Math 416 Lecture Note: Week 1

Daesung Kim

1 Vector spaces

Definition 1.1. A vector space V over $\mathbb{R}$ is a set on which two operations (addition $+$ and scalar multiplication $\cdot$) are well-defined (meaning that for any $x, y \in V$, $x + y \in V$, and for any $x \in V$ and $c \in \mathbb{R}$, $c \cdot x \in V$) with the following properties

(1) $x + y = y + x$ for all $x, y \in V$,

(2) $(x + y) + z = x + (y + z)$ for all $x, y, z \in V$,

(3) there exists $0 \in V$ such that $x + 0 = x$ for all $x \in V$,

(4) for each $x \in V$, there exists y such that $x + y = 0$,

(5) $1 \cdot x = x$ for all $x \in V$,

(6) $(ab) \cdot x = a \cdot (b \cdot x)$ for all $a, b \in \mathbb{R}$ and $x \in V$,

(7) $a \cdot (x + y) = a \cdot x + b \cdot y$ for all $a \in \mathbb{R}$ and $x, y \in V$,

(8) $(a + b) \cdot x = a \cdot x + b \cdot x$ for all $a, b \in \mathbb{R}$ and $x \in V$.

Remark 1.2. Note that a vector space can be defined over not only $\mathbb{R}$ but also $\mathbb{C}$ or $\mathbb{Q}$. More generally, it can be replaced by a field. For further information, see [FIS, Appendix C].

Example 1.3. For each $n \in \mathbb{N}$, $\mathbb{R}^n = \{(x_1, \cdots, x_n) : x_i \in \mathbb{R}\}$ with the operations

$$(x_1, \cdots, x_n) + (y_1, \cdots, y_n) = (x_1 + y_1, \cdots, x_n + y_n), \quad c(x_1, \cdots, x_n) = (cx_1, \cdots, cx_n)$$

is a vector space over $\mathbb{R}$. Note that $0 = (0, \cdots, 0)$ and

$$(x_1, \cdots, x_n) + (-x_1, \cdots, -x_n) = (0, \cdots, 0)$$

for each $(x_1, \cdots, x_n) \in \mathbb{R}^n$.

Example 1.4. Let $\mathcal{M}_{m\times n}(\mathbb{R})$ be the collection of all $(m \times n)$ matrices with real entries. For $A \in \mathcal{M}_{m\times n}(\mathbb{R})$, we simply denote by $A = (A_{ij})$ where A_{ij} is the (i, j)-th entry of A. Define addition and scalar multiplication component-wisely, that is,

$$A + B = (A_{ij}) + (B_{ij}) = (A_{ij} + B_{ij}), \quad cA = c(A_{ij}) = (cA_{ij}).$$

Then, $\mathcal{M}_{m\times n}(\mathbb{R})$ is a vector space over $\mathbb{R}$. Note that the zero vector is $O = (0)$ and for each $A \in \mathcal{M}_{m\times n}(\mathbb{R})$, let $B = (-A_{ij})$ then $A+B = O$. Note also that we say two matrices $A, B \in \mathcal{M}_{m\times n}(\mathbb{R})$ are equal if $A_{ij} = B_{ij}$ for all i, j.

Example 1.5. Let S be a nonempty set and $\mathcal{F}(S, \mathbb{R})$ the collection of all functions $f : S \to \mathbb{R}$. We say two functions $f, g \in \mathcal{F}(S, \mathbb{R})$ are equal if $f(s) = g(s)$ for all $s \in S$. Define two operations by

$$(f + g)(s) = f(s) + g(s), \quad (cf)(s) = cf(s)$$

for all $s \in S$, $f, g \in \mathcal{F}(S, \mathbb{R})$, and $c \in \mathbb{R}$. Then, $\mathcal{F}(S, \mathbb{R})$ is a vector space over $\mathbb{R}$. Note that the zero vector in $\mathcal{F}(S, \mathbb{R})$ is a function 0 such that $0(s) = 0$. For each $f \in \mathcal{F}(S, \mathbb{R})$, define $-f \in \mathcal{F}(S, \mathbb{R})$ by $(-f)(s) = -f(s)$, then $f + (-f) = 0$.

Example 1.6. Let $\mathcal{P}(\mathbb{R})$ be the collection of all polynomials $p(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$ where $n \in \mathbb{N}$ and $a_i \in \mathbb{R}$. The degree of a polynomial is the highest exponent of x. We use the notation $\deg(p) = n$ if $p(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$. Note that polynomials are a special case of functions. Define addition and scalar multiplication as in Example 1.4. Then, $\mathcal{P}(\mathbb{R})$ is a vector space over $\mathbb{R}$.

Example 1.7. Let $\mathcal{S}(\mathbb{R})$ be the set of all sequences $\{a_n\}_{n\in\mathbb{N}} = \{a_1, a_2, \cdots, a_n, \cdots\}$ with $a_i \in \mathbb{R}$. To each sequence $\{a_n\}_{n\in\mathbb{N}}$, we associate a function $a : \mathbb{N} \to \mathbb{R}$ in a sense that $a(n) = a_n$ for each $n \in \mathbb{N}$. Thus, $\mathcal{S}(\mathbb{R})$ is a special case of $\mathcal{F}(S, \mathbb{R})$ where $S = \mathbb{N}$.

Theorem 1.8. *Let V be a vector space over $\mathbb{R}$. If $x, y, z \in V$ and $x + z = y + z$, then $x = y$.*

Proof. It follows that

$$\begin{aligned}
x &= x + 0 \qquad \text{(by (3))} \\
&= x + (z + (-z)) \qquad \text{(by (4))} \\
&= (x + z) + (-z) \qquad \text{(by (2))} \\
&= (y + z) + (-z) \qquad \text{(by the hypothesis)} \\
&= y + (z + (-z)) \qquad \text{(by (2))} \\
&= y + 0 \qquad \text{(by (4))} \\
&= y \qquad \text{(by (3))}.
\end{aligned}$$

□

Corollary 1.9. *Let V be a vector space over $\mathbb{R}$. Then $0 \in V$ is unique.*

Proof. Homework. □

Corollary 1.10. *Let V be a vector space over $\mathbb{R}$. For each $x \in V$, the vector y that satisfies $x + y = 0$ is unique.*

Proof. Homework. □

Remark 1.11. Since such a vector y is unique for each x, we use the notation $-x$ for y.

Theorem 1.12.

(i) *$0 \cdot x = 0$ for all $x \in V$.*

(ii) *$(-a)x = -(ax) = a(-x)$ for all $a \in \mathbb{R}$ and $x \in V$.*

(iii) *$a \cdot 0 = 0$ for all $a \in \mathbb{R}$.*

Proof. (i) Since we have

$$\begin{aligned}
0 \cdot x + 0 \cdot x &= (0 + 0) \cdot x \qquad \text{(by (8))} \\
&= 0 \cdot x \\
&= 0 \cdot x + 0 \qquad \text{(by (3))},
\end{aligned}$$

it follows from Theorem 1.8 that $0 \cdot x = 0$.

(ii) Since $-(ax)$ is unique by Corollary 1.10, it suffices to show that

$$ax + (-a)x = ax + a(-x) = 0.$$

First, we have

$$\begin{aligned}
ax + (-a)x &= (a + (-a))x \qquad \text{(by (8))} \\
&= 0x \\
&= 0 \qquad \text{(by (i))}.
\end{aligned}$$

Also, we see

$$\begin{aligned} ax + a(-x) &= a(x + (-x)) \qquad \text{(by (7))} \\ &= a0 \\ &= 0 \qquad \text{(by (iii))}. \end{aligned}$$

(iii) This is similar to (i). Since

$$\begin{aligned} a0 + a0 &= a(0 + 0) \qquad \text{(by (7))} \\ &= a0 \\ &= a0 + 0 \qquad \text{(by (3))}, \end{aligned}$$

it follows from Theorem 1.8 that $a0 = 0$.

□

2 Subspaces

Definition 2.1. Let V be a vector space over $\mathbb{R}$ and W a subset of V. We say that W is a subspace of V is W is a vector space over $\mathbb{R}$ with the same operations $+$ and $\cdot$ as in V. We use the notation $W \leq V$.

Example 2.2. Trivial examples are $V \leq V$ and $\{0\} \leq V$.

Remark 2.3. Let V be a vector space over $\mathbb{R}$, then there are two operations $+ : V \times V \to V$ and $\cdot : \mathbb{R} \times V \to V$ under which V is closed. In addition, there are 8 properties

(1) $x + y = y + x$ for all $x, y \in V$,

(2) $(x + y) + z = x + (y + z)$ for all $x, y, z \in V$,

(3) there exists $0 \in V$ such that $x + 0 = x$ for all $x \in V$,

(4) for each $x \in V$, there exists y such that $x + y = 0$,

(5) $1 \cdot x = x$ for all $x \in V$,

(6) $(ab) \cdot x = a \cdot (b \cdot x)$ for all $a, b \in \mathbb{R}$ and $x \in V$,

(7) $a \cdot (x + y) = a \cdot x + b \cdot y$ for all $a \in \mathbb{R}$ and $x, y \in V$,

(8) $(a + b) \cdot x = a \cdot x + b \cdot x$ for all $a, b \in \mathbb{R}$ and $x \in V$.

Let W be a subset of V. In order for W to be a vector space, it is required that W is closed under $+$ and $\cdot$, that is, $+ : W \times W \to W$ and $\cdot : \mathbb{R} \times W \to W$. If we assume these, W automatically satisfies the properties (1)-(8) except (3) and (4).

Theorem 2.4. *Let V be a vector space over $\mathbb{R}$ and W a subset of V. Then, $W \leq V$ if and only if the following hold.*

(i) $0 \in W$ *(here 0 is the zero vector for V).*

(ii) *If $x, y \in W$, then $x + y \in W$.*

(iii) *If $x \in W$ and $c \in \mathbb{R}$, then $cx \in W$.*

Proof. ($\Rightarrow$): Since W is a vector space, (ii) and (iii) are satisfied. It suffices to show only (i). Indeed, W has its zero vector, say 0_W because it is a vector space. We need to justify that it is equal to the zero vector for V, say 0_V. By definition, we have $x + 0_V = x$ for all $x \in V$. In particular, $x + 0_V = x$ for all $x \in W$, which implies that 0_V is another zero vector for W. Since the zero vector is unique (Corollary 1.9), we obtain $0_V = 0_W$.

($\Leftarrow$): As we have seen in the remark, the properties (1),(2),(5)-(8) hold. The property (3) is okay by (i). Thus we need to show the property (4). Let $x \in W$, then it follows from part (ii) of Theorem 1.12 and the hypothesis (iii) that $-x = (-1)x \in W$, which finishes the proof. □

Example 2.5. The set of complex numbers $\mathbb{C}$ is a vector space over $\mathbb{R}$ with the standard addition and multiplication. Then $\mathbb{R} \leq \mathbb{C}$.

Example 2.6. Let $V = \mathbb{R}^n$ and $W = \{(x_1, \cdots, x_{n-1}, 0) : x_i \in \mathbb{R}\} \subset V$, then $W \leq V$. However, if we let $U = \{(x_1, \cdots, x_{n-1}, 1) : x_i \in \mathbb{R}\} \subset V$, then U is not a subspace.

Example 2.7. A matrix $A \in \mathcal{M}_{m\times n}(\mathbb{R})$ is a diagonal matrix if $A_{ij} = 0$ for all $i \neq j$. Then, the set of all diagonal matrices is a subspace of $\mathcal{M}_{m\times n}(\mathbb{R})$.

Definition 2.8. The trace of $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ is the sum of diagonal entries of A, that is,

$$\mathrm{tr}(A) = \sum_{i=1}^{n} A_{ii} = A_{11} + A_{22} + \cdots + A_{nn}.$$

Example 2.9. Let $V = \mathcal{M}_{n\times n}(\mathbb{R})$ and

$$W = \{A \in \mathcal{M}_{n\times n}(\mathbb{R}) : \mathrm{tr}(A) = 0\},$$

then it is a subspace of V.

Example 2.10. The degree of a polynomial is the highest exponent of x. We use the notation $\deg(p) = n$ if $p(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$. Let $V = \mathcal{P}(\mathbb{R})$ be the set of all polynomials with real coefficients and

$$W = \{p(x) \in \mathcal{P}(\mathbb{R}) : \deg(p) \leq n\}.$$

Then W is a subspace of V.

Theorem 2.11. *Let V be a vector space over $\mathbb{R}$ and $U, W \leq V$. Then, $U \cap W \leq V$.*

Proof. □

Theorem 2.12. *Let V be a vector space over $\mathbb{R}$ and $U, W \leq V$. We define*

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

Then, $U + W \leq V$

Proof. Homework. □

Remark 2.13. Note that $U + W$ is the smallest subspace of V containing both U and W. If $U \cap W = \{0\}$ and $U + W = V$, then we denote by $V = U \oplus W$.

Definition 2.14. Let $A \in \mathcal{M}_{m\times n}(\mathbb{R})$. The transpose of A, denoted by A^t, is an $n \times m$ matrix $(A^t)_{ij} = A_{ji}$.

Example 2.15. Let $V = \mathcal{M}_{n\times n}(\mathbb{R})$ and

$$W = \{A \in \mathcal{M}_{n\times n}(\mathbb{R}) : A^t = A\}.$$

We say $A \in W$ is symmetric. Then, $W \leq V$. To see this,

(i) For the zero matrix O, $O^t = O$, that is, $O \in W$.

(ii) Let $A, B \in W$, then $(A + B)^t = A^t + B^t = A + B$. Thus $A + B \in W$.

(iii) Let $A \in W$ and $c \in \mathbb{R}$, then $(cA)^t = cA^t = cA$, which yields $cA \in W$.

Similarly, $U = \{A \in \mathcal{M}_{n\times n}(\mathbb{R}) : A^t = -A\}$ is a subspace of V. Every matrix in U is called skew-symmetric. Note that $U \cap W = \{0\}$ and $U + W = V$ (Exercise!) so that $V = U \oplus W$.

References

[FIS] Freidberg, Insel, and Spence, *Linear Algebra*, 4th edition, 2002.

[Bee] Beezer, *A First Course in Linear Algebra*, Version 3.5, 2015.

Department of Mathematics, University of Illinois at Urbana-Champaign
E-mail address: daesungk@illinois.edu