MATH 1564: Linear Algebra with Abstract Vector Spaces

Frank Qiang

Georgia Institute of Technology
Fall 2022

Contents

Chapter 1

Elementary Set Theory

1.1 Introduction to Sets

Informally, a *set* is a collection of things. We can write a set by enumerating its elements as

$$A = \{1, 2, 3\},$$

or we can simply describe its elements by writing

$$B = \{x \mid x \text{ is a real number}\}.$$

The "|" symbol is read as "such that," and the colon ":" is sometimes used to mean the same thing. Some commonly used sets of numbers are:

- $\mathbb{R}$, the set of real numbers,
- $\mathbb{Q}$, the set of rational numbers,
- $\mathbb{C}$, the set of complex numbers,
- $\mathbb{Z}$, the set of integers,
- and $\mathbb{N}$, the set of natural numbers.

1.2 Set Membership and Subsets

We write $x \in A$ to mean that x is an element of A and $x \notin A$ otherwise. For example,

$$1 \in \{1, 2, 3\}, \quad 4 \notin \{1, 2, 3\}, \quad \text{and} \quad \pi \notin \mathbb{Z}.$$

We write $\varnothing$ to denote the empty set, and when the scope of a problem is understood, we write $\mathcal{U}$ to denote the universal set. In other words, $x \notin \varnothing$ and $x \in \mathcal{U}$ for any x.

For two sets A and B, we say that A is a *subset* of B if every element of A is also an element of B. We write $A \subseteq B$ or $A \subset B$ when this is the case. For example,

$$\{1\} \notin \{1, 2, 3\} \quad \text{but} \quad \{1\} \subseteq \{1, 2, 3\}.$$

Note that $\varnothing \subseteq A$ for any set A (also $\varnothing \subseteq \varnothing$ but $\varnothing \notin \varnothing$).

Equality for sets is defined in terms of subsets. We say that $A = B$ if $A \subseteq B$ and $B \subseteq A$. Note that sets are inherently unordered and contain unique elements, so that

$$\{1, 2, 3\} = \{3, 2, 1\} \quad \text{and} \quad \{1, 2, 2, 3\} = \{1, 2, 3\}.$$

We write $A \subsetneq B$ when A is a *proper* subset of B, i.e., $A \subseteq B$ but $A \neq B$.

1.3 The Cartesian Product

For two sets A and B, their *Cartesian product* is defined as

$$A \times B = \{(x, y) \mid x \in A, y \in B\}.$$

This is technically not a set operation, but rather shorthand for the description of a new set. We can also define the Cartesian product of n sets $A_1, \ldots, A_n$ as

$$A_1 \times \cdots \times A_n = \{(x_1, \ldots, x_n) \mid x_1 \in A_1, \ldots, x_n \in A_n\}.$$

Note that by this definition, $A \times B \times C \neq A \times (B \times C)$. We also write A^n to mean $A \times \cdots \times A$ (n times). As an example, the set $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$ is the usual 2-D Cartesian plane.

1.4 Set Operations

Let A and B be two sets. We can define the following operations on A and B:

- the *union* $A \cup B = \{x \mid x \in A \text{ or } x \in B\}$,
- the *intersection* $A \cap B = \{x \mid x \in A \text{ and } x \in B\}$,
- the *set difference* $A \setminus B = \{x \mid x \in A \text{ and } x \notin B\}$, we sometimes read this as "A take away B,"
- and when a universal set $\mathcal{U}$ is understood, the *complement* $A^c = \mathcal{U} \setminus A$.

For example, if $\mathcal{U} = \mathbb{R}^2$ and $A = \{(x, y) \mid x^2 + y^2 < 1\}$, then $A^c = \{(x, y) \mid x^2 + y^2 \geq 1\}$, i.e. the set of all points outside the unit circle and on its edge.

Proposition 1.1. *For any two sets A and B, we have $A = (A \cap B) \cup (A \setminus B)$.*

Proof. We need to show that LHS $\subseteq$ RHS and RHS $\subseteq$ LHS.

To see that LHS $\subseteq$ RHS, take any element $x \in \text{LHS} = A$. If $x \in A \setminus B$, then $x \in (A \cap B) \cup (A \setminus B)$ and we're done. Otherwise $x \notin A \setminus B$ and so $x \in B$ since we assumed that $x \in A$. Then $x \in A \cap B$ and thus $x \in (A \cap B) \cup (A \setminus B)$. So $x \in \text{RHS}$ and thus LHS $\subseteq$ RHS as desired.

For RHS $\subseteq$ LHS, take any element $x \in \text{RHS} = (A \cap B) \cup (A \setminus B)$. If $x \in A \setminus B$, then $x \in A$ by definition. Otherwise $x \in A \cap B$ and so $x \in A$ also. Thus $x \in \text{LHS}$ and RHS $\subseteq$ LHS as desired. □

Remark. This is related to the *law of total probability*, which states that for any two events A and B,

$$\Pr(A) = \Pr(A \cap B) + \Pr(A \cap B^c) = \Pr(A|B)\Pr(B) + \Pr(A|B^c)\Pr(B^c).$$

1.5 Set Laws

The *distributive laws* say that for any three sets A, B, and C,

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C) \quad \text{and} \quad A \cap (B \cup C) = (A \cap B) \cup (A \cap C).$$

The so-called *de Morgan's laws* say that for any two sets A and B,

$$(A \cup B)^c = A^c \cap B^c \quad \text{and} \quad (A \cap B)^c = A^c \cup B^c.$$

Notice that de Morgan's laws allow us to switch from unions to intersections and vice versa.

Chapter 2

Matrices and Linear Systems

2.1 Motivation

Consider the following system of equations:

$$\begin{cases} 3x + 2y - 3z = 0 \\ x + 5y = 2. \end{cases}$$

We're interested in determining whether such a system has a solution, and if so, how many solutions it has and what they are. In general, we're interested in systems of the form

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

Here $x = (x_1, x_2, \ldots, x_n)$ are the unknowns and $b = (b_1, b_2, \ldots, b_m)$ is the right-hand side of the system. We can also express the coefficients as a block of numbers in the following way:

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}.$$

2.2 Matrices

The block of numbers A from the previous section is precisely the general form of what we call a *matrix*. We say that a matrix has dimensions $m \times n$ if it has m rows and n columns. We also sometimes write A_{ij} to refer to the entry in the ith row and jth column of A.

2.2.1 Operations on Matrices

For a scalar $\lambda \in \mathbb{R}$, we can define the *scalar multiplication* of λ and A as

$$\lambda A = \begin{pmatrix} \lambda a_{11} & \lambda a_{12} & \cdots & \lambda a_{1n} \\ \lambda a_{21} & \lambda a_{22} & \cdots & \lambda a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda a_{m1} & \lambda a_{m2} & \cdots & \lambda a_{mn} \end{pmatrix}.$$

In other words, we simply multiply each entry of A by λ. For two matrices of the same size, we define *matrix addition* as simply addition component-wise. For example,

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} + \begin{pmatrix} 2 & 1 & 2 \\ 6 & 4 & 5 \end{pmatrix} = \begin{pmatrix} 3 & 3 & 5 \\ 10 & 9 & 11 \end{pmatrix}.$$

More generally, we have $(A+B)_{ij} = A_{ij} + B_{ij}$. For *matrix multiplication*, we have the formula

$$(AB)_{ij} = \sum_{k=1}^{n} A_{ik} B_{kj}.$$

In other words, we take a row of A and a column of B, multiply them component-wise, and repeat for each row and column. Notice that for this to work, the number of columns of A must equal the number of rows of B. In particular, the product of a $m \times n$ matrix and a $n \times k$ matrix is a $m \times k$ matrix. For example,

$$\begin{pmatrix} 1 & 0 & 2 & 0 & 3 \\ 2 & 1 & 3 & 1 & 4 \\ -1 & 2 & -1 & 3 & -1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 2 & 1 \\ 0 & 0 \\ 0 & 0 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 10 & 14 \\ 16 & 21 \\ 0 & -4 \end{pmatrix}.$$

Also note that matrix multiplication is *not* commutative, i.e. $AB \neq BA$ in general. In the specific example above, the product BA is not even defined!

2.3 Solving Linear Systems

Given an $m \times n$ matrix of coefficients A, a $n \times 1$ matrix of unknowns x, and a $m \times 1$ matrix b of right-hand side values, notice that we have a precise correspondence

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \quad\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m. \end{cases} \qquad \longleftrightarrow \qquad Ax = b.$$

The notation $Ax = b$ gives us a convenient way to express the linear system on the left, and it carries the exact same information with the way we defined matrix multiplication. We also like to work with

$$[A|b] = \left(\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right),$$

which is called the *augmented matrix* of the linear system.

Chapter 3

Vector Spaces

3.1 Motivation

The usual type of vector that we're familiar with is of the form $(x_1, \ldots, x_n) \in \mathbb{R}^n$. In this way, we can add two vectors with an addition operation[1] $+ : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ and multiply a vector by a scalar (from $\mathbb{R}$) with a scalar multiplication operation $\cdot : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$, both defined component-wise. We would like to generalize this structure to a more abstract setting.

3.2 Abstract Vector Spaces

Definition 3.1 (Vector space). A *vector space* $(V, +_v, \cdot_v,)_K$ over a field[a] K is a set V equipped with a binary addition operation $+_v : V \times V \to V$ and a scalar multiplication[b] operation $\cdot_v : K \times V \to V$ that satisfy the following 10 axioms:

(i) *Closure under addition*: For all $x, y \in V$, $x +_v y \in V$.

(ii) *Additive commutativity*: For all $x, y \in V$, $x +_v y = y +_v x$.

(iii) *Additivity associativity*: For all $x, y, z \in V$, $(x +_v y) +_v z = x +_v (y +_v z)$.

(iv) *Existence of additive identity*: There exists $0_v \in V$ such that $0_v +_v x = x$ for all $x \in V$.

(v) *Existence of additive inverse*: For all $x \in V$, there exists $(-x) \in V$ such that $x +_v (-x) = 0_v$.

(vi) *Closure under scalar multiplication*: For all $\alpha \in K$ and $x \in V$, $\alpha \cdot_v x \in V$.

(vii) *Distributive property with scalars*: For all $\alpha \in K$ and $x, y \in V$, $\alpha(x +_v y) = \alpha x +_v \alpha y$.

(viii) *Multiplicative associativity*: For all $\alpha, \beta \in K$ and $x \in V$, $\alpha(\beta x) = (\alpha\beta)x$.

(ix) *Distributive property with vectors*: For all $\alpha, \beta \in K$ and $x \in V$, $(\alpha + \beta)x = \alpha x +_v \beta x$.

(x) *Existence of multiplicative identity*: There exists $1 \in K$ such that $1 \cdot_v x = x$ for all $x \in V$.

[a] We call K the *base field* of V.

[b] We will sometimes just write αx instead of $\alpha \cdot_v x$ in these axioms for notational clarity.

Note that we will often omit the subscripts on $+_v$ and $\cdot_v$ when the vector space is clear from context. We often also just write V to refer to $(V, +_v, \cdot_v)_K$ when the base field and operations are clear from context.

[1] The notation $f : A \to B$ means that f is a function from A to B, i.e. it has domain A and codomain B.

We sometimes say that a vector space is *abstract* if it is not $\mathbb{R}^n$, which we sometimes refer to as *concrete*. It is not too important in our context to define what a field is, we will almost always take $K = \mathbb{R}$ or $\mathbb{C}$.

Just from these axioms, we can already prove some immediate (perhaps obvious) properties.

Proposition 3.1. *Let V be a vector space and $0_v \in V$ be its additive identity. Then $0 \cdot x = 0_v$ for any $x \in V$.*

Proof. By the distributive property, we have

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

Adding the additive inverse of $0x$ to both sides, we obtain $0_v = 0x$ as desired. □

Proposition 3.2. *The additive identity of a vector space is unique.*

Proof. Let V be a vector space and $0_v, \widetilde{0}_v \in V$ both be additive identities. Then

$$0_v = \widetilde{0}_v + 0_v = 0_v + \widetilde{0}_v = \widetilde{0}_v$$

by commutativity and the additive identity property. So the additive identity is unique. □

Chapter 4

Linear Transformations

Chapter 5

Inner Product Spaces

Chapter 6

Determinants

Chapter 7

Eigenvalues and Eigenvectors

Chapter 8

The Real Spectral Theorem