Advanced Linear Algebra

Chapter 2: Vector Spaces and Linear Equations

Jin-Ho Chung
School of IT Convergence
University of Ulsan

Table of Contents

Definition of Vector Space (1)

Definition

A nonempty set V is called *a vector space over a field F* if the following conditions are satisfied:

1) (closure) $\alpha + \beta \in V \quad \forall \alpha, \beta \in V$;
2) (associativity) $(\alpha + \beta) + \gamma = \alpha + (\beta + \gamma) \quad \forall \alpha, \beta, \gamma \in V$;
3) (additive identity) $\exists \mathbf{0} \in V$ such that $\alpha + \mathbf{0} = \mathbf{0} + \alpha = \alpha \quad \forall \alpha \in V$;
4) (additive inverse) $\forall \alpha \in V \ \ \exists (-\alpha) \in V$ such that $\alpha + (-\alpha) = (-\alpha) + \alpha = \mathbf{0}$;
5) (commutativity) $\alpha + \beta = \beta + \alpha \quad \forall \alpha, \beta \in V$;

Definition of Vector Space (2)

Definition (contd.)

6) (closure) $a\alpha \in V \quad \forall a \in F,\ \forall \alpha \in V$;
7) (associativity) $(ab)\alpha = a(b\alpha) \quad \forall a, b \in F,\ \forall \alpha \in V$;
8) (identity) $1\alpha = \alpha \quad \forall \alpha \in V$;
9) (distributivity)

$$a(\alpha + \beta) = a\alpha + a\beta \quad \forall a \in F,\ \alpha, \beta \in V$$
$$(a + b)\alpha = a\alpha + b\alpha, \quad \forall a, b \in F,\ \forall \alpha \in V.$$

Examples of Vector Spaces (1)

Example 1 - The set of complex numbers

$$\mathbb{C} = \{a + bi \,|\, a, b \in \mathbb{R}\}.$$

The set $\mathbb{C}$ is a vector space of dimension ( ) and over ( ).

Example 2 - A set of functions

Let $C(-\infty, \infty)$ be the set of all continuous functions defined on the real line with two operations

$$\begin{aligned} (f+g)(x) &\triangleq f(x) + g(x); \\ (cf)(x) &\triangleq c(f(x)). \end{aligned}$$

Examples of Vector Spaces (2)

Example 3 - The set of polynomials

Let P_n be the set of all polynomials over $\mathbb{R}$ of degree $\leq n$, that is,

$$P_n = \left\{ \sum_{i=0}^{n} a_i x^i \;\middle|\; a_i \in \mathbb{R} \right\}.$$

For $p(x) = \sum_{i=0}^{n} p_i x^i$ and $q(x) = \sum_{i=0}^{n} q_i x^i$, we define

$$p(x) + q(x) = \sum_{i=0}^{n} (p_i + q_i) x^i;$$

$$cp(x) = \sum_{i=0}^{n} (cp_i) x^i, \quad c \in \mathbb{R}.$$

Properties of a Vector Space

1) $c\mathbf{0} = \mathbf{0} \quad \forall c \in F$;
2) $0\mathbf{u} = \mathbf{0} \quad \forall \mathbf{u} \in V$;
3) $c\mathbf{u} = \mathbf{0} \quad \Rightarrow \quad c = 0$ or $\mathbf{u} = \mathbf{0}$;
4) $(-1)\mathbf{u} = -\mathbf{u}$.

Subspaces

Definition

A subset W of a vector space V over F is a subspace of V if and only if W is itself a vector space over F with respect to the same operations as in V.

Test for a subspace

A subset W of a vector space V over F is a subspace of V, provided that

1) if $\mathbf{x}, \mathbf{y} \in W$, then $\mathbf{x} + \mathbf{y} \in W$;
2) if $\mathbf{x} \in W$ and $c \in F$, then $c\mathbf{x} \in W$.

Note: V and $\{\mathbf{0}\}$ are trivial subspaces.

Generated Subspaces

Definition

Let $S = \{\mathbf{x}_1, \ldots, \mathbf{x}_k\}$ be a finite subset of a vector space V over F. We define *the subspace of V generated by S* as

$$\mathcal{L}(S) = \left\{ \sum_{i=1}^{k} a_i \mathbf{x}_i \;\middle|\; a_i \in F \right\}.$$

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

- The set of all linear combinations of columns of A, $\mathcal{C}(A)$ is called *the column space of A*;
- The set of all linear combinations of rows of A, $\mathcal{R}(A)$ is called *the row space of A*.
- The null space $\mathcal{N}(A)$ of A is defined as

$$\{\mathbf{x} \in F^n \mid A\mathbf{x} = \mathbf{0}\}.$$

Table of Contents

Column Space

Definition

For an $m \times n$ matrix A over $\mathbb{R}$, its column space $\mathcal{C}(A)$ is defined as

$$\mathcal{C}(A) = \{a_1\mathbf{c}_1 + \cdots + a_n\mathbf{c}_n \,|\, a_1, ..., a_n \in \mathbb{R}\}$$

where $\mathbf{c}_1, ..., \mathbf{c}_n$ are column vectors of A.

- The column space can be also written as

$$\mathcal{C}(A) = \{\mathbf{b} \,|\, \mathbf{b} = A\mathbf{x}, \quad \mathbf{x} \in \mathbb{R}^n\},$$

and so $\mathcal{C}(A)$ is the range of the linear transformation $\mathbf{x} \mapsto A\mathbf{x}$.

Row Space

Definition

For an $m \times n$ matrix A over $\mathbb{R}$, its row space $\mathcal{R}(A)$ is defined as

$$\mathcal{R}(A) = \{b_1\mathbf{r}_1 + \cdots + b_m\mathbf{r}_m \mid b_1, \ldots, b_m \in \mathbb{R}\}$$

where $\mathbf{r}_1, \ldots, \mathbf{r}_m$ are row vectors of A.

Null Space

Definition

For an $m \times n$ matrix A over $\mathbb{R}$, its null space $\mathcal{N}(A)$ is defined as

$$\mathcal{N}(A) = \{\mathbf{x} \mid A\mathbf{x} = \mathbf{0}.\}$$

- Note that $\mathcal{N}(A) = \{\mathbf{0}\}$ if and only if the linear transformation is one-to-one.

Table of Contents

Solvable System

Theorem

The system $A\mathbf{x} = \mathbf{b}$ is solvable if and only if $\mathbf{b} \in \mathcal{C}(A)$.

It is clear since if there is a solution $\mathbf{x}$ to the system

$$[\mathbf{a}_1 \ \mathbf{a}_2 \ \cdots \ \mathbf{a}_n] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix},$$

then

$$\mathbf{b} = \sum_{i=1}^{n} x_i \mathbf{a}_i.$$

General Linear System

Matrix Form

A system of m linear equations in n unknowns can be characterized as

$$\underbrace{A}_{m\times n}\underbrace{\mathbf{x}}_{n\times 1} = \underbrace{\mathbf{b}}_{m\times 1}$$

There are three possible cases.

1) No solution exists;
2) A unique solution exists;
3) Many solutions exist. (In general, infinite.)

General Solutions

Theorem

The general solution to a system of linear equations $A\mathbf{x} = \mathbf{b}$ is of the form

$$\mathbf{x} = \mathbf{x}_p + \mathbf{x}_h \tag{1}$$

where $\mathbf{x}_h \in \mathcal{N}(A)$ and $\mathbf{x}_p$ is any particular solution to $A\mathbf{x} = \mathbf{b}$. Conversely, any vector of the form (1) is a solution to the system.

$\Rightarrow$ Suppose that $A\mathbf{z} = \mathbf{b}$ and $A\mathbf{x}_p = \mathbf{b}$.
Then, $A(\mathbf{z} - \mathbf{x}_p) = \mathbf{0}$, so $\mathbf{z} - \mathbf{x}_p \in \mathcal{N}(A)$.
That is, any solution $\mathbf{z}$ is of the form $\mathbf{x}_p + \mathbf{y}$ for some $\mathbf{y} \in \mathcal{N}(A)$.
Therefore, every solution is of the form (1).

$\Leftarrow$ Trivial.

Echelon Form

Definition

A rectangular matrix is said to be in echelon form if it satisfies the following three conditions:

1) The nonzero rows come first and the first nonzero element in any nonzero row is the pivot;
2) Below each pivot is a column of zeros obtained by elimination;
3) Each pivot lies to the right of the pivot in the row above (staircase pattern).

Note that every matrix can be reduced to echelon form by using elementary row operations, that is,

$$PA = LU.$$

Example of Solving (1)

Step 1) Augmented matrix:

$$A\mathbf{x} = \mathbf{b} \quad \rightarrow \quad \left[A \,\vdots\, \mathbf{b}\right] \triangleq A_a.$$

Step 2) Elementary row operations:

$$\begin{bmatrix} 0 & 2 & -1 & 3 & b_1 \\ -1 & 1 & 2 & 0 & b_2 \\ 1 & 1 & -3 & 3 & b_3 \\ 1 & 5 & 5 & 9 & b_4 \end{bmatrix} \rightarrow \begin{bmatrix} -1 & 1 & 2 & 0 & b_2 \\ 0 & 2 & -1 & 3 & b_1 \\ 1 & 1 & -3 & 3 & b_3 \\ 1 & 5 & 5 & 9 & b_4 \end{bmatrix}$$

$$\rightarrow \begin{bmatrix} -1 & 1 & 2 & 0 & b_2 \\ 0 & 2 & -1 & 3 & b_1 \\ 0 & 0 & 10 & 0 & b_4 + b_2 - 3b_1 \\ 0 & 0 & 0 & 0 & b_3 + b_2 - b_1 \end{bmatrix} \triangleq U_a.$$

Example of Solving (2)

- Since elementary row operations do not change the solutions,

$$A\mathbf{x} = \mathbf{b} \quad \Leftrightarrow \quad U\mathbf{x} = \mathbf{b}'.$$

- Note that $EA = U \quad \Rightarrow \quad EA_a = U_a$.

- $U\mathbf{x} = \mathbf{b}'$ has a solution if and only if $b_3 + b_2 - b_1 = 0$.

- The variables x_i corresponding to the columns of U with nonzero pivot are called *pivot variables* (or basic variables). The other variables are called *free variables*.
  (In this example, x_1, x_2, x_3 are pivot variables, and x_4 is a free variable.)

Particular Solution

- Since x_4 is free, we set $x_4 = 0$. Then, $U\mathbf{x} = \mathbf{b}'$ becomes

$$\begin{bmatrix} -1 & 1 & 2 & 0 & \vdots & 2 \\ 0 & 2 & -1 & 3 & \vdots & 4 \\ 0 & 0 & 10 & 0 & \vdots & -10 \\ 0 & 0 & 0 & 0 & \vdots & 0 \end{bmatrix}.$$

- After solving the system, we get

$$\mathbf{x}_p = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} -5/2 \\ 3/2 \\ -1 \\ 0 \end{bmatrix}.$$

Homogeneous Solution

- Consider the equation $U\mathbf{x} = \mathbf{0}$, that is,

$$\left[\begin{array}{cccc:c} -1 & 1 & 2 & 0 & 0 \\ 0 & 2 & -1 & 3 & 0 \\ 0 & 0 & 10 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right].$$

- Then, it is easily checked that $x_3 = 0$ and $x_1 = x_2 = -\frac{3}{2}x_4$.
- Therefore, for an arbitrary x_4, the homogeneous solution is given by

$$\mathbf{x}_h = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = x_4 \begin{bmatrix} -3/2 \\ -3/2 \\ 0 \\ 1 \end{bmatrix}.$$

General Solution

The general solution is given by $\mathbf{x}_p + \mathbf{x}_h$.

Step 1) $A\mathbf{x} = \mathbf{b} \;\Rightarrow\; U\mathbf{x} = \mathbf{b}'$.

Step 2) When the number of pivots is r, $U\mathbf{x} = \mathbf{b}'$ is solvable if and only if the last $m - r$ components of $\mathbf{b}'$ are zero.

Step 3) Setting free variables to zero, find a particular solution $\mathbf{x}_p$.

Step 4) Find a homogeneous solution by expressing the basic variables in terms of the free variables.

Step 5) Find the general solution by

$$\mathbf{x}_p + \mathbf{x}_h.$$

Another Example

- Echelon form:

$$\left[\begin{array}{cccc:c} 1 & 3 & 3 & 2 & 1 \\ 2 & 6 & 9 & 7 & 5 \\ -1 & -3 & 3 & 4 & 5 \end{array}\right] \to \left[\begin{array}{cccc:c} 1 & 3 & 3 & 2 & 1 \\ 0 & 0 & 3 & 3 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right].$$

- By setting $x_2 = x_4 = 0$, we get

$$\mathbf{x}_p = [-2\ 0\ 1\ 0]^T.$$

- By setting $x_2 = 1, x_4 = 0$, we get

$$\mathbf{x}_{h_1} = [-3\ 1\ 0\ 0].$$

- By setting $x_2 = 0, x_4 = 1$, we get

$$\mathbf{x}_{h_2} = [1\ 0\ \ -1\ 1].$$

Example

Solve the system:

$$\left[\begin{array}{ccccc:c} 1 & 2 & -1 & 0 & 1 & 1 \\ 2 & 4 & 1 & 6 & 1 & 5 \\ 3 & 6 & -3 & 1 & 0 & 5 \end{array}\right].$$

Table of Contents

Linearly Independent Vectors

Definition

The vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ in V are said to be *linearly independent over F* if

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k = \mathbf{0}$$

implies $c_i = 0$ for all i. Otherwise, they are called *linearly dependent*.

- The most fundamental example is

$$[1\ 0\ 0]^T,\ [0\ 1\ 0]^T,\ [0\ 0\ 1]^T.$$

- If $A\mathbf{x} = \mathbf{0}$ has only one solution $\mathbf{x} = \mathbf{0}$, the columns of A are linearly independent.

Examples

- Two matrices:

$$\begin{bmatrix} 3 & -6 & 2 & -1 \\ -2 & 4 & 1 & 3 \\ 0 & 0 & 1 & 1 \\ 1 & -2 & 1 & 0 \end{bmatrix} \text{ and } \begin{bmatrix} -1 & 1 & 2 & 0 \\ 0 & 2 & -1 & 3 \\ 0 & 0 & 10 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

- Let V be the set of all polynomials of degree≤ 4. Consider the vectors

$$1, x, x^2, x^4.$$

Spanning

Definition

A set of vectors $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_l$ is said to *span a vector space V* if every vector in V can be expressed as a linear combination of the l vectors, that is,

$$\exists c_1, \ldots, c_l \in F \ \text{ such that } \mathbf{v} = \sum_{i=1}^{l} c_i \mathbf{w}_i \quad \forall \mathbf{v} \in V.$$

The space V is denoted by

$$V = \mathcal{L}(\mathbf{w}_1, \ldots, \mathbf{w}_l) \text{ or } V =< \mathbf{w}_1, \ldots, \mathbf{w}_l > .$$

Examples

1) Consider the three vectors

$$\mathbf{w}_1 = [1\ 0\ 0]^T,\ \mathbf{w}_2 = [0\ 1\ 0]^T,\ \mathbf{w}_3 = [-3\ 0\ 0]^T.$$

Note that

$< \mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3 > = < \mathbf{w}_1, \mathbf{w}_2 > = < \mathbf{w}_2, \mathbf{w}_3 > = (x, y)$-plane,

$< \mathbf{w}_1, \mathbf{w}_3 > = < \mathbf{w}_1 > = < \mathbf{w}_3 > = x$-axis.

2) Note that

$$< 1, x, x^2 > = < 1, x, x^2 + x > = P_2.$$

Basis

Definition

If S is a set of linearly independent vectors in V which spans V over F, S is called a basis of V over F. The number of vectors in S is called the *dimension* of V.

- Linear independency implies uniqueness of linear combinations.
- Spanning property means existence of linear combinations.
- In general, bases are not unique.
- The set $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ is called the standard basis. For example,

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

Properties

1) Given a basis for a vector space V, every vector in V can be uniquely represented as a linear combination of vectors in the basis.

2) In solving a system of linear equations,
   - Since $A\mathbf{x} = \mathbf{0} \;\Leftrightarrow\; U\mathbf{x} = \mathbf{0}$, we have $\mathcal{N}(A) = \mathcal{N}(U)$.
   - A and U have the same row space.
   - A and U have different column spaces.
   - The rows of U containing the nonzero pivots form a basis for the row space of A.
   - The columns of A corresponding to the columns of U that contain the nonzero pivots form a basis for the column space of A.

Ranks

Row and Column Ranks

The row (resp. column) rank of a matrix A is defined by the dimension of the row (resp. column) space of A.

1) (The row rank of A)= (The column rank of A)

2) The rank of A is defined to be equal to its row or column rank. It is also equivalent to the number of nonzero pivots in the echelon form and the number of basic variables in solving $A\mathbf{x} = \mathbf{0}$.

3) For an $m \times n$ matrix A,

$$\mathrm{rank}(A) + \dim \mathcal{N}(A) = n.$$

Table of Contents

Linear Transformation

Definition

A linear transformation T from a vector space V into a vector space W is a mapping which assigns a unique vector $\mathbf{y} \in W$ to each vector $\mathbf{x} \in V$, such that

(a) $T(\mathbf{x}_1 + \mathbf{x}_2) = T(\mathbf{x}_1) + T(\mathbf{x}_2)$ for any $\mathbf{x}_1, \mathbf{x}_2 \in V$;

(b) $T(c\mathbf{x}) = cT(\mathbf{x})$ for any $\mathbf{x} \in V$ and any $c \in F$.

- Note that matrix multiplication satisfies the rule of linearity.
- The kernel of T is the set of all $\mathbf{x} \in V$ such that $T\mathbf{x} = \mathbf{0}$. It is the null space if T is a matrix multiplication.

Examples in (X, Y)-Plane

1. Rotation with θ:

$$Q_\theta = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}.$$

2. Projection into the θ-Line:

$$P_\theta = \begin{bmatrix} \cos^2\theta & \cos\theta\sin\theta \\ \cos\theta\sin\theta & \sin^2\theta \end{bmatrix}.$$

3. Reflection with respect to the θ-Line:

$$P_\theta = \begin{bmatrix} 2\cos^2 - 1 & 2\cos\theta\sin\theta \\ 2\cos\theta\sin\theta & 2\sin^2\theta - 1 \end{bmatrix}.$$