IB Linear Algebra

Ishan Nath, Michaelmas 2022

Based on Lectures by Prof. Pierre Raphael

November 14, 2022

Contents

1 Vector Spaces and Subspaces

Let F be an arbitrary field.

Definition 1.1 (F vector space)**.** A F vector space is an abelian group $(V, +)$ equipped with a function

$$\begin{aligned} F \times V &\to V \\ (\lambda, v) &\mapsto \lambda v \end{aligned}$$

such that

- $\lambda(v_1 + v_2) = \lambda v_1 + \lambda v_2$,
- $(\lambda_1 + \lambda_2)v = \lambda_1 v + \lambda_2 v$,
- $\lambda(\mu v) = (\lambda\mu)v$,
- $1 \cdot v = v$.

We know how to

- Sum two vectors
- Multiply a vector $v \in V$ by a scalar $\lambda \in F$.

Example 1.1.

(i) Take $n \in \mathbb{N}$, then F^n is the set of column vectors of length n with elements in F. We have

$$v \in F^n, v = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, x_i \in F,$$

$$v + w = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} + \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} = \begin{pmatrix} v_1 + w_1 \\ \vdots \\ v_n + w_n \end{pmatrix},$$

$$\lambda v = \begin{pmatrix} \lambda v_1 \\ \vdots \\ \lambda v_n \end{pmatrix}.$$

Then F^n is a F vector space.

(ii) For any set X, take

$$\mathbb{R}^X = \{f : X \to \mathbb{R}\}.$$

Then $\mathbb{R}^X$ is an $\mathbb{R}$ vector space.

(iii) Take $M_{n,m}(F)$, the set of $n \times m$ F valued matrices. Then $M_{n,m}(F)$ is a F vector space.

Remark. The axiom of scalar multiplication implies that for all $v \in V$, $0 \cdot v = \mathbf{0}$.

Definition 1.2 (Subspace)**.** Let V be a vector space over F. A subset U of V is a vector subspace of V (denoted $U \leq V$) if

- $0 \in U$,
- $(u_1, u_2) \in U \times U$ implies $u_1 + u_2 \in U$,
- $(\lambda, u) \in F \times U$ implies $\lambda u \in U$.

Note if V is an F vector space, and $U \leq V$, then U is an F vector space.

Example 1.2.

(i) Take $V = \mathbb{R}^{\mathbb{R}}$, the space of functions $f : \mathbb{R} \to \mathbb{R}$. Let $\mathcal{C}(\mathbb{R})$ be the space of continuous function $f : \mathbb{R} \to \mathbb{R}$. Then $\mathcal{C}(\mathbb{R}) \leq \mathbb{R}^{\mathbb{R}}$.

(ii) Take the elements of $\mathbb{R}^3$ which sum up to t. This is a subspace if and only if $t = 0$.

Note that the union of two subspaces is generally not a subspace, as it is usually not closed under addition.

Proposition 1.1. *Let V be an F vector space, and $U, W \leq V$. Then $U \cap W \leq V$.*

Proof: Since $0 \in U, 0 \in W$, $0 \in U \cap W$. Now consider $(\lambda, \mu) \in F^2$, and $(v_1, v_2) \in (U \cap W)^2$. Take $\lambda_1 v_1 + \lambda_2 v_2$. Since $u_1, v_1 \in U$, this is in U. Similarly, it is in W. So it is in $U \cap W$, and $U \cap W \leq V$.

Definition 1.3 (Sum of subspaces)**.** Let V be an F vector space. Let $U, W \leq V$. Then the **sum** of U and W is the set

$$U + W = \{u + w \mid (u, w) \in U \times W\}.$$

Proof: Note $0 = 0 + 0 \in U + W$. Take $\lambda_1 f + \lambda_2 g$, where $f, g \in U + W$. Then we can write $f = f_1 + f_2, g = g_1 + g_2$, where $f_1, g_1 \in U$, $f_2, g_2 \in W$.

Then

$$\lambda_1 f + \lambda_2 g = \lambda_1(f_1 + f_2) + \lambda_2(g_1 + g_2) = (\lambda_1 f_1 + \lambda_2 g_1) + (\lambda_1 f_2 + \lambda_2 g_2) \in U + W.$$

Remark. $U + W$ is the smallest subspace of V which contains both U and W.

1.1 Subspaces and Quotients

Definition 1.4 (Quotient)**.** Let V be an F vector space. Let $U \leq V$. The quotient space V/U is the abelian group V/U equipped with the scalar product multiplication

$$\begin{aligned} F \times V/U &\to V/U \\ (\lambda, v + U) &\mapsto \lambda v + U \end{aligned}$$

Proposition 1.2. *V/U is an F vector space.*

2 Spans, Linear Independence and the Steinitz Exchange Lemma

Definition 2.1 (Span of a family of vectors)**.** Let V be a F vector space. Let $S \subset B$ be a subset. We define

$$\begin{aligned}\langle S\rangle &= \{\text{finite linear combinations of elements of } S\}\\ &= \left\{\sum_{\delta\in J}\lambda_\delta v_\delta, v_\delta \in S, \lambda_\delta \in F, J \text{ finite}\right\}.\end{aligned}$$

By convention, we let $\langle\emptyset\rangle = \{0\}$.

Remark. $\langle S'\rangle$ is the smallest vector subspace which contains S.

Example 2.1.

Take $V = \mathbb{R}^3$, and

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

Then we have

$$\langle S'\rangle = \left\{\begin{pmatrix}a\\b\\2b\end{pmatrix}, (a,b) \in \mathbb{R}^2\right\}.$$

Take $V = \mathbb{R}^n$, and let e_i be the i'th basis vector. Then $V = \langle e_1, \dots, e_n\rangle$.

Take X a set, and $V = \mathbb{R}^X$. Let $S_x : X \to \mathbb{R}$, such that $y \mapsto 1$ if $x = y$, otherwise $y \mapsto 0$. Then

$$\langle (S_x)_{x\in X}\rangle = \{f \in \mathbb{R}^X \mid f \text{ has finite support}\}.$$

Definition 2.2. Let V be a F vector space. Let S' be a subset of V. We may say that S **spans** V if $\langle S\rangle = V$.

Definition 2.3 (Finite dimension)**.** Let V be a F vector space. We say that V is **finite dimensional** if it is spanned by a finite set.

Example 2.2.

Consider $P[x]$, the polynomials over $\mathbb{R}$, and $P_n[x]$, the polynomials over $\mathbb{R}$ with degree $\leq n$. Then since

$$\langle 1, x, \ldots, x^n \rangle = P_n[x],$$

$P_n[x]$ is finite dimensional, however $P[x]$ is not.

Definition 2.4 (Independence)**.** We say that $(v_1, \ldots, v_n)$, elements of V are **linearly independent** if

$$\sum_{i=1}^{n} \lambda_i v_i = 0 \implies \lambda_i = 0 \,\forall i.$$

Remark.

1. We also say that the family $(v_1, \ldots, v_n)$ is **free**.
2. Equivalently, $(v_1, \ldots, v_n)$ are not linearly independent if one of these vectors is a linear combination of the remaining $(n-1)$.
3. If (v_i) is free, then $v_i = 0$ for all i.

Definition 2.5 (Basis)**.** A subset S of V is a **basis** of V if and only if

(i) $\langle S' \rangle = V$,

(ii) S is linearly independent.

Remark. A subset S that generates V is a generating family, so a basis S is a free generating family.

Example 2.3.

For $V = \mathbb{R}^n$, then (e_i) is a basis of V.

If $V = \mathbb{C}$, then for $F = \mathbb{C}$, $\{1\}$ is a basis.

If $V = P[x]$, then $S = \{x^n, n \geq 0\}$ is a basis for V.

Lemma 2.1. *V is a F vector space. Then $(v_1, \ldots, v_n)$ is a basis of V if and only if any vector $v \in V$ has a unique decomposition*

$$v = \sum_{i=1}^{n} \lambda_i v_i.$$

Remark. We call $(\lambda_1, \dots, \lambda_n)$ the coordinates of v in the basis $(v_1, \dots, v_n)$.

Proof: Since $\langle v_1, \dots, v_n \rangle = V$, we must have

$$v = \sum_{i=1}^{n} \lambda_i v_i$$

for some λ_i. Now assume

$$\begin{aligned} v = \sum_{i=1}^{n} \lambda_i v_i &= \sum_{i=1}^{n} \lambda_i' v_i, \\ \implies \sum_{i=1}^{n} (\lambda_i - \lambda_i') v_i &= 0. \end{aligned}$$

Since v_i are free, $\lambda_i = \lambda_i'$.

Lemma 2.2. *If $(v_1, \dots, v_n)$ spans V, then some subset of this family is a basis of V.*

Proof: If $(v_1, \dots, v_n)$ are linearly independent, we are done. Otherwise assume they are not independent, then by possibly reordering the vectors, we have

$$v_n \in \langle v_1, \dots, v_{n-1} \rangle.$$

Then we have $V = \langle v_1, \dots, v_n \rangle = \langle v_1, \dots, v_{n-1} \rangle$. By iterating, we must eventually get to an independent set.

Theorem 2.1 (Steinitz Exchange Lemma)**.** *Let V be a finite dimensional vector space over F. Take*

(i) $(v_1, \dots, v_m)$ *free,*

(ii) $(w_1, \dots, w_n)$ *generating.*

Then $m \le n$, and up to reordering, $(v_1, \dots, v_m, w_{m+1}, \dots, w_n)$ spans V.

Proof: We use induction. Suppose that we have replaced l of the w_i, reordering if necessary, so

$$\langle v_1, \dots, v_l, w_{l+1}, \dots, w_n \rangle = V.$$

If $m = l$, we are done. Otherwise, $l < m$. Then since these vectors span V, we have

$$v_{l+1} = \sum_{i \leq l} a_i v_i + \sum_{i > l} \beta_i w_i.$$

Since $(v_1, \ldots, v_{l+1})$ is free, some of the β_i are non-zero. Upon reordering, we may let $\beta_{l+1} \neq 0$. Then,

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

Hence,

$$\begin{aligned} V &= \langle v_1, \ldots, v_l, w_{l+1}, \ldots, w_n \rangle = \langle v_1, \ldots, v_l, v_{l+1}, w_{l+1}, \ldots, w_n \rangle \\ &= \langle v_1, \ldots, v_{l+1}, w_{l+2}, \ldots, w_n \rangle. \end{aligned}$$

Iterating this process, we eventually get $l = m$, which then proves $m \leq n$.

3 Basis, Dimension and Direct Sums

Corollary 3.1. *Let V be a finite dimensional vector space over F. Then any two bases of V have the same number of vectors, called the* ***dimension*** *of V.*

Proof: Take $(v_1, \dots, v_n), (w_1, \dots, w_m)$ bases of V.

(i) As (v_i) is free and (w_i) is generating, $n \le m$.

(ii) As (w_i) is free and (v_i) is generating, $m \le n$.

So $m = n$.

Corollary 3.2. *Let V be a vector space over F with dimension $n \in \mathbb{N}$.*

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

(ii) *Any spanning set of vectors has at least n elements, with equality if and only if it is a basis.*

Proof: Exercise (fill this in).

Proposition 3.1. *Let U, W be finite dimensional subspaces of 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 $(v_1, \dots, v_l)$ a basis of $U \cap W$. Since $U \cap W \le U$, we can extend this to a basis $(v_1, \dots, v_l, u_1, \dots, u_m)$ of U, and a basis $(v_1, \dots, v_l, w_1, \dots, w_n)$ of W. Then we show $(v_1, \dots, v_l, u_1, \dots, u_m, w_1, \dots, w_n)$ is a basis of $U + W$.

It is clearly a generating family, so we will show it is free. Suppose

$$\sum_{i=1}^{l} \alpha_i v_i + \sum_{=1}^{m} \beta_i u_i + \sum_{i=1}^{n} \gamma_i w_i = 0.$$

Then we get

$$\sum_{i=1}^{n} \gamma_i w_i \in U \cap W,$$

implying that

$$\sum_{i=1}^{l} s_i v_i = \sum_{i=1}^{} \gamma_i w_i.$$

But since $(v_1, \dots, w_n)$ is a basis of W, we get $\gamma_i = 0$. Similarly, $\beta_i = 0$. Thus,

$$\sum_{i=1}^{l} \alpha_i v_i = 0.$$

Since (v_i) is a basis of $U \cap W$, $\alpha_i = 0$.

Proposition 3.2. *Let V be a finite dimensional vector space over F. Let $U \leq V$. Then U and V/U are both finite dimensional and*

$$\dim V = \dim U + \dim(V/U).$$

Proof: Let $(u_1, u_2, \dots, u_l)$ be a basis of U. As $U \leq V$, we can extend this to a basis $(u_1, \dots, u_l, w_{l+1}, \dots, w_n)$ of V. Then we show that $(w_{l+1} + U, \dots, w_n + U)$ is a basis of V/U. (Fill this in).

Remark. If $U \leq V$, then we say U is proper if $U \neq V$. Then for finite dimensions, U proper implies $\dim U < \dim V$, as $\dim(V/U) > 0$.

Definition 3.1 (Direct sum)**.** Let V be a vector space over F, and $U, W \leq V$. We say $V = U \oplus W$ if and only if any element of $v \in V$ can be uniquely decomposed as $v = u + w$ for $u \in U, w \in W$.

Remark. If $V = U \oplus W$, we say that W is a complement of U in V. There is no uniqueness of such a complement.

In the sequel, we use the following notation. Let $\mathcal{B}_1 = \{u_1, \dots, u_l\}$ and $\mathcal{B}_2 = \{w_1, \dots, w_m\}$ be collections of vectors. Then

$$\mathcal{B}_1 \cup \mathcal{B}_2 = \{u_1, \dots, u_l, w_1, \dots, w_m\}$$

with the convention that $\{v\} \cup \{v\} = \{v, v\}$.

Lemma 3.1. *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 $\mathcal{B}_1$ of U, $\mathcal{B}_2$ of W, the union $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2$ is a basis of V.*

Proof: We show (ii) implies (i). Let $V = U + W$, then clearly U, W generate

V. We only need to show uniqueness. Suppose $u_1 + w_1 = u_2 + w_2$. Then

$$u_1 - u_2 = w_2 - w_1 \in U \cap W = \{0\}.$$

Hence $u_1 = u_2$ and $w_1 = w_2$, as required.

Now we show (i) implies (iii). Let $\mathcal{B}_1$ be a basis of U, and $\mathcal{B}_2$ a basis of W. Then $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2$ generates $U + W = V$, and $\mathcal{B}$ is free, as if $\sum \lambda_i v_i = u + w = 0$, then $0 = 0 + 0$ uniquely, so $u = 0, w = 0$, giving $\lambda_i = 0$ for all i.

Finally, we show (iii) implies (ii). Let $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2$. Then since $\mathcal{B}$ is a basis of V,

$$v = \sum_{u_i \in \mathcal{B}_1} \lambda_i u_i + \sum_{w_i \in \mathcal{B}_2} \lambda_i w_i = u + w.$$

Now if $v \in U \cap W$,

$$v = \sum_{u \in \mathcal{B}_1} \lambda_u u = \sum_{w \in \mathcal{B}_2} \lambda_w w.$$

This gives

$$\sum_{u \in \mathcal{B}_1} \lambda_u u - \sum_{w \in \mathcal{B}_2} \lambda_w w = 0.$$

Since $\mathcal{B}_1 \cup \mathcal{B}_2$ is free, we get $\lambda_u = \lambda_w = 0$, so $U \cap W = \{0\}$.

Definition 3.2. Let V be a vector space over F, and $V_1, \ldots, V_l \leq V$. Then

(i) The sum of the subspaces is

$$\sum_{i=1}^{l} V_i = \{v_1 + \cdots + v_l \mid v_j \in V_J, 1 \leq j \leq l\}.$$

(ii) The sum is direct:

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

if and only if

$$v_1 + \cdots + v_l = v_1' + \cdots + v_l' \implies v_1 = v_1', \ldots, v_l = v_l'.$$

Proof: Exercise.

Proposition 3.3. *The following are equivalent:*

(i)
$$\sum_{i=1}^{l} V_i = \bigoplus_{i=1}^{l} V_i,$$

(ii)
$$\forall i, V_i \cap \left(\sum_{j<i} V_i \right) = \{0\},$$

(iii) *For any basis* $\mathcal{B}_i$ *of* V_i,
$$\mathcal{B} = \bigcup_{i=1}^{l} \mathcal{B}_i \textit{ is a basis of } \sum_{i=1}^{l} V_l.$$

4 Linear maps, Isomorphisms and the Rank-Nullity Theorem

Definition 4.1 (Linear map)**.** Let V, W be vector spaces over F. A map $\alpha : V \to W$ is **linear** if and only if for all $\lambda_1, \lambda_2 \in F$ 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 4.1.

(i) Take an $m \times n$ matrix M, Then we can take the linear map $\alpha : \mathbb{R}^m \to \mathbb{R}^n$ defined by $X \mapsto MX$.

(ii) Take the linear map $\alpha : \mathcal{C}[0,1] \to \mathcal{C}[0,1]$ by

$$f \mapsto \alpha(f)(x) = \int_0^x f(t)\,\mathrm{d}t.$$

(iii) Fix $x \in [a,b]$. Then we can take a linear map $\mathcal{C}[a,b] \to \mathbb{R}$ by $f \mapsto f(x)$.

Remark. Let U, V, W be F-vector spaces.

(i) The identity map $\mathrm{id}_V : V \to V$ by $x \mapsto x$ is a linear map.

(ii) If $U \to V$ is β linear, and $V \to W$ is α linear, then $U \to W$ is linear by $\alpha \circ \beta$.

Lemma 4.1. *Let V, W be F-vector spaces, and $\mathcal{B}$ a basis of V. Let $\alpha_0 : \mathcal{B} \to W$ be any map, then there is a unique linear map $\alpha : V \to W$ extending α_0.*

Proof: For $v \in V$, we can write

$$v = \sum_{i=1}^n \lambda_i v_i,$$

where $\mathcal{B} = (v_1, \dots, v_n)$. Then by linearity, we must have

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

This is unique as $\mathcal{B}$ is a basis.

Remark. This is true in infinite dimensions as well.

Often, to define a linear map, we define its value on a basis and extend by linearity. As a corollary, if $\alpha_1, \alpha_2 : V \to W$ are linear and agree on a basis of V, they are equal.

Definition 4.2 (Isomorphism)**.** Let V, W be vector spaces over F. A map $\alpha : V \to W$ is called an **isomorphism** if and only if α is linear and bijective. If such an α exists, we say $V \cong W$.

Remark. If $\alpha : V \to W$ is an isomorphism, then $\alpha^{-1} : W \to V$ is linear. Indeed, for $w_1, w_2 \in W \times W$, let $w_1 = \alpha(v_1), w_2 = \alpha(v_2)$. Then,

$$\begin{aligned}
\alpha^{-1}(\lambda_1 w_1 + \lambda_2 w_2) &= \alpha^{-1}(\lambda_1 \alpha(v_1) + \lambda_2 \alpha(v_2)) \\
&= \alpha^{-1}(\alpha(\lambda_1 v_1 + \lambda_2 v_2)) \\
&= \lambda_1 v_1 + \lambda_2 v_2 \\
&= \lambda_1 \alpha^{-1}(v_1) + \lambda_2 \alpha^{-1}(v_2).
\end{aligned}$$

Lemma 4.2. *Congruence is an equivalence relation on the class of all vector spaces of F:*

(i) $\mathrm{id}_V : V \to V$ *is an isomorphism.*

(ii) $\alpha : V \to W$ *is an isomorphism implies* $\alpha^{-1} : W \to V$ *is an isomorphism.*

(iii) *If* $\alpha : U \to V$ *is an isomorphism,* $\beta : V \to W$ *is an isomorphism, then* $\beta \circ \alpha : U \to W$ *is an isomorphism.*

Proof: Exercise.

Theorem 4.1. *If V is a vector space over F of dimension n, then $V \cong F^n$.*

Proof: Let $\mathcal{B} = (v_1, \dots, v_n)$ be a basis of V. Then take

$$\begin{gathered}
\alpha : V \to F^n \\
v = \sum_{i=1}^{n} \lambda_i v_i \mapsto \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}
\end{gathered}$$

as an isomorphism.

Remark. In this way, choosing a basis of V is like choosing an isomorphism from V to F^n.

Theorem 4.2. *Let V, W be vector spaces over F with finite dimension. Then $V \cong W$ if and only if* $\dim V = \dim W$.

Proof: If $\dim V = \dim W$, then $V \cong F^n \cong W$, so $V \cong W$.

Otherwise, let $\alpha : V \to W$ be an isomorphism, and $\mathcal{B}$ a basis of V. Then we show $\alpha(\mathcal{B})$ is a basis of W.

- $\alpha(\mathcal{B})$ spans V from the surjectivity of α.
- $\alpha(\mathcal{B})$ is free from the injectivity of α.

Hence $\dim V = \dim W$.

Definition 4.3 (Kernal and Image)**.** Let V, W be vector spaces over F. Let $\alpha : V \to W$ be a linear map. We define

(i) $\operatorname{Ker} \alpha = \{v \in V \mid \alpha(v) = 0\}$, the kernel of α.

(ii) $\operatorname{Im}(\alpha = \{w \in W \mid \exists v \in V, \alpha(v) = w\}$, the image of α.

Lemma 4.3. $\operatorname{Ker} \alpha$ *is a subspace of V, and* $\operatorname{Im} \alpha$ *is a subspace of W.*

Proof: Let $\lambda_1, \lambda_2 \in F$, and $v_1, v_2 \in \operatorname{Ker} \alpha$. Then

$$\alpha(\lambda_1 v_1 + \lambda_2 v_2) = \lambda_1 \alpha(v_1) + \lambda_2 \alpha(v_2) = 0.$$

So $\lambda_1 v_1 + \lambda_2 v_2 \in \operatorname{Ker} \alpha$.

Now if $w_1 = \alpha(v_1), w_2 = \alpha(v_2)$, then

$$\lambda_1 w_1 + \lambda_2 w_2 = \lambda_1 \alpha(v_1) + \lambda_2 \alpha(v_2) = \alpha(\lambda_1 v_1 + \lambda_2 v_2).$$

Hence $\lambda_1 w_1 + \lambda_2 w_2 \in \operatorname{Im} \alpha$.

Example 4.2.

Consider $\alpha : \mathcal{C}^\infty(\mathbb{R}) \to \mathcal{C}^\infty(\mathbb{R})$, given by

$$f \mapsto \alpha(f) = f'' + f.$$

Then α is linear, and

$$\operatorname{Ker} \alpha = \{f \in \mathcal{C}^\infty(\mathbb{R}) \mid f'' + f = 0\} = \langle \sin t, \cos t \rangle.$$

Remark. If $\alpha : V \to W$ is linear, then α is injective if and only if $\operatorname{Ker} \alpha = \{0\}$, as

$$\alpha(v_1) = \alpha(v_2) \iff \alpha(v_1 - v_2) = 0.$$

Theorem 4.3. *Let V, W be vector spaces over F, and $\alpha : V \to W$ linear. Then*

$$\begin{aligned} V/\operatorname{Ker}\alpha &\to \operatorname{Im}\alpha \\ v + \operatorname{Ker}\alpha &\mapsto \alpha(v) \end{aligned}$$

is an isomorphism.

Proof: We proceed in steps.

- $\bar{\alpha}$ is well defined: Note if $v + \operatorname{Ker}\alpha = v' + \operatorname{Ker}\alpha$, then $v - v' \in \operatorname{Ker}\alpha$, so $\alpha(v - v') = 0$. Hence $\alpha(v) = \alpha(v')$.
- $\bar{\alpha}$ is linear: This follows from linearity of α.
- $\bar{\alpha}$ is a bijection: First, if $\bar{\alpha}(v + \operatorname{Ker}\alpha) = 0$, then $\alpha(v) = 0$, so $v \in \operatorname{Ker}\alpha$, hence $v + \operatorname{Ker}\alpha = 0 + \operatorname{Ker}\alpha$, so α is injective. Then $\bar{\alpha}$ is surjective from the definition of the image.

Definition 4.4 (Rank and Nullity)**.** We define the rank $r(\alpha) = \operatorname{rank}(\alpha) = \dim \operatorname{Im}\alpha$, and the nullity $n(\alpha) = \operatorname{null}(\alpha) = \dim \operatorname{Ker}\alpha$.

Theorem 4.4 (Rank-nullity theorem)**.** *Let U, V be vector spaces over F, with* $\dim U < \infty$*, and let $\alpha : U \to V$ be a lninear map. Then,*

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

Proof: We have proven that $U/Ker\alpha \cong \operatorname{Im}\alpha$, but we have already proven $\dim U/Ker\alpha = \dim U - r(\alpha)$, which proves the theorem.

Lemma 4.4. *Let V, W be vector spaces over F of equal finite dimension. Let $\alpha : V \to W$ be a linear map. Then the following are equivalent:*

- *α is injective,*
- *α is surjective,*
- *α is an isomorphism.*

This follows immediately from the rank-nullity theorem.

5 Linear maps and Matrices

Definition 5.1. If V, W are vector spaces over F, then

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

Proposition 5.1. *$L(V, W)$ is a vector space over F with*

$$(\alpha_1 + \alpha_2)(v) = \alpha_1(v) + \alpha_2(v),$$

$$(\lambda\alpha)(v) = \lambda\alpha 9v).$$

Moreover, if V and W are finite dimensional, then so is $L(V, W)$, and

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

Definition 5.2. An $m \times n$ matrix over F is an array with m rows and n columns with entries in F, $A = (a_{ij})$. Define

$$M_{m,n}(F) = \{\text{set of } m \times n \text{ matrices over } F\}.$$

Proposition 5.2. *$M_{m,n}(F)$ is a vector space over F, and* $\dim M_{m,n}(F) = mn$

Proof: Let E_{ij} be the matrix with $a_{xy} = \delta_{xi}\delta_{yj}$. Then (E_{ij}) is a basis of $M_{m,n}(F)$, as

$$N = (a_{ij}) = \sum_{i,j} a_{ij} E_{ij},$$

and (E_{ij}) is free.

If V, W are vector spaces over F, and $\alpha : V \to W$ is a linear map, we take a basis $\mathcal{B} = (v_1, \dots, v_n)$ of V, and $\mathcal{C} = (w_1, \dots, w_m)$ of W. Let $v \in V$, then

$$v = \sum_{i=1}^{n} \lambda_i v_i \sim \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} \in F^n.$$

We let this isomorphism from V to F^n be $[v]_{\mathcal{B}}$. Similarly, we can obtain $[w]_{\mathcal{B}}$ for $w \in W$.

Definition 5.3. We define a matrix of α with respect to a basis $\mathcal{B}, \mathcal{C}$ as

$$[\alpha]_{\mathcal{B},\mathcal{C}} = ([\alpha(v_1)]_{\mathcal{C}}, [\alpha(v_2)]_{\mathcal{C}}, \dots, [\alpha(v_n)]_{\mathcal{C}}).$$

By definition, if $[\alpha]_{\mathcal{B},\mathcal{C}} = (a_{ij})$, then

$$\alpha(v_j) = \sum_{i=1}^{m} a_{ij} w_i.$$

Lemma 5.1. *If $v \in V$, then*

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

or equivalently,

$$(\alpha(v))_i = \sum_{j=1}^{n} a_{ij} \lambda_j.$$

Proof: Let $v \in V$, then

$$v = \sum_{j=1}^{n} \lambda_j v_j.$$

Then

$$\begin{aligned} \alpha(v) &= \alpha\left(\sum_{j=1}^{n} \lambda_j v_j\right) = \sum_{j=1}^{n} \lambda_j \alpha(v_j) \\ &= \sum_{j=1}^{n} \lambda_j \sum_{i=1}^{n} a_{ij} w_i = \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij} \lambda_j\right) w_i. \end{aligned}$$

Lemma 5.2. *If $U \to V$ is linear under β, $V \to W$ linear under α, then $U \to W$ is linear under $\alpha \to W$. Let $\mathcal{A}$ be a basis of U, $\mathcal{B}$ a basis of V, and $\mathcal{C}$ a basis of W. Then*

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

Proof: Let $A = [\alpha]_{\mathcal{B},\mathcal{C}}$, $B = [\beta]_{\mathcal{A},\mathcal{B}}$. Pick $u_l \in A$. Then

$$\begin{aligned} (\alpha \circ \beta)(u_l) &= \alpha(\beta(u_l)) = \alpha\left(\sum_j b_{jl} v_j\right) \\ &= \sum_j b_{jl} \alpha(v_j) = \sum_j b_{jl} \sum_i a_{ij} w_i \\ &= \sum_i \left(\sum_j a_{ij} b_{jl}\right) w_i. \end{aligned}$$

Proposition 5.3. *If V and W are vector spaces over F, and $\dim V = n$, $\dim W = m$, then $L(V,W) \cong M_{m,n}(F)$, so $\dim L(V,W) = m \times n$.*

Proof: Fix $\mathcal{B}, \mathcal{C}$ bases of V and W. We show

$$\begin{aligned} \theta : L(V,W) &\to M_{m,n}(F) \\ \alpha &\mapsto [\alpha]_{\mathcal{B},\mathcal{C}} \end{aligned}$$

is an isomorphism.

- θ is linear: $[\lambda_1\alpha_1 + \lambda_2\alpha_2]_{\mathcal{B},\mathcal{C}} = \lambda_1[\alpha_1]_{\mathcal{B},\mathcal{C}} + \lambda_2[\alpha_2]_{\mathcal{B},\mathcal{C}}$.
- θ is surjective: Consider $A = (a_{ij})$. Consider the map

  $$\alpha : v_j \mapsto \sum_{i=1}^{m} a_{ij} w_i.$$

  This can be extended by linearity, and $[\alpha]_{\mathcal{B},\mathcal{C}} = A$.
- θ is injective: If $[\alpha]_{\mathcal{B},\mathcal{C}} = 0$, then $\alpha = 0$ for all v.

Remark. If $\mathcal{B}, \mathcal{C}$ are bases of V, W and $\varepsilon_{\mathcal{B}} : v \mapsto [v]_{\mathcal{B}}$, $\varepsilon_{\mathcal{C}} : w \mapsto [w]_{\mathcal{C}}$, then the following diagram commutes:

$$\begin{tikzcd} V \arrow[r, "\alpha"] \arrow[d, "\varepsilon_{\mathcal{B}}"] & W \arrow[d, "\varepsilon_{\mathcal{C}}"] \\ F^n \arrow[r, "{[\alpha]_{\mathcal{B},\mathcal{C}}}"] & F^m \end{tikzcd}$$

6 Change of Basis and Equivalent Matrices

Let $\alpha : V \to W$ with $\mathcal{B}$ and $\mathcal{C}$ bases of V, W. Then

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

If $Y \leq V$, we can take $\mathcal{B}$ a basis of V, such that $(v_1, \ldots, v_k, v_{k+1}, \ldots, v_n)$ is a basis of V, and $(v_1, \ldots, v_k)$ is a basis $\mathcal{B}'$ of Y, and $(v_{k+1}, \ldots, v_n)$ is a basis $\mathcal{B}''$.

Then if $Z \leq W$, we can take a basis $\mathcal{C}$ of W $(w_1, \ldots, w_l, w_{l+1}, \ldots, w_m)$, such that $(w_1, \ldots, w_l)$ is a basis $\mathcal{C}'$ of Z, and $(w_{l+1}, \ldots, w_m)$ is a basis $\mathcal{C}''$. Then

$$[\alpha]_{\mathcal{B},\mathcal{C}} = \begin{pmatrix} A & B \\ 0 & C \end{pmatrix}.$$

Then we can show that

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

if $\alpha(Y) \leq Z$. Moreover, we can show α induces a homomorphism

$$\begin{aligned} \bar{\alpha} : V/Y &\to W/Z \\ v + Y &\mapsto \alpha(v) + Z \end{aligned}$$

This is well-defined as $\alpha(v) \in Z$ for $v \in Y$, and $[\bar{\alpha}]_{\mathcal{B}'',\mathcal{C}''} = C$.

6.1 Change of Basis

Consider $\alpha : V \to W$, where V has two bases $\mathcal{B} = \{v_1, \ldots, v_n\}$ and $\mathcal{B}' = \{v_1', \ldots, v_n'\}$ and W has two bases $\mathcal{C} = \{w_1, \ldots, w_n\}$ and $\mathcal{C}' = \{w_1', \ldots, w_m'\}$. We aim to find the relation between $[\alpha]_{\mathcal{B},\mathcal{C}}$ and $[\alpha]_{\mathcal{B}',\mathcal{C}'}$.

Definition 6.1. The change of basis matrix from $\mathcal{B}'$ to $\mathcal{B}$ is $P = (p_{ij})$ given by

$$P = ([v_1']_{\mathcal{B}}, \ldots, [v_n']_{\mathcal{B}}) = [\mathrm{id}]_{\mathcal{B}',\mathcal{B}}.$$

Lemma 6.1. $[v]_{\mathcal{B}} = P[v]_{\mathcal{B}'}$.

Proof: In general $[\alpha(v)]_{\mathcal{C}} = [\alpha]_{\mathcal{B},\mathcal{C}}[v]_{\mathcal{B}}$. If $P = [\mathrm{id}]_{\mathcal{B}',\mathcal{B}}$, then

$$[v]_{\mathcal{B}} = [\mathrm{id}(v)]_{\mathcal{B}} = [\mathrm{id}]_{\mathcal{B}',\mathcal{B}}[v]_{\mathcal{B}'} = P[v]_{\mathcal{B}'}.$$

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

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

and similarly.

Note while we know $[v]_{\mathcal{B}} = P[v]_{\mathcal{B}'}$, to compute a vector in $\mathcal{B}'$, we have $[v]_{\mathcal{B}'} = P^{-1}[v]_{\mathcal{B}}$. This is hard to do.

Similarly, we can also change basis $\mathcal{C}$ to $\mathcal{C}'$ in W. In this case, the change of basis matrix $Q = [\mathrm{id}]_{\mathcal{C}',\mathcal{C}}$ is $m \times m$ and invertible.

Now given $\alpha : V \to W$, we wish to find how $[\alpha]_{\mathcal{B},\mathcal{C}}$ and $[\alpha]_{\mathcal{B}',\mathcal{C}'}$.

Proposition 6.1. *If $A = [\alpha]_{\mathcal{B},\mathcal{C}}$, $A' = [\alpha]_{\mathcal{B}',\mathcal{C}'}$, $P = [\mathrm{id}]_{\mathcal{B}',\mathcal{B}}$, $Q = [\mathrm{id}]_{\mathcal{C}'}, \mathcal{C}$, then*

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

Proof: Combining the facts we know, we get

$$[\alpha(v)]_{\mathcal{C}} = Q[\alpha(v)]_{\mathcal{C}'} = Q[a]_{\mathcal{B}',\mathcal{C}'}[v]_{\mathcal{B}'} = QA'[v]_{\mathcal{B}'}.$$

But we also know

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

But since this is true for any $v \in V$, we get $QA' = AP$, so $A' = Q^{-1}AP$.

Definition 6.2 (Equivalent matrices)**.** Two matrices $A, B \in M_{m,n}(F)$ are equivalent if $A' = Q^{-1}AP$, where $Q \in M_{m,m}$ and $P \in M_{n,n}$ are invertible.

Remark. This defines an equivalence relation on $M_{m,n}(F)$, as

- $A = I_m^{-1}AI_n$,
- If $A' = Q^{-1}AP$, then $A = (Q^{-1})^{-1}A'P^{-1}$,
- If $A' = Q^{-1}AP$, $A'' = (Q')^{-1}A'P'$, then $A'' = (QQ')^{-1}A(PP')$.

Proposition 6.2. *Let V, W be vector spaces over F, with $\dim_F V = n$, $\dim_F W = m$. Let $\alpha : V \to W$ be a linear map. Then there exists $\mathcal{B}$, $\mathcal{C}$ bases of V, W such that*

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

Proof: Choose $\mathcal{B}$ and $\mathcal{C}$ wisely. Fix $r \in \mathbb{N}$ such that $\dim \operatorname{Ker} \alpha = n - r$. Let $N(\alpha) = \operatorname{Ker}(\alpha) = \{x \in V \mid \alpha(x) = 0\}$. Fix any basis of $N(x)$, $(v_{r+1}, \ldots, v_n)$, and extend it to a basis $\mathcal{B} = (v_1, \ldots, v_r, v_{r+1}, \ldots, v_n)$.

We claim that $(\alpha(v_1), \ldots, \alpha(v_r))$ is a basis of $\operatorname{Im} \alpha$.

- First, if $v = \sum \lambda_i v_i$, then
$$\alpha(v) = \sum_{i=1}^{n} \lambda_i \alpha(v_i) = \sum_{i=1}^{r} \lambda_i \alpha(v_i).$$
Let $y \in \operatorname{Im} \alpha$, so then
$$y = \sum_{i=1}^{r} \lambda_i \alpha(v_i).$$
So $y \in \langle \alpha(v_1), \ldots, \alpha(v_r) \rangle$.
- Now, suppose that it is not free, so
$$\sum_{i=1}^{r} \lambda_i \alpha(v_i) = 0.$$
Then we get
$$\alpha\left(\sum_{i=1}^{r} \lambda_i v_i\right) = 0,$$
so
$$\sum_{i=1}^{r} \lambda_i v_i \in \operatorname{Ker} \alpha.$$
Hence, we get that
$$\sum_{i=1}^{r} \lambda_i v_i = \sum_{i=1}^{n} \mu_i v_i.$$
But since $(v_1, \ldots, v_n)$ is a basis, $\lambda_i = \mu_i = 0$.

So we have $(\alpha(v_1), \ldots, \alpha(v_r))$ is a basis of $\operatorname{Im} \alpha$, and $(v_{r+1}, \ldots, v_n)$ is a basis of $\operatorname{Ker} \alpha$. Let $\mathcal{C} = (\alpha(v_1), \ldots, \alpha(v_r), w_{r+1}, \ldots, w_m)$. We get that
$$[\alpha]_{\mathcal{B},\mathcal{C}} = (\alpha(v_1), \ldots, \alpha(v_r), \alpha(v_{r+1}), \ldots, \alpha(v_n)) = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}.$$

Remark. This proves another proof of the rank-nullity theorem: $r(\alpha) + n(\alpha) = n$.

Corollary 6.1. *Any $m \times n$ matrix is equivalent to*
$$\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix},$$
where $r = \operatorname{rank}(\alpha)$.

Definition 6.3. For $a \in M_{m,n}(F)$, the column rank $r_c(A)$ of A is the dimension of the span of the column vectors of A in F^m. Similarly, the row rank is the column rank of A^T.

Remark. If α is a linear map represented by A with respect to one basis, the column rank A equals the rank of α.

Proposition 6.3. *Two matrices are equivalent if and only if $r_c(A) = r_c(A')$.*

Proof: If A and A' are equivalent then they coorespond to the same linear map α except in two different bases.

Conversely, if $r_c(A) = r_c(A') = r$, then both A and A' are equivalent to

$$\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix},$$

hence are equivalent.

Theorem 6.1. *$r_c(A) = r_c(A^T)$, so column rank equals row rank.*

Proof: If $r = r_c(A)$, then

$$Q^{-1}AP = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}.$$

Take the transpose, to get

$$(Q^{-1}AP)^T = P^T A^T (Q^{-1})^T = P^T A^T (Q^T)^{-1} = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}.$$

Hence $r_c(A^T) = r = r_c(A)$.

7 Elementary operations and Elementary Matrices

This is a special case of the change of basis formula, when $\alpha : V \to V$ is a map from a vector space to itself, called an endomorphism. Suppose $\mathcal{B} = \mathcal{C}$ and $\mathcal{B}' = \mathcal{C}'$, and P is the change of basis matrix from $\mathcal{B}'$ to $\mathcal{B}$. Then

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

Definition 7.1. Let A, A' be $n \times n$ matrices. We say that A and A' are similar if and only if $A' = P^{-1}AP$ for a square invertible matrix P.

Definition 7.2. The elementary column operations on an $m \times n$ matrix A are:

(i) Swap columns i and j;

(ii) Replace column i by λ times column i;

(iii) Add λ times column i to column j, for $i \neq j$.

The elementary row operations are analogously defined.

Note elementary operations are invertible, and all operations can be realized through the action of elementary matrices:

(i) For swapping columns i and j, we can take an identity matrix, but with $a_{ij} = a_{ji} = 1$, and $a_{ii} = a_{jj} = 0$.

(ii) For multiplying column i by λ, we can take an identity matrix but with $a_{ii} = \lambda$.

(iii) For adding λ times columns i to column j, we can take an identity matrix but with $a_{ij} = \lambda$.

An elementary columns (resp. row) operation can be done by multiplying A by the corresponding elementary matrix from the right (resp. left).

We will now show that any $m \times n$ matrix is equivalent to

$$\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}.$$

Start with a matrix A. If all entries are zero, we are done. Otherwise, pick $a_{ij} = \lambda \neq 0$. By swapping columns and rows, we can ensure $a_{11} = \lambda$. Multiplying column 1 by $1/\lambda$, we get $a_{11} = 1$. We can then clean out row 1 by subtracting a suitable multiply of column 1 from every row, and similarly from column 1. This gives us a matrix

$$\begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & \tilde{A} & \\ 0 & & & \end{pmatrix}.$$

Iterating with $\tilde{A}$, a strictly smaller matrix, eventually gives

$$\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} = Q^{-1}AP.$$

A variation of this is known as **Gauss' pivot algorithm**. If we only use row operations, we can reach the row-echelon form of the matrix:

- Assume that $a_{i1} \neq 0$ for some i.
- Swap rows i and 1.
- Divide first row by $\lambda = a_{i1}$.
- Use 1 in a_{11} to clean the first column.
- Iterate over all columns.

This procedure is what is usually done when solving a system of linear equations.

7.1 Representation of Square Invertible Matrix

Lemma 7.1. *If A is an $n \times n$ square invertible matrix, then we can obtain I_n using either only row or column elementary operations.*

Proof: We prove for column operations; row operations are analogous. We proceed by induction on the number of rows.

- Suppose that we could write A in the form
$$\begin{pmatrix} I_h & 0 \\ * & * \end{pmatrix}.$$
Then we want to obtain the same structure as we go from h to $h+1$.
- We show there exists $j > h$ such that $\lambda = a_{h+1,j} \neq 0$. Otherwise, the row rank is less than n, as the first $h+1$ rows are linearly dependent. Hence $\operatorname{rank} A < n$.
- We swap columns $h+1$ and j, so $\lambda = a_{h+1,h+1} \neq 0$, and then divide by

λ.

- Finally, we can use the 1 in $a_{h+1,h+1}$ to clear out the rest of the $(h+1)$'st row.

This gives $AE_1 \dots E_c = I_n$, or $A^{-1} = E_1 \dots E_c$. This is an algorithm for computing A^{-1}.

Proposition 7.1. *Any invertible square matrix is a product of elementary matrices.*

8 Dual Spaces and Dual Maps

Definition 8.1. V is a F-vector space. We say V^* is the dual of V if

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

If $\alpha : V \to F$ is linear, then we say α is a linear form.

Example 8.1.

(i) $\operatorname{tr} : M_{n,n}(F) \to F$ is a linear map, so $\operatorname{tr} \in M_{n,n}^*(F)$.

(ii) Let $f : [0,1] \to \mathbb{R}$ by $x \mapsto f(x)$, and $Tf : \mathcal{C}^\infty([0,1], \mathbb{R}) \to \mathbb{R}$ by

$$\phi \mapsto \int_0^1 f(x)\phi(x)\,\mathrm{d}x.$$

Then Tf is a linear form.

Lemma 8.1. *Let V be a vector space over F with a finite basis $\mathcal{B} = \{e_1, \dots, e_n\}$. Then there exists a basis for V^* given by $\mathcal{B}^* = \{\varepsilon_1, \dots, \varepsilon_n\}$, with*

$$\varepsilon_j \left(\sum_{i=1}^n a_i e_i \right) = a_j.$$

Then $\mathcal{B}^*$ is the dual basis of $\mathcal{B}$.

Remark. If we define the Kronecker symbols

$$\delta_{ij} = \begin{cases} 1 & i = j, \\ 0 & \text{otherwise,} \end{cases}$$

then we can equivalently define

$$\varepsilon_j \left(\sum_{i=1}^n a_i e_i \right) = a_j \iff \varepsilon_j(e_i) = \delta_{ij}.$$

Proof: Let $(\varepsilon_1, \dots, \varepsilon_n)$ be defined as above.

We prove (ε_i) are free. Indeed, suppose

$$\sum_{j=1}^{n} \lambda_j \varepsilon_j = 0 \implies \sum_{j=1}^{n} \lambda_j e_j(e_i) = 0 \implies \lambda_i = 0.$$

Now we show (ε_i) generates V^*. Pick $\alpha \in V^*$, then for $x \in V$, we have

$$\alpha(x) = \alpha\left(\sum_{j=1}^{n} \lambda_j e_j\right) = \sum_{j=1}^{n} \lambda_j \alpha(e_j).$$

On the other hand, consider the linear form

$$\sum_{j=1}^{n} \alpha(e_j)\varepsilon_j \in V^*.$$

Then we have

$$\begin{aligned} \sum_{j=1}^{n} \alpha(e_j)\varepsilon_j(x) &= \sum_{j=1}^{n} \alpha(e_j)\varepsilon_j\left(\sum_{k=1}^{n} \lambda_k e_k\right) = \sum_{j=1}^{n} \alpha(e_j) \sum_{k=1}^{n} \lambda_k \varepsilon_j(e_k) \\ &= \sum_{j=1}^{n} \alpha(e_j)\lambda_j = \alpha(x). \end{aligned}$$

Hence (ε_i) generates V^*.

Corollary 8.1. *If V is finite dimensional, then* $\dim V^* = \dim V$.

This is very different in infinite dimensions.

Remark. It is sometimes convenient to think of V^* as the space of row vector of length n over F. If $(e_1, \ldots, e_n)$ is a basis of v such that $x = \sum x_i e_i$ and $(\varepsilon_1, \ldots, \varepsilon_n)$ is a basis of V^* such that $\alpha = \sum \alpha_i \varepsilon_i$, then

$$\begin{aligned} \alpha(x) &= \sum_{i=1}^{n} \alpha_i \varepsilon_i \left(\sum_{j=1}^{n} x_j e_j\right) = \sum_{i=1}^{n} \alpha_i \sum_{j=1}^{n} x_j \varepsilon_i(e_j) = \sum_{i=1}^{n} \alpha_i x_i \\ &= \begin{pmatrix} \alpha_1 & \cdots & \alpha_n \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}. \end{aligned}$$

This gives a scalar product structure on V^*.

Definition 8.2. If $U \leq V$, we define the annihilator of U by

$$U^\circ = \{\alpha \in V^* \mid \alpha(u) = 0 \; \forall u \in U\}.$$

Lemma 8.2.

(i) $U^\circ \leq V^*$.

(ii) *If* $U \leq V$ *and* $\dim V < \infty$*, then* $\dim V = \dim U + \dim U^\circ$.

Proof: Suppose $\alpha, \alpha' \in U^\circ$. Then for all $u \in U$,

$$(\alpha + \alpha')(u) = \alpha(u) + \alpha'(u) = 0,$$

and for all $\lambda \in F$, $(\lambda\alpha)(u) = \lambda\alpha(u) = 0$. Hence $U^\circ \leq V^*$.

Now let $U \leq V$, and $\dim V = n$. Let $(e_1, \ldots, e_k)$ be a basis of U and complete it to a basis $\mathcal{B} = (e_1, \ldots, e_k, e_{k+1}, \ldots, e_n)$ of V. Let $(\varepsilon_1, \ldots, \varepsilon_n)$ be the dual basis of $\mathcal{B}$. Then I claim $U^\circ = \langle \varepsilon_{k+1}, \ldots, \varepsilon_n \rangle$.

Indeed, pick $i > k$, then $\varepsilon_i(e_k) = \delta_{ik} = 0$, so $\varepsilon_i \in U^\circ$. Now let $\alpha \in U^\circ$. Then $(\varepsilon_1, \ldots, \varepsilon_n)$ is a basis of V^* implies $\alpha = \sum \alpha_i \varepsilon_i$. But $\alpha \in U^\circ \implies \alpha(e_i) = 0$, which gives $\alpha_i = 0$ for $i \leq k$. Hence $\alpha \in \langle \varepsilon_{k+1}, \ldots, \varepsilon_n \rangle$.

Definition 8.3. Let V, W be vector spaces over F, and let $\alpha \in L(V, W)$. Then the map

$$\begin{aligned} \alpha^* : W^* &\to V^* \\ \varepsilon &\mapsto \varepsilon \circ \alpha \end{aligned}$$

is an element of $L(W^*, V^*$. This is known as the dual map of α.

Proof: $\varepsilon \circ \alpha : V \to F$ is linear due to the linearity of ε and α. Hence $\varepsilon \circ \alpha \in V^*$.

We show α^* is linear. Let $\theta_1, \theta_2 \in W^*$. Then,

$$\alpha^*(\theta_1 + \theta_2) = (\theta_1 + \theta_2)(\alpha) = \theta_1 \circ \alpha + \theta_2 \circ \alpha = \alpha^*(\theta_1) + \alpha^*(\theta_2).$$

Similarly, if $\lambda \in F$, then

$$\alpha^*(\lambda\theta) = \lambda\alpha^*(\theta).$$

Hence $\alpha^* \in L(W^*, V^*)$.

Proposition 8.1. *Let V, W be finite dimensional spaces over F with bases $\mathcal{B}, \mathcal{C}$. Let $\mathcal{B}^*, \mathcal{C}^*$ be the dual bases for V^*, W^*. Then*

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

Proof: Let $\mathcal{B} = (b_1, \dots, b_n), \mathcal{C} = (c_1, \dots, c_m)$, $\mathcal{B}^* = (\beta_1, \dots, \beta_n), \mathcal{C}^* = (\gamma_1, \dots, \gamma_m)$. Say $[\alpha]_{\mathcal{B},\mathcal{C}} = A = (a_{ij})$. Recall $\alpha^* : W^* \to V^*$, so let us compute

$$\alpha^*(\gamma_r)(b_s) = \gamma_r \circ \alpha(b_s) = \gamma_r\left(\sum_t a_{ts} c_t\right) = \sum_t a_{ts} \gamma_r(c_t) = a_{rs}.$$

Say that

$$[\alpha^*]_{\mathcal{C}^*,\mathcal{B}^*} = \begin{pmatrix} \alpha^*(\gamma_1) & \cdots & \alpha^*(\gamma_m) \end{pmatrix} = (m_{ij}).$$

Then we can find that

$$\alpha^*(\gamma_r) = \sum_{i=1}^n m_{ir} \beta_i,$$

so

$$\alpha^*(\gamma_r)(b_s) = m_{sr}.$$

This gives $a_{rs} = m_{sr}$, as desired.

9 Properties of the Dual Map

Recall if V, W are vector spaces over F, and $\alpha \in L(V, W)$, then we can construct a dual map

$$\begin{aligned} \alpha^* : W^* &\to V^* \\ \varepsilon &\mapsto \varepsilon \circ \alpha \end{aligned}$$

Moreover, if $\mathcal{B}, \mathcal{C}$ are bases of V and W, and $\mathcal{B}^*$, $\mathcal{C}^*$ are the dual bases of $\mathcal{B}$ and $\mathcal{C}$ respectively, then

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

Now if $\mathcal{E} = (e_1, \ldots, e_n)$ is a basis of V and $\mathcal{F} = (f_1, \ldots, f_n)$ is another basis of V, then consider the change of basis matrix

$$P = [\mathrm{id}]_{\mathcal{F}, \mathcal{E}}.$$

Consider $\mathcal{E}^* = (\varepsilon_1, \ldots, \varepsilon_n)$ and $\mathcal{F}^* = (\eta_1, \ldots, \eta_n)$.

Lemma 9.1. *The change of basis matrix from $\mathcal{F}^*$ to $\mathcal{E}^*$ is*

$$(P^{-1})^T.$$

Proof: We have

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

9.1 Properties of the Dual Map

Lemma 9.2. *Let V, W be vector spaces over F. Let $\alpha \in L(V, W)$ and $\alpha^* \in L(W^*, V^*)$. Then*

(i) $\mathrm{Ker}(\alpha^*) = (\mathrm{Im}\,\alpha)^\circ$. *Hence α^* is injective if and only if α is surjective.*

(ii) $\mathrm{Im}\,\alpha^* \le (\mathrm{Ker}\,\alpha)^\circ$ *with equality if V, W are finite dimensional. Hence in this case, α^* is injective if and only if α is injective.*

There are many problems where the understanding of α^* is simpler than the understanding of α.

Proof:

(i) Let $\varepsilon \in W^*$. Then $\varepsilon \in \mathrm{Ker}\,\alpha^* \iff \alpha^*(\varepsilon) = 0$. But $\alpha^*(\varepsilon) = \varepsilon(\alpha)$, so for all x,

$$\varepsilon(\alpha)(x) = \varepsilon(\alpha(x)) = 0.$$

This holds if and only if $\varepsilon \in (\operatorname{Im}\alpha)^\circ$.

(ii) We will first show that

$$\operatorname{Im}\alpha^* \leq (\operatorname{Ker}\alpha)^\circ.$$

Indeed, if $\varepsilon \in \operatorname{Im}\alpha^*$, then $\varepsilon = \alpha^*(\phi)$, so for all $u \in \operatorname{Ker}\alpha$,

$$\varepsilon(u) = \alpha^*(\phi)(u) = \phi \circ \alpha(u) = \phi(0) = 0.$$

Hence $\varepsilon \in (\operatorname{Ker}\alpha)^\circ$. In finite dimension, we can compare the dimension of $\operatorname{Im}\alpha^*$ and $(\operatorname{Ker}\alpha)^\circ$. Indeed,

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

Hence, we get

$$\dim(\operatorname{Im}\alpha^*) = r(\alpha^*) = r(\alpha) = \dim V - \dim\operatorname{Ker}\alpha = \dim[(\operatorname{Ker}\alpha)^\circ].$$

Since the dimensions are the same, we get $\operatorname{Im}\alpha^* = (\operatorname{Ker}\alpha)^\circ$.

9.2 Double Dual

If V is a vector space over F, then $V^* = L(V, F)$.

We define the **bidual** as

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

This is a very important space in infinite dimension. In general, there is no obvious connection between V and V^*. However, there is a large class of function spaces such that

$$V \cong V^{**}.$$

This is known as a reflexive space.

Example 9.1.

For $p > 2$, define

$$L^p(\mathbb{R}) = \left\{ f : \mathbb{R} \to \mathbb{R} \,\middle|\, \int_{\mathbb{R}} |f(x)|^p \,\mathrm{d}x < \infty \right\}.$$

This is an example of a reflexive space.

In general, there is a canonical embedding of V into V^{**}. Indeed, pick $v \in V$. We define

$$\begin{aligned} \hat{v} : V^* &\to F \\ \varepsilon &\mapsto \varepsilon(v) \end{aligned}$$

Then this is linear, as

$$\hat{v}(\lambda_1\varepsilon_1 + \lambda_2\varepsilon_2) = (\lambda_1\varepsilon_1 + \lambda_2 + \varepsilon_2)(v) = \lambda_1\varepsilon_1(v) + \lambda_2\varepsilon_2(v) = \lambda_1\hat{v}(\varepsilon_1) + \lambda_2\hat{v}(\varepsilon_2).$$

Theorem 9.1. *If V is a finite dimensional vector space over F, then the hat map $v \mapsto \hat{v}$ is an isomorphism.*

In infinite dimension, under certain assumption (e.g. Banach space) we can show that the hat map is injective.

Proof: If V is finite dimensional, then first note that for $v \in V$, $\hat{v} \in V^{**}$. We show the hat map is linear: for $v_1, v_2 \in V$, $\lambda_1, \lambda_2 \in F$ and $\varepsilon \in V^*$,

$$\lambda_1\widehat{v_1 + \lambda_2}v_2(\varepsilon) = \varepsilon(\lambda_1 v_1 + \lambda_2 v_2) = \lambda_1\varepsilon(v_1) + \lambda_1\varepsilon_2(v_2) = \lambda_1\hat{v}_1(\varepsilon) + \lambda_2\hat{v}_2(\varepsilon).$$

Now we show the hat map is injective. Let $e \in V \setminus \{0\}$. Then extend to a basis $(e, e_2, \ldots, e_n)$. Let $(\varepsilon, \varepsilon_2, \ldots, \varepsilon_n)$ be the dual basis. Then

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

Hence $\hat{e} \neq \{0\}$, so the hat map is injective.

Finally, we show the hat map is an isomorphism. We already know $\dim V = \dim V^*$, and as a result $\dim V^* = \dim V^{**}$. Thus, since the hat map is injective, it is an isomorphism.

Lemma 9.3. *Let V be a finite dimensional vector space over K, and let $U \leq V$. Then*

$$\hat{U} = U^{\circ\circ}.$$

Hence after identification of V and V^{**}, we get

$$U = U^{\circ\circ}.$$

Proof: We will show $U \leq U^{\circ\circ}$. Indeed, let $u \in U$. Then for all $\varepsilon \in U^\circ$, $\varepsilon(u) = 0$. So for all $\varepsilon \in U^\circ$, $\hat{u}(\varepsilon) = \varepsilon(u) = 0$. Hence $\hat{u} \in U^{\circ\circ}$, so $\hat{U} \subset U^{\circ\circ}$.

But then we can compute dimension to find

$$\dim U^{\circ\circ} = \dim V - \dim U^\circ = \dim U,$$

proving this lemma.

Remark. If $T \leq V^*$, then

$$T^\circ = \{v \in V \mid \theta(v) = 0, \forall \theta \in T\}.$$

Lemma 9.4. *Let V be a finite dimensional vector space over K. Let $U_1, U_2 \leq V$. Then,*

(i) $(U_1 + U_2)^\circ = U_1^\circ \cap U_2^\circ$,

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

Proof:

(i) Let $\theta \in V^*$, then

$$\begin{aligned} \theta \in (U_1 + U_2)^\circ &\iff \theta(u_1 + u_2) = 0 \iff \theta(u) = 0 \, \forall u \in U_1 \cup U_2 \\ &\iff \theta \in U_1^\circ \cap U_2^\circ. \end{aligned}$$

Hence $(U_1 + U_2)^\circ = U_1^\circ \cap U_2^\circ$.

(ii) Looking at (i), we can take the annihilator of everything to get

$$(U_1 \cap U_2)^\circ = (U_1^\circ + U_2^\circ)^{\circ\circ} = U_1^\circ + U_2^\circ.$$

10 Bilinear Forms

Definition 10.1. Let U, V be vector spaces over K. Then

$$\phi : U \times V \to K$$

is a **bilinear form** if it is linear in both components.

Example 10.1.

(i) Take $V \times V^* \to K$ by $(v, \theta) \mapsto \theta(v)$.

(ii) The scalar product on $U = V = \mathbb{R}^n$ is $\psi : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ by

$$\left(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \right) \mapsto \sum_{i=1}^{n} x_i y_i.$$

(iii) If $U = V = \mathcal{C}([0,1], \mathbb{R})$, then we can define

$$\phi(f, g) = \int_0^1 f(t) g(t) \, \mathrm{d}t.$$

This can be thought of as an infinite dimensional scalar product.

Definition 10.2. Let $\mathcal{B} = (e_1, \ldots, e_m)$ be a basis of U, and $\mathcal{C} = (f_1, \ldots, f_n)$ be a basis of V. If $\phi : U \times V \to F$ is a bilinear form, then the matrix of ϕ with respect to $\mathcal{B}$ and $\mathcal{C}$ is

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

Lemma 10.1.

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

Proof: Let

$$u = \sum_{i=1}^{m} \lambda_i e_i, \quad v = \sum_{j=1}^{n} \mu_i j f_j.$$

Since ϕ is a bilinear form,

$$\begin{aligned}\phi(u,v) = \phi\left(\sum_{i=1}^m \lambda_i e_i, \sum_{j=1}^n \mu_j e_j\right) &= \sum_{i=1}^m \sum_{j=1}^n \lambda_i \mu_j \phi(e_i, f_j) \\ &= [u]_{\mathcal{B}}^T [\phi]_{\mathcal{B},\mathcal{C}} [v]_{\mathcal{C}}.\end{aligned}$$

Remark. $[\phi]_{\mathcal{B},\mathcal{C}}$ is the only matrix satisfying this property.

Definition 10.3. $\phi : U \times V \to K$ a bilinear form determines two linear maps:

$$\begin{aligned}\phi_L &: U \to V^* \\ \phi_L(u) &: V \to K \\ & \quad v \mapsto \phi(u,v) \\ \phi_R &: V \to U^* \\ \phi_R(v) &: U \to K \\ & \quad u \mapsto \phi(u,v)\end{aligned}$$

Lemma 10.2. *Let $\mathcal{B} = (e_1, \ldots, e_m)$ a basis of U, and $\mathcal{B}^* = (\varepsilon_1, \ldots, \varepsilon_m)$ a dual basis of U^*, Similarly, let $\mathcal{C} = (f_1, \ldots, f_n)$ be a basis of V, and $\mathcal{C}^*(\eta_1, \ldots, \eta_n)$ a dual basis of V^*.*

Let $A = [\phi]_{\mathcal{B},\mathcal{C}}$. Then,

$$\begin{aligned}[\phi_R]_{\mathcal{C},\mathcal{B}^*} &= A, \\ [\phi_L]_{\mathcal{B},\mathcal{C}^*} &= A^T.\end{aligned}$$

Proof: We have $\phi_L(e_i, f_j) = \phi(e_i, f_j) = A_{ij}$, and so

$$\phi_L(e_i) = \sum A_{ij} \eta_j.$$

Similarly, $\phi_R(f_j)(e_i) = \phi(e_i, f_j) = A_{ij}$, so

$$\phi_R(f_j) = \sum A_{ij} \varepsilon_i.$$

This naturally gives our result.

Definition 10.4. Let $\operatorname{Ker} \phi_L$ be the left kernel of ϕ, and $\operatorname{Ker} \phi_R$ be the right kernel of ϕ.

We say that ϕ is non-degenerate if $\operatorname{Ker}\phi_L = \{0\}$ and $\operatorname{Ker}\phi_R = \{0\}$. Otherwise, we say that ϕ is degenerate.

Lemma 10.3. *Let U, V be finite dimensional, $\mathcal{B}, \mathcal{C}$ bases of U and V, and $\phi : U \times V \to K$ a bilinear form. Let $A = [\phi]_{\mathcal{B},\mathcal{C}}$.*

Then ϕ is non-degenerate if and only if A is invertible.

Corollary 10.1. *If ϕ is non-degenerate, then* $\dim U = \dim V$.

Proof: ϕ is non-degenerate if and only if $\operatorname{Ker}\phi_L = \{0\}$ and $\operatorname{Ker}\phi_R = \{0\}$. But this implies $\operatorname{null}(A^T) = 0$ and $\operatorname{null}(A) = 0$, hence by rank-nullity theorem, we must have $\operatorname{rank}(A^T) = \dim U$, and $\operatorname{rank}(A) = \dim V$. But this gives A invertible and $\dim U = \dim V$.

Remark. Taking $\phi : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ by the scalar product, then ϕ is non-degenerate, as in the standard basis $\mathcal{B}$,

$$[\phi]_{\mathcal{B},\mathcal{B}} = I_n.$$

Corollary 10.2. *When U and V are finite dimensional, then choosing a non-degenerate bilinear form $\phi : U \times V \to K$ is equivalent to choosing an isomorphism $\phi_L : U \to V^*$.*

Definition 10.5. If $T \subset U$, we define

$$T^\perp = \{v \in V \mid \phi(t, v) = 0 \,\forall t \in T\}.$$

Similarly, if $S \subset V$, then

$${}^\perp S = \{u \in U \mid \phi(u, s) = 0 \,\forall s \in S\}.$$

Proposition 10.1. *Let $\mathcal{B}, \mathcal{B}'$ be two bases of U, and $P = [\mathrm{id}]_{\mathcal{B}',\mathcal{B}}$, and $\mathcal{C}, \mathcal{C}'$ two bases of V, and $Q = [\mathrm{id}]_{\mathcal{C}',\mathcal{C}}$, then if $\phi : U \times V \to K$ is a bilinear form, then*

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

Proof: We have

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

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

Definition 10.6. The rank of ϕ ($\operatorname{rank}\phi$) is the rank of any matrix representing ϕ.

This is true as $\operatorname{rank}(P^T AQ) = \operatorname{rank} A$, if P and Q are invertible.

Note we could have equivalently defined $\operatorname{rank}\phi = \operatorname{rank}\phi_L = \operatorname{rank}\phi_R$.

11 Determinant and Traces

Definition 11.1. If $A \in M_n(K)$, we define the trace of A as

$$\operatorname{tr} A = \sum_{i=1}^{n} A_{ii}.$$

Remark. The map $M_n(K) \to K$ by $A \mapsto \operatorname{tr} A$ is linear.

Lemma 11.1. $\operatorname{tr}(AB) = \operatorname{tr}(BA)$.

Proof:

$$\operatorname{tr}(AB) = \sum_{i=1}^{n}\left(\sum_{j=1}^{n} a_{ij}b_{ji}\right) = \sum_{j=1}^{n}\sum_{i=1}^{n} b_{ji}a_{ij} = \operatorname{tr}(BA).$$

Corollary 11.1. *Similar matrices have the same trace, as*

$$\operatorname{tr}(P^{-1}AP) = \operatorname{tr}(APP^{-1}) = \operatorname{tr}(A).$$

Definition 11.2. If $\alpha : V \to V$ is linear, we can define

$$\operatorname{tr} \alpha = \operatorname{tr}([\alpha]_{\mathcal{B}})$$

in any basis $\mathcal{B}$.

Lemma 11.2. *If $\alpha : V \to V$ with $\alpha^* : V^* \to V^*$ the dual map,*

Proof:

$$\operatorname{tr} \alpha = \operatorname{tr}([\alpha]_{\mathcal{B}}) = \operatorname{tr}([\alpha]_{\mathcal{B}}^T) = \operatorname{tr}([\alpha^*]_{\mathcal{B}^*}.$$

12 Determinants

12.1 Permutations and Transposition

We define S_n as the symmetric group, the permutations of $X = \{1, \dots, n\}$.

The transposition $\tau_{k,\ell} \in S_n$ for $k \neq l$ is $\tau_{k,\ell} = (k, \ell)$.

Then we know any permutation σ can be decomposed as a product of transpositions:

$$\sigma = \prod_{i=1}^{n_\sigma} \tau_i.$$

The signature is a map

$$\begin{aligned} \varepsilon : S_n &\to \{\pm 1\} \\ \sigma &\mapsto \begin{cases} 1 & n_\sigma \text{ even}, \\ -1 & n_\sigma \text{ odd}. \end{cases} \end{aligned}$$

Definition 12.1. For $A \in M_n(K)$, and $A = (a_{ij})$, we define the determinant of A as

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

Example 12.1.

For $n = 2$, we have

$$\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_{11}a_{22} - a_{12}a_{21}.$$

Lemma 12.1. *If $A = (a_{ij})$ is an upper (or lower) triangular matrix with 0 on the diagonal, then $\det A = 0$.*

Proof:

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

For the summands not to be 0, we need $\sigma(j) < j$ for all $j \in \{1, \dots, n\}$. But this is impossible for all $\sigma \in S_n$, so all summands are 0, and $\det A = 0$.

Similarly, if A is upper-triangular, not necessarily with 0's on the diagonal, then the summands are non-zero only if $\sigma(j) \leq j$ for all $j \in \{1, \dots, n\}$.

By induction and the fact σ is a permutation, we get $\sigma(j) = j$ for all $j \in \{1, \dots, n\}$ and the only term that doesn't vanish is $a_{11}a_{22}\cdots a_{nn}$. Hence

$$\det \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} = \prod_{i=1}^{n} \lambda_i = \det \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ * & & \lambda_n \end{pmatrix}.$$

Lemma 12.2. $\det A = \det(A^T)$.

Proof:

$$\begin{aligned} \det A &= \sum_{\sigma \in S_n} \varepsilon(\sigma) a_{\sigma(1)1} \cdots a_{\sigma(n)n} = \sum_{\sigma \in S_n} \varepsilon(\sigma) \prod_{i=1}^{n} a_{\sigma(i)i} \\ &= \sum_{\sigma \in S_n} \varepsilon(\sigma) \prod_{i=1}^{n} a_{i\sigma^{-1}(i)}. \end{aligned}$$

Now remember $1 = \varepsilon(\sigma\sigma^{-1}) = \varepsilon(\sigma)\varepsilon(\sigma^{-1})$, so $\varepsilon(\sigma^{-1}) = \varepsilon(\sigma)$. Hence

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

Our definition of $\det A$ has seemingly come out of nowhere. We want some reason to take this as our definition.

Definition 12.2. A volume form on K^n is a function

$$\underbrace{K^n \times \cdots \times K^n}_{n \text{ times}} \to K,$$

such that

(i) It is multilinear, so for any $1 \le i \le n$, and all $(v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_n) \in (K^n)^{n-1}$, we want the map

$$\begin{aligned} K^n &\to K \\ v &\mapsto d(v_1, \dots, v_{i-1}, v, v_{i+1}, \dots, v_n) \end{aligned}$$

to be linear.

(ii) It is alternate, so if $v_i = v_j$ for some $i \neq j$, then

$$d(v_1, \dots, v_n) = 0.$$

Then we want to show that there is in fact only one volume form on $K^n \times \cdots \times K^n$ given by the determinant: If $A = (a_{ij}) = (A^{(1)} \mid \dots \mid A^{(n)})$, then we denote

$$\det A = \det(A^{(1)}, \dots, A^{(n)}).$$

Lemma 12.3. *$K^n \times \cdots \times K^n \to K$ by $(A^{(1)}, \dots, A^{(n)}) \mapsto \det A$ is a volume form.*

Proof:

(i) Firstly, this map is multilinear. Pick $\sigma \in S_n$. Then the individual summands $\prod_{i=1}^n a_{\sigma(i)i}$ are multilinear, as there is only one term from each column appearing in this expression.

Since the sum of multilinear maps is multilinear, det is multilinear.

(ii) Now we show the map is alternate. Assume $k \neq \ell$, and $A^{(k)} = A^{(\ell)}$. Then we want to show $\det A = 0$. Let $\tau = (k, \ell)$ be a transposition. Then note $A^{(k)} = A^{(\ell)} \iff a_{ij} = a_{i\tau(j)}$ for all $i \in \{1, \dots, n\}$.

We can decompose $S_n = A_n \cup \tau A_n$. Here A_n is the alternating group, which are the permutations with an even number of transpositions, and τA_n are the permutations with an odd number of transpositions. Thus,

$$\begin{aligned}
\det A &= \sum_{\sigma \in S_n} \varepsilon(\sigma) \prod_{i=1}^n a_{i\sigma(i)} = \sum_{\sigma \in A_n} \prod_{i=1}^n a_{i\sigma(i)} + \sum_{\sigma \in \tau A_n} - \prod_{i=1}^n a_{i\sigma(i)} \\
&= \sum_{\sigma \in A_n} \prod_{i=1}^n a_{i\sigma(i)} - \sum_{\sigma \in A_n} a_{i\tau\sigma(i)} = \sum_{\sigma \in A_n} \left(\prod_{i=1}^n a_{i\sigma(i)} - \prod_{j=1}^n a_{i\sigma(i)} \right) \\
&= 0.
\end{aligned}$$

Lemma 12.4. *Let d be a volume form. Then swapping two entries changes the sign, so*

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

Proof:

$$\begin{aligned}
0 &= d(v_1, \dots, v_i + v_j, \dots, v_i + v_j, \dots, v_n) \\
&= d(v_1, \dots, v_i, \dots, v_i, \dots, v_n) + \quad + d(v_1, \dots, v_i, \dots, v_j, \dots, v_n) \\
&\quad + d(v_1, \dots, v_j, \dots, v_i, \dots, v_n) + d(v_1, \dots, v_j, \dots, v_j, \dots, v_n) \\
&= d(v_1, \dots, v_i, \dots, v_j, \dots, v_n) + d(v_1, \dots, v_j, \dots, v_i, \dots, n).
\end{aligned}$$

Corollary 12.1. *If $\sigma \in S_n$, and d is a volume form, then*

$$d(v_{\sigma(1)}, \dots, v_{\sigma(n)}) = \varepsilon(\sigma) d(v_1, \dots, v_n).$$

This follows as σ is a product of transpositions.

Theorem 12.1. *Let d be a volume form on K^n, and let $A = (A^{(1)}, \dots, A^{(n)})$. Then,*

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

Hence, up to a constant, det is the only volume form on K^n.

Proof:

$$\begin{aligned}
d(A^{(1)}, \dots, A^{(n)}) &= d\left(\sum_{i=1}^{n} a_{i1} e_1, \dots, A^{(n)}\right) = \sum_{i=1}^{n} a_{i1} d(e_1, A^{(2)}, \dots, A^{(n)}) \\
&= \sum_{i=1}^{n} a_{i1} \left(e_1, \sum_{j=1}^{n} a_{j2} e_2, \dots, A^{(n)}\right) \\
&= \sum_{i=1}^{n} \sum_{j=1}^{n} a_{i1} a_{j2} d(e_i, e_j, \dots, A^{(n)}) \\
&= \sum_{1 \le i_k \le n} \left(\prod_{k=1}^{n} a_{i_k k}\right) d(e_{i_1}, e_{i_2}, \dots, e_{i_n}).
\end{aligned}$$

This last term is non-zero if and only if all i_k are distinct, meaning there exists $\sigma \in S_n$ such that $i_k = \sigma(k)$. This means

$$\begin{aligned} d(A^{(1)}, \dots, A^{(n)}) &= \sum_{\sigma \in S_n} \prod_{k=1}^{n} a_{\sigma(k)k} d(e_{\sigma(1)}, \dots, e_{\sigma(n)}) \\ &= \sum_{\sigma \in S_n} \left[\prod_{k=1}^{n} a_{\sigma(k)k} \right] \varepsilon(\sigma) d(e_1, \dots, e_n) \\ &= d(e_1, \dots, e_n) \det A. \end{aligned}$$

Corollary 12.2. det *is the only volume form such that* $d(e_1, \dots, e_n) = 1$.

13 Some Properties of Determinants

Lemma 13.1. *If $A, B \in M_n(F)$, then*

$$\det(AB) = (\det A)(\det B).$$

Proof: Pick A. Consider the map $d_a : \underbrace{K^n \times \cdots \times K^n}_{n} \to K$ defined by

$$d_A(v_1, \ldots, v_n) = \det(Av_1, \ldots, Av_n).$$

Then d_A is multilinear and alternate, as $v_i \mapsto Av_i$ is linear, and $v_i = v_j \implies Av_i = Av_j$. Thus, there exists C such that

$$d_A(v_1, \ldots, v_n) = C \det(v_1, \ldots, v_n).$$

Computing on the canonical basis,

$$d_A(e_1, \ldots, e_n) = \det(Ae_1, \ldots, Ae_n) = \det(A_1, \ldots, A_n) = \det A.$$

Hence, $C = \det A$.

Now observe $AB = ((AB)_1, \ldots, (AB)_n)$, so

$$\det(AB) = \det(AB_1, \ldots, AB_n) = (\det A) \det(B_1, \ldots, B_n) = (\det A)(\det B).$$

Definition 13.1. For $A \in M_n(K)$, we say that

(i) A is singular if $\det A = 0$,

(ii) A is non-singular if $\det A \neq 0$.

Lemma 13.2. *A is invertible implies A is non-singular.*

Proof: If A is invertible, then there exists A^{-1} such that $AA^{-1} = A^{-1}A = I_n$. Thus

$$(\det A)(\det A^{-1}) = \det(AA^{-1}) = \det I_n = 1,$$

so $\det A \neq 0$.

Remark. This also prove $\det A^{-1} = (\det A)^{-1}$.

Theorem 13.1. *Let $A \in M_n(K)$. Then the following are equivalent:*

(i) *A is invertible;*

(ii) *A is non-singular;*

(iii) $r(A) = n$.

Proof: We have already seen (i) $\iff$ (iii), from rank nullity, and we have just shown (i) $\implies$ (ii). Thus it suffices to show (ii) $\implies$ (iii). Indeed, assume $r(a) < n$.

Then $r(A) < n \iff \dim \operatorname{span}\{c_1, \ldots, c_n\} < n$, so there exists $(\lambda_1, \ldots, \lambda_n) \neq (0, \ldots, 0)$ such that

$$\sum_{i=1}^{n} \lambda_i c_i = 0.$$

Pick j with $\lambda_j \neq 0$. Then,

$$c_j = \frac{1}{\lambda_j} \sum_{i \neq j} \lambda_i c_i.$$

This gives

$$\det A = \det(c_1, \ldots, c_j, \ldots, c_n) = \det\left(c_1, \ldots, \frac{-1}{\lambda_j} \sum_{i \neq j} \lambda_i c_i, \ldots, c_n\right) = 0.$$

Hence by contrapositive, (ii) $\implies$ (iii).

Remark. This gives a sharp criterion for invertibility of a linear system of n equations with n unknowns.

13.1 Determinant of linear maps

Lemma 13.3. *Conjugate matrices have the same determinant.*

Proof:

$$\det(P^{-1}AP) = \det P^{-1} \det A \det P = \det A \det(P^{-1}P) = \det A.$$

Definition 13.2. For $\alpha : V \to V$ linear, we define $\det \alpha = \det([\alpha]_{\mathcal{B}})$.

Theorem 13.2. $\det : L(V, V) \to K$ *satisfies*

(i) $\det \mathrm{id} = 1$*;*

(ii) $\det(\alpha \circ \beta) = (\det \alpha)(\det \beta)$*;*

(iii) $\det\alpha \neq 0$ *if and only if* α *is invertible, and then* $\det(\alpha^{-1}) = (\det\alpha)^{-1}$.

Proof: Pick a basis $\mathcal{B}$ and express in terms of $[\alpha]_{\mathcal{B}}, [\beta]_{\mathcal{B}}$.

13.2 Determinant of Block Matrices

Lemma 13.4. *For* $A \in M_k(K)$*,* $B \in M_\ell(K)$*, and* $C \in M_{k,\ell}(K)$*, let*

$$M = \begin{pmatrix} A & C \\ 0 & B \end{pmatrix} \in M_n(K).$$

Then, $\det M = (\det A)(\det B)$.

Proof: We know that

$$\det M = \sum_{\sigma \in S_n} \varepsilon(\sigma) \prod_{i=1}^{n} m_{\sigma(i)i}.$$

Observe, that $m_{\sigma(i)i} = 0$ if $i \leq k$, and $\sigma(i) > k$. Hence,we only need to sum over $\sigma \in S_n$ such that

(i) For all $j \in [1,k]$, $\sigma(j) \ni [1,k]$;

(ii) For all $j \in [k+1,n], \sigma(j) \in [k+1,n]$.

In other words, we restrict σ to $\sigma_1 : \{1,\dots,k\} \to \{1,\dots,k\}$ and $\sigma_2 : \{k+1,\dots,n\} \to \{k+1,\dots,n\}$. Hence

$$m_{\sigma(j)j} = \begin{cases} a_{\sigma_1(j)j} & j \leq k, \\ b_{\sigma_2(j)(j)} & j \geq k. \end{cases}.$$

We know $\varepsilon(\sigma) = \varepsilon(\sigma_1)\varepsilon(\sigma_2)$. So

$$\begin{aligned}
\det M = \sum_{\sigma \ni S_n} \varepsilon(\sigma) \prod_{i=1}^{n} m_{\sigma(i)i} &= \sum_{\substack{\sigma_1 \in S_k, \\ \sigma_2 \in S_\ell}} \varepsilon(\sigma_1 \circ \sigma_2) \prod_{i=1}^{k} a_{\sigma_1(i)i} \prod_{j=k+1}^{n} b_{\sigma_2(j)j} \\
&= \left(\sum_{\sigma_1 \in S_k} \varepsilon(\sigma_1) \prod_{i=1}^{k} a_{\sigma_1(i)i} \right) \left(\sum_{\sigma_2 \in S_\ell} \varepsilon(\sigma_2) \prod_{j=k+1}^{n} b_{\sigma_2(j)j} \right) \\
&= (\det A)(\det B).
\end{aligned}$$

Corollary 13.1. *If $A_1, \ldots, A_k$ are square matrices, then*

$$\det \begin{pmatrix} A_1 & * & \cdots & * \\ 0 & A_2 & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_k \end{pmatrix} = (\det A_1) \cdots (\det A_k).$$

This follows from induction on the number of matrices. In particular, if A is upper-triangular with λ_i on the diagonals, then $\det A = \prod \lambda_i$.

However, note that in general,

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

Remark. In 3 dimensions, $(a, b, c) \mapsto (a \times b) \cdot c$ is a volume form. Thus we can show $\det(a, b, c) = (a \times b) \cdot c$.

14 Adjugate Matrix

Observe by swapping two columns in $A = (A^{(1)}, \ldots, A^{(n)})$, the determinant alternates parity. Using the fact that $\det A = \det(A^T)$, we can see that swapping two rows also changes the parity of the determinant.

14.1 Column expansion and Adjugate Matrix

It is hard to compute the determinant using our current definitions. Using column expansion, we can reduce the computation of $n \times n$ determinants to $(n-1) \times (n-1)$ determinants.

Definition 14.1. Let $A \in M_n(K)$. Pick $i, j \in \{1, \ldots, n\}$. We define $A_{\widehat{ij}} \in M_{n-1}(K)$, obtained by removing the i'th row and j'th column from A.

Lemma 14.1. *Let $A \in M_n(K)$.*

(i) *Pick $1 \leq j \leq n$, then:*

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

(ii) *Pick $1 \leq i \leq n$, then:*

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

Proof: We prove expansion with respect to the j'th column. Then row expansion will follow by taking the transpose. First, we can write $A = (A^{(1)}, \ldots, A^{(j)}, \ldots, A^{(n)})$. Then,

$$A^{(j)} = \sum_{i=1}^{n} a_{ij} e_i.$$

Hence we get

$$\det\left(A^{(1)}, \ldots, \sum_{i=1}^{n} a_{ij} e_i, \ldots, A^{(n)}\right) = \sum_{i=1}^{n} a_{ij} \det(A^{(1)}, \ldots, e_i, \ldots, A^{(n)}).$$

Now, we can compute:

$$\begin{aligned}
\det(A^{(1)}, \ldots, e_i, \ldots, A^{(n)}) &= \det \begin{pmatrix} & & 0 & & \\ & & \vdots & & \\ A^{(1)} & \cdots & 1 & \cdots & A^{(n)} \\ & & \vdots & & \\ & & 0 & & \end{pmatrix} \\
&= (-1)^{j-1} \det \begin{pmatrix} 0 & & & & & & \\ \vdots & & & & & & \\ 1 & A^{(1)} & \cdots & A^{(j-1)} & A^{(j+1)} & \cdots & A^{(n)} \\ \vdots & & & & & & \\ 0 & & & & & & \end{pmatrix} \\
&= (-1)^{i-1}(-1)^{j-1} \det \begin{pmatrix} 1 & a_{i1} & \cdots & a_{ij-1} & a_{ij+1} & \cdots & a_{in} \\ 0 & & & & & & \\ \vdots & & & A_{\widehat{ij}} & & & \\ 0 & & & & & & \end{pmatrix} \\
&= (-1)^{i+j} \det(A_{\widehat{ij}}).
\end{aligned}$$

Combining these facts,

$$\det A = \sum_{i=1}^{n} a_{ij} \det(A^{(1)}, \ldots, a_i, \ldots, A^{(n)}) = \sum_{i=1}^{n} a_{ij} (-1)^{i+j} \det_{\widehat{ij}}.$$

Definition 14.2. Let $A \in M_n(K)$. The adjugate matrix $\operatorname{adj} A$ is the $n \times n$ matrix with (i, j) entry given by $(-1)^{i+j} \det(A_{\widehat{ji}})$.

Theorem 14.1. *Let $A \in M_n(K)$. Then,*

$$(\operatorname{adj} A)A = (\det A)I_n.$$

In particular, when A is invertible,

$$A^{-1} = \frac{1}{\det A} \operatorname{adj} A.$$

Proof: From what we have just proven,

$$\det A = \sum_{i=1}^{n} (-1)^{i+j} (\det A_{\widehat{ij}}) a_{ij} = \sum_{i=1}^{n} (\operatorname{adj} A)_{ji} a_{ij} = (\operatorname{adj}(A)A)_{jj}.$$

Now for $j \neq k$, we have

$$\begin{aligned}
0 &= \det(A^{(1)}, \dots, A^{(k)}, \dots, A^{(k)}, \dots, A^{(n)}) \\
&= \det\left(A^{(1)}, \dots, \sum_{i=1}^{n} a_{ik} e_i, \dots, A^{(k)}, \dots, A^{(n)}\right) \\
&= \sum_{i=1}^{n} a_{ik} \det(A^{(1)}, \dots, e_i, \dots, A^{(n)}) \\
&= \sum_{i=1}^{n} (\operatorname{adj} A)_{ji} a_{ik} = ((\operatorname{adj} A)A)_{jk}.
\end{aligned}$$

Proposition 14.1. *Let $A \in M_n(K)$ be invertible, and $b \in K^n$. 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 obtained by replacing the i'th column of A by b.

Algorithmically, this avoids computing A^{-1}.

Proof: If A is invertible, then there exists unique $x \in K^n$ with $Ax = b$. Let x be this solution, then

$$\begin{aligned}
\det(A_{\hat{i}b}) &= \det(A^{(1)}, \dots, A^{(i-1)}, b, A^{(i+1)}, \dots, A^{(n)}) \\
&= \det(A^{(1)}, \dots, A^{(i-1)}, Ax, A^{(i+1)}, \dots, A^{(n)}) \\
&= \det\left(A^{(1)}, \dots, A^{(i-1)}, \sum_{j=1}^{n} x_j A^{(j)}, A^{(i+1)}, \dots, A^{(n)}\right) \\
&= x_i \det(A^{(1)}, \dots, A^{(i-1)}, A^{(i)}, A^{(i+1)}, \dots, A^{(n)}) = x_i \det A.
\end{aligned}$$

Inverting, this gives

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

15 Eigenvectors, Eigenvalues and Triangular Matrices

Here, we set up towards our goal of the diagonalisation of endomorphisms. Let V be a vector space over K, and $\dim V = n < +\infty$. Then recall $\alpha : V \to V$ linear is an endomorphism of V.

We want to find a basis $\mathcal{B}$ of V such that in this basis,

$$[\alpha]_{\mathcal{B}} = [\alpha]_{\mathcal{B},\mathcal{B}}$$

has a "nice" form.

Recall that for another basis $\mathcal{B}'$ of V, the change of basis matrix satisfies

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

Equivalently, given a matrix $A \in M_n(K)$, we want to find whether it is conjugate to a matrix with a "simple" form.

Definition 15.1.

(i) $\alpha \in L(V)$ is *diagonalisable* if there exists a basis $\mathcal{B}$ of V such that $[\alpha]_{\mathcal{B}}$ is diagonal.

(ii) $\alpha \in L(V)$ is *triangulable* if there exists a basis $\mathcal{B}$ of V such that $[\alpha]_{\mathcal{B}}$ is triangular:
$$[\alpha]_{\mathcal{B}} = \begin{pmatrix} \lambda_1 & * & \cdots & * \\ 0 & \lambda_2 & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

Remark. A matrix is diagonalisable (resp. triangulable) if and only if it is conjugate to a diagonal (resp. triangular) matrix.

Definition 15.2.

(i) $\lambda \in K$ is an *eigenvalue* of $\alpha \in L(V)$ if there exists $v \in V \setminus \{0\}$ such that $\alpha(v) = \lambda v$.

(ii) $v \in V$ is an *eigenvector* of $\alpha \in L(V)$ if and only if $v \neq 0$ and there exists $\lambda \in K$ such that $\alpha(v) = \lambda v$.

(iii) $V_\lambda = \{v \in V \mid \alpha(v) = \lambda v\} \leq V$ is the *eigenspace* associated to $\lambda \in K$.

Lemma 15.1. *Let $\alpha \in L(V)$ and $\lambda \in K$. Then*

$$\lambda \text{ is an eigenvalue of } \alpha \iff \det(\alpha - \lambda \,\mathrm{id}) = 0.$$

Proof: If λ is an eigenvalue, then we have a chain of equalities

$$\begin{aligned} &\lambda \text{ eigenvalue} \\ \iff &\exists v \in V \setminus \{0\}, \alpha(v) = \lambda v \\ \iff &\exists v \in V \setminus \{0\}, (\alpha - \lambda \,\mathrm{id})(v) = 0 \\ \iff &\ker(\alpha - \lambda \,\mathrm{id}) \neq \{0\} \\ \iff &r(\alpha - \lambda \,\mathrm{id}) < n \\ \iff &\det(\alpha - \lambda \,\mathrm{id}) = 0. \end{aligned}$$

Remark. If $\alpha(v_j) = \lambda v_j$, and $v_j \neq 0$, then we can complete to a basis of V $(v_1, \dots, v_j, \dots, v_n)$, such that

$$[\alpha]_{\mathcal{B}}(A^{(1)}, \dots, e_j, \dots, A^{(n)}).$$

15.1 Polynomials

We will look at how polynomials interact with $\alpha \in L(V)$. First, if K is a field, and

$$f(t) = a_n t^n + a_{n-1} t^{n-1} + \cdots + a_1 t + a_0,$$

with $a_i \in K$, then n is the largest exponent such that $a_n \neq 0$. We say $n = \deg f$. Then, we can easily show

$$\deg(f+g) \leq \max\{\deg f, \deg g\}, \quad \deg(fg) = \deg f + \deg g.$$

Define $K[t]$ as the ring of polynomials with coefficients in K. Then λ is a root of $f(t) \iff f(\lambda) = 0$.

Lemma 15.2. *If λ is a root of f, then $t - \lambda$ divides f.*

Proof: Write $f(t) = a_n t^n + \cdots + a_1 t + a_0$, then $f(\lambda) = a_n \lambda^n + \cdots + a_1 \lambda + a_0 = 0$. Hence,

$$\begin{aligned} f(t) = f(t) - f(\lambda) &= a_n(t^n - \lambda^n) + \cdots + a_1(t - \lambda) \\ &= a_n(t-\lambda)(t^{n-1} + \cdots + \lambda^{n-1}) + \cdots + a_1(t-\lambda) \\ &= (t-\lambda)g(t). \end{aligned}$$

Corollary 15.1. *A non-zero polynomial of degree n has at most n roots.*

This follows from induction of degree, and the above lemma.

Corollary 15.2. *If f_1, f_2 are polynomials of degree less than n, such that $f_1(t_i) = f_2(t_i)$ for at least n values (t_i), then $f_1 = f_2$.*

This follows from the above corollary on $f_1 - f_2$.

Theorem 15.1. *Any $f \in \mathbb{C}[t]$ of positive degree has a complex root (hence exactly* $\deg f$ *roots when counted with multiplicity).*

This will be proved in complex analysis.

Definition 15.3. Let $\alpha \in L(V)$. The *characteristic polynomial* of α is

$$\chi_\alpha(t) = \det(A - t\,\mathrm{id}).$$

Remark. We can visualise

$$A - t\,\mathrm{id} = \begin{pmatrix} a_{11} - t & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - t & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} - t \end{pmatrix}.$$

The fact that $\det(A - t\,\mathrm{id})$ is a polynomial of degree n comes from the definition of det.

Moreover, notice that conjugate matrices have the same characteristic polynomial:

$$\det(P^{-1}AP - \lambda\,\mathrm{id}) = \det(P^{-1}(A - \lambda\,\mathrm{id})P) = \det(A - \lambda\,\mathrm{id}).$$

Hence $\chi_\alpha(t) = \det(A - \lambda\,\mathrm{id})$ does not depend on the basis $\mathcal{B}$ in which we express α.

Theorem 15.2. *$\alpha \in L(V)$ is triangulable if and only if χ_α can eb written as a product of linear factors over K:*

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

In particular, over $K = \mathbb{C}$, any matrix is triangulable.

Proof: Suppose α is triangulable. Then in some basis, we have

$$[\alpha]_{\mathcal{B}} = \begin{pmatrix} \lambda_1 & * & \cdots & * \\ 0 & \lambda_2 & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}.$$

Then we can expand

$$\chi_\alpha(t) = \begin{pmatrix} \lambda_1 - t & * & \cdots & * \\ 0 & \lambda_2 - t & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n - t \end{pmatrix} = \prod_{i=1}^{n}(\lambda_i - t).$$

For the backwards direction, we argue by induction on $n = \dim V$. If $n = 1$, then the conclusion is obvious. So suppose $n > 1$.

By assumption let $\chi_\alpha(t)$ have a root λ. Then note $\chi_\alpha(\lambda) = 0 \iff \lambda$ is an eigenvalue of α. Let $U = V_\lambda$ be the associated eigenspace, and note $\{0\} \subsetneq U$.

Let $(v_1, \ldots, v_k)$ be a basis of U, and complete to a basis $(v_1, \ldots, v_k, v_{k+1}, \ldots, v_n)$ of V. Let $\operatorname{span}(v_{k+1}, \ldots, v_n) = W$, then $V = U \oplus W$. In $\mathcal{B}$, we have

$$[\alpha]_{\mathcal{B}} = \begin{pmatrix} \lambda I_k & * \\ 0 & C \end{pmatrix}.$$

α induces an endomorphism

$$\bar{\alpha} : V/U \to V/U.$$

Then $C = [\bar{\alpha}]_{\bar{\mathcal{B}}}$, where $\bar{\mathcal{B}} = (v_{k+1} + U, \ldots, v_n + U)$. Then, as this is a block product,

$$\det(\alpha - t\,\mathrm{id}) = \det \begin{pmatrix} (\lambda - t)\,\mathrm{id} & * \\ 0 & C - t\,\mathrm{id} \end{pmatrix}$$

$$= (\lambda - t)^k \det(C - t\,\mathrm{id}) = c\prod_{i=1}^{n}(t - \lambda_i).$$

From uniqueness of factorisation, we can determine

$$\det(C - t\,\mathrm{id}) = \tilde{c} \prod_{i=k+1}^{n} (t - \tilde{\lambda}_i).$$

Hence, by induction (as $\dim V/U < \dim V$), there is a basis $\check{\mathcal{B}} = (\check{v}_{k+1}, \ldots, \check{v}_n)$ of W where $[\mathcal{C}]_{\check{\mathcal{B}}}$ is triangular.

Hence letting $\hat{\mathcal{B}} = (v_1, \ldots, v_k, \check{v}_{k+1}, \ldots, \check{v}_n)$, $[\alpha]_{\hat{\mathcal{B}}}$ is triangular.

Lemma 15.3. *If V is n-dimensional over $K = \mathbb{R}, \mathbb{C}$, and $\alpha \in L(V)$, then if $\chi_\alpha(t) = (-1)^n t^n + c_{n-1} t^{n-1} + \cdots + c_0$, we have*

$$c_0 = \det A = \det \alpha, \quad c_{n-1} = (-1)^{n-1} \operatorname{tr} A.$$

Proof: We known $\chi_\alpha(t) = \det(a - t \operatorname{id})$, so $\chi_\alpha(0) = \det \alpha = c_0$.

Say that $K = \mathbb{C}$. We known α is triangulable over $\mathbb{C}$, so

$$\begin{aligned}
\chi_\alpha(t) = \det \begin{pmatrix} a_1 - t & \cdots & * \\ \vdots & \ddots & \vdots \\ 0 & \cdots & a_n - t \end{pmatrix} &= \prod_{i=1}^n (a_i - t) \\
&= (-1)^n t^n + c_{n-1} t^{n-1} + \cdots + c_0, \\
c_{n-1} = (-1)^{n-1} \sum_{i=1}^n a_i &= (-1)^{n-1} \operatorname{tr} \alpha.
\end{aligned}$$

16 Diagonalisation Matrix and Minimal Polynomial

Definition 16.1. Pick $p(t)$, a polynomial over K, with $p(t) = a_n t^n + \cdots + a_1 t + a_0$ Hence if $A \in M_n(K)$, then $A^m \in M_n(K)$, and we define

$$p(A) = a_n A^n + \cdots + a_1 A + a_n \operatorname{id} \in M_n(K).$$

Similarly, for $\alpha \in L(V)$, we can define

$$p(\alpha) = a_n \alpha^n + \cdots + a_1 \alpha + a_n \operatorname{id}.$$

Theorem 16.1. *Let V be a vector space over K, with $\dim V < \infty$, and $\alpha \in L(V)$.*

Then α is diagonalizable if and only if there exists a polynomial which is a product of distinct linear factors such that $p(\alpha) = 0$.

In other words α is diagonalizable if and only if there exist distinct $(\lambda_1, \dots, \lambda_k)$, with $\lambda_j \in K$, such that

$$p(\alpha) = (\alpha - \lambda_1 \operatorname{id}) \cdots (\alpha - \lambda_k \operatorname{id}) = 0.$$

Proof: Suppose α is diagonalizable with $\lambda_1, \dots, \lambda_k$ the distinct polynomial. Let $p(t) = \prod (t - \lambda_i)$, and let $\mathcal{B}$ be a basis of V made of eigenvectors of α. Then for $v \in \mathcal{B}$, we have $\alpha(v) = \lambda_i v$ for some $i \in \{1, \dots, k\}$, so

$$(\alpha - \lambda_i \operatorname{id})(v) = 0.$$

Hence

$$p(\alpha) = \left[\prod_{j=1}^{k} (\alpha - \lambda_j \operatorname{id}) \right] (v) = 0$$

for all $v \in \mathcal{B}$. But since $\mathcal{B}$ is a basis, then by linearity, for all $v \in K$, $p(\alpha)(v) = 0$, so $p(\alpha) = 0$.

Now suppose $p(\alpha) = 0$ for

$$p(t) = \prod_{i=1}^{k} (t - \lambda_i),$$

where $\lambda_i \neq \lambda_j$. Let $V_{\lambda_i} = \operatorname{Ker}(\alpha - \lambda_i \operatorname{id})$. Then we claim

$$V = \bigoplus_{i=1}^{k} V_{\lambda_i}.$$

Indeed, let

$$q_j(t) = \prod_{\substack{i=1 \\ i\neq j}}^{k} \left(\frac{t-\lambda_i}{\lambda_j - \lambda_i} \right).$$

Then we have

$$q_j(\lambda_i) = \begin{cases} 1 & i = j, \\ 0 & i \neq j. \end{cases}$$

Consider the polynomial

$$q(t) = \sum_{j=1}^{k} q_j(t).$$

Then $\deg q_j \leq j-1$, so $\deg q \leq k-1$. On the other hand, $q(\lambda_i) = 1$ for all i. Hence the polynomial $[q(t)-1]$ degree less than or equal to $k-1$, and has at least k roots, so for all t, $q(t) = 1$. Thus, we have

$$q_1(t) + \cdots + q_k(t) = 1.$$

Define the projector

$$\mathbb{T}_j = q_j(\alpha) \in L(V).$$

Then we have

$$\sum_{j=1}^{k} \mathbb{T}_j = \sum_{j=1}^{k} q_j(\alpha) = \left(\sum_{j=1}^{k} q_j \right)(\alpha) = \mathrm{id}$$

This means for any vector $v \in V$,

$$v = q(\alpha)(v) = \sum_{j=1}^{k} \mathbb{T}_j(v) = \sum_{j=1}^{k} q_j(\alpha)(v).$$

Now if we pick $j \in \{1, \ldots, k\}$, then

$$(\alpha - \lambda_j \,\mathrm{id}) q_j(\alpha)(v) = \frac{1}{\prod_{i\neq j}(\lambda_j - \lambda_i)} p(\alpha)(v) = 0.$$

Thus for all $j \in \{1, \ldots, k\}$, $(\alpha - \lambda_j \,\mathrm{id})\mathbb{T}_j(v) = 0$, so $\mathbb{T}_j(v) \in V_{\lambda_j}$ for all v. Now for all $v \in V$,

$$v = \sum_{j=1}^{k} \mathbb{T}_j(v) \implies V = \sum_{j=1}^{k} V_{\lambda_j}.$$

Now we prove the sum is direct. Indeed, let $v \in V_{\lambda_j} \cap (\sum_{i \neq j} V_{\lambda_i})$. Then since $v \in V_{\lambda_j}$,

$$\mathbb{T}_j(v) = q_j(\alpha)(v) = \prod_{i \neq j} \frac{\alpha - \lambda_i \operatorname{id}}{\lambda_i - \lambda_j}(v) = \prod_{i \neq j} \frac{(\lambda_j - \lambda_i)v}{\lambda_j - \lambda_i} = v.$$

Now if $v \in \sum_{i \neq j} V_{\lambda_i}$, then note for $v \in V_{\lambda_i}$, then $\alpha(v) = \lambda_i v$ so

$$\mathbb{T}_j(\alpha) = q_j(\alpha)(v) = \prod_{i \neq j} \frac{\alpha - \lambda_i \operatorname{id}}{\lambda_j - \lambda_i}(v) = 0.$$

Hence if $v \in V_{\lambda_j} \cap (\sum_{i \neq j} V_{\lambda_i})$, then $v = 0$. Hence V is a direct sum of eigenspaces, meaning it can be diagonalized.

Remark. We have actually proved the following: If $\lambda_1, \ldots, \lambda_k$ are distinct eigenvalues of α then

$$\sum_{i=1}^{k} V_{\lambda_i} = \bigoplus_{i=1}^{k} V_{\lambda_i}.$$

This means that the only way digitalization fails is if the sum of the eigenspaces is a proper subspace of V.

Example 16.1.

For $A \in M_n(K)$, with A having finite order m, then A is diagonalizable, as A is a root of

$$t^m - 1 = \prod_{j=1}^{m} (t - \zeta_m^j).$$

Theorem 16.2. *For* $\dim V < +\infty$ *and* $\alpha, \beta \in L(V)$ *diagonalizable, then* α, β *are simultaneously diagonalizable if and only if* α *and* β *commute.*

Proof: First, if α, β are simultaneously diagonalizable, then there is a basis of V in which

$$[\alpha]_{\mathcal{B}} = D_1, \quad [\beta]_{\mathcal{B}} = D_2.$$

Since D_1 and D_2 diagonal, $D_1 D_2 = D_2 D_1$, so $\alpha\beta = \beta\alpha$.

Now suppose α, β are both diagonalizable and $\alpha\beta = \beta\alpha$. Let $\lambda_1, \ldots, \lambda_k$ be the k distinct eigenvalues of α. Then we can write

$$V = \bigoplus_{i=1}^{k} V_{\lambda_i}$$

where V_{λ_i} is the eigenspace associated to λ_i. We claim that V_{λ_i} is stable by β. Indeed, if $v \in V_{\lambda_i}$, then

$$\alpha(\beta(v)) = \beta(\alpha(v)) = \beta(\lambda_i v) = \lambda_i \beta(v).$$

Hence $\beta(v) \in V_{\lambda_i}$. Now we use the criterion for diagonalizability: if β is diagonalizable, then there exists a polynomial with distinct linear factors such that $p(\beta) = 0$.

Since $\beta|_{V_{\lambda_j}}$ is an endomorphism and $p(\beta|_{V_{\lambda_j}}) = 0$, $B|_{V_{\lambda_j}}$ is diagonalizable. Let $\mathcal{B}_j$ be a basis for which $\beta|_{V_{\lambda_j}}$ is diagonal.

Then, since V is the sum of V_{λ_j}, $(\mathcal{B}_1, \dots, \mathcal{B}_k) = \mathcal{B}$ is a basis of V in which both α and β are in diagonal form.

16.1 Minimal Polynomials

Proposition 16.1 (Euclidean Algorithm for Polynomials). *Let a, b be polynomials over K, with $b \neq 0$. Then there exist polynomials q, r over K with $\deg r < \deg b$ and $a = qb + r$.*

Definition 16.2. Let V be a finite-dimensional vector space over K, and let $\alpha \in L(V)$. The *minimal polynomial* m_α of α is the unique non-zero polynomial with smallest degree such that $m_\alpha(\alpha) = 0$.

The existence and uniqueness of a minimal polynomial can be seen as such. If $\dim V = n$, then we known $\dim L(V) = n^2$, so $(\mathrm{id}, \alpha, \dots, \alpha^{n^2})$ cannot be free. Hence there is some combination

$$a_{n^2}\alpha^{n^2} + \cdots + a_1\alpha + a_0 = 0.$$

Hence the existence of a minimal polynomial is shown.

Now suppose $p(\alpha) = 0$. We show that $m_\alpha \mid p$. Indeed, from the Euclidean algorithm, we can find q, r such that $p = m_\alpha q + r$, with $\deg r < \deg m_\alpha$. Since $p(\alpha) = m_\alpha(\alpha) = 0$, we must have $r(\alpha) = 0$.

Hence by minimality of m_α, $r = 0$, and $m_\alpha \mid p$. But this implies uniqueness, as if m_1, m_2 are both polynomials of smallest degree, then $m_1 \mid m_2$ and $m_2 \mid m_1$, so they are equal up to a constant factor.

Example 16.2.

If $V = \mathbb{R}^2$, take

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

Let $p(t) = (t-1)^2$, then $p(A) = p(B) = 0$. So their minimal polynomial is only $t-1$ or $(t-1)^2$. From this, we can check $m_A = t - 1$and $m_B = (t-1)^2$. In particular, B is not diagonalizable.

17 Cayley-Hamilton Theorem

Theorem 17.1 (Cayley-Hamilton Theorem)**.** *Let V be a finite dimensional K vector space, and $\alpha \in L(V)$ with characteristic polynomial $\chi_\alpha(t) = \det(\alpha - t\,\mathrm{id})$. Then*

$$\chi_\alpha(\alpha) = 0.$$

As a corollary, $m_\alpha \mid \chi_\alpha$.

Proof: We solve over $K = \mathbb{C}$. Take a basis $\mathcal{B} = \{v_1, \ldots, v_n\}$ for which $[\alpha]_\mathcal{B}$ is triangular, i.e.

$$[\alpha]_\mathcal{B} = \begin{pmatrix} a_1 & & * \\ & \ddots & \\ 0 & & a_n \end{pmatrix},$$

and let $U_j = \langle v_1, \ldots, v_j \rangle$. Then, $(\alpha - a_j\,\mathrm{id})U_j \le U_{j-1}$, due to the triangular form. Now we known $\chi_\alpha(t) = \prod(a_i - t)$, so

$$\begin{aligned} &(\alpha - a_1\,\mathrm{id}) \cdots (\alpha - a_{n-1}\,\mathrm{id})(\alpha - a_n\,\mathrm{id})V \\ \le &(\alpha - a_1\,\mathrm{id}) \cdots (\alpha - a_{n-1}\,\mathrm{id})U_{n-1} \\ &\vdots \\ \le &(\alpha - a_1\,\mathrm{id})U_1 \\ = &0. \end{aligned}$$

Hence $\chi_\alpha(\alpha) = 0$.

Definition 17.1 (Multiplicity)**.** For a finite-dimensional vector space V and $\alpha \in L(V)$, let λ be an eigenvalue of λ. Then

$$\chi_\alpha(t) = (t - \lambda)^{a_\lambda} q(t),$$

where a_λ is the *algebraic multiplicity* of λ, and the *geometric multiplicity* of λ is $\dim \mathrm{Ker}(\alpha - \lambda\,\mathrm{id})$.

Remark. If λ is an eigenvalue, then $\alpha - \lambda\,\mathrm{id}$ is singular, so $\det(\alpha - \lambda\,\mathrm{id}) = \chi_\alpha(\lambda) = 0$.

Lemma 17.1. *For an eigenvalue λ of $\alpha \in L(V)$, then $1 \le g_\lambda \le a_\lambda$.*

Proof: Immediately, $g_\lambda = \dim \mathrm{Ker}(\alpha - \lambda\,\mathrm{id}) \ge 1$, as $\alpha - \lambda\,\mathrm{id}$ is singular. So we show $g_\lambda \le a_\lambda$.

Indeed, let $(v_1, \ldots, v_{g_\lambda})$ be a basis of $V_\lambda = \mathrm{Ker}(\alpha - \lambda\,\mathrm{id})$, and complete to a basis $\mathcal{B} = (v_1, \dots, v_{g_\lambda}, v_{g_\lambda+1}, \dots, v_n)$ of V. Then,

$$[\alpha]_{\mathcal{B}} = \begin{pmatrix} \lambda \operatorname{id}_{g_\lambda} & * \\ 0 & A_1 \end{pmatrix},$$

$$\implies \det[\alpha - t \operatorname{id}] = \det \begin{pmatrix} (\lambda - t) \operatorname{id}_{g_\lambda} & * \\ 0 & A_1 - t \operatorname{id} \end{pmatrix} = (\lambda - t)^{g_t} \chi_{A_1}(t).$$

Hence $g_\lambda \leq a_\lambda$.

Lemma 17.2. *For λ an eigenvalue of $\alpha \in L(V)$, let c_λ be the multiplicity of λ as a root of m_α. Then $1 \leq c_\lambda \leq a_\lambda$.*

Proof: From Cayley-Hamilton, $m_\alpha \mid \chi_\alpha$, immediately giving $c_\lambda \leq a_\lambda$. Now note $c_\lambda \geq 1$, as there exists a non-zero eigenvector v of λ. Hence,

$$m_\alpha(\alpha)(v) = (m_\alpha(\lambda))v = 0,$$

so $m_\alpha(\lambda) = 0$, and $c_\lambda \geq 1$.

Example 17.1.

Take the matrix

$$A = \begin{pmatrix} 1 & 0 & -2 \\ 0 & 1 & 1 \\ 0 & 0 & 2 \end{pmatrix}.$$

Since A is triangular $\chi_A(t) = (t-1)^2(t-2)$. Hence m_A is either $(t-1)^2(t-2)$ or $(t-1)(t-2)$. We can check that $(A-I)(A-2I) = 0$, so $m_A = (t-1)(t-2)$, and A is diagonalizable.

Index