Dilla University

Department of Mathematics

Advanced Linear Algebra Lecture Note

by

Dereje Kifle (PhD)

Contents

Advanced Linear Algebra Worksheet I

1. Let V be a vector space and $W \subseteq V$ be a subspace. Define a relation $\sim_W$ on V as follows:
$$v_1 \sim_W v_2 \text{ if and only if } v_1 - v_2 \in W.$$
Show that $\sim_W$ is an equivalence relation on V. We write the equivalence classes as
$$[v_1] = \{v_2 \in V \mid v_1 - v_2 \in W\} = v_1 + W.$$
Set $V/W = \{v + W \mid v \in V\}$. Addition and scalar multiplication on V/W are defined as follows. Let $v_1, v_2 \in V$ and $c \in F$. Define
$$(v_1 + W) + (v_2 + W) = (v_1 + v_2) + W;$$
$$c(v_1 + W) = cv_1 + W.$$
Show that V/W is an F-vector space.

2. Let V be an F -vector space of dimension n. Let $\tau \in \mathcal{L}(V)$ such that $\tau^2 = 0$. Prove that the image of τ is contained in the kernel of τ and hence the dimension of the image of τ is at most $\frac{n}{2}$.

Chapter 1

Vector Spaces

1.1 Vector Spaces

Definition 1.1.1. Let F be a field. A *vector space* over F is a nonempty set V together with two operations:

- addition: assigns to each pair $(u, v) \in V \times V$ a vector $u + v \in V$.
- scalar multiplication: assigns to each pair $(r, u) \in F \times V$ a vector ru in V.

Furthermore, the following properties must be satisfied:

- *Associativity of addition*: For all vectors $u, v, w \in V$, $u + (v + w) = (u + v) + w$.
- *Commutativity of addition*: For all vectors $u, v \in V$, $u + v = v + u$.
- *Existence of zero*: There is a zero vector $0 \in V$ with the property that $0 + u = u + 0 = u$ for all vectors $u \in V$.
- *Existence of additive inverses*: For each vector $u \in V$, there is a vector in V, denoted by $-u$, with the property that $u + (-u) = (-u) + u = 0$.
- *Properties of scalar multiplication*: For all scalars $a, b \in F$ and for all vectors $u, v \in V$,

$$\begin{aligned} a(u + v) &= au + av \\ (a + b)u &= au + bu \\ (ab)u &= a(bu) \\ 1u &= u \end{aligned}$$

In the above definition

- Elements of F (resp. V) are referred to as *scalars* (resp. *vectors*).
- The first four properties are equivalent to $(V, +)$ is an abelian group.
- V is sometimes called an *F-space*.
- If $F = \mathbb{R}$ (resp. $\mathbb{C}$), then V is a *real* (resp. *complex*) vector space.

1.2 Examples of a vector space

1) Let F be a field. The set V_F of all functions from F to F is a vector space over F, under the operations of ordinary addition and scalar multiplication of functions:

$$(f+g)(x) = f(x) + g(x), \text{ and } (af)(x) = a(f(x)).$$

2) The set $M_{m\times n}(F)$ of all $m \times n$ matrices with entries in a field F is a vector space over F, under the operations of matrix addition and scalar multiplication.

1.3 Subspaces, Linear combinations and Generators

Most algebraic structures contain substructures.

Definition 1.3.1. A *subspace* of a vector space V is a subset S of V that is a vector space in its own right under the operations obtained by restricting the operations of V to S. To indicate that S is a subspace of V, we use the notation $S \leq V$. If S is a subspace of V but $S \neq V$, we say that S is a proper subspace of V and it is denoted by $S < V$. The zero subspace of V is $\{0\}$.

Definition 1.3.2. Let S be a nonempty subset of a vector space V. A *linear combination* (L.C) of vectors in S is an expression of the form

$$a_1v_1 + \ldots + a_nv_n$$

where $v_1 \ldots v_n \in S$ and $a_1, \ldots, a_n \in F$. The scalars a_i are called the *coefficients* of the linear combination. A L.C is trivial if every coefficient a_i is zero. Otherwise, it is non trivial.

Theorem 1.3.3. *A non-empty subset S of a vector space V is a subspace of V if and only if S is closed under addition and scalar multiplication or equivalently, S is closed under linear combinations, that is,*

$$a, b \in F, u, v \in S \Longrightarrow au + bv \in S.$$

Example 1.3.4. Consider the vector space $V(n,2)$ of all binary n-tuples, that is, n-tuples of 0's and 1's. The weight $\mathcal{W}(v)$ of a vector $v \in V(n,2)$ is the number of non-zero coordinates in v. Let E_n be the set of all vectors in V of even weight. Then $E_n \leq V(n,2)$.

Proof. For vectors $u, v \in V(n,2)$, show that

$$\mathcal{W}(u+v) = \mathcal{W}(u) + \mathcal{W}(v) - 2\mathcal{W}(u \cap v) \tag{1.1}$$

where $u \cap v$ is the vector in $V(n,2)$ whose i^{th} component is the product of the i^{th} components of u and v, that is, $(u \cap v)_i = u_i \cdot v_i$. Let u and v be elements of E_n. Then by definition $\mathcal{W}(u)$ and $\mathcal{W}(v)$ are even which by (1.1) implies $\mathcal{W}(u+v)$ is even, that is, $u+v \in E_n$. Let $a \in \mathbb{F}_2$ and let $u \in E_n$. Clearly, $\mathcal{W}(au)$ is even which implies $au \in E_n$. Thus $E_n \leq V(n,2)$, known as the even weight subspace of $V(n,2)$. □

Definition 1.3.5. The subspace *spanned (or generated)* by a nonempty set S of vectors in V is the set of all linear combinations of vectors from S:

$$\langle S \rangle = \text{Span}(S) = \left\{ \sum_{i=1}^{n} r_i v_i \;\middle|\; r_i \in F, v_i \in S \right\}.$$

When $S = \{v_1, \ldots, v_n\}$ is a finite set, we use the notation $\langle v_1, \ldots, v_n \rangle$ or $\text{span}(v_1, \ldots, v_n)$. A set S of vectors in V is said to be span V, or generates V, if $V = \text{Span}(S)$.

Any superset of a spanning set is also a spanning set and all vector spaces have spanning set since V spans itself.

1.4 Linear Dependence and Independence of Vectors

Definition 1.4.1. Let V be a vector space. A nonempty set S of vectors in V is linearly independent (L.I) if for any distinct vectors $s_1, \ldots, s_n$ in S

$$a_1 s_1 + \ldots + a_n s_n = 0 \Rightarrow a_i = 0 \text{ for all } i.$$

In other words, S is L.I if the only L.C of vectors from S that is equal to 0 is the trivial L.C, all of whose coefficients are 0. If S is not L.I, it is said to be linearly dependent (LD).

A L.I set of vectors cannot contain the zero vector, since $1 \cdot 0 = 0$ violates the condition of linear independence.

Definition 1.4.2. Let S be a nonempty set of vectors in V. To say that a nonzero vector $v \in V$ is an *essentially unique L.C* of the vectors in S is to say that, up to the order of terms, there is one and only one way to express v as a L.C $v = \sum_{i=1}^{n} a_i s_i$ where the s_i's are distinct vectors in S and the coefficients a_i are nonzero. More explicitly, $v \neq 0$ is an essentially unique L.C of vectors in S if $v \in \langle S \rangle$ and if whenever

$$v = a_1 s_1 + \ldots + a_n s_n \text{ and } v = b_1 t_1 + \ldots + b_m t_m$$

where the s_i's are distinct and t_i's are distinct and all coefficients are nonzero, then $m = n$ and after a reindexing of the $b_i t_i$'s if necessary, we have $a_i = b_i$ and $s_i = t_i$ for all $i = 1, \ldots, n$.

Theorem 1.4.3. *Let $S \neq \{0\}$ be a nonempty set of vectors in V. The following are equivalent:*

(a) S is L.I.

(b) Every nonzero vector $v \in \operatorname{span}(S)$ is an essentially unique L.C of the vectors in S.

(c) No vector in S is a L.C of other vectors in S.

Proof. (a) $\Rightarrow$ (b) Suppose that

$$0 \neq v = a_1 s_1 + \ldots + a_n s_n \text{ and } v = b_1 t_1 + \ldots + b_m t_m$$

where the s_i's are distinct and t_i's are distinct and the coefficients are nonzero. By subtracting and grouping s's and t's that are equal, we can write

$$\begin{aligned} 0 = &\left(a_{i_1} - b_{i_1}\right) s_{i_1} + \ldots + \left(a_{i_k} - b_{i_1}\right) s_{i_k} \\ &+ a_{i_{k+1}} s_{i_{k+1}} + \ldots + a_{i_n} s_{i_n} - b_{i_{k+1}} t_{i_{k+1}} - \ldots - b_{i_m} t_{i_m} \end{aligned}$$

(a) $\Rightarrow m = n = k$ and $a_{i_u} = b_{i_u}$ and $s_{i_u} = t_{i_u}$ for all $u = 1, \ldots, k$.

(b) $\Rightarrow$ (c) and (c) $\Rightarrow$ (a) is left as an exercise. □

1.5 Direct sum and direct product of subspaces

Definition 1.5.1. Let $V_1, \ldots, V_n$ be vector spaces over a field F. The *external direct sum* of $V_1, \ldots, V_n$, denoted by $V_1 \boxplus \ldots \boxplus V_n$ is the vector space V whose elements are ordered n-tuples:

$$V = \{(v_1, \ldots, v_n) \mid v_i \in V_i, i = 1, \ldots, n\}$$

with componentwise operations

$$\begin{aligned}(u_1, \ldots, u_n) + (v_1, \ldots, v_n) &= (u_1 + v_1, \ldots, u_n + v_n) \text{ and} \\ r(u_1, \ldots, u_n) &= (ru_1, \ldots, ru_n) \qquad \text{for all } r \in F.\end{aligned}$$

Example 1.5.2. The vector space F^n is the external direct sum of n copies of F, that is, $F^n = F \boxplus \ldots \boxplus F$ where there are n summands on the right hand side.

The above construction can be generalized to any collection of vector spaces by generalizing the idea that an ordered n-tuple $(v_1, \ldots, v_n)$ is just a function

$$\begin{aligned}f : \{1, \ldots, n\} &\to \bigcup V_i, \\ i &\mapsto f(i)\,.\end{aligned}$$

Definition 1.5.3. Let $\mathcal{F} = \{V_i \mid i \in I\}$ be any family of vector spaces over F. The *direct product* of $\mathcal{F}$ is the vector space

$$\prod_{i \in I} V_i = \left\{ f \; : \; I \to \bigcup V_i \;\middle|\; f(i) \in V_i \right\}$$

thought of as a subspace of the vector space of all functions from I to $\bigcup V_i$.

Note that

$$\prod_{i \in I} V_i = \{v = (v_i)_{i \in I} \;\mid\; v_i \in V_i\} = \left\{ f \; : \; I \to \bigcup V_i \;\middle|\; f(i) \in V_i \right\}.$$

If we define addition and scalar multiplication by

$$\begin{aligned}v + w &= \left(f \; : \; I \to \bigcup V_i\right) + \left(g \; : \; I \to \bigcup V_i\right) \\ &= \left(f + g \; : \; I \to \bigcup V_i\right) \text{ and} \\ av &= a\left(f \; : \; I \to \bigcup V_i\right) \\ &= \left(af \; : \; I \to \bigcup V_i\right)\end{aligned}$$

or by

$$\begin{aligned}(v_i)_{i \in I} + (w_i)_{i \in I} &= (v_i + w_i)_{i \in I} \text{ and} \\ a(v_i)_{i \in I} &= (av_i)_{i \in I}\end{aligned}$$

Then the direct product $\prod_{i \in I} V_i$ is a vector space over F.

Definition 1.5.4. Let $\mathcal{F} = \{V_i \mid i \in I\}$ be any family of vector spaces over F. The support of a function $f : I \to \bigcup V_i$ is the set

$$\text{support}(f) = \{i \in I \mid f(i) \neq 0\}.$$

We say that f has *finite support* if $f(i) = 0$ for all but a finite number of $i \in I$.

Definition 1.5.5. Let $\mathcal{F} = \{V_i \mid i \in I\}$ be any family of vector spaces over F. The *external direct sum* of the family $\mathcal{F}$ is the vector space

$$\bigoplus_{i \in I}^{\text{ext}} V_i = \left\{ f : I \to \bigcup V_i \mid f(i) \in V_i, f \text{ has finite support} \right\}.$$

thought of as a subspace of the vector space of all functions from I to $\bigcup V_i$.

If $V_i = V$ for all $i \in I$,

- we denote the set of all functions from I to V by V^I, and
- we denote the set of all functions in V^I that have finite support by $\left(V^I\right)_0$.

In this case, we have

$$\prod_{i \in I} V = V^I \quad \text{and} \quad \bigoplus_{i \in I}^{\text{ext}} V = \left(V^I\right)_0.$$

Definition 1.5.6. A vector space V is the *internal direct sum* of a family $\mathcal{F} = \{S_i \mid i \in I\}$ of subspaces of V, written

$$V = \bigoplus \mathcal{F} \text{ or } V = \bigoplus_{i \in I} S_i$$

if the following hold:

(1) (*Join of the family*) V is the sum (join) of the family $V = \sum_{i \in I} S_i$

(2) (*Independence of the family*) For each $i \in I$,

$$S_i \bigcap \left(\sum_{j \neq i} S_j \right) = \{0\}.$$

In this case,

- each S_i is called a *direct summand* of V.
- if $\mathcal{F} = \{S_1, \ldots, S_n\}$ is a finite family, the direct sum is often written $V = S_1 \oplus \ldots \oplus S_n$.

- if $V = S \oplus T$, then T is called a *complement* of S in V.

If S and T are subspaces of V, then we may always say that the sum $S + T$ exists. However, to say that the direct sum of S and T exists or to write $S \oplus T$ is to imply that $S \cap T = \{0\}$. Thus, while the sum of two subspaces always exists, the direct sum of two subspaces does not always exist. Similar statements apply to families of subspaces of V.

Theorem 1.5.7. *Let $\mathcal{F} = \{V_i \mid i \in I\}$ be any family of vector spaces over F. The following are equivalent:*

*(1) (*Independence of the family*) For each $i \in I$,*

$$S_i \bigcap \left(\sum_{j \neq i} S_j \right) = \{0\}.$$

*(2) (*Uniqueness of expression for 0*) The zero vector cannot be written as a sum of nonzero vectors from distinct subspaces of $\mathcal{F}$.*

*(3) (*Uniqueness of expression*) Every nonzero vector $v \in V$ has a unique, except for order of terms, expression as a sum*

$$v = s_1 + \ldots + s_n$$

of nonzero vectors from distinct subspaces in $\mathcal{F}$.

Hence, a sum

$$V = \sum_{i \in I} S_i$$

is direct if and only if any one of (1)-(3) holds.

Proof. (1) $\Rightarrow$ (2) Suppose that (2) fails, that is,

$$0 = s_{j_1} + \ldots + s_{j_n}$$

where the nonzero vectors s_{j_i}'s are from distinct subspaces of S_{j_i}. Then $n > 1$ and, hence,

$$-s_{j_1} = s_{j_2} \ldots + s_{j_n}$$

which violates (1).
(2) $\Rightarrow$ (3) If (2) holds and

$$v = s_1 + \ldots + s_n = t_1 + \ldots + t_n$$

where the terms are nonzero and both the s_i's and the t_i's belong to distinct subspaces in $\mathcal{F}$. Then

$$0 = s_1 + \ldots + s_n - t_1 - \ldots - t_n.$$

Now, by collecting terms from the same subspaces, we may write

$$\begin{aligned} 0 &= (s_{i_1} - t_{i_1}) + \ldots + (s_{i_k} - t_{i_k}) \\ &\quad + s_{i_{k+1}} + \ldots + s_{i_n} - t_{i_{k+1}} - \ldots - t_{i_m}. \end{aligned}$$

Then (2) implies that $m = n = k$ and $s_{i_u} = t_{i_u}$ for all $u = 1, \ldots, k$.
(3) $\Rightarrow$ (1)

$$0 \neq v \in S_i \bigcap \left(\sum_{j \neq i} S_j \right) \Rightarrow v = s_i \in S_i \text{ and } s_i = s_{j_1} + \ldots + s_{j_n}$$

where $s_{j_k} \in S_{j_k}$ are nonzero which violates (3). □

Example 1.5.8. Let $A = \{(x, 0) \in \mathbb{R}^2 \mid x \in \mathbb{R}\}$ and let $B = \{(0, y) \in \mathbb{R}^2 \mid y \in \mathbb{R}\}$. Then $\mathbb{R}^2 = A \oplus B$ since $A \cap B = \{0\}$ and $\mathbb{R}^2 = A + B$. Any element (x, y) of $\mathbb{R}^2$ can be written as

$$(x, y) = (x, 0) + (0, y).$$

Proposition 1.5.9. *Suppose U and W are subspaces of the vector space V over a field F. Consider the map*

$$\alpha : U \oplus W \to V$$

defined by $\alpha(u, w) = u + w$. Then

- *α is injective if and only if $U \cap W = \{0\}$.*
- *α is surjective if and only if $U \cup W$ spans V.*

Example 1.5.10. Let $A = \{(x, 0) \in \mathbb{R}^2 \mid x \in \mathbb{R}\}$ and let $C = \{(y, y) \in \mathbb{R}^2 \mid y \in \mathbb{R}\}$. Then $\mathbb{R}^2 = A \oplus C$. To see this, note that the map

$$\begin{aligned} \alpha : A \oplus B &\to \mathbb{R}^2 \\ (x, y) &\mapsto x + y \end{aligned}$$

is injective since $A \cap C = \{0\}$. Moreover, α is a surjective map since any element (x, y) of $\mathbb{R}^2$ can be written as

$$(x, y) = \underbrace{(x - y, 0)}_{\in A} + \underbrace{(y, y)}_{\in C}.$$

Thus, by the above proposition $A \cup C$ spans $\mathbb{R}^2$.

Example 1.5.11. Let $A \in \mathcal{M}_n$ be a matrix. Then A can be written in the form

$$A = \frac{1}{2}(A + A^t) + \frac{1}{2}(A - A^t) = B + C \tag{1.2}$$

where A^t is the transpose of A. Verify that B is symmetric and C is skew-symmetric. Thus (1.2) is a decomposition of A as a sum of a symmetric matrix ($A^t = A$) and a skew-symmetric matrix ($A^t = -A$).

Exercise 1.5.12. Show that the sets Sym and SkewSym of all symmetric and skew-symmetric matrices in $\mathcal{M}_n$ are subspaces of $\mathcal{M}_n$.

Thus, we have

$$\mathcal{M}_n = \mathrm{Sym} + \mathrm{SkewSym}.$$

Furthermore, if $S, S' \in \mathrm{Sym}$ and $T, T' \in \mathrm{SkewSym}$ such that $S + T = S' + T'$, then the matrix

$$U = S - S' = T - T' \in \mathrm{Sym} \cap \mathrm{SkewSym}.$$

Hence, provided that $\mathrm{char}(F) \neq 2$, we must have $U = 0$. Thus,

$$\mathcal{M}_n = \mathrm{Sym} \oplus \mathrm{SkewSym}.$$

1.6 Bases of a Vector Space

Theorem and Definition 1.6.1. Let S be a set of vectors in V. The following are equivalent:

(i) S is L.I and spans V.

(ii) Every nonzero vector $v \in V$ is an essentially unique L.C of vectors in S.

(iii) S is a minimal spanning set, that is, S spans V but any proper subset of S does not span V.

(iv) S is a maximal L.I set, that is, S is L.I but any proper superset of S is not L.I.

A set of vectors in V that satisfies any (and hence all) of these conditions is called a basis for V.

Proof. (i) $\longleftrightarrow$ (ii) by Theorem 1.4.3.

(i) $\Rightarrow$ (iii) By given S is L.I and a spanning set, $V = \mathrm{span}(S)$. Suppose that any proper subset S' of S spans V. Let $s \in S - S'$. Since $s \in V$, s is a L.C of the vectors in S' which is a contradiction to the fact that S is L.I.

(iii) $\Rightarrow$ (i) If S is a minimal spanning set, then it must be L.I. For if not, some vector $s \in S$ would be a L.C of the other vectors in S, $S - \{s\}$. Then $S - \{s\}$ would be a proper spanning subset of S which is not possible.

(i) $\Leftrightarrow$ (iv): **exercise** $\square$

Example 1.6.2.

(1) Find a basis of the subspace of $\mathbb{R}^3$ given by

$$V = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \;\middle|\; x - 2y + 5z = 0 \right\}.$$

Solution: Let $v = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ be any vector in V. Then

$$\begin{aligned} v = \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2y - 5z \\ y \\ z \end{pmatrix} &= \begin{pmatrix} 2y \\ y \\ 0 \end{pmatrix} + \begin{pmatrix} -5z \\ 0 \\ z \end{pmatrix} \\ &= y \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} + z \begin{pmatrix} -5 \\ 0 \\ 1 \end{pmatrix}, \quad y, z \in \mathbb{R}. \end{aligned}$$

This shows that the set

$$\{u, v\} = \left\{ \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -5 \\ 0 \\ 1 \end{pmatrix} \right\}$$

spans V. It is easy to see that the set $\{u, v\}$ is L.I. Thus it is a basis for the subspace V of $\mathbb{R}^3$.

(2) The set $S = \left\{ \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\}$ is a basis of $\mathbb{R}^2$.

(3) The i^{th} standard vector in F^n is the vector e_i that has 0's in all coordinate positions except the i^{th}, where it has a 1. Thus,

$$e_1 = (1, \ldots, 0),\ e_2 = (0, 1, \ldots, 0), \ldots, e_n = (0, \ldots, 0, 1).$$

The set $\{e_1, \ldots, e_n\}$ is called the standard basis for F^n.

Theorem 1.6.3. *Let V be a nonzero vector space. Let I be a L.I set in V and let S be a spanning set in V containing I. Then there is a basis $\mathcal{B}$ for V for which $I \subset \mathcal{B} \subset S$. In particular,*

(1) Any vector space, except the zero space $\{0\}$, has a basis.

(2) Any L.I set in V is contained in a basis.

(3) Any spanning set in V contains a basis.

1.7 Dimension of a Vector Space

The following theorem says that if a vector space V has a finite spanning set S, then the size of any linearly independent set cannot exceed the size of S.

Theorem 1.7.1. *Let V be a vector space and assume that the vectors $v_1, \ldots, v_n$ are L.I and the vectors $s_1, \ldots, s_m$ span V. Then $n \leq m$.*

Corollary 1.7.2. *If V has a finite spanning set, then any two bases of V have the same size.*

Theorem 1.7.3. *If V is a vector space, then any two bases for V have the same cardinality.*

Definition 1.7.4. A vector space V is *finite-dimensional* if it is the zero space or if it has a finite basis. All other vector spaces are *infinite-dimensional.* The *dimension* of the a non-zero vector space V is the cardinality of any basis for V.

(a) The dimension of the zero space is 0.

(b) If a vector space V has a basis of cardinality k, we say that V is *k-dimensional* and write $\dim(V) = k$.

(c) If S is a subspace of V, then $\dim(S) \leq \dim(V)$. If in addition $\dim(S) = \dim(V) < \infty$, then $S = V$.

Theorem 1.7.5. *Let V be a vector space.*

1) If $\mathcal{B}$ is a basis for V and if $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2$ and $\mathcal{B}_1 \cap \mathcal{B}_2 = \emptyset$, then $V = \langle \mathcal{B}_1 \rangle \oplus \langle \mathcal{B}_2 \rangle$.

2) Let $V = S \oplus T$. If $\mathcal{B}_1$ is a basis for S and $\mathcal{B}_2$ is a basis for T, then $\mathcal{B}_1 \cap \mathcal{B}_2 = \emptyset$ and $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2$ is a basis for V.

Proof. 1) If $\mathcal{B}_1 \cap \mathcal{B}_2 = \emptyset$ and $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2$ is a basis for V, then $0 \notin \mathcal{B}_1 \cup \mathcal{B}_2$. But, if a nonzero vector $v \in \langle \mathcal{B}_1 \rangle \cap \langle \mathcal{B}_2 \rangle$, then $\mathcal{B}_1 \cap \mathcal{B}_2 \neq \emptyset$, a contradiction. Hence, $\{0\} = \langle \mathcal{B}_1 \rangle \cap \langle \mathcal{B}_2 \rangle$. Furthermore, since $\mathcal{B}_1 \cup \mathcal{B}_2$ is a basis for V and for $\langle \mathcal{B}_1 \rangle + \langle \mathcal{B}_2 \rangle$, we must have $V = \langle \mathcal{B}_1 \rangle + \langle \mathcal{B}_2 \rangle$. Thus, $V = \langle \mathcal{B}_1 \rangle \oplus \langle \mathcal{B}_2 \rangle$.

2) If $V = S \oplus T$, then $S \cap T = \{0\}$. Since $0 \notin \mathcal{B}_1 \cup \mathcal{B}_2$, we have $\mathcal{B}_1 \cap \mathcal{B}_2 = \emptyset$. Let $v \in V$. Then v has the form

$$a_1 u_1 + \ldots + a_n u_n + b_1 v_1 + \ldots + b_m v_m$$

for $u_1, \ldots, u_n \in \mathcal{B}_1$ and $v_1, \ldots, v_m \in \mathcal{B}_2$ which implies $v \in \langle \mathcal{B}_1 \cup \mathcal{B}_2 \rangle$ and thus $\mathcal{B}_1 \cup \mathcal{B}_2$ is a basis for V by Theorem 1.6.1. □

Theorem 1.7.6. *Let S and T be subspaces of a vector space V. Then*

$$\dim(S) + \dim(T) = \dim(S + T) + \dim(S \cap T).$$

In particular, if T is any complement of S in V, then

$$\dim(S) + \dim(T) = \dim(V) = \dim(S \oplus T).$$

Proof. Suppose that $\mathcal{B} = \{v_i \mid i \in I\}$ is a basis for $S \cap T$. Extend this to a basis $\mathcal{A} \cup \mathcal{B}$ for S and to a basis $\mathcal{B} \cup \mathcal{C}$ for T, where $\mathcal{A} = \{u_j \mid j \in J\}$ and $\mathcal{C} = \{w_k \mid k \in K\}$, $\mathcal{A} \cap \mathcal{B} = \emptyset$ and $\mathcal{C} \cap \mathcal{B} = \emptyset$.

Claim: $\mathcal{A} \cup \mathcal{B} \cup \mathcal{C}$ is a basis for $S + T$.

Clearly, $\langle \mathcal{A} \cup \mathcal{B} \cup \mathcal{C} \rangle = S + T$. It remains to show that the set $\mathcal{A} \cup \mathcal{B} \cup \mathcal{C}$ is L.I. To see this, suppose to the contrary that

$$\alpha_1 v_1 + \ldots + \alpha_n v_n = 0$$

where $v_i \in \mathcal{A} \cup \mathcal{B} \cup \mathcal{C}$ and $\alpha_i \neq 0$ for all i. Then there must be vectors $v_i \in \mathcal{A} \cap \mathcal{C}$ since $\mathcal{A} \cup \mathcal{B}$ and $\mathcal{B} \cup \mathcal{C}$ are L.I. Now, isolating the terms involving the vectors from $\mathcal{A}$, say $v_1, \ldots, v_k$ without loss of generality, on one side of the equality shows that there is a nonzero vector in $x \in \mathcal{A} \cap \langle \mathcal{B} \cup \mathcal{C} \rangle$.

That is,

$$\begin{aligned} x = \underbrace{a_1 v_1 + \ldots + a_k v_k}_{\in\, \text{span}(A)} &= \underbrace{a_{k+1} v_{k+1} + \ldots + a_n v_n}_{\in\, \text{span}(\mathcal{B} \cup \mathcal{C})} \\ \Rightarrow x &\in \text{span}(\mathcal{A}) \cap \text{span}(\mathcal{B} \cup \mathcal{C}) \subset S \cap T = \langle \mathcal{B} \rangle \quad (\text{span}(\mathcal{A}) \subset S) \\ \Rightarrow x &\in \langle \mathcal{A} \rangle \cap \langle \mathcal{B} \rangle = \{0\} \\ \Rightarrow x &= 0, \text{ a contradiction.} \end{aligned}$$

Hence, $\mathcal{A} \cup \mathcal{B} \cup \mathcal{C}$ is L.I and a basis for $S + T$. Now,

$$\begin{aligned}
\dim(S) + \dim(T) &= |\mathcal{A} \cup \mathcal{B}| + |\mathcal{B} \cup \mathcal{C}| \\
&= |\mathcal{A}| + |\mathcal{B}| + |\mathcal{B}| + |\mathcal{C}| \\
&= |\mathcal{A}| + |\mathcal{B}| + |\mathcal{C}| + \dim(S \cap T) \\
&= \dim(S + T) + \dim(S \cap T),
\end{aligned}$$

as desired. □

Chapter 2

Linear Transformations

2.1 Linear Transformations

Roughly speaking, a linear transformation is a function from one vector space to another that preserves the vector space operations.

Definition 2.1.1. Let V and W be vector spaces over a field F. A function $\tau : V \to W$ is a *linear transformation (L.T)* if

$$\tau(u+v) = \tau(u) + \tau(v) \text{ and } \tau(ru) = r\tau(u)$$

for all scalars $r \in F$ and vectors $u, v \in V$. The set of all linear transformations from $V \to W$ is denoted by $\mathcal{L}(V, W)$.

- A L.T from V to V is called a *linear operator* on V. The set of all linear operators on V is denoted by $\mathcal{L}(V)$.
- A linear operator on a real vector space is called a *real operator* and a linear operator on a complex vector space is called a *complex operator*.
- A L.T from V to the base field F (thought of as a vector space over itself) is called a *linear functional* on V. The set of all linear functions on V is denoted by V^* and called the *dual space* of V.

Definition 2.1.2. The following terms are also employed:

- **homomorphism** for L.T denoted also by $\mathrm{Hom}(V, W)$;
- **endomorphism** for L. operator denoted also by $\mathrm{End}(V)$;
- **monomorphism (embedding)** for injective L.T;
- **epimorphism** for surjective L.T;

- **isomorphism (invertible L.T)** for bijective L.T $\tau \in \mathcal{L}(V, W)$. In this case, we write $V \cong W$ to say that V and W are isomorphic. The set of all linear isomorphisms from V to W is denoted $\mathrm{GL}(V, W)$.

- **automorphism** for bijective L. operator. The set of all automorphisms of V is denoted $\mathrm{Auto}(V)$ or $\mathrm{GL}(V)$.

Example 2.1.3.

⊚ The derivative $D : V \to V$ is a linear operator on the vector space V of all infinitely differentiable functions on $\mathbb{R}$.

⊚ Let $V = \mathbb{R}^2$ and let $W = \mathbb{R}$. Define $L : V \to W$ by $f(v, w) = vw$. Is L a L.T?

⊚ The integral operator $\tau : F[x] \to F[x]$ defined by

$$\tau(f) = \int_0^x f(t)dt$$

is a linear operator on $F[x]$.

⊚ Let $V = \mathbb{R}^2$ and let $W = \mathbb{R}^3$. Define $L : V \to W$ by $L(v, w) = (v, w - v, w)$. Is L a L.T?

⊚ Let A be an $m \times n$ matrix over F. The function

$$\begin{aligned} \tau_A : F^n &\to F^m, \\ v &\mapsto Av, \end{aligned}$$

where all vectors are written as column vectors, is a L.T from $F^n \to F^m$.

Note:

⊖ The set $\mathcal{L}(V, W)$ is a vector space in its own right.

⊖ The *identity transformation*, $I_V : V \to V$, given by $I_V(x) = x$ for all $x \in V$. Clearly, since $I_V(av + bu) = av + bu = aI_V(u) + bI_V(v)$, I_V is L.T.

⊖ The *zero transformation*, $\tau_0 : V \to W$, given by $\tau_0(x) = 0$ for all $x \in V$, is a L.T.

⊖ If $\tau \in \mathcal{L}(V)$ such that $\tau^2 = \tau$, we call τ an *idempotent operator.*

2.2 Basic properties of Linear Transformations

In the following we collect a few facts about linear transformations:

Theorem 2.2.1. *Let τ be a L.T from a vector space V into a vector space W. Then*

i) $\tau(0) = 0$.

ii) $\tau(-v) = -\tau(v)$ *for all* $v \in V$.

iii) $\tau(u - v) = \tau(u) - \tau(v)$ *for all* $u, v \in V$.

iii) $\tau\left(\sum_{k=1}^{n} a_k v_k\right) = \sum_{k=1}^{n} a_k \tau(v_k)$ *for all* $v_1, \ldots, v_k \in V$.

Theorem 2.2.2. *Let V and W be vector spaces over over a field F and let $\mathcal{B} = \{v_i \mid i \in I\}$ is a basis for V. Then for any $\tau \in \mathcal{L}(V, W)$, we have* $\operatorname{im}(\tau) = \langle \tau(\mathcal{B}) \rangle$.

Theorem 2.2.3.

a) *The set $\mathcal{L}(V, W)$ is a vector space under ordinary addition of functions and scalar multiplication of functions by elements of F.*

b) *If $\sigma \in \mathcal{L}(U, V)$ and $\tau \in \mathcal{L}(V, W)$, then the composition $\tau\sigma$ is in $\mathcal{L}(U, W)$.*

c) *If $\tau \in \mathcal{L}(V, W)$ is bijective, then $\tau^{-1} \in \mathcal{L}(W, V)$.*

Proof. b) Since for all scalars $r, s \in F$ and vectors $u, v \in U$

$$\begin{aligned} \tau\sigma(ru + sv) &= \tau(r\sigma(u) + s\sigma(v)) && (\sigma \in \mathcal{L}(U, V)) \\ &= r(\tau\sigma(u)) + s(\tau\sigma(v)) && (\tau \in \mathcal{L}(V, W)) \\ &\Rightarrow \ \tau\sigma \in \mathcal{L}(U, W). \end{aligned}$$

c) Let $\tau : V \to W$ be a bijective L.T. Then $\tau^{-1} : W \to V$ is a well-defined function and since any two vectors w_1 and w_2 in W have the form $w_1 = \tau v_1$ and $w_2 = \tau v_2$, we have

$$\begin{aligned} \tau^{-1}(rw_1 + sw_2) &= \tau^{-1}(r\tau v_1 + s\tau v_2) \\ &= \tau^{-1}(\tau(rv_1 + sv_2)) \\ &= rv_1 + sv_2 \\ &= r\tau^{-1}(w_1) + s\tau^{-1}(w_2) \\ &\Rightarrow \ \tau^{-1} \in \mathcal{L}(V, W). \end{aligned}$$

□

One of the easiest ways to define a L.T is to give its values on a basis.

Theorem 2.2.4. *Let V and W be vector spaces and let $\mathcal{B} = \{v_i \mid i \in I\}$ be a basis for V. Then we can define a L.T $\tau \in \mathcal{L}(V, W)$ by specifying the values of $\tau(v_i)$ arbitrarily for all $v_i \in \mathcal{B}$ and extending τ to V by linearity, that is,*

$$\tau(a_1 v_1 + \ldots + a_n v_n) = a_1 \tau(v_1) + \ldots + a_n \tau(v_n).$$

This process defines a unique L.T, that is, if $\tau, \sigma \in \mathcal{L}(V, W)$ satisfying $\tau(v_i) = \sigma(v_i)$ for all $v_i \in \mathcal{B}$, then $\tau = \sigma$.

Note that if $\tau \in \mathcal{L}(V, W)$ and if S is a subspace of V, then the restriction $\tau|S$ of τ to S is a L.T from S to W.

2.3 The Kernel and Image of a L.T

Definition 2.3.1. Let $\tau \in \mathcal{L}(V, W)$.

- ⊙ The subspace
$$\ker(\tau) = \{v \in V \mid \tau(v) = 0\}$$
is called the *kernel* of τ.

- ⊙ The subspace
$$\mathrm{im}(\tau) = \{\tau(v) \in W \mid v \in V\}$$
is called the *image* of τ.

- ⊙ The dimension of $\ker(\tau)$ is called the *nullity* of τ and is denoted by $\mathrm{null}(\tau)$.

- ⊙ The dimension of $\mathrm{im}(\tau)$ is called the *rank* of τ and is denoted by $\mathrm{rk}(\tau)$.

Remark and Exercise 2.3.2.

- $\ker(\tau)$ is a subspace of V.
- $\mathrm{im}(\tau)$ is a subspace of W.

Theorem 2.3.3. *Let $\tau \in \mathcal{L}(V, W)$. Then*

1) τ is surjective if and only if $\mathrm{im}(\tau) = W$.

2) τ is injective if and only if $\ker(\tau) = \{0\}$.

Proof. 1) is clear. 2) Observe that,

$$\tau(v) = \tau(u) \Leftrightarrow \tau(v-u) = 0 \Leftrightarrow u - v \in \ker(\tau) = \{0\}$$

which implies $u = v$ and, hence, τ is injective. Conversely, suppose τ is injective and $u \in \ker(\tau)$. Then $\tau(u) = 0 = \tau(0)$ and, hence, $u = 0$. □

Theorem 2.3.4. *Let $\tau \in \mathcal{L}(V, W)$ be an isomorphism. Let $S \subset V$. Then*

a) S spans V if and only if $\tau(S) = \{\tau(u) \mid u \in S\}$ spans W.

b) S is L.I in V if and only if $\tau(S)$ is L.I in W.

c) S is a basis for V if and only if $\tau(S)$ is a basis for W.

Proof. a) $V = \langle S \rangle \Leftrightarrow W = \mathrm{im}(\tau) = \tau(\langle S \rangle) = \langle \tau(S) \rangle$ (since $\tau \in \mathrm{GL}(V, W)$).

b) By given S is L.I. For any $s_1, \ldots, s_n \in S$, we have

$$\sum_{i=1}^{n} a_i s_i = 0 \Leftrightarrow a_i = 0 \text{ for all } i,$$

which implies

$$\begin{aligned} \tau\left(\sum_{i=1}^{n} a_i s_i\right) &= \sum_{i=1}^{n} a_i \tau(s_i) = 0 = \tau(0) \\ &\Rightarrow \sum_{i=1}^{n} a_i s_i = 0 \ \ (\tau \in \mathrm{GL}(V, W)) \\ &\Rightarrow a_1 = \ldots = a_n = 0 \ \ (S \text{ is L.I }) \\ &\Rightarrow \tau(S) \text{ is L.I} \quad (\text{ since this is true for all } s_i \in S). \end{aligned}$$

Conversely, if $\tau(S)$ is L.I we have for any $\tau(s_1), \ldots, \tau(s_n) \in \tau(S)$

$$\begin{aligned} 0 = \sum_{i=1}^{n} a_i \tau(s_i) = \tau\left(\sum_{i=1}^{n} a_i s_i\right) &= \tau(0) \\ &\Rightarrow \sum_{i=1}^{n} a_i s_i = 0 \ \ (\tau \in \mathrm{GL}(V, W)) \\ &\Rightarrow a_1 = \ldots = a_n = 0 \ \ (\tau(S) \text{ is L.I }) \\ &\Rightarrow S \text{ is L.I}. \end{aligned}$$

c) S is a basis for V iff, by a) and b), $\tau(S)$ is L.I in W and $W = \langle \tau(S) \rangle$ which implies $\tau(S)$ is a basis for W.

□

Isomorphisms Preserve Bases

An isomorphism can be characterized as a L.T $\tau : V \to W$ that maps a basis for V to a basis for W.

Theorem 2.3.5. *A L.T $\tau \in \mathcal{L}(V, W)$ is an isomorphism if and only if there is a basis $\mathcal{B}$ for V for which $\tau(\mathcal{B})$ is a basis for W. In this case, τ maps any basis of V to a basis of W.*

Proof. $\tau \in \mathrm{GL}(V, W) \Rightarrow \tau$ is bijective. Thus by Theorem 2.2.2 $\tau(\mathcal{B})$ is a basis for W. Conversely, if $\tau(\mathcal{B})$ is a basis for W, then for all $v \in V$, there exist unique elements $a_1, \ldots, a_n \in F$ and $u_1, \ldots, u_n$ such that $u = a_1u_1 + \ldots + a_nu_n$. Therefore,

$$\begin{aligned} 0 = \tau(u) &= a_1\tau(u_1) + \ldots + a_n\tau(u_n) \\ &\Rightarrow a_1 = \ldots = a_n = 0 \\ &\Rightarrow \ker(\tau) = \{0\} \\ &\Rightarrow \tau \text{ is injective.} \end{aligned}$$

Since $W = \langle \tau(\mathcal{B}) \rangle$, we have for all $w \in W$ there exist unique elements $a_1, \ldots, a_n \in F$ such that

$$w = a_1\tau(u_1) + \ldots + a_n\tau(u_n) = \tau(a_1u_1 + \ldots + a_nu_n).$$

So there exists $u = a_1u_1 + \ldots + a_nu_n \in V$ such that $w = \tau(u) \in \tau(V) = \mathrm{im}(\tau)$ which implies $W \subset \mathrm{im}(\tau)$. Clearly, $\mathrm{im}(\tau) \subset W$ and, hence, τ is surjective. Thus τ is bijective. □

Isomorphisms Preserve Dimension

The following theorem says that, upto isomorphism, there is only one vector space of any given dimension over a given field.

Theorem 2.3.6.

(i) *Let V and W be vector spaces over F. Then $V \cong W$ if and only if dim$(V) = \dim(W)$.*

(ii) *If n is a natural number, then any n-dimensional vector space over F is isomorphic to F^n.*

Proof. (i) $V \cong W \Rightarrow \exists \tau \in \mathrm{GL}(V, W)$. Thus $\mathcal{B}$ is a basis for V implies $\tau(\mathcal{B})$ is a basis for W and $\dim(V) = |\mathcal{B}| = |\tau(\mathcal{B})| = \dim(W)$. Conversely, if $\dim(V) = |\mathcal{B}_1| = |\mathcal{B}_2| = \dim(W)$, where $\mathcal{B}_1$ (resp. $\mathcal{B}_2$) is a basis for V (resp. W), then $\exists \tau \in \mathrm{GL}(\mathcal{B}_1, \mathcal{B}_2)$.

Extending τ to V by linearity defines a unique $\tau \in \mathcal{L}(V, W)$ by Theorem 2.2.4 and τ is an isomorphism because it is surjective and injective, that is, $\mathrm{im}(\tau) = W$ and $\ker(\tau) = \{0\}$.

(ii) Clear by (i). □

2.4 The Rank-Nullity Theorem

Lemma 2.4.1. *If V and W are vector spaces over a field F and $\tau \in \mathcal{L}(V, W)$, then any complement of the kernel τ is isomorphic to the range of τ, that is,*

$$V = \ker(\tau) \oplus \ker(\tau)^c \Rightarrow \ker(\tau)^c \cong \mathrm{im}(\tau)$$

where $\ker(\tau)^c$ is any complement of $\ker(\tau)$.

Proof. $V = \ker(\tau) \oplus \ker(\tau)^c \Rightarrow \dim(V) = \dim\left(\ker(\tau)\right) + \dim\left(\ker(\tau)^c\right)$. Let τ^c be the restriction of τ to $\ker(\tau)^c$. That is,

$$\tau^c : \ker(\tau)^c \to \mathrm{im}(\tau).$$

We claim that the map τ^c is bijective. □

To see this, note that the map τ^c is injective since

$$\ker(\tau^c) = \ker(\tau) \cap \ker(\tau)^c = \{0\}.$$

Clearly, $\mathrm{im}(\tau^c) \subset \mathrm{im}(\tau)$. For the reverse inclusion, if $\tau(v) \in \mathrm{im}(\tau)$, then since $v = u + w$ for $u \in \ker(\tau)$ and $w \in \ker(\tau)^c$, we have

$$\tau(v) = \tau(u) + \tau(w) = \tau(w) = \tau^c(w) \in \mathrm{im}(\tau^c).$$

Thus $\mathrm{im}(\tau^c) = \mathrm{im}(\tau)$ which implies

$$\tau^c : \ker(\tau)^c \to \mathrm{im}(\tau)$$

is an isomorphism.

Theorem 2.4.2 (Rank-Nullity Theorem). *Let V and W be vector spaces over a field F and let $\tau \in \mathcal{L}(V, W)$. Then*

$$\dim(\ker(\tau)) + \dim(\mathrm{im}(\tau)) = \dim(\mathrm{V})$$

or in other notation

$$\mathrm{rk}(\tau) + \mathrm{null}(\tau) = \dim(\mathrm{V})$$

Proof.

$$\begin{aligned}\dim(V) &= \dim(\ker(\tau)) + \dim(\ker(\tau)^c)\\ &= \dim(\ker(\tau)) + \dim(\operatorname{im}(\tau)) \quad \text{(Lemma 2.4.1)}\\ &= \operatorname{null}(\tau) + \operatorname{rk}(\tau)\end{aligned}$$

which completes the proof. □

Corollary 2.4.3. *Let V and W be vector spaces over a field F and $\tau \in \mathcal{L}(V, W)$. If $\dim(V) = \dim(W)$, then the following are equivalent:*

i) τ is injective.

ii) τ is surjective.

iii) $rk(\tau) = \dim(V)$.

Proof. By the Rank-Nullity Theorem, $\operatorname{rank}(\tau) + \operatorname{null}(\tau) = \dim(V)$ and , we have

$$\begin{aligned}\tau \text{ is 1-1} &\overset{\text{Thm 2.3.3}}{\Leftrightarrow} \ker(\tau) = \operatorname{null}(\tau) = \{0\}\\ &\overset{\text{R-N Thm}}{\Leftrightarrow} \dim(\operatorname{im}(\tau)) = \operatorname{rk}(\tau) = \dim(V) \overset{\text{assu.}}{=} \dim(V)\\ &\Leftrightarrow \operatorname{im}(\tau) = W\\ &\Leftrightarrow \tau \text{ is onto which completes the proof.}\end{aligned}$$

□