Introduction To Linear Algebra
Supplementary Note

MATH 240, Spring 2024

By

Shin Eui Song

The University of Maryland
College Park

Contents

CHAPTER 0

Introduction

This is a supplementary note for MATH240, taught at the University of Maryland during Spring 2024, based on the textbook written by Lay, Lay, and Mcdonald. The notes will be updated as the semester proceeds. The main purpose of this note is to summarize and supplement the textbook with an emphasis on the theoretical aspect. While some computational examples may be included, it is recommended to refer to the textbooks for detailed examples.

This note is not the definitive study guide for the course. To excel in this course, or college-level math in general, here are some tips:

(i) Understand the definitions (highlighted in red in this note) and have an example for each definitions. It greatly helps if you also know counter-examples.

(ii) Solve and understand all homework problems and worksheet problems provided during discussion sections.

(iii) Don't hesitate to ask for help when you get stuck. You can always work with your peers or come to office hours.

0.1. Notations

In this course, a number will means a real number or a complex number. We denote by $\mathbb{R}$ the set of real numbers and $\mathbb{C}$ the set of complex numbers. Instead of saying a *is a real number*, we can write $a \in \mathbb{R}$. Similarly, the phrase a *is a complex number* can be replaced by $a \in \mathbb{C}$.

A proposition will always mean a mathematical statement that is true. If a proposition is *important*, we will call them a theorem instead.

0.2. Short Introduction

Linear algebra is a branch of mathematics concerning solving linear equations and matrices. An efficient way of studying a collection of linear equations was to expressing it as a matrix, consequently leading to the development of matrix algebras.

Though started out as a pure mathematics, linear algebra has evolved to find myriad real-life applications. Search engine ranking (Section 10.2), optimization problems (Chapter 9), error correcting codes, machine learning (Section 6.6), and quantum computing, and more are some of the examples. The essence lies in the recongnition that data can often be represented as arrays of numbers, i.e. matrices.

CHAPTER 1

Linear Equations in Linear Algebra

In this chapter, we will learn about collections of degree-one equations with multiple variables called linear systems.

$$\{ \text{ linear systems } \} \leftrightarrow \{ \text{ matrices } \}$$

For a linear system, we can associate a matrix called an augmented matrix which contains all important data of a system. (Section 1.1) We will perform certain matrix operations to turn augmented matrix into a simpler form called the (reduced) row echelon form. It turns out that the row reduction algorithm, or changing a matrix into its reduced row echelon form, gives an explicit description of the solution set of a linear system. (Section 1.2) Then we introduce vectors in real Euclean space $\mathbb{R}^n$ and express linear systems as a *matrix equation* of the form $A\mathbf{x} = \mathbf{b}$. (Section 1.3 and 1.4) Next, we will use the vector notations to give explicit and geometric description of solution sets. (Section 1.5)

1.1. Systems of Linear Equations

Definition 1.1.1. A linear equation is an equation (of variables $x_1, \ldots, x_n$) of the form

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b$$

where $a_1, \ldots, a_n, b$ are (real or complex) numbers. The numbers $a_1, \ldots, a_n$ in front of the variables, are called coefficients.

Definition 1.1.2.

(i) A collection of one or more linear equations

$$\mathscr{L} : \left\{ \begin{array}{rcl} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n & = & b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n & = & b_2 \\ \vdots & \vdots & \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n & = & b_m \end{array} \right. \quad (1)$$

is called a linear system (or a system of linear equations). In this note, we will sometimes denote systems of linear equations by caligraphic letters $\mathscr{L}$.

(ii) A solution of a linear system $\mathscr{L}$ is a list $(s_1, \ldots, s_n)$ of (real or complex) numbers that makes all the equations true when substituted for $x_1, \ldots, x_n$. The set of all solutions of a linear system is called a solution set. Two linear systems are equivalent if they have the same solution set.

Proposition 1.1.3. *A system of linear equations has either*
a) no solution *b) exactly one solution* *c) infinitely many solutions.*

Definition 1.1.4. A system of linear equations is said to be consistent if it has a solution and is said to be inconsistent if it has no solution.

There is a more compact way of writing all the essential informations of linear systems.

Definition 1.1.5. A rectangular array A of numbers

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

is called a matrix. If the number of rows of A is m and the number of columns of A of n, then $m \times n$ is called the size of A and we call A an $m \times n$ matrix.

Definition 1.1.6. Given a system of equations $\mathscr{L}$ as in equation (1), we can associate two matrices.

$$A_{\mathscr{L}} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \qquad A_{\mathscr{L}}^{+} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_n \end{bmatrix}$$

Here $A_{\mathscr{L}}$ is called the coefficient matrix of the system $\mathscr{L}$, and $A_{\mathscr{L}}^{+}$ is called the augmented matrix of the system $\mathscr{L}$. We will often simply write A instead of $A_{\mathscr{L}}$ or $A_{\mathscr{L}}^{+}$.

How do you solve a linear system (in other words, how do you find the solution set)? The rough idea is to replace one system with an equivalent system that is easier to solve. The following are three basic operations to simplify a linear system and their matrix counterparts. One can show that elementary row operations do not change solutions sets (see Problem 1.1 or Proposition 1.1.9).

Definition 1.1.7. The following operations on linear systems or matrices are called elementary row operations.

(1) replace one equation by the sum of itself and a multiple of another equation,

(2) interchange two equations,

(3) multiply all the terms in an equation by a nonzero constants

(1) replace one row by the sum of itself and a multiple of another row,

(2) interchange two rows,

(3) multiply all the entries in a row by a nonzero constants

For a step-by-step illustration of solving a linear system using elementary row operations, see **Example 1** of section 1.1. Similar methods in **Example 3** of section 1.1 can be used to show that a linear system is inconsistent.

Definition 1.1.8. Two matrices are called row equivalent if there is a sequence of elementary row operations that transforms one matrix into the other.

The following proposition explains why the augmented matrix of a linear system contains all the essential information of a linear system. In other words, we can work with matrices instead of linear systems to solve the systems. To see why the proposition is true, see Problem 1.1.

Proposition 1.1.9. *Two linear systems have the same solution set if and only if the augmented matrices of these two linear systems are row equivalent.*

1.2. Row Reduction and Echelon Forms

In **Example 1** of section 1.1, we have the following row equivalent matrices

$$A_1 = \begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ 5 & 0 & -5 & 10 \end{bmatrix} \mapsto A_2 = \begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 1 & -4 & 4 \\ 0 & 0 & 30 & -30 \end{bmatrix} \mapsto A_3 = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{bmatrix}. \tag{2}$$

The zero entries on the lower-left blocks of A_2 and A_3 look like a staircase. Using the French word échelon which means *steplike*, we make the following definition.

Definition 1.2.1.

(a) The leftmost nonzero entry is called the leading entry.

(b) A matrix is in row echelon form if

  (i) All rows with at least one nonzero entry is above any rows of all zeroes.

  (ii) All entries in a column below a leading entry are zeroes.

  (iii) All leading entry is on the right of the leading entry of every row above.

(c) A matrix is in row reduced echelon form if

(i) All leading entry is 1.

(ii) All columns containing a leading entry has zeores in all its other entries.

Though there are also column echelon form and column reduced echelon form, we simply say echelon form and reduced echelon form in this course. The following definitions are made mainly for bookkeeping purposes.

Definition 1.2.2.

(a) A matrix that is in echelon form is called a echelon matrix. Similarly, a matrix that is in reduced echelon form is called a reduced echelon matrix.

(b) A echelon matrix U that is row equivalent to a matrix A is called an echelon form of A. Similarly, one can define a reduced echelon form of A.

(c) A matrix is said to be row reduced into another matrix if it is transformed by elementary row operations. You may regard it as a *verb form* of row equivalent.

Example 1.2.3. Consider the three matrices A_1, A_2, and A_3 in (2) above. The matrix A_2 is a echelon matrix, and A_3 is a reduced echelon matrix. Since both A_2 and A_3 are row equivalent to A_1, A_2 is a echelon form of A_1, and A_3 is a reduced echelon form of A_2.

Note that given a nonzero matrix A, one can have (infinitely) many echelon forms of A. To see this, given a echelon form of A, simply multiply all the entries by a nonzero constants (ER3) to obtain another echleon form of A. However,

Theorem 1.2.4 (Uniqueness of the Reduced Echelon Form)**.** *Each matrix is row equivalent to one and only one reduced echelon matrix.*

The proof is given in Appendix A at the end of the textbook.

Definition 1.2.5. A pivot position in A is the location in A that corresponds to the leading entry that is 1 in the reduced echelon form of A. A column of A that contains a pivot position is called a pivot column. A pivot is a nonzero number in a pivot position that may be used to create zeros via row operations.

Finding pivots are explained in **Example 2** of Section 1.2. Note that pivots are not the entries of A in the pivot positions. Next **Example 3** of Section 1.2 shows the row reduction algorithm to transform any matrix A into its (reduced) echelon form.

Definition 1.2.6. Let $\mathscr{L}$ be a linear system of m equations with n variables and A be its associated augmented matrix. Then let U be the reduced echelon form of A. If the ith column of the augmented matrix is a pivot column, then we say that the variable x_i is a basic variable. On the other hands, if the ith column is not a pivot column, we say that x_i is a free variable. This means that x_is can have any values, and the basic variables are determined by the free variables.

1.3. Vector Equations

A vector is an ordered list of numbers. In this section, we will focus only on column vectors, i.e. a matrix with one column. For example,

$$\mathbf{v}_1 = \begin{bmatrix} 3 \\ 7 \\ -2 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} \pi \\ \sqrt{2} \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 1 \\ 2 \\ \vdots \\ 20 \end{bmatrix}$$

are all columns vectors. The set of vectors with n entries are denoted by $\mathbb{R}^n$. Therefore, $\mathbf{v}_1 \in \mathbb{R}^3$, $\mathbf{v}_2 \in \mathbb{R}^2$, and $\mathbf{v}_3 \in \mathbb{R}^{20}$. Note that two vectors are equal to each other if their entries are the same. In particular, $\begin{bmatrix} 1 \\ 2 \end{bmatrix} \neq \begin{bmatrix} 2 \\ 1 \end{bmatrix}$.

Given two vectors

$$\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}, \quad \mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

and a number (or a scalar) $c \in \mathbb{R}$, we can define the following operations.

a) (sum)

$$\mathbf{u} + \mathbf{v} = \begin{bmatrix} u_1 + v_1 \\ u_2 + v_2 \\ \vdots \\ u_n + v_n \end{bmatrix}$$

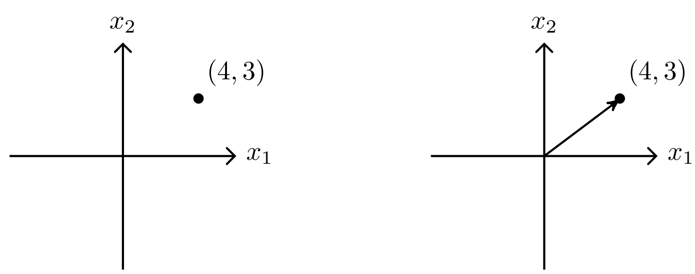
b) (scalar multiplication)

$$c\mathbf{u} = \begin{bmatrix} cu_1 \\ cu_2 \\ \vdots \\ cu_n \end{bmatrix}$$

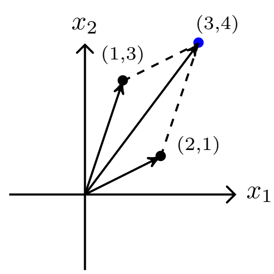
The column vectors in $\mathbb{R}^n$ have a geometric interpretation. We view the vector

$$\begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}$$

as the point $(u_1, \ldots, u_n)$ in the n-dimensional space also denoted $\mathbb{R}^n$. So we can view the column vectors in $\mathbb{R}^2$ as a point in the x_1x_2-plane (similar to xy-plane) as below.



Then vector sums $\mathbf{u} + \mathbf{v}$ is geometrically the fourth vertex of the parallelogram formed by $\mathbf{u}$ and $\mathbf{v}$. This is called the parallelogram rule for addition.



The vector sum and scalar multiplication satisfies a similar algebraic properties as real or complex numbers do.

Proposition 1.3.1 (Algebraic Properties of $\mathbb{R}^n$). *For all $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^n$, and all scalars $c, d \in \mathbb{R}$,*

a) $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$

b) $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$

c) $\mathbf{u} + \mathbf{0} = \mathbf{0} + \mathbf{u} = \mathbf{u}$

d) $\mathbf{u} + (-\mathbf{u}) = -\mathbf{u} + \mathbf{u} = \mathbf{0}$ *where* $-\mathbf{u} = (-1)\mathbf{u}$

e) $c(\mathbf{u} + \mathbf{v}) = c\mathbf{u} + c\mathbf{v}$

f) $(c + d)\mathbf{u} = c\mathbf{u} + c\mathbf{v}$

g) $c(d\mathbf{u}) = (cd)\mathbf{u}$

h) $1\mathbf{u} = \mathbf{u}$.

Definition 1.3.2. Let $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p \in \mathbb{R}^n$ and $c_1, \ldots, c_p \in \mathbb{R}$. Then the vector

$$\mathbf{y} = c_1\mathbf{v}_1 + \cdots + c_p\mathbf{v}_p$$

is called a linear combination of $\mathbf{v}_1, \ldots, \mathbf{v}_p$ with weights $c_1, \ldots, c_p$. The set of all linear combinations of $\mathbf{v}_1, \ldots, \mathbf{v}_p$ is denoted by $\mathrm{Span}\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ and is called the subset of $\mathbb{R}^n$ spanned (or generated) by $\mathbf{v}_1, \ldots, \mathbf{v}_p$.

With this new language, we now have three ways to describe linear equations.

$$\{\text{ Augmented Matrix }\} \leftrightarrow \{\text{ Linear Systems }\} \leftrightarrow \{\text{ Vector Equations }\}$$

To see this, let $x_1, \ldots, x_n$ be variables. We have

$$\{\text{ Linear Systems }\} \quad \leftrightarrow \quad \{\text{ Vector Equations }\}$$

$$\begin{array}{l} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2 \\ \vdots \qquad\qquad\qquad\quad \vdots \;=\; \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m \end{array} \quad \leftrightarrow \quad \mathbf{a}_1x_1 + \mathbf{a}_2x_2 + \cdots + \mathbf{a}_nx_n = \mathbf{b}$$

$$\mathbf{a}_1 = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix}, \ldots, \mathbf{a}_n = \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix}, \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

Therefore, the linear system has a solution if and only if $\mathbf{b}$ is a linear combination of $\mathbf{a}_1, \ldots, \mathbf{a}_n$.

1.4. The Matrix Equation $A\mathbf{x} = \mathbf{b}$

$$\{\text{ Vector Equations }\} \quad \leftrightarrow \quad \{\text{ Matrix Equation }\}$$

$$\mathbf{a}_1x_1 + \mathbf{a}_2x_2 + \cdots + \mathbf{a}_nx_n = \mathbf{b} \quad \leftrightarrow \quad A\mathbf{x} = \mathbf{b}$$

$$\mathbf{a}_1 = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix}, \ldots, \mathbf{a}_n = \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix}, \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} \leftrightarrow A = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{bmatrix}, \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

Theorem 1.4.1 (Theorem 4 on page 39). *Let A be an $m \times n$ matrix. Then the following statements are logically equivalent. That is, for a particular A, either they are all true statements or they are all false.*

a. *For each $\mathbf{b} \in \mathbb{R}^m$, the equation $A\mathbf{x} = \mathbf{b}$ has a solution.*

b. *Each $\mathbf{b}$ in $\mathbb{R}^m$ is a linear combination of the columns of A.*

c. *The columns of A span $\mathbb{R}^m$.*

d. *A has a pivot position in every row.*

Proposition 1.4.2 (Theorem 5 on page 41). *If A is $m \times n$ matrix, $\mathbf{u}$ and $\mathbf{v}$ are vectors in $\mathbb{R}^n$, and c is a scalar, then:*

a. $A(\mathbf{u} + \mathbf{v}) = A\mathbf{u} + A\mathbf{v}$;

b. $A(c\mathbf{u}) = c(A\mathbf{u})$.

1.5. Solution Sets of Linear Systems

Definition 1.5.1. A homogeneous linear system is a linear system which can be written in the form $A\mathbf{x} = 0$. Then the zero vector 0 is always a solution to the linear system $A\mathbf{x} = 0$, and will be called the trivial solution. For a homogeneous linear system, a solution that is nonzero (i.e. not trivial) is called a nontrivial solution.

Proposition 1.5.2. *The homogeneous equation $A\mathbf{x} = 0$ has a nontrivial solution if and only if the equation has at least one free variable.*

Definition 1.5.3. A solution is in parametric vector form if it is of the form

$$\mathbf{x} = s_1\mathbf{v}_1 + \cdots + s_n\mathbf{v}_n.$$

General solutions of a homogeneous linear system $A\mathbf{x} = 0$ can be written in parametric vector form. For arbitrary linear system $A\mathbf{x} = 0$, general solutions of the form

$$\mathbf{x} = \mathbf{p} + s_1\mathbf{v}_1 + \cdots + s_n\mathbf{v}_n$$

will be still called in parametric vector form.

To write the general solutions of $A\mathbf{x} = 0$ in parametric vector form, write the general solutions as in 1.2. For example

$$\begin{cases} x_1 = & 3 + 7x_3 + 6x_4 \\ x_2 = & 1 - 2x_3 \\ x_3 & \text{is free} \\ x_4 & \text{is free} \\ x_5 = & x_4 \end{cases} \leftrightarrow \mathbf{x} = \begin{bmatrix} 3 + 7x_3 + 6x_4 \\ 1 - 2x_3 \\ x_3 \\ x_4 \\ x_4 \end{bmatrix}$$

Then we can *separate* each free variables from the general solution as follows

$$\mathbf{x} = \underbrace{\begin{bmatrix} 3 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}}_{=\mathbf{p}} + x_3 \underbrace{\begin{bmatrix} 7 \\ -2 \\ 1 \\ 0 \\ 0 \end{bmatrix}}_{=\mathbf{v}_1} + x_4 \underbrace{\begin{bmatrix} 6 \\ 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}}_{=\mathbf{v}_2}$$

Theorem 1.5.4. *Suppose the equation $A\mathbf{x} = \mathbf{b}$ is consistent for some given $\mathbf{b}$, and let $\mathbf{p}$ be a solution. Then the solution set of $A\mathbf{x} = \mathbf{b}$ is the set of all vectors of the form $\mathbf{w} = \mathbf{p} + \mathbf{v}_h$, where $\mathbf{v}_h$ is any solution of the homogeneous equation $A\mathbf{x} = 0$.*

1.7. Linear Independence

Definition 1.7.1. The set of vectors $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ be a set in $\mathbb{R}^n$ is linearly independent if the vector equation

$$x_1\mathbf{v}_1 + \cdots + x_p\mathbf{v}_p = 0$$

has only the trivial solution. The set $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ is linearly dependent if there exists weights $c_1, \ldots, c_p$, not all zero, such that

$$c_1\mathbf{v}_1 + \cdots + c_p\mathbf{v}_p = 0. \tag{3}$$

The equation above is (3) is called linear dependence relation.

Proposition 1.7.2. *Each linear dependence relation among the columns of A corresponds to a nontrivial solution of $A\mathbf{x} = 0$. Therefore, the columns of a matrix A are linearly indpendent if and only if the equation $A\mathbf{x} = 0$ has only the trivial solution.*

Proposition 1.7.3.

(a) *The set $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ is linearly dependent if one of them is a zero vector.*

(b) *The set $\{\mathbf{v}\}$ is linearly independent if $\mathbf{v} \neq 0$ and is linearly dependent if $\mathbf{v} = 0$.*

(c) *The set $\{\mathbf{v}_1, \mathbf{v}_2\}$ is linearly dependent if and only if $\mathbf{v}_1$ and $\mathbf{v}_2$ are scalar multiple of each other.*

Theorem 1.7.4 (Theorem 7 and 8 on page 62 and 63)**.**

(a) *Let $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ be a set of vectors. Then S is linearly dependent if and only if at least one of the vectors in S is a linear combination of the others. To be precise, $\mathbf{v}_1 \neq 0$, and there exists $\mathbf{v}_j$ such that $\mathbf{v}_j$ is a linear combination of $\mathbf{v}_1, \ldots, \mathbf{v}_{j-1}$.*

(b) *Let $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ be a vectors in $\mathbb{R}^n$. If $p > n$, then S is linearly dependent.*

1.8. Introduction to Linear Transformations

Definition 1.8.1.

(a) Let A and B be a set. Then a function $f : A \to B$ is a rule that assigns to element $a \in A$ to a unique element $f(a) \in B$. To be more precise, every element $a \in A$ must be sent to an element in B and to only one element. We use the notation $a \mapsto f(a)$. The set A is called the domain, and B is called the codomain. The element $f(a)$ is called the image of a under f. The set of all images $f(a)$ is called the range of f.

(b) A function of the form $T : \mathbb{R}^n \to \mathbb{R}^m$ is often called transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$.

A transformation associated to matrix multiplication A is called a matrix transformation and is defined by $\mathbf{x} \mapsto A\mathbf{x}$. If $\mathbf{a}_1, \dots, \mathbf{a}_n$ are column vectors of A, then the image $A\mathbf{x}$ of $\mathbf{x}$

$$A\mathbf{x} = x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n$$

is a linear combination of $\mathbf{a}_1, \dots, \mathbf{a}_n$. In other words, the range of T is $\operatorname{Span}(\mathbf{a}_1, \dots, \mathbf{a}_n)$.

Definition 1.8.2. A transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ is linear if for all $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ and $c \in \mathbb{R}$

(a) $T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$,

(b) $T(c\mathbf{u}) = cT(\mathbf{u})$.

By Proposition 1.4.2, matrix transformations are linear transformations. Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation, then

(a) $T(0) = 0$

(b)
$$T(c_1\mathbf{v}_1 + \cdots + c_p\mathbf{v}_p) = c_1T(\mathbf{v}_1) + \cdots + c_pT(\mathbf{v}_p)$$
for all $\mathbf{v}_1, \dots, \mathbf{v}_p \in \mathbb{R}^n$ and $c_1, \dots, c_p \in \mathbb{R}$.

Example 1.8.3. Let $T_r : \mathbb{R}^n \to \mathbb{R}^n$ be a linear transformation defined by $\mathbf{x} \mapsto r\mathbf{x}$ for $r \in \mathbb{R}$. Then T_r is called a dilation since it *stretches* or *contracts* the length of the vector and perserves its direction. If $r = 1$, then $T_1(\mathbf{x}) = \mathbf{x}$ is called the identity transformation. A dilation is, in fact, a matrix multiplication. Let

$$A_r = \begin{bmatrix} r & 0 & \cdots & 0 \\ 0 & r & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & r \end{bmatrix}$$

Then $T_r(\mathbf{x}) = A_r\mathbf{x}$. If $r = 1$, there is a special name for $A_1 = I_n$ and is called the $n \times n$ identity matrix.

1.9. The Matrix of a Linear Transformation

In fact, we can show that every linear transformation is actually a matrix transformation.

Theorem 1.9.1. *Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. Then there exists a unique matrix A such that*

$$T(\mathbf{x}) = A\mathbf{x}$$

for all $\mathbf{x}$ in $\mathbb{R}^n$. In fact, A is the $m \times n$ matrix whose jth column is the vector $T(\mathbf{e}_j)$, where $\mathbf{e}_j$ is the jth column of the identity matrix in $\mathbb{R}^n$:

$$A = \begin{bmatrix} T(\mathbf{e}_1) & \cdots & T(\mathbf{e}_n) \end{bmatrix}$$

The matrix A is called the standard matrix for the linear transformation T.

Example 1.9.2. The transformation $T_\theta : \mathbb{R}^2 \to \mathbb{R}^2$ where $T(\mathbf{x})$ is the rotation of $\mathbf{x}$ by θ radian counterclockwise is a linear transformation. Then one can show that the standard matrix is

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

Definition 1.9.3.

(a) A mapping $T : \mathbb{R}^n \to \mathbb{R}^m$ is said to be onto $\mathbb{R}^m$ if each $\mathbf{b}$ in $\mathbb{R}^m$ is the image of at least one $\mathbf{x}$ in $\mathbb{R}^n$. In other words, for every $\mathbf{b}$ in $\mathbb{R}^m$, there exists $\mathbf{x}$ in $\mathbb{R}^n$ such that $T(\mathbf{x}) = \mathbf{b}$.

(b) A mapping $T : \mathbb{R}^n \to \mathbb{R}^m$ is said to be one-to-one if each $\mathbf{b}$ in $\mathbb{R}^m$ is the image of at most one $\mathbf{x}$ in $\mathbb{R}^n$. In other words, if $T(\mathbf{x}) = T(\mathbf{y})$, then $\mathbf{x} = \mathbf{y}$.

Theorem 1.9.4. *Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation and A be its standard matrix.*

(a) *T is one-to-one if and only if $T(\mathbf{x}) = 0$ implies $\mathbf{x} = 0$.*

(b) *T is onto if and only if the columns of A span $\mathbb{R}^m$.*

(c) *T is one-to-one if and only if the columns of A are linearly independent.*

Appendix: Extra Theoretical Problems

PROBLEM 1.1. Let $\mathscr{L} = \{L_1, \ldots, L_m\}$ be a linear system where each $L_i : a_{i1}x_1 + \cdots + a_{in}x_n = b_i$ is a linear equation for $i = 1, \ldots, m$. Let $S_i \subset \mathbb{R}^n$ be the solution set of L_i. Then, by definition, the solution set $S_{\mathscr{L}}$ of $\mathscr{L}$ is the intersection

$$S_{\mathscr{L}} = S_1 \cap \cdots \cap S_m$$

of all S_i for $i = 1, \ldots, m$. Prove that the new linear system $\mathscr{L}'$ after a sequence of elementary row operations (ER1) - (ER3) in Definition 1.1.7 have the same solution set as $\mathscr{L}$.

CHAPTER 2

Matrix Algebra

CHAPTER 3

Determinant

CHAPTER 4

Vector Spaces

CHAPTER 5

Eigenvalues and Eigenvectors

CHAPTER 6

Orthogonality and Least Squares

CHAPTER 7

Symmetric Matrices and Quadratic Forms