Math 416 Lecture Note: Week 2

Daesung Kim

1 Linear Combination

Definition 1.1. Let V be a vector space over $\mathbb{R}$ and S a nonempty subset of V. A vector $v \in V$ is a linear combination of vectors in S if there exist $n \in \mathbb{N}$, $v_1, \cdots, v_n \in S$, and $a_1, \cdots, a_n \in \mathbb{R}$ such that

$$v = a_1 v_1 + \cdots + a_n v_n.$$

In this case, v is also called a linear combination of $v_1, \cdots, v_n$.

Example 1.2. Let $V = \mathbb{R}^3$, $v_1 = (1, 0, -1)$, $v_2 = (0, 1, 0)$, and $S = \{v_1, v_2\}$. Then v_1, v_2, $v_1 + v_2$, and $v_1 - v_2$ are linear combinations of S.

Definition 1.3. Let V be a vector space over $\mathbb{R}$ and S a nonempty subset of V. The span of S is the set of all linear combinations of S, denoted by

$$\mathrm{Span}(S) = \{a_1 v_1 + \cdots + a_n v_n : n \in \mathbb{N}, a_i \in \mathbb{R}, v_i \in V, i = 1, 2, \cdots, n\}.$$

Example 1.4. Let $V = \mathbb{R}^3$, $v_1 = (1, 0, -1)$, $v_2 = (0, 1, 0)$, and $S = \{v_1, v_2\}$. Every linear combination of S can be written as $av_1 + bv_2$ for $a, b \in \mathbb{R}$. Since

$$av_1 + bv_2 = (a, 0, -a) + (0, b, 0) = (a, b, -a),$$

the collection of all linear combinations of S forms a plane given by $x + z = 0$.

Example 1.5. Let V be a vector space and $v \in V$. Then, $\mathrm{Span}(\{v\}) = \{av : a \in \mathbb{R}\}$.

Theorem 1.6. *Let V be a vector space over $\mathbb{R}$ and S a nonempty subset of V.*

(i) $\mathrm{Span}(S) \leq V$.

(ii) $\mathrm{Span}(S)$ *is the smallest subspace of V containing S.*

Proof. (i): Since V is closed under addition and scalar multiplication, $\mathrm{Span}(S) \subseteq V$. Since $0 \cdot x = 0$ (see [FIS, Theorem 1.2, (a)]), $0 \in \mathrm{Span}(S)$. If $x, y \in \mathrm{Span}(S)$, then

$$\begin{aligned} x &= a_1 v_1 + \cdots + a_m v_m \\ y &= b_1 w_1 + \cdots + b_n w_n \end{aligned}$$

for some $a_i, b_j \in \mathbb{R}$ and $v_i, w_j \in S$. Then,

$$x + y = a_1 v_1 + \cdots + a_m v_m + b_1 w_1 + \cdots + b_n w_n \in \mathrm{Span}(S).$$

If $x \in \mathrm{Span}(S)$ and $c \in \mathbb{R}$, then

$$x = a_1 v_1 + \cdots + a_m v_m$$

for some $a_i \in \mathbb{R}$ and $v_i \in S$. Then,

$$cx = c(a_1 v_1 + \cdots + a_m v_m) = (ca_1)v_1 + \cdots + (ca_m)v_m \in \mathrm{Span}(S).$$

By [FIS, Theorem 1.3], $\mathrm{Span}(S)$ is a subspace of V.

(ii): Let $U \leq V$ and $S \subseteq U$. It suffices to show that $\mathrm{Span}(S) \subseteq U$. Let $x \in \mathrm{Span}(S)$, then

$$x = a_1 v_1 + \cdots + a_m v_m$$

for some $a_i \in \mathbb{R}$ and $v_i \in S$. Since $v_i \in U$ and U is closed under addition and scalar multiplication, $x \in U$. □

Definition 1.7. Let V be a vector space over $\mathbb{R}$ and S a nonempty subset of V. We say that S generates V is $\mathrm{Span}(S) = V$.

Example 1.8. Let $n \in \mathbb{N}$ and V be the set of all $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ such that $A_{ij} = 0$ if $i > j$. Such a matrix is called upper triangular matrix. Note that V is a subspace of $\mathcal{M}_{n\times n}(\mathbb{R})$. Consider the case $n = 2$. Let

$$S = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\},$$

then one can see that S generates V. (Homework)

Example 1.9. Let $V = \mathbb{R}^3$, $v_1 = (1, 1, -1)$, $v_2 = (-1, 1, 2)$, $S = \{v_1, v_2\}$, and $W = \mathrm{Span}(S)$. Let us find an equation for W of the form $ax + by + cz = 0$. Since $v_1, v_2 \in W$, we have

$$\begin{cases} a + b - c = 0 \\ -a + b + 2c = 0. \end{cases}$$

By solving the system of equations, one sees $a = \frac{3c}{2}$ and $b = -\frac{c}{2}$. Letting $c = 2$, the equation is $3x - y + 2z = 0$. Note that there are infinitely many solutions for the system of equations.

Example 1.10. Let $V = \mathbb{R}^3$, $v_1 = (1, 1, -1)$, $v_2 = (-1, 1, 2)$, $S = \{v_1, v_2\}$, and $W = \mathrm{Span}(S)$. A question is whether $w = (5, 1, -7)$ belongs to W. To see this, it is enough to find $a, b \in \mathbb{R}$ such that $w = av_1 + bv_2$. That is, we need to solve

$$\begin{cases} a - b = 5 \\ a + b = 1 \\ -a + 2b = -7. \end{cases}$$

By solving the system of equations, one sees $a = 3$ and $b = -2$. Note that $(3, -2)$ is the only solution.

Example 1.11. Let $V = \mathbb{R}^3$, $v_1 = (1, 1, -1)$, $v_2 = (-1, 1, 2)$, $S = \{v_1, v_2\}$, and $W = \mathrm{Span}(S)$. A question is whether $w = (6, 1, -7)$ belongs to W. To see this, it is enough to find $a, b \in \mathbb{R}$ such that $w = av_1 + bv_2$. That is, we need to solve

$$\begin{cases} a - b = 5 \\ a + b = 1 \\ -a + 2b = -7. \end{cases}$$

By solving the system of equations, one sees that there is no solution, which means that w does not belong to W.

Definition 1.12. A system of linear equations is a collection of m equations with n variables $x_1, \cdots, x_n$ of the form

$$(*) \begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases}$$

where $a_{ij}, b_i \in \mathbb{R}$ for $i \in \{1, 2, \cdots, m\}$, $j \in \{1, 2, \cdots, n\}$.

To solve a system of linear equations, we use the following procedures.

(1) Swap two equations.

(2) Multiply one equation by a nonzero $c \in \mathbb{R}$.

(3) Add one equation to another.

Example 1.13. Consider the following system of linear equations

$$\begin{cases} 3x - 7y + 4z = 10 \\ x - 2y + z = 3 \\ 2x - y - 2z = 6. \end{cases}$$

First, switch the first and the second equation. Consider the following system of linear equations

$$\begin{cases} x - 2y + z = 3 \\ 3x - 7y + 4z = 10 \\ 2x - y - 2z = 6. \end{cases}$$

Then, add (-3) times of the first equation to the second and add (-2) times of the first equation to the third

$$\begin{cases} x - 2y + z = 3 \\ -y + z = 1 \\ 3y - 4z = 0. \end{cases}$$

Add (-2) times of the second equation to the first and add 3 times of the second equation to the third

$$\begin{cases} x - z = 1 \\ -y + z = 1 \\ -z = 3. \end{cases}$$

Thus the solution is $(x, y, z) = (-2, -4, -3)$.

2 A System of Linear Equations

We consider a system of linear equations

$$(*) \begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases}$$

where $a_{ij}, b_i \in \mathbb{R}$ for $i \in \{1, 2, \cdots, m\}$, $j \in \{1, 2, \cdots, n\}$. We associate it to a coefficient matrix, a constant vector, and an augmented matrix

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}, \quad b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}, \quad (A|b) = \begin{pmatrix} a_{11} & \cdots & a_{1n} & b_1 \\ \vdots & \ddots & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} & b_m \end{pmatrix}.$$

Then the system of linear equations $(*)$ can be written as $Ax = b$ where

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

Example 2.1. Consider the following system of linear equations

$$\begin{cases} 3x - 7y + 4z = 10 \\ x - 2y + z = 3 \\ 2x - y - 2z = 6 \end{cases}.$$

Then, its associated coefficient matrix, constant vector, and augmented matrix are

$$A = \begin{pmatrix} 3 & -7 & 4 \\ 1 & -2 & 1 \\ 2 & -1 & -2 \end{pmatrix}, \quad b = \begin{pmatrix} 10 \\ 3 \\ 6 \end{pmatrix}, \quad M = (A, b) = \begin{pmatrix} 3 & -7 & 4 & 10 \\ 1 & -2 & 1 & 3 \\ 2 & -1 & -2 & 6 \end{pmatrix}.$$

For a matrix $M = (A, b)$, we use the notation

$$LS(M) = LS(A, b) = \begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases}.$$

2.1 Row operations

Row operations transform a matrix $M \in \mathcal{M}_{m\times n}(\mathbb{R})$ to another matrix with the same size. There are three types of operations.

(1) Swap two rows of M. ($R_i \leftrightarrow R_j$)

(2) Multiply one row by a nonzero constant $c \in \mathbb{R}$. ($R_i \to cR_i$)

(3) Add one row to another. ($R_i \to R_i + R_j$).

Note that combining (ii) and (iii), we often use the operation $R_i \to R_i + cR_j$.

Example 2.2.

$$M = (A, b) = \begin{pmatrix} 3 & -7 & 4 & 10 \\ 1 & -2 & 1 & 3 \\ 2 & -1 & -2 & 6 \end{pmatrix} \xrightarrow{R_1 \leftrightarrow R_2} \begin{pmatrix} 1 & -2 & 1 & 3 \\ 3 & -7 & 4 & 10 \\ 2 & -1 & -2 & 6 \end{pmatrix} \xrightarrow{R_2 \to R_2 - 3R_1} \begin{pmatrix} 1 & -2 & 1 & 3 \\ 0 & -1 & 1 & 1 \\ 2 & -1 & -2 & 6 \end{pmatrix}$$

Definition 2.3. Let $M, N \in \mathcal{M}_{m\times n}(\mathbb{R})$. A matrix M is row equivalent to N if there exists a sequence of row operations that turns M to N. We denote by $M \sim N$.

Proposition 2.4. *Let $M, N, L \in \mathcal{M}_{m\times n}(\mathbb{R})$. Then we have the following.*

(i) $M \sim M$.

(ii) *If $M \sim N$, then $N \sim M$.*

(iii) *If $M \sim N$ and $N \sim L$, then $M \sim L$.*

Proof. Homework. □

Theorem 2.5. *Let $M, N \in \mathcal{M}_{m\times n}(\mathbb{R})$. If $M \sim N$, then the solution set of $LS(M)$ is equal to that of $LS(N)$.*

Proof. By Proposition 2.4, it suffices to assume that N is obtain by a single row operation from M. If the operation is reordering (i) or multiplying (ii), then it is trivial. Suppose $M = (A, b)$ and N is obtained by $R_i \to R_i + R_j$. Then i-th and j-th equations of M and N are

$$(*) \begin{cases} a_{i1}x_1 + \cdots + a_{in}x_n = b_i \\ a_{j1}x_1 + \cdots + a_{jn}x_n = b_j \end{cases}$$

and

$$(**)\begin{cases}(a_{i1}+a_{j1})x_1+\cdots+(a_{in}+a_{jn})x_n=b_i+b_j\\ a_{j1}x_1+\cdots+a_{jn}x_n=b_j.\end{cases}$$

If $(x_1,\cdots,x_n)$ satisfies $(*)$, then it also does $(**)$. Vice versa. $\square$

Example 2.6. Consider the following system of linear equations

$$\begin{cases}3x-7y+4z=10\\ x-2y+z=3\\ 2x-y-2z=6\end{cases}.$$

Then, the associated augmented matrix is row-equivalent to

$$M=(A,b)=\begin{pmatrix}3&-7&4&10\\1&-2&1&3\\2&-1&-2&6\end{pmatrix}\xrightarrow{R_1\leftrightarrow R_2}\begin{pmatrix}1&-2&1&3\\3&-7&4&10\\2&-1&-2&6\end{pmatrix}\xrightarrow{R_2\to R_2-3R_1}\begin{pmatrix}1&-2&1&3\\0&-1&1&1\\2&-1&-2&6\end{pmatrix}$$

$$\xrightarrow{R_3\to R_3-2R_1}\begin{pmatrix}1&-2&1&3\\0&-1&1&1\\0&3&-4&0\end{pmatrix}\xrightarrow{R_2\to -R_2}\begin{pmatrix}1&-2&1&3\\0&1&-1&-1\\0&3&-4&0\end{pmatrix}$$

$$\xrightarrow{R_3\to R_3-3R_2}\begin{pmatrix}1&-2&1&3\\0&1&-1&-1\\0&0&-1&3\end{pmatrix}\xrightarrow{R_1\to R_1+2R_2}\begin{pmatrix}1&0&-1&1\\0&1&-1&-1\\0&0&-1&3\end{pmatrix}$$

$$\xrightarrow{R_3\to -2R_3}\begin{pmatrix}1&0&-1&1\\0&1&-1&-1\\0&0&1&-3\end{pmatrix}\xrightarrow[R_1\to R_1+R_3]{R_2\to R_2+R_3}\begin{pmatrix}1&0&0&-2\\0&1&0&-4\\0&0&1&-3\end{pmatrix}=N.$$

The corresponding system of linear equation $LS(N)$ is

$$\begin{cases}x_1=-2\\ x_2=-4\\ x_3=-3\end{cases}.$$

Check that $(-2,-4,-3)$ is also a solution to the original system of linear equation $LS(M)$.

2.2 Reduced Row-Echelon Form (RREF)

Definition 2.7. A matrix $A\in\mathcal{M}_{m\times n}(\mathbb{R})$ is in reduced row-echelon form (RREF) if

(1) All zero rows are at the bottom;

(2) The leftmost nonzero entry of each row is 1 (it is called *leading 1*);

(3) A leading 1 is the only nonzero entry in its column;

(4) If (i,j) and (s,t) are leading 1's and $i>s$, then $j>t$.

Example 2.8. Let $M\in\mathcal{M}_{4\times 6}(\mathbb{R})$ be row-equivalent to

$$N=(A,b)=\begin{pmatrix}1&2&0&3&0&4\\0&0&1&-4&0&3\\0&0&0&0&1&2\\0&0&0&0&0&0\end{pmatrix}.$$

Note that N is in reduced row-echelon form. Its associated system of linear equations is

$$\begin{cases} x_1 + 2x_2 + 3x_4 = 4 \\ x_3 - 4x_4 = 3 \\ x_5 = 2. \end{cases}$$

Note that every variable that corresponds to a leading 1 depends on the non-leading 1 variables. Set $x_2 = s$ and $x_4 = t$, then the solutions set of the system of linear equations is

$$\{(-2s - 3t + 4, s, 4t + 3, t, 2) : s, t \in \mathbb{R}\}.$$

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