IB Linear Algebra

Ishan Nath, Michaelmas 2022

Based on Lectures by Prof. Pierre Raphael

October 21, 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 $(\lambda_1, \ldots, \lambda_n)$ the coordinates of v in the basis $(v_1, \ldots, v_n)$.

Proof: Since $\langle v_1, \ldots, v_n \rangle = V$, 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 $(v_1, \ldots, v_n)$ spans V, then some subset of this family is a basis of V.*

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

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

Then we have $V = \langle v_1, \ldots, v_n \rangle = \langle v_1, \ldots, v_{n-1} \rangle$. 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) $(v_1, \ldots, v_m)$ *free,*

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

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

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

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

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, $V = \langle v_1, \ldots, v_l, w_{l+1}, \ldots, w_n \rangle = \langle v_1, \ldots, v_l, v_{l+1}, w_{l+1}, \ldots, w_n \rangle = \langle v_1, \ldots, v_{l+1}, w_{l+2}, \ldots, w_n \rangle$. 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 $(v_1, \ldots, v_n), (w_1, \ldots, w_m)$ bases of V.

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

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

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 $(v_1, \ldots, v_l)$ a basis of $U \cap W$. Extend this to a basis $(v_1, \ldots, v_l, u_1, \ldots, u_m)$ of U, and a basis $(v_1, \ldots, v_l, w_1, \ldots, w_n)$ of W. Then we show $(v_1, \ldots, v_l, u_1, \ldots, u_m, w_1, \ldots, w_n)$ 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

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

But since $(v_1, \ldots, w_n)$ 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 $(u_1, u_2, \ldots, u_l)$ be a basis of U. Extend this to a basis $(u_1, \ldots, u_l, w_{l+1}, \ldots, w_n)$ of V. Then we show that $(w_{l+1} + U, \ldots, w_n + U)$ 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 $\mathcal{B}_1 = \{u_1, \ldots, u_l\}$ and $\mathcal{B}_2 = \{w_1, \ldots, w_m\}$ be collections of vectors. Then

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

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,*

$$\mathcal{B} = \bigcup_{i=1}^{l} \mathcal{B}_i \text{ is a basis of } \sum_{i=1}^{l} V_l.$$

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

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

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 $\mathcal{B} = (v_1, \ldots, v_n)$ of V, and $\mathcal{C} = (w_1, \ldots, w_m)$ 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

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

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 $(v_1, \dots, v_k, v_{k+1}, \dots, v_n)$ is a basis of V, and $(v_1, \dots, v_k)$ is a basis $\mathcal{B}'$ of Y, and $(v_{k+1}, \dots, v_n)$ is a basis $\mathcal{B}''$.

Then if $Z \leq W$, we can take a basis $\mathcal{C}$ of W $(w_1, \dots, w_l, w_{l+1}, \dots, w_m)$, such that $(w_1, \dots, w_l)$ is a basis $\mathcal{C}'$ of Z, and $(w_{l+1}, \dots, w_m)$ 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 $\mathcal{B} = \{v_1, \dots, v_n\}$ and $\mathcal{B}' = \{v_1', \dots, v_n'\}$ and W has two bases $\mathcal{C} = \{w_1, \dots, w_n\}$ and $\mathcal{C}' = \{w_1', \dots, w_m'\}$. 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

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

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

- $A = I_m^{-1} A I_n$,
- 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 $N(x)$, $(v_{r+1}, \dots, v_n)$, and extend it to a basis $\mathcal{B} = (v_1, \dots, v_r, v_{r+1}, \dots, v_n)$.

We claim that $(\alpha(v_1), \dots, \alpha(v_r))$ 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 $y \in \langle \alpha(v_1), \dots, \alpha(v_r) \rangle$.

- 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 $(v_1, \dots, v_n)$ is a basis, $\lambda_i = \mu_i = 0$.

So we have $(\alpha(v_1), \dots, \alpha(v_r))$ is a basis of $\operatorname{Im} \alpha$, and $(v_{r+1}, \dots, v_n)$ is a basis of $\operatorname{Ker} \alpha$. Let $\mathcal{C} = (\alpha(v_1), \dots, \alpha(v_r), w_{r+1}, \dots, w_m)$. We get that

$$[\alpha]_{\mathcal{B},\mathcal{C}} = (\alpha(v_1), \dots, \alpha(v_r), \alpha(v_{r+1}), \dots, \alpha(v_n)) = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}.$$

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.*

Index