \dots, v_r}_{\mathcal{B}}, v_{r+1}, \dots, v_n$ . Now  $\alpha(v_1), \dots, \alpha(v_r)$

is a basis for  $\text{im}(\alpha)$ .

– span:  $\alpha(v_1), \dots, \alpha(v_r), \underbrace{\alpha(v_{r+1})}_{=0}, \dots, \underbrace{\alpha(v_n)}_{=0}$  certainly span  $\text{im}(\alpha)$

– LI:

$$\begin{aligned} \sum_{i=1}^r \lambda_i \alpha(v_i) = \mathbf{0} &\Rightarrow \alpha \left( \underbrace{\sum_{i=1}^r \lambda_i v_i}_{\in \ker(\alpha)} \right) = \mathbf{0} \\ &\Rightarrow \sum_{i=1}^r \lambda_i v_i = \sum_{j=r+1}^n \mu_j v_j \text{ some } \mu_j \in \mathbb{F} \\ &\Rightarrow \text{as } v_1, \dots, v_n \text{ LI, } \lambda_i = \mu_j = 0 \forall i, j \end{aligned}$$

Extend  $\alpha(v_1), \dots, \alpha(v_r)$  to a basis of  $W$ , say  $\mathcal{C}$ . By construction,

$$[\alpha]_{\mathcal{B}, \mathcal{C}} = \begin{pmatrix} I_r & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{pmatrix}$$

Remark: didn't need to assume in the proof that  $r = r(\alpha)$ . Can think of this as giving a different proof of the r-n theorem.  $\square$

**Corollary.** Any  $m \times n$  matrix is equivalent to  $\begin{pmatrix} I_r & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{pmatrix}$  for some  $r$ .

**Definition.** Let  $A \in M_{m,n}(\mathbb{F})$ . The *column rank* of  $A$  is the dimension of the subspace of  $\mathbb{F}^m$  spanned by the columns of  $A$ . The *row rank* of  $A$  is the column rank of  $A^T$  (the dimension of the subspace of  $\mathbb{F}^n$  spanned by the row vectors of  $A$ ).

Note: if  $\alpha$  is a linear map represented by  $A$  wrt. any choice of basis, then  $r(\alpha) = r(A)$ , ie column rank = rank.

**Proposition.** Two  $m \times n$  matrices  $A, A'$  are equivalent iff  $r(A') = r(A)$ .

*Proof.* ( $\Leftarrow$ ) Both  $A$  and  $A'$  are equivalent to  $\begin{pmatrix} I_r & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{pmatrix}$ ,  $r = r(A') = r(A)$

(this is a transitive relation)

( $\Rightarrow$ ) Let  $\alpha$  be the linear map:  $\alpha : \mathbb{F}^n \rightarrow \mathbb{F}^m$  represented by  $A$  wrt. the standard basis  $A' = Q^{-1}AP$ .  $P$  and  $Q$  invertible, so  $A'$  represents  $\alpha$  wrt. two other bases.  $r(\alpha)$  is defined in a basis invariant way, so  $r = r(\alpha) = r(A) = r(A')$   $\square$

**Theorem.**  $r(A) = r(A^T)$  ("row rank = column rank").

*Proof.*  $Q^{-1}AP = \begin{pmatrix} I_r & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{pmatrix}_{m,n}$  where  $Q, P$  invertible

Take transpose of whole equation:

$$\begin{pmatrix} I_r & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{pmatrix}_{n,m} = (Q^{-1}AP)^T = P^T A^T (Q^T)^{-1}$$

so  $A^T$  equiv to  $\begin{pmatrix} I_r & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{pmatrix}_{n,m}$ . Thus  $r(A) = r(A^T)$ .  $\square$

$V = W$ ,  $\mathcal{C} = \mathcal{B}$ , other basis  $\mathcal{B}'$ .  $P$  change of basis matrix from  $\mathcal{B}$  to  $\mathcal{B}'$ ,  $\alpha \in L(V, V)$ .

$$[\alpha]_{\mathcal{B}', \mathcal{B}'} = P^{-1}[\alpha]_{\mathcal{B}, \mathcal{B}}P$$

**Definition.**  $A, A' \in M_{n,n}(\mathbb{F})$ ,  $A, A'$  are *similar* (or *conjugate*) if  $A' = P^{-1}AP$  for some invertible  $P$ .

## 2.5 Elementary Matrices and Operations

**Definition.** *Elementary column operators* on an  $m \times n$  matrix  $A$ :

- (i) swap columns  $i$  and  $j$  (wlog  $i \neq j$ )
- (ii) replace column  $i$  by  $\lambda$  (column  $i$ ),  $\lambda \neq 0$
- (iii) add  $\lambda$ (column  $i$ ) to column  $j$ ,  $i \neq j$ ,  $\lambda \neq 0$ .

*Elementary row operators* analogous (replace ‘column’ by ‘row’)

Note: all of these operations are reversible.

Corresponding elementary matrices: effect of performing the column operations on  $I_n = n \times n$  id. For row operations,  $I_m$ .

$$(i) \begin{pmatrix} 1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & 1 & & & & & & \\ & & & 0 & & & 1 & & \\ & & & & 1 & & & & \\ & & & & & \ddots & & & \\ & & & & & & 1 & & \\ & & 1 & & & & & 0 & 1 \\ & & & & & & & & 1 \\ & & & & & & & & & \ddots \\ & & & & & & & & & & 1 \end{pmatrix}$$

The zeros appear in row  $i$ , row  $j$ .

$$(ii) \begin{pmatrix} 1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & 1 & & & & & & \\ & & & \lambda & & & & & \\ & & & & 1 & & & & \\ & & & & & \ddots & & & \\ & & & & & & 1 & & \end{pmatrix}$$

with  $\lambda$  in the  $i^{\text{th}}$  row

- (iii)  $I_n + \lambda E_{ij}$ , where  $E_{ij}$  is defined as 1 in the  $(i, j)$  position and 0 everywhere else.

An elementary column operation on  $A \in M_{m,n}(\mathbb{F})$  can be performed by multiplying  $A$  by the corresponding elementary matrix on the right.

**Exercise.**

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

For row operations, multiply on the left

**Exercise.**

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

**Theorem.** Any  $m \times n$  matrix is equivalent to

$$\begin{pmatrix} I_r & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{pmatrix}$$

for some  $r$

*Proof.* Start with  $A$ . If all entries of  $A$  are 0, we're done ( $r = 0$ ). If not, some  $a_{ij} = \lambda \neq 0$ .

- swap rows  $1, i$
- swap columns  $1, j$
- multiply column 1 by  $\frac{1}{\lambda}$

to get 1 in position  $(1, 1)$ . Now

- add  $(-a_{12})(\text{column } 1)$  to column 2.
- Similarly clear out all other entries in row 1.
- Also use row operations to clear out all other entries in column 1

Upshot: get

$$\begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & \tilde{A} & & \\ 0 & & & \end{pmatrix}, \tilde{A} \in M_{m-1, n-1}(\mathbb{F})$$

$$\text{Now iterate, to get } \begin{pmatrix} I_r & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{pmatrix} = \underbrace{E'_1, \dots, E'_k}_{\text{elem row}} \overset{Q^{-1}}{A} \underbrace{E_1, \dots, E_k}_P$$

Row/column ops are reversible  $\Rightarrow$  elem matrices are invertible.  $Q : m \times m$  invertible,  $P : n \times n$  invertible.

□

Variations:

If you use elementary row operations, can get the row echelon form of a matrix.

$$\begin{pmatrix} 1 & 0 & 0 & 0 & a \\ & & 1 & 0 & b \\ & & & 1 & c \end{pmatrix}$$

How? Assume  $a'_{i1} = \lambda \neq 0$  some  $i$ .

- swap rows 1 and  $i \Rightarrow$  get  $\lambda$  in  $(1, 1)$
- divide row 1 by  $\lambda \Rightarrow 1$  in  $(1, 1)$
- use (iii)-type operation to clear out rest of column 1, then move on to second column etc.

**Lemma.** If  $A$  is  $n \times n$  invertible, we can obtain  $I_n$  by using only elementary row operations (or elementary column operations).

*Proof.* Induction on number of rows

Suppose we have

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

There exists  $j > k$  with  $a_{k+1,j} = \lambda \neq 0$

If not,

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

with 1 in the  $(k+1)$ th entry would not lie in the span of the column vectors, which would contradict invertability

- Swap columns  $k+1$  and  $j$
- Divide column  $k+1$  by  $\lambda$
- Use type 3 operators to clear the other entries of the  $(k+1)$ th row.
- now proceed inductively

□

Upshot

$$AE_1E_2 \cdots E_l = I_n \Rightarrow A^{-1} = E_1E_2 \cdots E_l$$

one recipe for inverses.

**Proposition.** Any invertible matrix can be written as a product of elementary matrices.

### 3 Dual Spaces and Dual Maps

#### 3.1 Dual Vector Spaces

**Definition.** Let  $V$  be a vector space over  $\mathbb{F}$ . The *dual vector space*  $V^*$  of  $V$

$$V^* = L(V, \mathbb{F}) = \{ \alpha : V \rightarrow \mathbb{F} \text{ linear} \}$$

$V^*$  is a vector space over  $\mathbb{F}$ . Its elements are sometimes called linear functionals.

**Example.**  $V = \mathbb{R}^3$ ,

$$\theta : V \rightarrow \mathbb{R}$$

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} \rightarrow a - c \quad \theta \in V^*$$

**Example.**

$$t_n : M_{m,n}(\mathbb{F}) \rightarrow \mathbb{F}$$

$$A \mapsto \sum_i A_{ii}, \quad t_n \in (M_{m,n}(\mathbb{F}))^*$$

**Lemma.** Let  $V$  be a vector space over  $\mathbb{F}$  with finite basis  $\mathcal{B} = \{e_1, \dots, e_n\}$ . Then there is a basis for  $V^*$ , given by  $\mathcal{B}^* = \{\varepsilon_1, \dots, \varepsilon_n\}$  where

$$\varepsilon_j \left( \underbrace{\sum_{i=1}^m a_i e_i}_{\in V} \right) = a_j \quad 1 \leq j \leq n$$

$\mathcal{B}^*$  is called the dual basis to  $\mathcal{B}$

*Proof.* – LI

$$\begin{aligned} \sum_{j=1}^n \lambda_j \varepsilon_j = \mathbf{0} &\Rightarrow \left( \sum_{j=1}^n \lambda_j \varepsilon_j \right) e_i = \mathbf{0} \\ &= \sum_j \lambda_j \underbrace{\varepsilon_j(e_i)}_{\delta_{ij}} \end{aligned}$$

$$\Rightarrow \lambda_i = 0 \quad \forall i = 1, \dots, n$$

– Span: If  $\alpha \in V^*$ , then  $\alpha = \sum_{i=1}^n \alpha(e_i) \varepsilon_i$   
 (“linear maps are determined by their action on a basis”)

□

**Corollary.** If  $V$  is finite dim, then  $\dim V = \dim V^*$



Remark: Sometimes useful to think about  $(\mathbb{F}^n)^*$  as the space of row vectors of length  $n$  over  $\mathbb{F}$ . Suppose

$V$  basis  $e_1, \dots, e_n$

$V^*$  dual basis  $\varepsilon_1, \dots, \varepsilon_n$

$$x = \sum x_i e_i \in V$$

$$\alpha = \sum a_i \varepsilon_i \in V^*$$

$$\alpha(x) = \sum_{i=1}^n \alpha_i x_i = (a_1 \quad \dots \quad a_n) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

**Definition.** If  $U \subseteq V$ ,

$$U^0 = \{\alpha \in V^* \text{ st. } \alpha(u) = 0 \text{ for all } u \in U\}$$

is the *annihilator* of  $U$

**Lemma.** (i)  $U^0 \leq V^*$

(ii) If  $U \leq V$  and  $\dim V = n < \infty$ , then

$$\dim V = \dim U + \dim U^0$$

*Proof.* (i)  $0 \in U^0$ . If  $\alpha, \alpha' \in U^0$ , then  $(\alpha + \alpha')(u) = \alpha(u) + \alpha'(u) = 0 + 0 = 0$ , for  $u \in U$  thus  $\alpha + \alpha' \in U^0$

Similarly,  $\lambda\alpha \in U^0$  for any  $\lambda \in \mathbb{F}$

(ii) Let  $e_1, \dots, e_k$  be a basis for  $U$ . Extend to a basis for  $V$ .  $e_1, \dots, e_k, e_{k+1}, \dots, e_n$ .

Let  $\mathcal{B}^*$  be the dual basis to this.  $\varepsilon_1, \dots, \varepsilon_n$

Claim:  $\varepsilon_{k+1}, \varepsilon_{k+2}, \dots, \varepsilon_n$  is a basis for  $U^0$

- If  $i > k$ ,  $\varepsilon_i(e_j) = 0$  where  $j \leq k$ , so  $\varepsilon_i$  (for  $i > k$ ) is in  $U^0$ .
- LI comes from the fact that  $\mathcal{B}^*$  is a basis. (So any subset of it is LI).
- Span? If  $\alpha \in U^0$ , then  $\sum_{i=1}^n \alpha_i \varepsilon_i$ , some  $a_i \in \mathbb{F}$ .

$$\left( \sum_{i=1}^n a_i \varepsilon_i \right) (e_j) = 0 \Rightarrow a_j = 0, \text{ any } j \leq k$$

where  $e_j$  is a basis element for  $U$ , for  $j \leq k$

$$\Rightarrow \alpha \in \langle \varepsilon_{k+1}, \dots, \varepsilon_n \rangle$$

□

### 3.2 Dual Maps

**Lemma.** Let  $V, W$  be vector spaces over  $\mathbb{F}$ . Let  $\alpha \in L(V, W)$ . Then the map

$$\alpha^* : W^* \rightarrow V^*$$

$$\varepsilon \mapsto \varepsilon \circ \alpha \text{ is linear}$$

$$V \xrightarrow{\alpha} W \xrightarrow{\varepsilon} F$$

We'll call  $\alpha^*$  the *dual* of  $\alpha$ .

*Proof.* –  $\varepsilon \circ \alpha$  is linear, so in  $V^*$ .

–  $\alpha^*$  linear? Fix  $\theta_1, \theta_2 \in W^*$

$$\begin{aligned} \alpha^*(\theta_1 + \theta_2) &= (\theta_1 + \theta_2) \circ \alpha \\ &= \theta_1 \circ \alpha + \theta_2 \circ \alpha \\ &= \alpha^* \theta_1 + \alpha^* \theta_2 \end{aligned}$$

Similarly,  $\alpha^*(\lambda\theta) = \lambda\alpha^*\theta$  □

**Proposition.** Let  $V, W$  be v-spaces over  $\mathbb{F}$ , with basis  $\mathcal{B}, \mathcal{C}$  respectively. Let  $\mathcal{B}^*, \mathcal{C}^*$  be the dual basis. Consider  $\alpha \in L(V, W)$  with dual  $\alpha^*$ .

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

*Proof.* Say  $\mathcal{B} = \{b_1, \dots, b_n\}$ ,  $\mathcal{C} = \{c_1, \dots, c_n\}$

$$\mathcal{B}^* = \{\beta_1, \dots, \beta_n\}, \mathcal{C}^* = \{\gamma_1, \dots, \gamma_n\}$$

$$\text{and } [\alpha]_{\mathcal{B}, \mathcal{C}} = (a_{ij}) \quad m \times n$$

$$\begin{aligned} \alpha^*(\gamma_r)(b_s) &= \gamma_r \circ \alpha(b_s) \\ &= \gamma_r(\alpha(b_s)) \\ &= \gamma_r\left(\sum_t a_{ts} c_t\right) \\ &= \sum_t a_{ts} \gamma_r(c_t) \\ &= a_{rs} \\ &= \left(\sum_i a_{ri} \beta_i\right)(b_s) \end{aligned}$$

$$\Rightarrow \alpha^*(\gamma_r) = \sum_i a_{ri} \beta_i$$

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

□

Let  $V$  be a finite dim  $\mathbb{F}$  vector space.

Bases  $\varepsilon = \{e_1, \dots, e_n\}$ ,  $F = \{f_1, \dots, f_n\}$

Dual bases  $\varepsilon^* = \{\varepsilon_1, \dots, \varepsilon_n\}$ ,  $\mathcal{F} = \{\eta_1, \dots, \eta_n\}$

And let us consider  $P = [\text{id}]_{\mathcal{F}\mathcal{E}}$

**Lemma.** Change of basis matrix from  $\mathcal{F}^*$  to  $\mathcal{E}^*$  is  $(P^{-1})^T$

*Proof.*

$$[\text{id}]_{\mathcal{F}^*, \mathcal{E}^*} = [\text{id}]_{\mathcal{E}\mathcal{F}}^T = ([\text{id}]_{\mathcal{F}\mathcal{E}}^{-1})^T$$

□

CAUTION:  $V \simeq V^*$  only if  $V$  is finite dimensional.

Let  $V = \mathcal{P}$ , the space of all real polynomials, with basis

$$p_j, j = 0, 1, 2, \dots \quad p_j(t) = t^j$$

Ex sheet 2 Q 9:

$$P^* \simeq \mathbb{R}^{\mathbb{N}}$$

$$\varepsilon \mapsto (\varepsilon(p_0), \varepsilon(p_1), \dots)$$

Ex sheet 1, Q3 g)  $P \not\simeq \mathbb{R}^{\mathbb{N}}$  does NOT have a countable basis

**Lemma.** Let  $V, W$  be vector spaces over  $\mathbb{F}$ . Fix  $\alpha \in L(V, W)$ , let  $\alpha^* \in L(W^*, V^*)$  be the dual map. Then

- (i)  $N(\alpha^*) = (\text{Im}(\alpha))^0$  ie.  $\alpha^*$  injective iff  $\alpha$  is surjective
- (ii)  $\text{Im}(\alpha^*) \leq (N(\alpha))^0$ , with equality if  $V$  and  $W$  are finite dimensional. ie.  $\alpha^*$  surjective iff  $\alpha$  is injective

*Proof.* (i) Let  $\varepsilon \in W^*$ . Then

$$\begin{aligned} \varepsilon \in N(\alpha^*) &\iff \alpha^* \varepsilon = 0 \\ &\iff \varepsilon \circ \alpha = 0 \\ &\iff \varepsilon(u) = 0 \text{ for all } u \in \text{Im } \alpha \\ &\iff \varepsilon \in (\text{Im}(\alpha))^0 \end{aligned}$$

- (ii) Let  $\varepsilon \in \text{Im } \alpha^*$ . Then  $\varepsilon = \alpha^* \varphi$ , for some  $\varphi \in W^*$ .

For any  $u \in N(\alpha)$ ,

$$\begin{aligned} \varepsilon(u) &= (\alpha^* \varphi)(u) \\ &= (\varphi \circ \alpha)(u) \\ &= \varphi(\alpha(u)) \\ &= \varphi(0) \\ &= 0 \end{aligned}$$

So  $\varepsilon \in N(\alpha^0)$

Now use the fact that  $\dim V, \dim W$  are finite.

$$\begin{aligned}
 \dim(\text{Im}(\alpha^*)) &= r(\alpha^*) \\
 &= r(\alpha) \quad \text{as } r(A) = r(A^T) \\
 &= \dim V - \dim N(\alpha) \quad \text{by R-N} \\
 &= \dim(N(\alpha))^0
 \end{aligned}$$

□

### 3.3 Double Duals

**Definition.** Let  $V$  be an  $\mathbb{F}$  vector space,  $V^* = L(V, \mathbb{F})$  dual of  $V$ . Then the *double dual* of  $V$  is the dual of  $V^*$ , given by

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

**Theorem.** If  $V$  is a finite dimensional vector space over  $\mathbb{F}$ , then the map

$$\begin{aligned}
 \hat{\phantom{x}} : V &\rightarrow V^{**} \\
 v &\mapsto \hat{v}, \quad \hat{v}(\varepsilon) = \varepsilon(v)
 \end{aligned}$$

is an isomorphism

*Proof.* Firstly, for  $v \in V$ , the map  $\hat{v} : V^* \rightarrow \mathbb{F}$  is linear, so  $\hat{\phantom{x}}$  does indeed give a map from  $V$  to  $V^{**}$

–  $\hat{\phantom{x}}$  is linear. If  $v_1, v_2 \in V, \lambda_1, \lambda_2 \in \mathbb{F}, \varepsilon \in V^*$ .

$$\begin{aligned}
 \widehat{(\lambda_1 v_1 + \lambda_2 v_2)}(\varepsilon) &= \varepsilon(\lambda_1 v_1 + \lambda_2 v_2) \\
 &= \lambda_1 \varepsilon(v_1) + \lambda_2 \varepsilon(v_2) \\
 &= \lambda_1 \hat{v}_1(\varepsilon) + \lambda_2 \hat{v}_2(\varepsilon)
 \end{aligned}$$

–  $\hat{\phantom{x}}$  is injective: Let  $e \in V \setminus \{\mathbf{0}\}$ . Extend it to a basis of  $V$ , say  $e_1, e_2, \dots, e_n$ . Let  $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$  be the dual basis for  $V^*$ .

$\hat{e}(\varepsilon) = \varepsilon(e) = 1$ . So  $\hat{e} \neq 0$ .

Thus  $N(\hat{\phantom{x}}) = \{\mathbf{0}\}$ , so  $\hat{\phantom{x}}$  is injective.

–  $V$  is finite dim, so  $\dim V = \dim V^* = \dim V^{**}$ .

Thus  $\hat{\phantom{x}}$  is an isomorphism

□

**Lemma.** Let  $V$  be a finite dim vector space over  $\mathbb{F}$  and  $U \leq V$ .

Then  $\hat{U} = U^{00}$ , so after identification of  $V$  with  $V^{**}$ , we have that  $U^{00} = U$ .

*Proof.* – First show  $\hat{U} \leq U^{00}$ .

$$\begin{aligned} u \in U &\Rightarrow \varepsilon(u) = 0 \quad \forall \varepsilon \in U^0 \\ &= \hat{u}(\varepsilon) = 0 \\ &\Rightarrow \hat{u} \in (U^0)^0 = U^{00} \end{aligned}$$

$$\begin{aligned} \dim U^{00} &= \dim V^* - \dim U^0 \\ &= \dim V - \dim U^0 \\ &= \dim U \end{aligned}$$

Thus  $\hat{U} = U^{00}$

□

**Lemma.** Let  $V$  be a finite dim vector space of  $\mathbb{F}$ , Let  $U_1, U_2 \leq V$ . Then

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

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

*Proof.* (i) Let  $\theta \in V^*$

$$\begin{aligned} \theta \in (U_1 + U_2)^0 &\iff \theta(u_1 + u_2) = 0 \text{ for all } u_1 \in U_1, u_2 \in U_2 \\ &= \theta(u) = 0 \text{ for all } u \in U_1 \cap U_2 \\ &\theta \in U_1^0 \cap U_2^0 \end{aligned}$$

(ii) Apply annihilator to (i).

$$w_i = U_i^0 \quad u_i = W_i^0$$

$$\begin{aligned} (W_1^0 + W_2^0)^0 &= W_1 \cap W_2 \\ W_1^0 + W_2^0 &= (W_1 \cap W_2)^0 \end{aligned}$$

□

## 4 Bilinear Forms I

**Definition.** Let  $U, V$  be vector spaces over  $\mathbb{F}$ .

$$\varphi : U \times V \rightarrow \mathbb{F}$$

is *bilinear* or a *bilinear form* if its linear in both arguments

$$\varphi(u, -) : V \rightarrow \mathbb{F} \quad \in V^* \quad \forall u \in U$$

$$\varphi(-, v) : U \rightarrow \mathbb{F} \quad \in U^* \quad \forall v \in V$$

**Example.** (i)  $V \times V^* \rightarrow \mathbb{F}$  with  $(v, \theta) \mapsto \theta(v)$

(ii)  $U = V = \mathbb{R}^n$  with  $\varphi(x, y) = \sum_{i=1}^n x_i y_i$  for  $x \in U, y \in V$

(iii)  $A \in M_{m,n}(\mathbb{F})$  with  $\varphi : \mathbb{F}^m \times \mathbb{F}^n \rightarrow \mathbb{F}$ ,  $(u, v) \mapsto u^T A v$

(iv) (infinite dim)  $U = V = C([0, 1], \mathbb{R})$  with  $\varphi(f, g) = \int_0^1 f(t)g(t) dt$  for  $f \in U, g \in V$

**Definition.**  $\mathcal{B} = \{e_1, \dots, e_m\}$  basis for  $U$ ,

$\mathcal{C} = \{f_1, \dots, f_n\}$  basis for  $V$

$\varphi : U \times V \rightarrow \mathbb{F}$  bilinear,

The matrix of  $\varphi$  wrt  $\mathcal{B}$  and  $\mathcal{C}$

$$[\varphi]_{\mathcal{B}, \mathcal{C}} = (\varphi(e_i, f_j))$$

$m \times n$ ,  $i, j$ th entry

**Lemma.**

$$\varphi(u, v) = [u]_{\mathcal{B}}^T [\varphi]_{\mathcal{B}, \mathcal{C}} [v]_{\mathcal{C}}$$

*Proof.* Say  $u = \sum \lambda_i e_i$ ,  $v = \sum \mu_j f_j$

$$\begin{aligned} \varphi(u, v) &= \varphi\left(\sum \lambda_i e_i, \sum \mu_j f_j\right) \\ &= \sum_i \lambda_i \varphi(e_i, \sum_j \mu_j f_j) \\ &= \sum_{i,j} \lambda_i \varphi(e_i, f_j) \mu_j \end{aligned}$$

□

Note:  $[\varphi]_{\mathcal{B}, \mathcal{C}}$  is the unique representation with this property

Note:  $\varphi : U \times V \rightarrow \mathbb{F}$  bilinear, determines linear maps

$$\varphi_L : U \rightarrow V^* \quad \text{and} \quad \varphi_R : V \rightarrow U^*$$

$$\varphi_L(u)(v) = \varphi(u, v) \quad \text{and} \quad \varphi_R(v)(u) = \varphi(u, v)$$

**Lemma.**  $\mathcal{B} = \{e_1, \dots, e_m\}$  basis for  $U$ ,

dual  $\mathcal{B}^* = \{\varepsilon_1, \dots, \varepsilon_m\}$  basis for  $U^*$ .

Similarly,  $\mathcal{C} = \{f_1, \dots, f_n\}$  for  $V$ ,  $\mathcal{C}^* = \{\eta_1, \dots, \eta_n\}$  for  $V^*$

If  $[\varphi]_{\mathcal{B}, \mathcal{C}} = A$ , then  $[\varphi_R]_{\mathcal{C}, \mathcal{B}^*} = A$ ,  $[\varphi_L]_{\mathcal{B}, \mathcal{C}^*} = A^T$

*Proof.*

$$\begin{aligned}\varphi_L(e_i)(f_j) &= A_{ij} \Rightarrow \varphi_L(e_i) = \sum_j A_{ij} \eta_j \\ \varphi_R(f_j)(e_i) &= A_{ij} \Rightarrow \varphi_R(f_j) = \sum_i A_{ij} \varepsilon_i\end{aligned}$$

□

**Definition.** Left kernel of  $\varphi = \ker \varphi_L$ , Right kernel of  $\varphi = \ker \varphi_R$

**Definition.**  $\varphi$  is *non-degenerate* if  $\ker \varphi_L = 0$  and  $\ker \varphi_R = 0$ . Otherwise  $\varphi$  is *degenerate*

**Lemma.** Let  $U, \mathcal{B}, V, \mathcal{C}$  as before,

$$\varphi : U \times V \rightarrow \mathbb{F}$$

$$A = [\varphi]_{\mathcal{B}, \mathcal{C}}$$

assume  $\dim U, \dim V$  finite. Then

$$\varphi \text{ non-degenerate} \iff A \text{ invertible}$$

*Proof.*

$$\begin{aligned}\varphi \text{ non-degenerate} &\iff \ker \varphi_L = \mathbf{0} \text{ and } \ker \varphi_R = \{\mathbf{0}\} \\ &\iff n(A^T) = 0 \text{ and } n(A) = 0 \\ &\iff r(A^T) = \dim V \text{ and } r(A) = \dim U \\ &\iff A \text{ invertible} \quad (\text{and necessarily } \dim U = \dim V)\end{aligned}$$

□

**Corollary.** If  $\varphi$  is non-degenerate and  $U$  and  $V$  are finite, then

$$\dim U = \dim V$$

**Corollary.** When  $U$  and  $V$  are finite dim, choosing a non-degenerate bilinear form  $\varphi : U \times V \rightarrow \mathbb{F}$  is equivalent to picking an isomorphism  $\varphi_L : U \rightarrow V^*$

**Definition.** For  $T \subset U$ ,  $T^\perp = \{v \in V \mid \varphi(t, v) = 0 \forall t \in T\} \leq V$   
For  $S \subset T$ ,  ${}^\perp S = \{u \in U \mid \varphi(u, s) = 0 \forall s \in S\} \leq U$   
(Generalisation of annihilators)

**Proposition.**  $U$  bases  $\mathcal{B}, \mathcal{B}'$ ,  $P = [\text{id}]_{\mathcal{B}', \mathcal{B}}$   
 $V$  bases  $\mathcal{C}, \mathcal{C}'$  with  $Q = [\text{id}]_{\mathcal{C}', \mathcal{C}}$   
Let  $\varphi : U \times V \rightarrow \mathbb{F}$  bilinear. Then

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

*Proof.*

$$\begin{aligned}
 \varphi_{u,v} &= [u]_{\mathcal{B}}^T [\varphi]_{\mathcal{B},\mathcal{C}} [v]_{\mathcal{C}} \\
 &= (P[u]_{\mathcal{B}'})^T [\varphi]_{\mathcal{B},\mathcal{C}} (Q[v]_{\mathcal{C}'}) \\
 &= [u]_{\mathcal{B}'}^T [\varphi]_{\mathcal{B},\mathcal{C}} [v]_{\mathcal{C}'}
 \end{aligned}$$

□

**Definition.** The rank of  $\varphi$ ,  $r(\varphi)$  is the rank of any matrix representing it (well-def) by prev thm.

Note:  $r(\varphi) = r(\varphi_L) = r(\varphi_R)$



## 5 Determinant and Trace

### 5.1 Trace

**Definition.** For  $A \in M_n(\mathbb{F})$  (this is  $M_{n,n}(\mathbb{F})$ ), then

$$\text{tr}(A) = \sum_i A_{ii}$$

is the *trace* of  $A$ . This is a linear map.

**Lemma.** For  $A, B \in M_n(\mathbb{F})$ ,

$$\text{tr}(AB) = \text{tr}(BA)$$

*Proof.*

$$\begin{aligned} \text{tr}(AB) &= \sum_i \sum_j a_{ij} b_{ji} \\ &= \sum_j \sum_i b_{ji} a_{ij} \\ &= \text{tr}(BA) \end{aligned}$$

□

**Lemma.** Similar (= conjugate) matrices have the same trace.

*Proof.*  $B = P^{-1}AP$ ,  $A, B \in M_n(\mathbb{F})$ .

$$\begin{aligned} \text{tr}(B) &= \text{tr}(P^{-1}AP) \\ &= \text{tr}(APP^{-1}) \\ &= \text{tr}A \end{aligned}$$

□

**Definition.** If  $\alpha : V \rightarrow V$  linear, define  $\text{tr } \alpha = \text{tr}[\alpha]_{\mathcal{B}, \mathcal{B}}$ . By the above, this is well defined.

**Lemma.** Let  $\alpha : V \rightarrow V$  linear, and  $\alpha^* : V^* \rightarrow V^*$  its dual. Then

$$\text{tr } \alpha = \text{tr } \alpha^*$$

*Proof.*

$$\begin{aligned} \text{tr } \alpha &= \text{tr}[\alpha]_{\mathcal{B}, \mathcal{B}} \\ &= \text{tr}[\alpha]_{\mathcal{B}, \mathcal{B}}^T \\ &= \text{tr}[\alpha^*]_{\mathcal{B}^*, \mathcal{B}^*} \\ &= \text{tr } \alpha^* \end{aligned}$$

□

## 5.2 Determinants

$S_n$  = group of permutations of  $\{1, \dots, n\}$

Define  $\varepsilon_n : S_n \rightarrow \{-1, 1\}$  as

$$\varepsilon(\sigma) = \begin{cases} +1 & \text{if } \sigma \text{ product of even no. of transposes} \\ -1 & \text{if } \sigma \text{ product of odd no. of transposes} \end{cases}$$

**Definition.** Let  $A \in M_n(\mathbb{F})$ ,  $A = (a_{ij})$ . Then

$$\det(A) = \sum_{\sigma \in S} \varepsilon(\sigma) a_{\sigma(1)1} a_{\sigma(2)2} \cdots a_{\sigma(n)n}$$

There are  $n!$  summands, each is sign  $\times$  product of  $n$  elements (one for each row and each column).

Eg  $n = 2$ ,

$$\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \underbrace{a_{11}a_{22}}_{\sigma=\text{id}} - \underbrace{a_{12}a_{21}}_{\sigma=(12)}$$

**Lemma.** If  $A = (a_{ij})$  is an upper triangular matrix (ie.  $a_{ij} = 0$  if  $i > j$ ) then  $\det A = a_{11}a_{22} \cdots a_{nn}$ . Similar for lower triangular matrices (ie.  $a_{ij} = 0$  if  $i < j$ ).

*Proof.*

$$\det(A) = \sum_{\sigma \in S_n} \varepsilon(\sigma) a_{\sigma(1)1} \cdots a_{\sigma(n)n}$$

For a summand to be non-zero, need  $\sigma(j) \leq j \forall j$ . Thus  $\sigma = \text{id}$  □

**Lemma.**

$$\det(A) = \det(A^T)$$

*Proof.*

$$\begin{aligned} \det(A) &= \sum_{\sigma \in S_n} \varepsilon(\sigma) \times \prod_{i=1}^n a_{\sigma(i)i} \\ &= \sum_{\sigma \in S_n} \underbrace{\varepsilon(\sigma)}_{=\varepsilon^{-1}} \prod_{i=1}^n a_{i\sigma^{-1}(i)} \\ &= \sum_{\tau \in S_n} \varepsilon(\tau) \prod_{i=1}^n a_{i\tau(i)} \quad (\sigma^{-1} = \tau) \\ &= \det(A^T) \end{aligned}$$

□

**Definition.** A *volume form* on  $\mathbb{F}^n$  is a function:

$$d : \underbrace{\mathbb{F}^n \times \mathbb{F}^n \times \cdots \times \mathbb{F}^n}_{n \text{ copies}} \rightarrow \mathbb{F}$$

such that

(i)  $d$  is *multilinear*: for any  $i$  and  $v_1, \dots, v_{i-1}, v_i, v_{i+1}, \dots, v_n \in \mathbb{F}^n$ ,

$$d(v_1, v_2, \dots, v_{i-1}, -, v_{i+1}, \dots, v_n) \in (\mathbb{F}^n)^*$$

(ii)  $d$  is *alternating*: if  $v_i = v_j$  for  $i \neq j$ , then  $d(v_1, \dots, v_n) = 0$

Note that the notation we will use will look like

$$A = (a_{ij}) = (A^{(1)} \mid A^{(2)} \mid \dots \mid A^{(n)})$$

If  $\{e_i\}$  is the standard basis for  $\mathbb{F}^n$  then

$$I = (e_1 \mid \dots \mid e_n)$$

**Lemma.**

$$\det : \underbrace{\mathbb{F}^n \times \dots \times \mathbb{F}^n}_{n \text{ copies}} \rightarrow \mathbb{F}$$

$$(A^{(1)}, \dots, A^{(n)}) \mapsto \det(A) \text{ is a volume form}$$

*Proof.* (i) Multilinear: For any fixed  $\sigma \in S_n$ ,  $\prod_{i=1}^n a_{\sigma(i)}$  contains exactly one term from each column, and so is multilinear. Now use the fact that the sum of multilinear functions is multilinear.

(ii) Alternating: Suppose  $A^{(k)} = A^{(J)}$ , for  $J \neq k$ . Let  $\tau = (kJ)$  transposition.  $a_{ij} = a_{i\tau(j)} \forall i, j \in \{1, \dots, n\}$ ,  $S_n = A_n \sqcup \tau A_n$ , where  $\sqcup$  is disjoint union.

$$\begin{aligned} \det(A) &= \sum_{\sigma \in A_n} \prod a_{i\sigma(i)} - \sum_{\sigma \in A_n} \prod a_{i\tau(\sigma(i))} \\ &= \sum_{\sigma \in A_n} \prod a_{i\sigma(i)} - \sum_{\sigma \in A_n} \prod a_{i\sigma(i)} \\ &= 0 \end{aligned}$$

□

**Lemma.** Let  $d$  be a volume form. Then swapping two entries changes the sign.

$$d(v_1, \dots, v_i, \dots, v_j, \dots, v_n) = -d(v_1, \dots, v_j, \dots, v_i, \dots, v_n)$$

*Proof.*

$$\begin{aligned} 0 &= d(v_1, \dots, v_{i-1}, v_i + v_j, v_{i+1}, \dots, v_{j-1}, v_j + v_i, v_{j+1}, \dots, v_n) \\ &= d(v_1, \dots, v_i, \dots, v_i, \dots, v_n) + d(v_1, \dots, v_j, \dots, v_i, \dots, v_n) \\ &\quad + d(v_1, \dots, v_i, \dots, v_j, \dots, v_n) + d(v_1, \dots, v_j, \dots, v_j, \dots, v_n) \\ &= d(v_1, \dots, v_j, \dots, v_i, \dots, v_n) + d(v_1, \dots, v_i, \dots, v_j, \dots, v_n) \end{aligned}$$

□

**Corollary.** If  $\sigma \in S_n$ ,  $d(v_{\sigma(1)}, \dots, v_{\sigma(n)}) = \varepsilon(\sigma)d(v_1, \dots, v_n)$

**Theorem.** Let  $d$  be a volume form on  $\mathbb{F}^n$ .  $A = (A^{(1)} \mid \cdots \mid A^{(n)})$ . Then

$$d(A^{(1)}, \dots, A^{(n)}) = \det(A) \times d(e_1, \dots, e_n)$$

*Proof.*

$$\begin{aligned} d(A_1, \dots, A^n) &= d\left(\sum_{i=1}^n a_{ij}e_i, A^{(2)}, \dots, A^{(n)}\right) \\ &= \sum_{i=1}^n a_{i1}d(e_i, A^{(2)}, \dots, A^{(n)}) \\ &= \sum_i \sum_j a_{i1}a_{j2}d(e_i, e_j, A^{(3)}, \dots, A^{(n)}) \\ &= \sum_{i_1, \dots, i_n} \prod_{k=1}^n a_{ik} \underbrace{d(e_{i_1}, e_{i_2}, \dots, e_{i_n})}_{\substack{0 \text{ unless all of } i_k \text{ are distinct}^9}} \\ &= \sum_{i_1, \dots, i_n} \prod_{k=1}^n A_{\sigma(k)k} \underbrace{d(e_{\sigma(1)}, \dots, e_{\sigma(n)})}_{= \varepsilon(\sigma)d(e_1, \dots, e_n)} \end{aligned}$$

□

**Corollary.**  $\det$  is the unique volume form s.t.

$$d(e_1, \dots, e_n) = 1$$

Recall:

$$\det : \underbrace{\mathbb{F}^n \times \cdots \times \mathbb{F}^n}_{n \text{ copies}} \rightarrow \mathbb{F}$$

$$(v_1, \dots, v_n) \mapsto \det(v_1 \mid \cdots \mid v_n) \text{ is a volume form}$$

**Proposition.** Let  $A, B \in M_n(\mathbb{F})$ . Then  $\det(AB) = \det(A)\det(B)$

*Proof.* Let  $d_A : \underbrace{\mathbb{F}^n \times \cdots \times \mathbb{F}^n}_{n \text{ copies}} \rightarrow \mathbb{F}$ ,

$$(v_1, \dots, v_n) \mapsto \det(Av_1 \mid \cdots \mid Av_n)$$

- $d_A$  is multilinear:  $v_i \mapsto Av_i$  linear, and  $d$  multilinear
- $d_A$  is alternating:  $v_i = v_j \Rightarrow Av_i = Av_j$  and  $d$  is alternating

Thus  $d_A$  is a volume form.

$$\begin{aligned} d_A(Be_1, \dots, Be_n) &= \det B d_A(e_1, \dots, e_n) \quad (d_A \text{ a v.f.}) \\ &= \det B \det A \end{aligned}$$

$$\text{Also } d_A(Be_1, \dots, Be_n) = \det(ABe_1 \mid \cdots \mid ABe_n) = \det(AB).$$

□

**Definition.**  $A \in M_n(\mathbb{F})$  is singular if  $\det A = 0$ . Otherwise  $A$  is non-singular.

**Lemma.** if  $A$  is invertible, then  $A$  is non-singular,  $\det(A^{-1}) = \frac{1}{\det A}$

*Proof.*

$$\begin{aligned} 1 &= \det(I_n) \\ &= \det(AA^{-1}) \\ &= \det(A) \det(A^{-1}) \\ &\Rightarrow \det(A) \neq 0, \det(A^{-1}) = (\det(A))^{-1} \end{aligned}$$

□

**Theorem.** Let  $A \in M_{m,n}(\mathbb{F})$ . TFAE:

- (i)  $A$  is invertible
- (ii)  $A$  is non-singular
- (iii)  $r(A) = n$

*Proof.* – (i)  $\Rightarrow$  (ii) done

- (ii)  $\Rightarrow$  (iii): Suppose that  $r(A) < n$ . By rank-nullity,  $n(A) > 0$ , so  $\exists \lambda \in \mathbb{F}^n \setminus \{\mathbf{0}\}$  st.  $A\lambda = \mathbf{0}$ . Say  $\lambda = (\lambda_i)$ , say  $\lambda_k \neq 0$ . Have  $\sum_{i=1}^n A^{(i)} \lambda_i = \mathbf{0}$ .

Let  $B = (e_1 \mid \cdots \mid e_{k-1} \mid \lambda \mid e_{k+1} \mid \cdots \mid e_n)$ .

$$\begin{aligned} AB \text{ has } k^{\text{th}} \text{ column zero} &\Rightarrow \det(AB) = 0 \\ &= \det(A) \det(B) \\ &= \det(A) \underbrace{\lambda_k}_{\neq 0} \end{aligned}$$

Thus  $\det A = 0$

- (iii)  $\Rightarrow$  (i) by rank-nullity

□

### 5.2.1 Determinants of Linear Maps

**Lemma.** Conjugate matrices have the same determinant.

*Proof.* Let  $B = P^{-1}AP$ . Then

$$\begin{aligned} \det B &= \det(P^{-1}AP) \\ &= \det(P^{-1}) \det(A) \det(P) \\ &= (\det P)^{-1} (\det A) (\det P) \\ &= \det A \end{aligned}$$

□

**Definition.** Let  $\alpha : V \rightarrow V$ ,  $V$  a finite-dim  $v$ -space. Define  $\det \alpha = \det[\alpha]_{\mathcal{B}, \mathcal{B}}$ , where  $\mathcal{B}$  is any basis for  $V$ . This is well-defined by the previous lemma.

**Theorem.**  $\det : L(V, V) \rightarrow \mathbb{F}$  satisfies:

- (i)  $\det(I_d) = 1$
- (ii)  $\det(\alpha \circ \beta) = \det(\alpha) \det(\beta)$
- (iii)  $\det(\alpha) \neq 0 \iff \alpha$  invertible, and if  $\alpha$  invertible then  $\det(\alpha^{-1}) = (\det \alpha)^{-1}$

### 5.2.2 Determinants of Block Triangular Matrices

**Lemma.**  $A \in M_k(\mathbb{F})$ ,  $B \in M_l(\mathbb{F})$ ,  $C \in M_{k,l}(\mathbb{F})$ .

$$\det \begin{pmatrix} A & C \\ 0 & B \end{pmatrix} = \det A \det B$$

*Proof.* set  $n = k + l$ . Let  $X = \begin{pmatrix} A & C \\ 0 & B \end{pmatrix} \in M_n(\mathbb{F})$ ,  $X = (x_{ij})$ .

$$\det(X) = \sum_{\sigma \in S_n} \varepsilon(\sigma) \prod_{i=1}^n x_{\sigma(i)i}$$

Note:  $x_{\sigma(i)i} = 0$  if  $i \leq k$  and  $\sigma(i) > k$ . Thus we're summing over all  $\sigma$  with

- (i) if  $j \in [1, k]$ ,  $\sigma(j) \in [1, k]$  AND
- (ii) if  $j \in [k+1, n]$ ,  $\sigma(j) \in [k+1, n]$

this means

- (i) get  $x_{\sigma(i)i} = \underbrace{x_{\sigma_1(i)i}}_{=a_{\sigma_1(i)i}}$  where  $\sigma_1 =$  restriction of  $\sigma$  to  $[1, k]$
- (ii) get  $x_{\sigma(i)i} = \underbrace{x_{\sigma_2(i)i}}_{=b_{\sigma_2(i)i}}$  where  $\sigma_2 =$  restriction of  $\sigma$  to  $[k+1, n]$ .

$$\sigma = \sigma_1 \sigma_2 \Rightarrow \varepsilon(\sigma) = \varepsilon(\sigma_1) \varepsilon(\sigma_2)$$

We get

$$\begin{aligned} \det X &= \left( \sum_{\sigma_1 \in S_k} \varepsilon(\sigma_1) \prod_{j=1}^k a_{\sigma_1(j)j} \right) \left( \sum_{\sigma_2 \in S_l} \varepsilon(\sigma_2) \prod_{j=k+1}^n b_{\sigma_2(j)j} \right) \\ &= \det A \det B \end{aligned}$$

□

**Corollary.** For square matrices  $A_1, \alpha, A_k$ , the upper-triangular matrix with  $A_1, \alpha, A_k$  along the diagonal has determinant  $= \prod_{i=1}^k \det A_i$ .

*Proof.* Apply lemma immediately. □

Caution: In general,

$$\det \begin{pmatrix} A & B \\ C & D \end{pmatrix} \neq \det A \det D - \det B \det C$$

Aside: Volume Interpretation of Determinants:

$\mathbb{R}^2$   $\det(\mathbf{u}|\mathbf{v})$  is the signed area of parallelogram made by extending  $\mathbf{u}$  and  $\mathbf{v}$ .

$\mathbb{R}^3$   $\det(\mathbf{u}|\mathbf{v}|\mathbf{w})$  = signed volume of parallelepiped.

There are analogous interpretations in higher dimensions.

### 5.2.3 Elementary Operations and Det

- (i)  $E_1$  swaps 2 columns/rows.  $\det E_1 = -1$
- (ii)  $E_2$  multiplies a column/row by  $\lambda \neq 0$ .  $\det E_2 = \lambda$
- (iii)  $E_3$  add  $\lambda$ (column  $i$ ) to column  $j$  (/rows).  $\det E_3 = 1$

One could prove properties of  $\det$  (eg  $\det(AB) = \det A \det B$ ) by using the factorisation of matrices into products of  $E_i$ .

### 5.2.4 Column Expansion and Adjugate Matrices

**Lemma.** Let  $A \in M_n(\mathbb{F})$ ,  $A = (a_{ij})$ . Define  $A_{\hat{i}\hat{j}}$  by deleting row  $i$  and col  $j$  from  $A$ . Then

- (i) for a fixed  $j$ ,

$$\det A = \sum_{i=1}^n (-1)^{i+j} a_{ij} \det A_{\hat{i}\hat{j}}$$

‘expansion in column  $j$ ’

- (ii) for a fixed  $i$ ,

$$\det A = \sum_{j=1}^n (-1)^{i+j} a_{ij} \det A_{\hat{i}\hat{j}}$$

‘expansion in row  $i$ ’

Remark: could use 1) to define determinants iteratively, starting with  $\det a = a$  for  $n = 1$ .

**Example.**

$$\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = a \begin{vmatrix} e & f \\ h & i \end{vmatrix} - b \begin{vmatrix} d & f \\ g & i \end{vmatrix} + c \begin{vmatrix} d & e \\ g & h \end{vmatrix}$$

*Proof.* We will only prove (i), and get (ii) by transposition

$$\begin{aligned}
 \det(A) &= \det(A^{(1)} \mid A^{(2)} \mid \cdots \mid \sum_{i=1}^n a_{ij} e_i \mid \cdots \mid A^{(n)}) \\
 &= \sum_{i=1}^n a_{ij} \det(A^{(1)} \mid \cdots \mid e_i \mid A^{(ji)} \mid \cdots \mid A^{(n)}) \\
 &= \sum_{i=1}^n \underbrace{a_{ij} (-1)^{(i-1)+(j-1)}}_{\text{row and col swaps}} \det \begin{pmatrix} 1 & & 0 \\ 0 & & A_{\hat{i}\hat{j}} \\ & & \end{pmatrix} = \sum_{i=1}^n a_{ij} (-1)^{i+j} \det(A_{\hat{i}\hat{j}})
 \end{aligned}$$

□

**Definition.** Let  $A \in M_n(\mathbb{F})$ . The *adjugate matrix* of  $A$ ,  $\text{adj}(A)$ , is the  $n \times n$  matrix

$$(\text{adj } A)_{ij} = (-1)^{i+j} \det(A_{\hat{i}\hat{j}})$$

**Theorem.** (i)

$$(\text{adj } A)A = (\det A)I = \begin{pmatrix} \det A & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \det A \end{pmatrix}$$

(ii) If  $A$  is invertible, then  $A^{-1} = \frac{1}{\det A} \text{adj}(A)$ .

*Proof.* (i)  $\det A = \sum_i (\text{adj } A)_{ji} a_{ij} = j^{\text{th}}, j^{\text{th}}$  entry of  $(\text{adj } A)A$

For  $j \neq k$ ,

$$\begin{aligned}
 0 &= \det(A^{(1)} \mid \cdots \mid \underbrace{A^k}_{j^{\text{th}} \text{ col}} \mid \cdots \mid A^k \mid \cdots \mid A^{(n)}) \\
 &= \sum_i (\text{adj } A)_{ji} a_{ik} \\
 &= j, k^{\text{th}} \text{ entry of } (\text{adj } A)A
 \end{aligned}$$

(ii) If  $A$  invertible, then  $\det A \neq 0$ , so  $I = \frac{\text{adj}(A)}{\det A} A$

□

### 5.3 Systems of Linear Equations

- $A\mathbf{x} = \mathbf{b}$  is  $m$  equations in  $n$  unknowns ( $A : m \times n$  and  $\mathbf{b} : m \times 1$  known,  $\mathbf{x} = (x_1, \dots, x_n) = n \times 1$  unknown)
- $A\mathbf{x} = \mathbf{b}$  has solution iff  $r(A) = r(A|b)$  where  $A|b$  is the augmented matrix:  $A$  with extra column  $b$  (ie. iff  $\mathbf{b}$  is a linear combo of columns in  $A$ ).
- The solution is unique iff  $r(A) = n$
- Special case:  $m = n$ . If  $A$  is non-singular then there is a unique solution  $\mathbf{x} = A^{-1}\mathbf{b}$ .



### 5.3.1 The Cramer Rule

If  $A \in M_n(\mathbb{F})$  invertible, the system  $A\mathbf{x} = \mathbf{b}$  has unique solution  $\mathbf{x} = (x_i)$ ,

$$x_i = \frac{\det(A_{\hat{i}\hat{b}})}{\det A}$$

where  $A_{\hat{i}\hat{b}}$  is obtained from  $A$  by deleting  $i^{\text{th}}$  column and replacing it with  $\mathbf{b}$ .

*Proof.* Assume that  $\mathbf{x}$  is a solution of  $A\mathbf{x} = \mathbf{b}$ .

$$\begin{aligned} \det(A_{\hat{i}\hat{b}}) &= \det(A^{(1)} \mid \cdots \mid A^{(i-1)} \mid \mathbf{b} \mid A^{(i+1)} \mid \cdots \mid A^{(n)}) \\ &= \det(A^{(1)} \mid \cdots \mid A\mathbf{x} \mid \cdots \mid A^{(n)}) \\ &= \sum_{j=1}^n x_j \det(A^{(1)} \mid \cdots \mid A^{(i-1)} \mid A^{(j)} \mid A^{(i+1)} \mid \cdots \mid A^{(n)}) \text{ as } A\mathbf{x} = \sum_j A^{(j)} x_j \\ &= x_i \det A \end{aligned}$$

□

**Corollary.** If  $A \in M_n(\mathbb{Z})$  ie.  $(n \times n)$  with integer entries, with  $\det A = \pm 1$ , then

–  $A^{-1} \in M_n(\mathbb{Z})$  also,

$$A^{-1} = \frac{\text{adj } A}{\pm 1} \quad \text{with adj } A \text{ entries in } \mathbb{Z}$$

–  $\mathbf{b} \in \mathbb{Z}^n$ , can solve  $A\mathbf{x} = \mathbf{b}$  for integer solution.

## 6 Endomorphisms

Let  $V$  be a vector space over  $\mathbb{F}$ ,  $\dim V = n < \infty$ , and  $\alpha \in L(V) = L(V, V)$ , with  $\mathcal{B} = \{v_1, \dots, v_n\}$  basis.

Problem: Choose  $\mathcal{B}$  st.  $[\alpha]_{\mathcal{B}} (= [\alpha]_{\mathcal{B}, \mathcal{B}})$  has "nice form".  $\mathcal{B}'$  other basis,  $P$  change of basis matrix  $[\alpha]_{\mathcal{B}} = P^{-1}[\alpha]_{\mathcal{B}'}P$ .

Problem:  $A \in M_n(\mathbb{F})$ , want  $A'$  conjugate to it which has a nice form.

**Definition.**  $\alpha \in L(V)$  is *diagonalisable* if there exists  $\mathcal{B}$  st.

$$[\alpha]_{\mathcal{B}} \text{ is diagonal}$$

A weaker possibility is

**Definition.**  $\alpha \in L(V)$  is *triangularisable* if  $\exists \mathcal{B}$  st.  $[\alpha]_{\mathcal{B}}$  is upper triangular

( $A \in M_n(\mathbb{F})$  is diagonalisable if its conjugate to a diagonal matrix, similarly for triangular. )

**Definition.** (i)  $\lambda \in \mathbb{F}$  is an *eigenvalue* of  $\alpha$  if there exists some  $v \in V \setminus \{\mathbf{0}\}$  st.  $\alpha(v) = \lambda v$

(ii)  $v \in V$  is an *eigenvector* for  $\alpha$  if  $\alpha(v) = \lambda v$  for some eigenvalue  $\lambda$

(iii)  $V_{\lambda} = \{v \in V \mid \alpha(v) = \lambda v\}$   $\lambda$ -eigenspace of  $\alpha$ , Note  $V_{\lambda} \leq V$

Shorthand: evector, eval, espace.

Remark:

(i)  $\lambda$  eval,  $\iff \alpha - \lambda \iota$  singular  $\iff \det(\alpha - \lambda \iota) = 0$ .

$$V_{\lambda} = \ker(\alpha - \lambda \iota)$$

Note  $\iota$  is the identity map.

(ii) If  $\alpha(v_j) = \lambda v_j$ , then  $j^{\text{th}}$  col of  $[\alpha]_{\mathcal{B}}$  is

$$\begin{pmatrix} 0 \\ \vdots \\ \lambda \\ \vdots \\ 0 \end{pmatrix}$$

( $j^{\text{th}}$  entry)

(iii)  $[\alpha]_{\mathcal{B}}$  diagonal  $\iff \mathcal{B}$  consists of evecs.

$[\alpha]_{\mathcal{B}}$  upper triangular  $\iff \alpha(v_j) \in \langle v_1, \dots, v_j \rangle$  for all  $j$ . In particular,  $v_1$  is an eigenvector.

### 6.1 Aside on Polynomials

$$F[t]\{\text{polys w/ coefficients in } \mathbb{F}\}$$

- $\deg(f + g) \leq \max(\deg f, \deg g)$
- $\deg 0 = -\infty$
- $\deg(fg) = \deg f + \deg g$
- If  $\lambda \in \mathbb{F}$  is a root of  $f \in F[t]$  (ie.  $f(\lambda) = 0$ ), then  $(t - \lambda)$  divides  $f$ :

$$f(t) = (t - \lambda)g(t), \quad \text{some } g(t) \in F[t]$$

- We say  $\lambda$  is a root of  $f \in F[t]$  with multiplicity  $e (\in \mathbb{N})$  if  $(t - \lambda)^e$  divides  $f$ , but  $(t - \lambda)^{e+1}$  does not.
- A poly of degree  $n$  has at most  $n$  roots, counted with multiplicity

**Theorem.** Fundamental Theorem of Algebra Any  $f \in \mathbb{C}[t]$  of positive degree has a root (hence  $\deg f$  roots.)

**Definition.** The characteristic polynomial of  $\alpha : \chi_\alpha(t) = \det(\alpha - tI)$ . ( $\alpha \in L(V), A \in M_n(\mathbb{F})$ ).

Conjugate matrices have same characteristic poly.

**Theorem.**  $\alpha$  triangulable iff  $\chi_\alpha(t)$  can be written as a product of linear factors over  $\mathbb{F}$ .

In particular, if  $\mathbb{F} = \mathbb{C}$ , every matrix is triangulable.

*Proof.* ( $\Rightarrow$ ) Suppose  $\alpha$  is triangulable, and represented by

$$\begin{pmatrix} a_1 & & * \\ & \ddots & \\ 0 & & a_n \end{pmatrix}$$

wrt. some basis.

$$\text{Then } \chi_\alpha(t) = \det \begin{pmatrix} a_1 - t & & * \\ & \ddots & \\ 0 & & a_n - t \end{pmatrix} = \prod_{i=1}^n (a_i - t)$$

( $\Leftarrow$ ) Induction on  $n = \dim V$

- $n = 0$  or  $1$  Done.
- Suppose  $n > 1$ , and the thm holds for all endomorphisms of spaces of smaller dimension.

By hypothesis,  $\chi_\alpha(t)$  has a root in  $\mathbb{F}$ , say  $\lambda$ . Let  $U := V_\lambda (\neq \{0\})$

$$\alpha(U) \subseteq U \Rightarrow \alpha \text{ induces } \bar{\alpha} : V/U \rightarrow V/U$$

Pick basis  $v_1 | \cdots | v_k$  for  $U$ , extend it to basis

$$\mathcal{B} = \{v_1, \dots, v_n\} \text{ for } V$$

wrt  $\mathcal{B}$ ,  $\alpha$  is represented by:

$$\begin{pmatrix} \lambda I_k & * \\ 0 & C \end{pmatrix}$$

where  $\lambda I_k$  is the matrix of  $\alpha$  restricted to  $U$ ,  $\alpha|_U$ , and  $C$  represents  $\bar{\alpha}$  wrt  $v_{k+1} + U, \alpha, v_n + U$ .

$$\begin{aligned} \chi_\alpha(t) &= \det(\alpha - tI) \\ &= (\lambda - t)^k \chi_{\bar{\alpha}}(t) \end{aligned}$$

Thus  $\chi_\alpha$  is *also* a product of linear factors. By induction hypothesis, (since  $\bar{\alpha}$  is acting on a lower dimensional vector space) there is a basis for  $V/U$ , say  $w_{k+1} + U, \dots, w_n + U$  wrt. which  $\bar{\alpha}$  is represented by an upper-triangular matrix, say  $T$ .

wrt  $v_1, \dots, v_k, w_{k+1}, \dots, w_n$ ,  $\alpha$  is represented by

$$\begin{pmatrix} \lambda I_k & * \\ 0 & T \end{pmatrix}$$

□

**Example.**  $\mathbb{F} = \mathbb{R}$ ,  $V = \mathbb{R}^2$ ,  $\alpha$  rotation

$$\begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix}$$

$\chi_\alpha(t) = t^2 - 2 \cos \theta t + 1$  NOT triangulable over  $\mathbb{R}$   
(Conjugate to a diagonal matrix over  $\mathbb{C}$ ).

**Lemma.** Let  $V$  be  $n$ -dim over  $\mathbb{F}$ ,  $\alpha \in L(V)$ .

$$\chi_\alpha(t) = (-1)^n t^n + c_{n-1} t^{n-1} + \dots + c_0$$

Then:

- $c_0 = \det \alpha$
- for  $\mathbb{F}$  in  $\mathbb{R}$  or  $\mathbb{C}$ ,  $c_{n-1} = (-1)^{n-1} \text{tr } \alpha$

*Proof.* -  $c_0 = \chi_\alpha(0) = \det(\alpha - 0) = \det \alpha$

- For  $\mathbb{F} = \mathbb{R}$ ,  $[\alpha]_{\mathcal{B}}$  can be thought of as a matrix over  $\mathbb{C}$  that happens to have real coeffs.

$$\chi_\alpha(t) = \det \begin{pmatrix} a_0 - t & & * \\ & \ddots & \\ 0 & & a_n - t \end{pmatrix} = \prod_{i=1}^n (a_i - t)$$

$$\sum_{i=1}^n a_i = \text{tr } \alpha$$

□