Dilla University

Department of Mathematics

Advanced Linear Algebra Lecture Note

by

Dereje Kifle (Dr.)

Contents

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.
3) $\mathbb{R}^n = \{(a_1, \ldots, a_n) \mid a_i \in \mathbb{R}\}$ is an n-dimensional vector space. The vector $(a_1, \ldots, a_n)$ is called an *n-tuple*.
4) $\mathbb{C}^n = \{(c_1, \ldots, c_n) \mid c_i \in \mathbb{R}\}$ is an n-dimensional vector space.
5) The set
$$P_n = \{f = \sum_{i=1}^{n} a_i x^i \mid a_i \in \mathbb{R} \text{ and } \deg f \leq n\}$$
of all polynomials (with coefficients in $\mathbb{R}$) of degree at most n is an $\mathbb{R}$-vector space.

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
$$a_1 v_1 + \ldots + a_n v_n$$

where $v_1 \dots v_n \in S$ and $a_1, \dots, a_n \in F$. 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 $V(n, 2)$ 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 $V(n, 2)$. □

Example 1.3.5. Let $V = \mathbb{R}^3$.

i) The subset $U = \{(a, b, 0) \in V \mid a, b \in R\}$ of V is a subspace of V.

ii) The subset $W = \{au + bv \mid a, b \in \mathbb{R} \text{ and } u, v \in V\}$ of V is a subspace of V.

Definition 1.3.6. 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:

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

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

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 (L.D).

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 $\quad v = \sum_{i=1}^{n} a_i s_i$ 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

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

(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. $\square$

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

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

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,

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