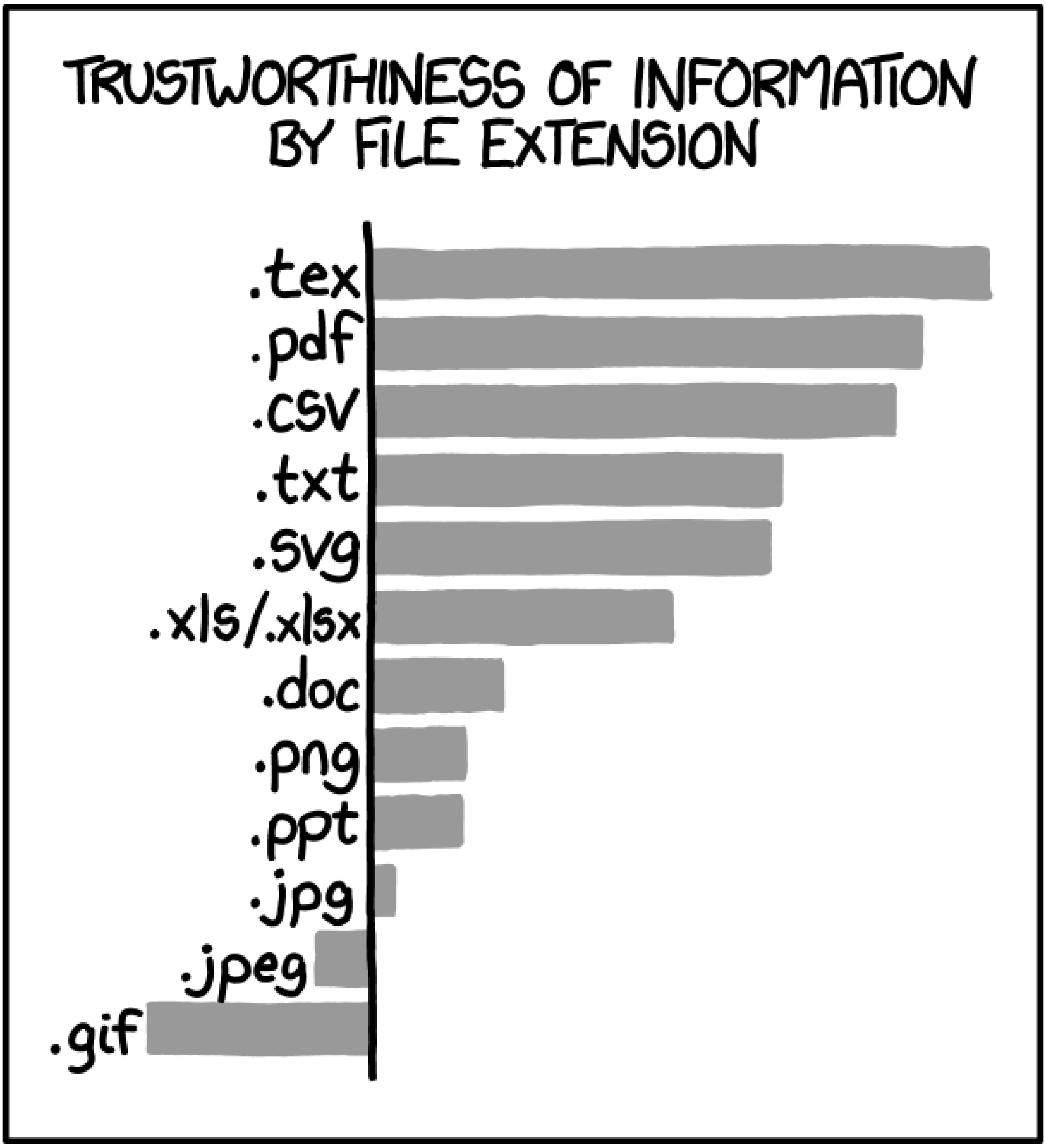
MATH2135: LINEAR ALGEBRA

ADITHYA BHASKARA
PROFESSOR: ROBIN DEELEY

TEXTBOOK: STEPHEN ANDRILLI & DAVID HECKER

UNIVERSITY OF COLORADO BOULDER


TRUSTWORTHINESS OF INFORMATION BY FILE EXTENSION
.tex
.pdf
.csv
.txt
.svg
.xls/.xlsx
.doc
.png
.ppt
.jpg
.jpeg
.gif

Contents

Preface

To the interested reader,

This document is a compilation of lecture notes taken during the Fall 2022 semester for MATH2135: Linear Algebra for Mathematics Majors at the University of Colorado Boulder. The course used *Elementary Linear Algebra*[1] by Stephen Andrilli and David Hecker as its primary text. As such, many theorems, definitions, and content may be quoted from the book. This course was taught by Robin Deeley, Ph. D.

Appendix A is provided as a nonexhaustive refresher on proof techniques commonly covered in a Discrete Mathematics course. Appendix A is a result of lecture notes compiled during the Summer 2022 semester for MATH2520: Discrete Mathematics at Front Range Community College. This course was taught by Kenneth M. Monks, Ph. D. The primary text was Oscar Levin's *Discrete Mathematics: An Open Introduction*[2], but Richard Hammack's *Book of Proof*[3] was also used as supplement.

Appendix B is provided as a nonexhaustive refresher on functions, as preparation for Chapter 5. Appendix B is a result of both content covered in *Elementary Linear Algebra* and MATH2520 lecture notes.

The author would like to provide the following resources for students currently taking a Linear Algebra course:

1. Sergei Treil's *Linear Algebra Done Wrong*.
2. Sheldon Axler's *Linear Algebra Done Right*.
3. Gilbert Strang's Linear Algebra Lectures From Fall 2011.
4. 3Blue1Brown's *Essence of Linear Algebra*.

While much effort has been put in to remove typos and mathematical errors, it is very likely that some errors, both small and large, are present. Please keep in mind that the author wrote this resource during his first semester of his undergraduate studies. If an error needs to be resolved, please contact Adithya Bhaskara at adithya.bhaskara@colorado.edu.

Finally, the author would like to dedicate this set of lecture notes to *Aidan Janney*, *Erika Sjöblom*, and *Tate McDonald*, three of the author's closest friends who plan to take Linear Algebra in the upcoming semester, at the time of writing.

Best Regards,
Adithya Bhaskara

`REVISED: December 14, 2022`

[1]Andrilli, S., & Hecker, D. (2016). *Elementary Linear Algebra* (5th ed.). Academic Press.
[2]Levin, O. (2019). *Discrete Mathematics: An Open Introduction* (3rd ed.). Oscar Levin.
[3]Hammack, R. (2018). *Book of Proof* (3rd ed.). Richard Hammack.

I go by the name of Vector. It's a mathematical term, represented by an arrow with both direction and magnitude. Vector! That's me, because I commit crimes with both direction and magnitude. Oh yeah!

Victor (Vector) Perkins

1
Vectors and Matrices

1.1 Lecture 1: August 22, 2022

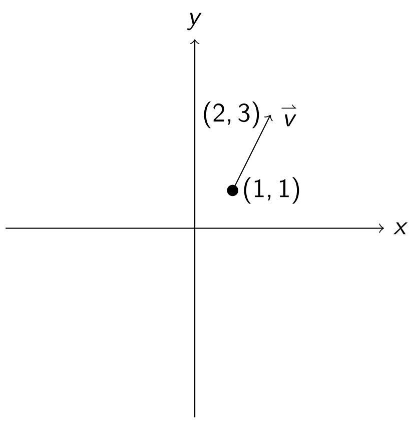
1.1.1 Notations, Definitions, and Conventions

Consider the following table for a very basic review of fundamental sets.

Symbol	Explanation
$\mathbb{N}$	Natural Numbers ($\mathbb{N} = \{1, 2, ...\}$)
$\mathbb{Z}$	Integers
$\mathbb{Q}$	Rationals
$\mathbb{R}$	Reals
$\mathbb{R}^n$	$\{[v_1, \dots, v_n] : v_1, \dots, v_n \in \mathbb{R}\}$

.

We will also note that there is an important distinction between vectors and points. Vectors describe "movement," whereas points describe location. For example, the vector $[1, 2]$ starting at the point $(1, 1)$ ends at the point $(2, 3)$. Consider the following diagram.



Consider the following definitions and statements.

Definition 1.1.1: The Zero Vector

We define

$$\vec{0} = [0, 0, \dots, 0].$$

Definition 1.1.2: Vector Equality

Two vectors $\vec{v} = [v_1, v_2, \dots, v_n]$ and $\vec{w} = [w_1, w_2, \dots, w_n]$ are equal if and only if

$$v_1 = w_1, v_2 = w_2, \dots, v_n = w_n.$$

Definition 1.1.3: Vector Magnitude

Given a $\vec{v} = [v_1, \dots, v_n] \in \mathbb{R}^n$, we define $||\vec{v}||$, the magnitude, or norm, of $\vec{v}$, as

$$||\vec{v}|| = \sqrt{v_1^2 + \cdots + v_n^2}.$$

Definition 1.1.4: Scalar Multiplication

Given a $\vec{v} = [v_1, \dots v_n] \in \mathbb{R}^n$, and a scalar $c \in \mathbb{R}$, we define scalar multiplication as

$$c\vec{v} = [cv_1, \dots, cv_n].$$

Consider the following theorem.

Theorem 1.1.1: Scalar Multiplication and Magnitude

Given a scalar $c \in \mathbb{R}$ and $\vec{v} \in \mathbb{R}^n$,

$$||c\vec{v}|| = |c|||\vec{v}||.$$

Proof.

$$\begin{aligned} ||c\vec{v}|| &= \sqrt{(cv_1)^2 + \cdots + (cv_n)^2} \\ &= \sqrt{c^2 v_1^2 + \cdots + c^2 v_n^2} \\ &= \sqrt{c^2}\sqrt{v_1^2 + \cdots + v_n^2} \\ &= |c|||\vec{v}||. \end{aligned}$$

The theorem is hence proved. □

Consider the following definitions.

Definition 1.1.5: Vector Direction

Two nonzero vectors $\vec{v}, \vec{w} \in \mathbb{R}^n$ are

1. in the same direction if there exists $c > 0$ such that $\vec{v} = c\vec{w}$.
2. in opposite directions if there exists $c < 0$ such that $\vec{v} = c\vec{w}$.

Definition 1.1.6: Unit Vectors

The vector $\vec{v} \in \mathbb{R}^n$ is a unit vector if and only if $||\vec{v}|| = 1$.

Consider the following theorem.

Theorem 1.1.2: Unit Vectors Represent Direction

Given a nonzero $\vec{v} \in \mathbb{R}^n$, there exists a unique unit vector in the same direction as $\vec{v}$.

Proof. Since $\vec{v} \neq \vec{0}$, $||\vec{v}|| \neq 0$. Let $\vec{u} = \frac{1}{||\vec{v}||}\vec{v}$, meaning that

$$\begin{aligned} ||\vec{u}|| &= \left|\left|\frac{1}{||\vec{v}||}\vec{v}\right|\right| \\ &= \left|\frac{1}{||\vec{v}||}\right| ||\vec{v}|| \\ &= 1. \end{aligned}$$

This means that $\vec{u}$ is a unit vector and, because $\frac{1}{||\vec{v}||}$ is a scalar, is in the same direction as $\vec{v}$. We have now shown existence. For uniqueness, assume $\vec{w}$ is a unit vector in the same direction as $\vec{v}$. Then, for a positive $c \in \mathbb{R}$, $\vec{w} = c\vec{v}$ and $||\vec{w}|| = 1$. That is,

$$||\vec{w}|| = |c|||\vec{v}|| = 1,$$

meaning that $|c| = \frac{1}{||\vec{v}||}$. Because $c > 0$, $|c| = c$, so $\vec{w} = \vec{u}$. □

In Theorem 1.1.2, we used the fact that for some $\vec{v} \in \mathbb{R}^n$, $\vec{v} \neq \vec{0} \implies ||\vec{v}|| \neq 0$. We will now provide a proof.

Theorem 1.1.3: Nonzero Vector Implies Nonzero Magnitudes

For some $\vec{v} = [v_1, \dots, v_n] \in \mathbb{R}^n$,

$$\vec{v} \neq \vec{0} \implies ||\vec{v}|| \neq 0.$$

Proof. We will proceed by proving the contrapositive. Suppose that $||\vec{v}|| = 0$, meaning that

$$||\vec{v}|| = \sqrt{v_1^2 + \cdots + v_n^2} = 0.$$

The only way this is true is if all of $v_1, \dots, v_n$ are zero, which, by definition, means that $\vec{v} = \vec{0}$. $\square$

1.2 Lecture 2: August 24, 2022

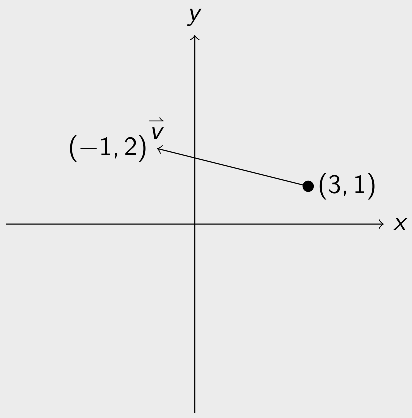
1.2.1 Vector Operations

As a review from the previous lecture, consider the following exercises.

Example 1.2.1: ☢☢ Vector Representing Movement

Find the vector representing the movement from $(3, 1)$ to $(-1, 2)$.

We simply find the change in the y coordinates and the change in the x coordinates to find the vector $[-4, 1]$, *starting from* $(3, 1)$. This is illustrated below.



Example 1.2.2: ☢☢ Finding a Unit Vector

Find the unit vector in the direction $[3, -1, -\pi]$.

We construct the unit vector by normalization. This produces $\frac{[3,-1,-\pi]}{\sqrt{3^2+(-1)^2+(-\pi)^2}}$.

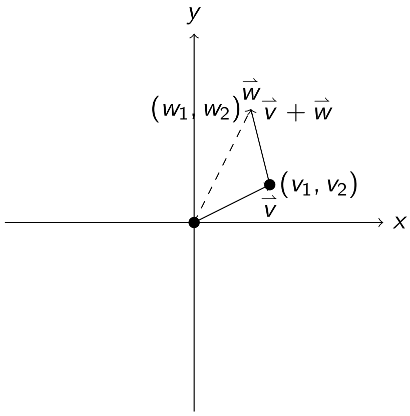
We will now define addition and subtraction with vectors and provide a few properties.

Definition 1.2.1: Addition and Subtraction With Vectors

Let $\vec{v} = [v_1, \dots, v_n]$ and $\vec{w} = [w_1, \dots, w_n]$ be vectors in $\mathbb{R}^n$. Then,

$$\vec{v} \pm \vec{w} = [v_1 \pm w_1, \dots, v_n \pm w_n].$$

Geometrically, we may visualize vector addition as



.

Theorem 1.2.1: Properties of Vector Addition and Scalar Multiplication

Let $\vec{u} = [u_1, \dots, u_n]$, $\vec{v} = [v_1, \dots, v_n]$, and $\vec{w} = [w_1, \dots, w_n]$ be vectors in $\mathbb{R}^n$. Let c and d be scalars. Then,

1. $\vec{u} + \vec{v} = \vec{v} + \vec{u}$
2. $\vec{u} + (\vec{v} + \vec{w}) = (\vec{u} + \vec{v}) + \vec{w}$
3. $\vec{0} + \vec{u} = \vec{u} + \vec{0} = \vec{u}$
4. $\vec{u} + (-\vec{u}) = (-\vec{u}) + \vec{u} = \vec{0}$
5. $c(\vec{u} + \vec{v}) = c\vec{u} + c\vec{v}$
6. $(c + d)\vec{u} = c\vec{u} + d\vec{u}$
7. $(cd)\vec{u} = c(d\vec{u})$
8. $1\vec{u} = \vec{u}$

Note that $\vec{0}$ is called the *identity element* for addition and $-\vec{u}$ is the *additive inverse element of* $\vec{u}$.

We present the proof of one of the components of Theorem 1.2.1.

Example 1.2.3: ☢☢ Commutativity of Addition

Let $\vec{v} = [v_1, \dots, v_n]$ and $\vec{w} = [w_1, \dots, w_n]$ be vectors in $\mathbb{R}^n$. Prove that

$$\vec{v} + \vec{w} = \vec{w} + \vec{v}.$$

Proof. Consider the following.

$$\begin{aligned}\vec{v} + \vec{w} &= [v_1, \dots, v_n] + [w_1, \dots, w_n] \\ &= [v_1 + w_1, \dots, v_n + w_n] \\ &= [w_1 + v_1, \dots w_n + v_n] \\ &= \vec{w} + \vec{v}.\end{aligned}$$

The proposition is hence proved. ☐

Theorem 1.2.2: 🛑 Scalar Multiplication Producing the Zero Vector

Let $\vec{v} \in \mathbb{R}^n$ and let c be a scalar. Then,

$$(c = 0 \vee \vec{v} = \vec{0}) \iff c\vec{v} = \vec{0}.$$

Proof. First, we wish to show that

$$(c = 0 \vee \vec{v} = \vec{0}) \implies c\vec{v} = \vec{0}.$$

Suppose $c = 0$. We have

$$\vec{0} = c\vec{v} = [cv_1, \dots, cv_n].$$

We wish to show that all of $cv_1, \dots, cv_n$ must be zero, no matter the components of $\vec{v}$. By basic arithmetic, zero multiplied by any other number is also zero, so if $c = 0$, $c\vec{v}$ must be $\vec{0}$, as all the components of $c\vec{v}$ are zero. Suppose $\vec{v} = \vec{0}$. We again have

$$\vec{0} = c\vec{v} = [cv_1, \dots, cv_n]$$

and wish to show that all of $cv_1, \dots, cv_n$ must be zero, no matter the value of c. If $\vec{v} = 0$, all components of $\vec{v}$ are zero, and again, zero multiplied by any other number is zero, so all components of $c\vec{v}$ are zero. We have now shown that, indeed, $(c = 0 \vee \vec{v} = \vec{0}) \implies c\vec{v} = \vec{0}$. Now, we wish to show that

$$c\vec{v} = \vec{0} \implies (c = 0 \vee \vec{v} = \vec{0}).$$

Suppose that $\vec{v} \neq \vec{0}$. Again, we have

$$\vec{0} = c\vec{v} = [cv_1, \dots, cv_n].$$

There must be some nonzero $v_1, \dots, v_n$. Let this nonzero number be n. The only way $cn = 0$ is if $c = 0$. ☐

1.3 Lecture 3: August 26, 2022

1.3.1 Matrices

Consider the following warm-up exercise.

Example 1.3.1: General Vector Movement

What is the formula for the vector $\vec{v}$ representing the movement from $A = (a_1, \dots, a_n)$ to $B = (b_1, \dots, b_n)$. Then, find the magnitude of $\vec{v}$.

We see that

$$\vec{v} = [b_1 - a_1, \dots, b_n - a_n],$$

starting from A. Then,

$$||\vec{v}|| = \sqrt{(b_1 - a_1)^2 + \cdots + (b_n - a_n)^2}.$$

Consider the following definitions.

Definition 1.3.1: Matrices

Let $m, n \in \mathbb{N}$. An $m \times n$ matrix is a rectangular array of real or complex numbers with m rows and n columns. Matrices are often denoted with capital letters. The elements are called entries, and are usually written as

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}.$$

The main diagonal of A consists of $a_{11}, a_{22}, a_{33}, \dots$.

Definition 1.3.2: Square Matrices

A matrix is square if and only if $m = n$.

Definition 1.3.3: Diagonal Matrices

A matrix is diagonal if and only if it is square and $a_{ij} = 0$ whenever $i \neq j$. That is, all elements not on the main diagonal are zero.

Definition 1.3.4: The Identity Matrix

The identity matrix A is an $n \times n$ matrix where

$$a_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}.$$

Definition 1.3.5: The Zero Matrix

The zero matrix A is an $m \times n$ matrix where $a_{ij} = 0$ for all i and j.

Definition 1.3.6: Upper Triangular Matrices

An upper triangular matrix is a matrix such that $a_{ij} = 0$ for all $i > j$.

Definition 1.3.7: Lower Triangular Matrices

An lower triangular matrix is a matrix such that $a_{ij} = 0$ for all $i < j$.

Definition 1.3.8: The Set of All $m \times n$ Matrices

The set $\mathcal{M}_{mn}$ is the set of all $m \times n$ matrices. We can distinguish between real and complex matrices with $\mathcal{M}_{mn}^{\mathbb{R}}$ and $\mathcal{M}_{mn}^{\mathbb{C}}$, respectively. By default, we take $\mathcal{M}_{mn} = \mathcal{M}_{mn}^{\mathbb{R}}$.

Definition 1.3.9: Matrix Addition

Given $A, B \in \mathcal{M}_{mn}$, we define $A \pm B$ to be the matrix in $\mathcal{M}_{mn}$ with entries $a_{ij} \pm b_{ij}$.

Consider the following example.

Example 1.3.2: Matrix Addition

Consider the following addition.

$$\begin{bmatrix} 3 & 1 & e \\ 0 & \pi & 5 \end{bmatrix} + \begin{bmatrix} -1 & 3 & e \\ 5 & 0 & -2 \end{bmatrix} = \begin{bmatrix} 2 & 4 & 2e \\ 5 & \pi & 3 \end{bmatrix}.$$

Consider the following definition.

Definition 1.3.10: Scalar Multiplication With Matrices

Given $A \in \mathcal{M}_{mn}$ and $c \in \mathbb{R}$, we define cA as the matrix with elements ca_{ij}.

Consider the following example.

Example 1.3.3: Scalar Multiplication With Matrices

Consider the following scalar multiplication.

$$5 \begin{bmatrix} 5 & 1 \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} 25 & 5 \\ -5 & 0 \end{bmatrix}.$$

Consider the following definition.

Definition 1.3.11: Matrix Transpose

Given $A \in \mathcal{M}_{mn}$, we define $A^T \in \mathcal{M}_{nm}$ to be the matrix with the (i, j) entry equal to the (j, i) entry of A.

Consider the following example.

Example 1.3.4: Matrix Transpose

Consider the following transpose.

$$\begin{bmatrix} 2 & -5 \\ 3 & -1 \\ \pi & -\pi \end{bmatrix}^T = \begin{bmatrix} 2 & 3 & \pi \\ -5 & -1 & -\pi \end{bmatrix}.$$

Consider the following definition.

Definition 1.3.12: Symmetric and Skew-Symmetric Matrices

Suppose $A \in \mathcal{M}_{nn}$. Then,

1. A is symmetric if and only if $A = A^T$
2. A is skew-symmetric if and only if $A = -A^T$.

Note that A being square is necessary for the above conditions, but this condition is not sufficient.

Example 1.3.5: A Symmetric Matrix

Let $A = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$. Is A symmetric?

We see that

$$A^T = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix},$$

so A is symmetric.

Example 1.3.6: A Skew-Symmetric Matrix

Let $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$. Is A symmetric?

We see that
$$A^T = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} = -A,$$
so A is skew-symmetric.

Consider the following theorems.

Theorem 1.3.1: Transpose Properties

Suppose $A, B \in \mathcal{M}_{mn}$ and $c \in \mathbb{R}$. Then,

1. $(A^T)^T = A$
2. $(A+B)^T = A^T + B^T$
3. $(cA)^T = cA^T$.

Proof. Consider the matrix
$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}.$$
We see that
$$A^T = \begin{bmatrix} a_{11} & \cdots & a_{m1} \\ \vdots & \ddots & \vdots \\ a_{1n} & \cdots & a_{mn} \end{bmatrix}.$$
We will take the transpose of A^T, that is $(A^T)^T$, to yield
$$\begin{aligned} (A^T)^T &= \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \\ &= A. \end{aligned}$$
We have proved the first property. Let the matrix
$$B = \begin{bmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{m1} & \cdots & b_{mn} \end{bmatrix},$$
and therefore,
$$A + B = \begin{bmatrix} a_{11} + b_{11} & \cdots & a_{1n} + b_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} + b_{m1} & \cdots & a_{mn} + b_{mn} \end{bmatrix}.$$

Then,

$$\begin{aligned}(A+B)^T &= \begin{bmatrix} a_{11}+b_{11} & \cdots & a_{m1}+b_{m1} \\ \vdots & \ddots & \vdots \\ a_{1n}+b_{1n} & \cdots & a_{mn}+b_{mn} \end{bmatrix} \\ &= \begin{bmatrix} a_{11} & \cdots & a_{m1} \\ \vdots & \ddots & \vdots \\ a_{1n} & \cdots & a_{mn} \end{bmatrix} + \begin{bmatrix} b_{11} & \cdots & b_{m1} \\ \vdots & \ddots & \vdots \\ b_{1n} & \cdots & b_{mn} \end{bmatrix} \\ &= A^T + B^T. \end{aligned}$$

Thus, we have proved the second property. Finally, consider the matrix

$$cA = \begin{bmatrix} ca_{11} & \cdots & ca_{1n} \\ \vdots & \ddots & \vdots \\ ca_{m1} & \cdots & ca_{mn} \end{bmatrix}.$$

Then,

$$\begin{aligned}(cA)^T &= \begin{bmatrix} ca_{11} & \cdots & ca_{m1} \\ \vdots & \ddots & \vdots \\ ca_{1n} & \cdots & ca_{mn} \end{bmatrix} \\ &= c\begin{bmatrix} a_{11} & \cdots & a_{m1} \\ \vdots & \ddots & \vdots \\ a_{1n} & \cdots & a_{mn} \end{bmatrix} \\ &= cA^T. \end{aligned}$$

The final property is hence proved. □

Consider the following definition.

Definition 1.3.13: Trace

Let $A \in \mathcal{M}_{nn}$. Then,

$$\operatorname{trace} A = \sum_{i=1}^{n} a_{ii}.$$

That is, the sum of the elements on the main diagonal.

Theorem 1.3.2: Sum and Difference of Matrices and Their Transpose

Suppose A is an $n \times n$ matrix. Then,

1. $A + A^T$ is symmetric.
2. $A - A^T$ is skew-symmetric.

Proof. Consider the transpose of $A + A^T$. That is,

$$\begin{aligned}(A + A^T)^T &= A^T + (A^T)^T \\ &= A^T + A \\ &= A + A^T.\end{aligned}$$

This means that $A + A^T$ is symmetric. Similarly, Consider the transpose of $A - A^T$. That is,

$$\begin{aligned}(A - A^T)^T &= A^T - (A^T)^T \\ &= A^T - A \\ &= -(A - A^T).\end{aligned}$$

This means that $A - A^T$ is skew-symmetric. □

Theorem 1.3.3: The Relation Between Square, Symmetric, and Skew-Symmetric

Suppose A is a square matrix. There exists a symmetric matrix S and a skew-symmetric matrix R such that

$$A = S + R.$$

Proof. Note that

$$2A = (A + A^T) + (A - A^T).$$

Dividing both sides by 2 produces

$$A = \frac{1}{2}(A + A^T) + \frac{1}{2}(A - A^T),$$

meaning $S = \frac{1}{2}(A+A^T)$ and $R = \frac{1}{2}(A-A^T)$. By Theorem 1.3.2, and basic properties of the transpose operation, S is symmetric and R is skew-symmetric, as desired. We have now shown existence, and will now show uniqueness. Suppose P is a symmetric matrix and Q is a skew-symmetric matrix, and $A = P + Q$. Then, $A^T = P^T + Q^T = P - Q$. Note that $A + A^T = 2P$ so $P = \frac{1}{2}(A + A^T) = S$. Then, $A - A^T = 2Q$ so $Q = \frac{1}{2}(A - A^T) = R$. □

1.4 Lecture 4: August 29, 2022

1.4.1 The Dot Product

We will now define another vector operation: the dot product. We will also provide some important properties.

Definition 1.4.1: The Dot Product

Let $\vec{v} = [v_1, \dots, v_n]$ and $\vec{w} = [w_1, \dots, w_n]$ be vectors in $\mathbb{R}^n$. The dot product, or inner product, of $\vec{v}$ and $\vec{w}$ is given by

$$\vec{v} \cdot \vec{w} = \sum_{k=1}^{n} v_k w_k.$$

Note that $\vec{v}$ and $\vec{w}$ are orthogonal if and only if $\vec{v} \cdot \vec{w} = 0$. Consider the following examples.

Example 1.4.1: Dot Product 1

Find $[1, \pi] \cdot [-1, \pi]$.

We see that $[1, \pi] \cdot [-1, \pi] = -1 + \pi^2$.

Example 1.4.2: Dot Product 2

Find $[1, 0] \cdot [0, 1]$.

We see that $[1, 0] \cdot [0, 1] = 0$.

Theorem 1.4.1: Properties of the Dot Product

Let $\vec{u} = [u_1, \dots, u_n]$, $\vec{v} = [v_1, \dots, v_n]$, and $\vec{w} = [w_1, \dots, w_n]$ be vectors in $\mathbb{R}^n$. Let c be a scalar. Then,

1. $\vec{u} \cdot \vec{v} = \vec{v} \cdot \vec{u}$
2. $\vec{u} \cdot \vec{u} = ||\vec{u}||^2 \geq 0$
3. $\vec{u} \cdot \vec{u} = 0 \iff \vec{u} = \vec{0}$
4. $c(\vec{u} \cdot \vec{v}) = (c\vec{u}) \cdot \vec{v} = \vec{u} \cdot (c\vec{v})$
5. $\vec{u} \cdot (\vec{v} + \vec{w}) = (\vec{u} \cdot \vec{v}) + (\vec{u} \cdot \vec{w})$
6. $(\vec{u} + \vec{v}) \cdot \vec{w} = (\vec{u} \cdot \vec{w}) + (\vec{v} \cdot \vec{w})$

Consider the following example.

Example 1.4.3: ☢☢ Proving a Dot Product Property

Let $\vec{v} = [v_1, \dots, v_n]$ and $\vec{w} = [w_1, \dots, w_n]$ be vectors in $\mathbb{R}^n$. Let c be a scalar. Prove that

$$c(\vec{v} \cdot \vec{w}) = (c\vec{v}) \cdot \vec{w}.$$

Proof. Consider the following transitive chain of equality:

$$\begin{aligned} c(\vec{v} \cdot \vec{w}) &= c(v_1 w_1 + \cdots + v_n w_n) \\ &= (cv_1 w_1 + \cdots + cv_n w_n) \\ &= ((cv_1)w_1 + \cdots + (cv_n)w_n) \\ &= (c\vec{v}) \cdot \vec{w}. \end{aligned}$$

The proposition is hence proved. □

Consider the following theorem regarding the angle between two vectors.

Theorem 1.4.2: ☻ The Angle Between Two Vectors

Given nonzero $\vec{v}, \vec{w} \in \mathbb{R}^n$,

$$\vec{v} \cdot \vec{w} = ||\vec{v}||||\vec{w}|| \cos\theta,$$

where θ is the angle between the two vectors.

Proof. Consider two unit vectors $\hat{v}$ and $\hat{w}$ where

$$\hat{v} = [\cos\alpha, \sin\alpha], \quad \hat{w} = [\cos\beta, \sin\beta].$$

The angle between the two vectors is $\theta = \beta - \alpha$. The dot product between the two vectors is

$$\begin{aligned} \hat{v} \cdot \hat{w} &= \cos\alpha\cos\beta + \sin\beta\sin\alpha \\ &= \cos(\beta - \alpha) \\ &= \cos\theta. \end{aligned}$$

Now, if we consider $\vec{v} = c_1\hat{v}$ and $\vec{w} = c_2\hat{w}$, meaning that $||\vec{v}|| = c_1$ and $||\vec{w}|| = c_2$, we simply scale the above result with $\hat{v}$ and $\hat{v}$ by the magnitudes of $\vec{v}$ and $\vec{w}$ to produce

$$\vec{v} \cdot \vec{w} = ||\vec{v}||||\vec{w}|| \cos\theta,$$

The proposition is hence proved, but it is also of note to realize that the same proposition can be proved with the law of cosines. □

We will now state the famous Cauchy-Schwarz Inequality.

Theorem 1.4.3: The Cauchy-Schwarz Inequality

If $\vec{v}, \vec{w} \in \mathbb{R}^n$,

$$|\vec{v} \cdot \vec{w}| \leq ||\vec{v}||||\vec{w}||.$$

Proof. Consider the following lemma. If $\vec{v}, \vec{w} \in \mathbb{R}^n$, with $||\vec{v}|| = ||\vec{w}|| = 1$,

$$-1 \leq \vec{v} \cdot \vec{w} \leq 1.$$

Proof. We will start by showing that $-1 \leq \vec{v} \cdot \vec{w}$. Consider $||\vec{v} + \vec{w}||^2 \geq 0$. We have that

$$\begin{aligned} ||\vec{v} + \vec{w}||^2 &= (\vec{v} + \vec{w}) \cdot (\vec{v} + \vec{w}) \\ &= \vec{v} \cdot \vec{v} + \vec{v} \cdot \vec{w} + \vec{w} \cdot \vec{v} + \vec{w} \cdot \vec{w} \\ &= ||\vec{v}||^2 + 2\vec{v} \cdot \vec{w} + ||\vec{w}||^2 \\ &= 1 + 2\vec{v} \cdot \vec{w} + 1 \\ &= 2 + 2\vec{v} \cdot \vec{w} \\ &\geq 0. \end{aligned}$$

We then solve the resulting inequality which yields

$$2\vec{v} \cdot \vec{w} \geq -2 \implies \vec{v} \cdot \vec{w} \geq -1.$$

Similarly, consider that $||\vec{v} - \vec{w}||^2 \geq 0$. We then have

$$\begin{aligned} ||\vec{v} - \vec{w}||^2 &= (\vec{v} - \vec{w}) \cdot (\vec{v} - \vec{w}) \\ &= ||\vec{v}||^2 - 2\vec{v} \cdot \vec{w} + ||\vec{w}||^2 \\ &= 2 - 2\vec{v} \cdot \vec{w} \\ &\geq 0. \end{aligned}$$

Then, we have

$$2 - 2\vec{v} \cdot \vec{w} \geq 0 \implies \vec{v} \cdot \vec{w} \leq 1$$

as desired. □

Consider the following cases. If $\vec{v} = \vec{0}$ or $\vec{w} = \vec{0}$, then, $\vec{v} \cdot \vec{w} = 0$ and $||\vec{v}||\vec{w}|| = 0$. Thus, in this case, $|\vec{v} \cdot \vec{w}| = ||\vec{v}||||\vec{w}|| = 0$. If $\vec{v} \neq \vec{0}$ and $\vec{w} \neq \vec{0}$, Consider the unit vectors $\frac{1}{||\vec{v}||}\vec{v}$ and $\frac{1}{||\vec{w}||}\vec{w}$. Hence, by the lemma,

$$-1 \leq \frac{1}{||\vec{v}||}\vec{v} \cdot \frac{1}{||\vec{w}||}\vec{w} \leq 1.$$

This implies that

$$-||\vec{v}||||\vec{w}|| \leq \vec{v} \cdot \vec{w} \leq ||\vec{v}||||\vec{w}||.$$

That is,

$$|\vec{v} \cdot \vec{w}| \leq ||\vec{v}||||\vec{w}||,$$

which is precisely the statement of the theorem. □

We will now state the famous Triangle Inequality, or Minkowski's Inequality.

Theorem 1.4.4: The Triangle Inequality

If $\vec{v}, \vec{w} \in \mathbb{R}^n$,

$$||\vec{v} + \vec{w}|| \leq ||\vec{v}|| + ||\vec{w}||.$$

Proof. Since $f(x) = \sqrt{x}$ is always increasing, we need only prove

$$||\vec{v} + \vec{w}||^2 \leq (||\vec{v}|| + ||\vec{w}||)^2.$$

We have

$$\begin{aligned} ||\vec{v} + \vec{w}||^2 &= ||\vec{v}||^2 + 2\vec{v} \cdot \vec{w} + ||\vec{w}||^2 \\ &\leq ||\vec{v}||^2 + 2|\vec{v} \cdot \vec{w}| + ||\vec{w}||^2 \\ &\leq ||\vec{v}||^2 + 2||\vec{v}||||\vec{w}|| + ||\vec{w}||^2 \\ &= (||\vec{v}|| + ||\vec{w}||)^2. \end{aligned}$$

The theorem is hence proved. □

We will now define the projection onto a vector.

Definition 1.4.2: The Projection Onto a Vector

Given $\vec{v}, \vec{w} \in \mathbb{R}^n$, where $\vec{w} \neq \vec{0}$, we conclude that the projection of $\vec{v}$ onto $\vec{w}$ is

$$\begin{aligned} \text{proj}_{\vec{w}} \vec{v} &= ||\vec{v}|| \cos\theta \frac{\vec{w}}{||\vec{w}||} \\ &= ||\vec{v}|| \frac{\vec{v} \cdot \vec{w}}{||\vec{v}||||\vec{w}||} \frac{\vec{w}}{||\vec{w}||} \\ &= \frac{\vec{v} \cdot \vec{w}}{||\vec{w}||^2} \vec{w} \\ &= \frac{\vec{v} \cdot \vec{w}}{\vec{w} \cdot \vec{w}} \vec{w}. \end{aligned}$$

Consider the following example.

Example 1.4.4: Projections

Given $\vec{v} = [x, y]$, find $\text{proj}_{[1,0]} \vec{v}$ and $\text{proj}_{[0,1]} \vec{v}$.

We see that

$$\text{proj}_{[1,0]} \vec{v} = [x, 0], \quad \text{proj}_{[0,1]} \vec{v} = [0, y].$$

Consider the following theorems.

Theorem 1.4.5: Sum of Parallel and Perpendicular Projections

Given $\vec{w} \neq 0$ where $\vec{v}, \vec{w} \in \mathbb{R}^n$,

$$\vec{v} = \operatorname{proj}_{\vec{w}} \vec{v} + (\vec{v} - \operatorname{proj}_{\vec{w}} \vec{v}).$$

The first addend returns a vector parallel to $\vec{w}$, and the second added returns a vector perpendicular to $\vec{w}$.

Proof. Consider the expansion of the right hand side, that is,

$$\begin{aligned}\operatorname{proj}_{\vec{w}} \vec{v} + (\vec{v} - \operatorname{proj}_{\vec{w}} \vec{v}) &= \frac{\vec{v} \cdot \vec{w}}{\vec{w} \cdot \vec{w}} \vec{w} + \left(\vec{v} - \frac{\vec{v} \cdot \vec{w}}{\vec{w} \cdot \vec{w}} \vec{w} \right) \\ &= \vec{v}.\end{aligned}$$

We have arrived at the desired result. $\square$

Theorem 1.4.6: Projections Depend on Lines

Given $\vec{w} \neq 0$ where $\vec{v}, \vec{w} \in \mathbb{R}^n$, and $c \in \mathbb{R}$ where $c \neq 0$,

$$\operatorname{proj}_{\vec{w}} \vec{v} = \operatorname{proj}_{c\vec{w}} \vec{v}.$$

Proof. Consider the expansion of the left hand side, that is,

$$\begin{aligned}\operatorname{proj}_{c\vec{w}} \vec{v} &= \frac{\vec{v} \cdot c\vec{w}}{c\vec{w} \cdot c\vec{w}} c\vec{w} \\ &= \frac{\vec{v} \cdot \vec{w}}{\vec{w} \cdot \vec{w}} \vec{w} \\ &= \operatorname{proj}_{\vec{w}} \vec{v}.\end{aligned}$$

The proposition is hence proved. $\square$

1.5 Lecture 5: August 31, 2022

1.5.1 Matrix Multiplication

Consider the following definition of matrix multiplication.

Definition 1.5.1: Matrix Multiplication

For $A \in \mathcal{M}_{mn}$ and $B \in \mathcal{M}_{np}$, such that

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \quad B = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1p} \\ b_{21} & b_{22} & \cdots & b_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{np} \end{bmatrix},$$

the matrix product $C = AB$ is defined such that $C \in \mathcal{M}_{mp}$ and is given by

$$C = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1p} \\ c_{21} & c_{22} & \cdots & c_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ c_{m1} & c_{m2} & \cdots & c_{mp} \end{bmatrix}$$

such that $c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj} = \sum_{k=1}^{n} a_{ik}b_{kj}$. The matrix product is only defined if the number of columns of the matrix on the left is equal to the number of the rows of the matrix on the right. Note that matrix multiplication is not commutative. We may also say that the entry in c_{ij} is the dot product of the ith row of A and the jth column of B.

Consider the following examples.

Example 1.5.1: Matrix Multiplication 1

Find $\begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 3 \\ -1 \end{bmatrix}$.

We see that

$$\begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 3 \\ -1 \end{bmatrix} = \begin{bmatrix} 5 \\ 2 \end{bmatrix}.$$

Example 1.5.2: Matrix Multiplication 2

Find $\begin{bmatrix} 3 & 1 & -1 \end{bmatrix} \begin{bmatrix} \pi \\ -3 \\ 7 \end{bmatrix}$.

We see that

$$\begin{bmatrix} 3 & 1 & -1 \end{bmatrix} \begin{bmatrix} \pi \\ -3 \\ 7 \end{bmatrix} = \begin{bmatrix} 3\pi - 10 \end{bmatrix}.$$

Example 1.5.3: Matrix Multiplication 3

Find $\begin{bmatrix} 3 & 1 & -1 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 1 & 7 \\ 1 & 3 \end{bmatrix}$.

We see that

$$\begin{bmatrix} 3 & 1 & -1 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 1 & 7 \\ 1 & -3 \end{bmatrix} = \begin{bmatrix} 6 & 13 \end{bmatrix}.$$

Example 1.5.4: Matrix Multiplication 4

Find $\begin{bmatrix} 3 & 1 \\ -1 & 5 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$.

We see that

$$\begin{bmatrix} 3 & 1 \\ -1 & 5 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 1 \\ -1 & 5 \end{bmatrix}.$$

Example 1.5.5: Matrix Multiplication 5

Find $\begin{bmatrix} 3 & 1 \\ -1 & 5 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$.

We see that

$$\begin{bmatrix} 3 & 1 \\ -1 & 5 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Consider the following properties of matrix multiplication.

Theorem 1.5.1: Properties of Matrix Multiplication

Suppose A, B, and C are matrices such that matrix multipication is well-defined. Then,

1. $A(BC) = (AB)C$
2. $A(B + C) = AB + AC$
3. $(A + B)C = AC + BC$
4. $c(AB) = (cA)B = A(cB)$.

We now ponder two questions.

1. Does $AB = BA$? No.
2. If $AB = 0$, does $A = 0$ or $B = 0$? No.

Consider the following definition.

Definition 1.5.2: Raising a Matrix to a Power

Let $A \in \mathcal{M}_{nn}$. Then Consider an $n \times n$ matrix A. Then, $A^0 = I_n$, $A^1 = A$, and for $k \geq 2$,

$$A^k = (A^{k-1})A.$$

2

Systems of Linear Equations

2.1 Lecture 6: September 2, 2022

2.1.1 Systems of Linear Equations

Consider the following definitions.

Definition 2.1.1: Linear Equations

A linear equation is an equation of the form

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

Definition 2.1.2: Systems of Linear Equations

A system of linear equations is a system of the form

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n &= b_1 \\ &\vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n &= b_m. \end{aligned}$$

2.2 Lecture 7: September 7, 2022

2.2.1 Systems of Linear Equations as Matrices

We may write systems of linear equations in terms of matrices as

$$A\begin{bmatrix}x_1\\ \vdots \\ x_n\end{bmatrix} = \begin{bmatrix}b_1\\ \vdots \\ b_m\end{bmatrix},$$

where A is the matrix with entries a_{ij}. We will use the convention that $X = [x_1, \dots, x_n]^T$ and $B = [b_1, \dots, b_m]^T$. Consider the following theorem.

Theorem 2.2.1: Characterizing Solutions of Linear Systems

A system of linear equations can either have

1. No solution.
2. One unique solution.
3. Infinitely many solutions.

2.2.2 Matrix Row Operations

Consider the following operations.

1. Multiplication of a row by a nonzero scalar. Notated as $c\langle r_1\rangle \to \langle r_1\rangle$.
2. Addition of a scalar multiple of one row to another. Notated as $\langle r_1\rangle + (c)\langle r_2\rangle \to \langle r_1\rangle$
3. Switching the elements of two rows. Notated as $\langle r_1\rangle \leftrightarrow \langle r_2\rangle$.

Consider the following examples.

Example 2.2.1: Row Operation 1

Consider the matrix $\begin{bmatrix}3 & 1 & -1\\ 1 & 0 & 1\\ -1 & 1 & 5\end{bmatrix}$. Find $4\langle 3\rangle \to \langle 3\rangle$.

We obtain

$$\begin{bmatrix}3 & 1 & -1\\ 1 & 0 & 1\\ -4 & 4 & 20\end{bmatrix}.$$

Example 2.2.2: ☢☢ Row Operation 2

Consider the matrix $\begin{bmatrix} 3 & 1 & -1 \\ 1 & 0 & 1 \\ -1 & 1 & 5 \end{bmatrix}$. Find $\langle 1\rangle + (-3)\langle 2\rangle \to \langle 1\rangle$.

We obtain

$$\begin{bmatrix} 0 & 1 & -4 \\ 1 & 0 & 1 \\ -1 & 1 & 5 \end{bmatrix}.$$

Example 2.2.3: ☢☢ Row Operation 3

Consider the matrix $\begin{bmatrix} 3 & 1 & -1 \\ 1 & 0 & 1 \\ -1 & 1 & 5 \end{bmatrix}$. Find $\langle 2\rangle \leftrightarrow \langle 3\rangle$.

We obtain

$$\begin{bmatrix} 3 & 1 & -1 \\ -1 & 1 & 5 \\ 1 & 0 & 1 \end{bmatrix}.$$

2.3 Lecture 8: September 9, 2022

2.3.1 Solving Linear Systems

Given a linear system of equations, we solve by the following steps.

1. Convert the linear system into the matrix equation $AX = B$, written $[A|B]$.
2. Use the three row operations to reduce $[A|B]$ to one with "lots of zeroes and ones."
3. Perform back substitution and analyze the solution set.

Consider the following examples.

Example 2.3.1: ☢☢ No Solution

Consider the matrix

$$\begin{bmatrix} 3 & -6 & 0 & 3 & | & 9 \\ -2 & 4 & 2 & -1 & | & -11 \\ 4 & -8 & 6 & 7 & | & -5 \end{bmatrix}.$$

By row operations, we obtain

$$\begin{bmatrix} 1 & -2 & 0 & 1 & | & 3 \\ 0 & 0 & 1 & \frac{1}{2} & | & -\frac{5}{2} \\ 0 & 0 & 0 & 0 & | & -2 \end{bmatrix}.$$

Looking at the last row, we see the equation $0 = -2$, which is not true. Hence, the system has no solution.

Example 2.3.2: ☢☢ Infinitely Many Solutions

Consider the matrix

$$\begin{bmatrix} 3 & 1 & 7 & 2 & | & 13 \\ 2 & -4 & 14 & -1 & | & -10 \\ 5 & 11 & -7 & 8 & | & 59 \\ 2 & 5 & -4 & -3 & | & 39 \end{bmatrix}.$$

By row operations, we obtain

$$\begin{bmatrix} 1 & \frac{1}{3} & \frac{7}{3} & \frac{2}{3} & | & \frac{13}{3} \\ 0 & 1 & -2 & \frac{1}{2} & | & 4 \\ 0 & 0 & 0 & 1 & | & -2 \\ 0 & 0 & 0 & 0 & | & 0 \end{bmatrix}.$$

We see that x_1, x_2, and x_4 are determined because their respective column has a 1 in the correct position. In contrast, x_3 is a free variable. To find the solution set, let $x_3 = c \in \mathbb{R}$ and solve for x_1, x_2, and x_4 in terms of c. We have $x_4 = -2$. Then to find x_2 we have

$$x_2 - 2x_3 + \frac{1}{2}x_4 = 4 \implies x_2 - 2 + \frac{1}{2}(-2) = 4 \implies x_2 = 2c + 5.$$

For x_1, we have

$$x_1 + \frac{1}{3}x_2 + \frac{7}{3}x_3 + \frac{2}{3}x_4 = \frac{13}{3} \implies x_1 = -3c + 4.$$

The solution set is then $\{(-3c + 4, 2c + 5, c, -2) : c \in \mathbb{R}\}$.

We generally agree that back substitution is not much fun. Consider the following example.

Example 2.3.3: ☢☢ No More Back Substitution

Note that the matrix

$$\begin{bmatrix} 3 & -3 & -2 & | & 23 \\ -6 & 4 & 3 & | & -40 \\ -2 & 1 & 1 & | & -12 \end{bmatrix}$$

reduces into

$$\begin{bmatrix} 1 & -1 & -\frac{2}{3} & | & \frac{23}{3} \\ 0 & 1 & \frac{1}{3} & | & -\frac{10}{3} \\ 0 & 0 & 1 & | & 2 \end{bmatrix}.$$

We may now "get rid of" -1, $-\frac{2}{3}$, and $\frac{1}{3}$. We perform $-\frac{1}{3}\langle 3\rangle + \langle 2\rangle \to \langle 2\rangle$ which produces

$$\begin{bmatrix} 1 & -1 & -\frac{2}{3} & | & \frac{23}{3} \\ 0 & 1 & 0 & | & -4 \\ 0 & 0 & 1 & | & 2 \end{bmatrix}.$$

Then, we will perform $\langle 1\rangle + \langle 2\rangle \to \langle 1\rangle$, yielding

$$\begin{bmatrix} 1 & 0 & -\frac{2}{3} & | & \frac{23}{3} - 4 \\ 0 & 1 & 0 & | & -4 \\ 0 & 0 & 1 & | & 2 \end{bmatrix}.$$

Finally, we will perform $\frac{2}{3}\langle 3\rangle + \langle 1\rangle \to 1$, obtaining

$$\begin{bmatrix} 1 & 0 & 0 & | & 5 \\ 0 & 1 & 0 & | & -4 \\ 0 & 0 & 1 & | & 2 \end{bmatrix}$$

which means $x_1 = 5$, $x_2 = -4$, and $x_3 = 2$.

Consider the following theorem.

Theorem 2.3.1: 🛑 Row Operations

Suppose $A \in \mathcal{M}_{mn}$ and $B \in \mathcal{M}_{np}$. Then,

1. If R is a row operation, $R(AB) = (R(A))B$.
2. If $R_1, \dots, R_n$ are row operations, $R_n(\dots(R_2(R_1(AB)))\dots) = (R_n(\dots(R_2(R_1(A)))\dots))B$.

Note that this result follows from the associativity of matrix multiplication, as any row operation can be represented by a multiplication of two matrices.

Consider the following examples of solving linear systems.

Example 2.3.4: ☢☢ Linear System 1

Solve the following system:

$$\left[\begin{array}{ccc|c} 2 & -1 & 1 & 0 \\ 1 & 3 & 4 & 0 \end{array}\right]$$

We first perform the row operation $\frac{1}{2}\langle 1\rangle \to \langle 1\rangle$, which produces

$$\left[\begin{array}{ccc|c} 1 & -\frac{1}{2} & \frac{1}{2} & 0 \\ 1 & 3 & 4 & 0 \end{array}\right].$$

Then, we perform $\langle 1\rangle - \langle 2\rangle \to \langle 2\rangle$. We obtain

$$\left[\begin{array}{ccc|c} 1 & -\frac{1}{2} & \frac{1}{2} & 0 \\ 0 & -\frac{7}{2} & -\frac{7}{2} & 0 \end{array}\right].$$

Next, we have $-\frac{2}{7}\langle 2\rangle \to \langle 2\rangle$. This yields

$$\left[\begin{array}{ccc|c} 1 & -\frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 1 & 1 & 0 \end{array}\right].$$

From here, let $x_3 = c$. Then, $x_2 = -c$. To find x_1, we use the equation

$$x_1 - \frac{1}{2}(-c) + \frac{1}{2}c = 0,$$

which implies that $x_1 = -c$. Thus, the solution set is $\{(-c, -c, c) : c \in \mathbb{R}\}$.

Example 2.3.5: ☢☢ Linear System 2

Solve the following system:

$$\begin{bmatrix} 1 & -2 & 1 & 2 & | & 1 \\ 1 & 1 & -1 & 1 & | & 2 \\ 1 & 7 & -5 & -1 & | & 3 \end{bmatrix}$$

First, we perform $\langle 1\rangle - \langle 2\rangle \to \langle 2\rangle$, yielding

$$\begin{bmatrix} 1 & -2 & 1 & 2 & | & 1 \\ 0 & -3 & 2 & 1 & | & -1 \\ 1 & 7 & -5 & -1 & | & 3 \end{bmatrix}.$$

Then, we perform $\langle 1\rangle - \langle 3\rangle \to \langle 3\rangle$. We obtain

$$\begin{bmatrix} 1 & -2 & 1 & 2 & | & 1 \\ 0 & -3 & 2 & 1 & | & -1 \\ 0 & -9 & 6 & 3 & | & -2 \end{bmatrix}.$$

Then, we have $-\frac{1}{3}\langle 2\rangle \to \langle 2\rangle$. This produces

$$\begin{bmatrix} 1 & -2 & 1 & 2 & | & 1 \\ 0 & 1 & -\frac{2}{3} & -\frac{1}{3} & | & \frac{1}{3} \\ 0 & -9 & 6 & 3 & | & -2 \end{bmatrix}.$$

Our final row operation is $\langle 3\rangle + 9\langle 2\rangle \to \langle 3\rangle$. This provides us with

$$\begin{bmatrix} 1 & -2 & 1 & 2 & | & 1 \\ 0 & 1 & -\frac{2}{3} & -\frac{1}{3} & | & \frac{1}{3} \\ 0 & 0 & 0 & 0 & | & 1 \end{bmatrix},$$

meaning that there is no solution to the system.

Example 2.3.6: ☢☢ Linear System 3

Solve the following system:

$$\left[\begin{array}{ccc|c} 1 & -1 & 2 & 1 \\ 2 & 0 & 2 & 1 \\ 1 & -3 & 4 & 2 \end{array}\right]$$

Our first row operation is $2\langle 1\rangle - \langle 2\rangle \to \langle 2\rangle$. This produces

$$\left[\begin{array}{ccc|c} 1 & -1 & 2 & 1 \\ 0 & -2 & 2 & 1 \\ 1 & -3 & 4 & 2 \end{array}\right].$$

Then, we have $\langle 1\rangle - \langle 3\rangle \to \langle 3\rangle$, providing

$$\left[\begin{array}{ccc|c} 1 & -1 & 2 & 1 \\ 0 & -2 & 2 & 1 \\ 0 & 2 & -2 & -1 \end{array}\right].$$

Next, we perform $\langle 2\rangle + \langle 3\rangle \to \langle 3\rangle$. We obtain

$$\left[\begin{array}{ccc|c} 1 & -1 & 2 & 1 \\ 0 & -2 & 2 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

We perform another row operation, $-\frac{1}{2}\langle 2\rangle \to \langle 2\rangle$. This yields

$$\left[\begin{array}{ccc|c} 1 & -1 & 2 & 1 \\ 0 & 1 & -1 & -\frac{1}{2} \\ 0 & 0 & 0 & 0 \end{array}\right].$$

Next, we have $\langle 2\rangle + \langle 1\rangle \to \langle 1\rangle$, which gives

$$\left[\begin{array}{ccc|c} 1 & 0 & 1 & \frac{1}{2} \\ 0 & 1 & -1 & -\frac{1}{2} \\ 0 & 0 & 0 & 0 \end{array}\right].$$

Let $x_3 = c$. Then, $x_2 = -\frac{1}{2} + c$ and $x_1 = \frac{1}{2} - c$. This means the solution set is $\left\{\left(\frac{1}{2} - c, -\frac{1}{2} + c, c\right) : c \in \mathbb{R}\right\}$.

2.4 Lecture 9: September 12, 2022

2.4.1 Formalizing Previous Notions: Part I

Consider the following definitions.

Definition 2.4.1: Row Echelon Form

A matrix A is in row echelon form if and only if

1. All rows consisting of only zeroes are at the bottom.
2. The leading coefficient, or the pivot, of a nonzero row is always strictly to the right of the leading coefficient of the row above it.

Definition 2.4.2: Reduced Row Echelon Form

A matrix A is in reduced row echelon form if and only if

1. The first nonzero entry in each row is one.
2. Each successive row has its first nonzero entry in a later column.
3. All entries above and below the first nonzero entry are zero.
4. All rows consisting of only zeroes are at the bottom.

Note that every matrix has a unique reduced row echelon form.

Consider the following theorems and definitions.

Theorem 2.4.1: Number of Solutions to a Linear System

Let $AX = B$ be a system of linear equations. Let C be the reduced row echelon form augmented matrix obtained by row reducing $[A|B]$. Then,

1. If there is a row of C having all zeroes to the left of the augmentation bar but with its last entry nonzero, $AX = B$ has no solution.
2. If not, and if one of the columns of C to the left of the augmentation bar has no nonzero pivot entry, $AX = B$ has an infinite number of solutions. The nonpivot columns correspond to (independent) variables that can take on any value, and the values of the remaining (dependent) variables are determined from those.
3. Otherwise $AX = B$ has a unique solution.

Consider the following definitions.

Definition 2.4.3: Homogeneous Systems

Given $A \in \mathcal{M}_{mn}$, the homogeneous system associated with A is

$$AX = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Theorem 2.4.2: Solutions to Homogeneous Systems

Given $A \in \mathcal{M}_{mn}$, the homogeneous system always has at least one solution, called the *trivial solution*. Namely,

$$x_1 = 0, \quad x_2 = 0, \quad \dots, \quad x_n = 0.$$

Also, consider the following.

1. If $m < n$, the solution set is infinite.
2. If $X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ and $X_\sim = \begin{bmatrix} x_{\sim 1} \\ \vdots \\ x_{\sim n} \end{bmatrix}$ are solutions,
$$cX + X_\sim$$
is a solution for any $c \in \mathbb{R}$.
3. If $AX = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$ and $A\hat{X} = B = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}$,
$$cX + \hat{X}$$
is a solution to $[A|B]$. Notice that (2) is a special case of (3).

Proof. Consider $A(cX + \hat{X})$. We wish to show that $A(cX + \hat{X}) = B$. We see that

$$\begin{aligned} A(cX + \hat{X}) &= A(cX) + A\hat{X} \\ &= cAX + A\hat{X} \\ &= \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} + \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix} \\ &= B, \end{aligned}$$

as desired. □

This process is analogous to solving homogeneous differential equations to solve nonhomogeneous differential equations.

Definition 2.4.4: Equivalence of Linear Systems

Two systems $[A|B]$ and $[A_\sim|B_\sim]$ are equivalent if and only if

$$AX = B \wedge A_\sim X = B_\sim.$$

That is, if they have the same solution sets.

Definition 2.4.5: Row Equivalence

A matrix A is row equivalent to a matrix B if B can be obtained by a finite number of row operations conducted on A.

For example, Gaussian Elimination and Gauss-Jordan Elimination produce matrices that are row equivalent to the original matrix.

One may ask: What is the relationship between these relations? We see that row equivalence implies system equivalence. But, two systems can have the same solution set, but have different sizes, making row equivalence impossible. For the latter case, consider two matrices of different sizes, but with an empty solution set. Recall the following definition from discrete mathematics.

Definition 2.4.6: Equivalence Relations

A relation $\sim$ on a set S is an equivalence relation on S if and only if $\sim$ is reflexive, symmetric, and transitive. That is, if

1. $\forall a \in S, a \sim a$.
2. $\forall a, b \in S, a \sim b \implies b \sim a$
3. $\forall a, b, c \in S, a \sim b \wedge b \sim c \implies a \sim c$.

Consider the following theorem.

Theorem 2.4.3: System Equivalence and Row Equivalence are Equivalence Relations

First, consider the following table.

Row Operation	Reverse Operation
$c\langle i\rangle \to \langle i\rangle$	$\frac{1}{c}\langle i\rangle \to \langle i\rangle$
$c\langle i\rangle + \langle j\rangle \to \langle j\rangle$	$-c\langle i\rangle + \langle j\rangle \to \langle j\rangle$
$\langle i\rangle \leftrightarrow \langle j\rangle$	$\langle i\rangle \leftrightarrow \langle j\rangle$

Proof. We will consider row equivalence first, and wish to show that row equivalence is reflexive, symmetric, and transitive. Reflexivity is trivial. We can simply not perform any row operations on a matrix A, and we are left with A. The above table can be used to show that row equivalence is symmetric. If a sequence of row operations is carried out on A and produces a matrix B, we can simply carry out the reverse operations on B to lead us back to A. For transitivity, if a sequence of row operations is carried out on A and leads to B, and a second sequence of row operations is performed on B and leads to C, we simply carry out the sequences, in sequence, on A to get us to C.

Now, we consider system equivalence. The system $[A|B]$, of course, has the same solution set as itself. If the system $[A|B]$ has the same solution set as $[C|D]$, $[C|D]$ has the same solution set as $[A|B]$. If $[A|B]$ has the same solution set as $[C|D]$ and $[C|D]$ has the same solution set as the system $[E|F]$, $[A|B]$ has the same solution set as $[E|F]$. □

2.5 Lecture 10: September 14, 2022

2.5.1 Formalizing Previous Notions: Part II

Consider the following formalization of our last discoveries.

Theorem 2.5.1: Row Equivalence Implies System Equivalence

If $[A|B]$ is row equivalent to $[C|D]$, $[A|B]$ is equivalent to $[C|D]$.

Theorem 2.5.2: Uniqueness of Reduced Row Echelon Form

Every matrix is row equivalent to a unique matrix in reduced row echelon form. Two matrices are row equivalent if and only if they have the same reduced row echelon form.

Definition 2.5.1: Rank

Given $A \in \mathcal{M}_{mn}$, $\operatorname{rank} A$ is the number of nonzero rows in the unique matrix that is row equivalent to A and is in reduced row echelon form.

Consider the following example.

Example 2.5.1: Rank

Consider

$$A = \begin{bmatrix} 3 & 1 & 0 & 1 \\ 0 & -2 & 12 & -5 \\ 2 & -3 & 22 & -14 \end{bmatrix}.$$

By row reduction, we have the matrix

$$\begin{bmatrix} 1 & 0 & 2 & -1 \\ 0 & 1 & -6 & 4 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

and see that $\operatorname{rank} A = 2$.

Consider the following theorem.

Theorem 2.5.3: Number of Solutions to Homogeneous Systems

If $A \in \mathcal{M}_{mn}$,

1. If $\operatorname{rank} A < n$, $AX = 0$ has an infinite solution set.
2. If $\operatorname{rank} A = n$, $AX = 0$ has only the trivial solution.

We will now define linear combinations of vectors.

Definition 2.5.2: Linear Combinations

Let $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k \in \mathbb{R}^n$. The vector $\vec{v}$ is a linear combination of $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k$ if and only if there are scalars $c_1, c_2, \dots, c_k$ such that

$$\vec{v} = c_1\vec{v}_1 + \cdots + c_k\vec{v}_k.$$

In general, $\{c\vec{v} : c \in \mathbb{R}\}$ is a line unless $\vec{v} = \vec{0}$. Also, $\{c_1\vec{v}_1 + c_2\vec{v}_2 : c_1, c_2 \in \mathbb{R}\}$ is usually a plane, but could be either a point or a line. This pattern is an introduction to the concept of linear independence, which will be elaborated on later in the text. Consider the following example.

Example 2.5.2: Is a Vector a Linear Combination of Others? 1

Let $\vec{v} = [1, 0]$, $\vec{v}_1 = [\pi, 1]$, and $\vec{v}_2 = [2, 1]$. Is $\vec{v}$ a linear combination of $\vec{v}_1$ and $\vec{v}_2$?

Notice that $\vec{v}_1 - \vec{v}_2 = [\pi - 2, 0]$. Then,

$$\frac{1}{\pi - 2}[\pi - 2, 0] = [1, 0] = \vec{v}.$$

We then have that

$$\vec{v} = \frac{1}{\pi - 2}[\pi, 1] - \frac{1}{\pi - 2}[2, 1].$$

Therefore, $\vec{v}$ is a linear combination of $\vec{v}_1$ and $\vec{v}_2$.

The above solution used a bit of trickery. Instead, given $\vec{v}_1, \dots, \vec{v}_k$, we form the equation

$$\begin{bmatrix}\vec{v}_1, \dots, \vec{v}_k\end{bmatrix}\begin{bmatrix}c_1 \\ \vdots \\ c_k\end{bmatrix} = \begin{bmatrix}\vec{v}\end{bmatrix}$$

and solve for the necessary constants.

Consider the following examples.

Example 2.5.3: Is a Vector a Linear Combination of Others? 2

Let $\vec{v} = [1, 0, 0]$, $\vec{v}_1 = [-4, 2, 0]$, and $\vec{v}_2 = [2, 1, 1]$. Is $\vec{v}$ a linear combination of $\vec{v}_1$ and $\vec{v}_2$?

Consider the system

$$\left[\begin{array}{cc|c} -4 & 2 & 1 \\ 2 & 1 & 0 \\ 0 & 1 & 0 \end{array}\right].$$

We first perform the row operation $-\frac{1}{4}\langle 1\rangle \to \langle 1\rangle$ to obtain

$$\left[\begin{array}{cc|c} 1 & -\frac{1}{2} & -\frac{1}{4} \\ 2 & 1 & 0 \\ 0 & 1 & 0 \end{array}\right].$$

Then, we have $2\langle 1\rangle - \langle 2\rangle \to \langle 2\rangle$, producing

$$\left[\begin{array}{cc|c} 1 & -\frac{1}{2} & -\frac{1}{4} \\ 0 & -2 & -\frac{1}{2} \\ 0 & 1 & 0 \end{array}\right].$$

Next, we will carry out $-\frac{1}{2}\langle 2\rangle \to \langle 2\rangle$ to yield

$$\left[\begin{array}{cc|c} 1 & -\frac{1}{2} & -\frac{1}{4} \\ 0 & 1 & \frac{1}{4} \\ 0 & 1 & 0 \end{array}\right].$$

We will then compute $\langle 2\rangle - \langle 3\rangle \to \langle 3\rangle$; we have

$$\left[\begin{array}{cc|c} 1 & -\frac{1}{2} & -\frac{1}{4} \\ 0 & 1 & \frac{1}{4} \\ 0 & 0 & \frac{1}{4} \end{array}\right].$$

There is no solution, so $\vec{v}$ is not a linear combination of $\vec{v}_1$ and $\vec{v}_2$.

Example 2.5.4: ☢☢ Is a Vector a Linear Combination of Others? 3

Let $\vec{v} = [14, -21, 7]$, $\vec{v}_1 = [2, -3, 1]$, and $\vec{v}_2 = [-4, 6, 2]$. Is $\vec{v}$ a linear combination of $\vec{v}_1$ and $\vec{v}_2$?

Consider the system

$$\left[\begin{array}{cc|c} 2 & -4 & 14 \\ -3 & 6 & -21 \\ 1 & 2 & 7 \end{array}\right].$$

We first perform the row operation $\frac{1}{2}\langle 1\rangle \to \langle 1\rangle$ to obtain

$$\left[\begin{array}{cc|c} 1 & -2 & 7 \\ -3 & 6 & -21 \\ 1 & 2 & 7 \end{array}\right].$$

Then, we have $3\langle 1\rangle + \langle 2\rangle \to \langle 2\rangle$, producing

$$\left[\begin{array}{cc|c} 1 & -2 & 7 \\ 0 & 0 & 0 \\ 1 & 2 & 7 \end{array}\right].$$

Next, we will carry out $\langle 2\rangle \leftrightarrow \langle 3\rangle$ to yield

$$\left[\begin{array}{cc|c} 1 & -2 & 7 \\ 1 & 2 & 7 \\ 0 & 0 & 0 \end{array}\right].$$

We will then compute $\langle 1\rangle + \langle 2\rangle \to \langle 2\rangle$; we have

$$\left[\begin{array}{cc|c} 1 & -2 & 7 \\ 2 & 0 & 14 \\ 0 & 0 & 0 \end{array}\right].$$

Then, we will execute $2\langle 1\rangle - \langle 2\rangle \to \langle 2\rangle$, and we obtain

$$\left[\begin{array}{cc|c} 1 & -2 & 7 \\ 0 & -4 & 0 \\ 0 & 0 & 0 \end{array}\right].$$

We then have, by $-\frac{1}{4}\langle 2\rangle \to \langle 2\rangle$,

$$\left[\begin{array}{cc|c} 1 & -2 & 7 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{array}\right].$$

Finally, we have the operation $\langle 1\rangle + 2\langle 2\rangle \to \langle 1\rangle$, which produces

$$\left[\begin{array}{cc|c} 1 & 0 & 7 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{array}\right].$$

Here, we see that $\vec{v} = 7\vec{v}_1$. We note that it would have been simple to conclude this based on the problem statement, but the method shown is the systematic algorithm for answering such questions.

Example 2.5.5: ☢☢ Is a Vector a Linear Combination of Others? 4

Let $\vec{v} = [14, -21, 7]$, $\vec{v}_1 = [2, -3, 1]$, and $\vec{v}_2 = [-4, 6, -2]$. Is $\vec{v}$ a linear combination of $\vec{v}_1$ and $\vec{v}_2$?

Consider the system

$$\left[\begin{array}{cc|c} 2 & -4 & 14 \\ -3 & 6 & -21 \\ 1 & -2 & 7 \end{array}\right].$$

By row reduction, we finally obtain

$$\left[\begin{array}{cc|c} 1 & -2 & 7 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{array}\right].$$

Here, the solution set is $\{(2c+7, c) : c \in \mathbb{R}\}$. There are thus infinitely many ways to express $\vec{v}$ as a linear combination of $\vec{v}_1$ and $\vec{v}_2$.

Consider the following definition.

Definition 2.5.3: 🛑 Row Space

Suppose $A \in \mathcal{M}_{mn}$. The row space of A is the subset of $\mathbb{R}^n$ consisting of the linear combinations of the rows of A.

Consider the following examples.

Example 2.5.6: ☢ Row Space 1

Consider

$$A = \begin{bmatrix} 1 & 2 \\ 5 & 10 \end{bmatrix}.$$

The row space of A is

$$\{c_1[1, 2] + c_2[5, 10] : c_1, c_2 \in \mathbb{R}\}.$$

In this case, the row space of A is a line. Generally, though, with two vectors, the row space will be a plane.

Example 2.5.7: ☢ Row Space 2

Consider

$$A = \begin{bmatrix} 1 & 3 \\ 5 & 10 \end{bmatrix}.$$

The row space of A is

$$\{c_1[1, 3] + c_2[5, 10] : c_1, c_2 \in \mathbb{R}\}.$$

In this case, the row space of A is a plane.

To determine if a vector is in the row space of a matrix A, we consider the system $[A^T|X]$. One may ask: why? Well, considering A instead of A^T would provide the wrong system of equations to solve. All we are doing when determining if a vector is in the row space of A is asking if the vector can be written as a linear combination of the rows of A. Consider the following example.

Example 2.5.8: ☢☢ Are Vectors in the Row Space?

Consider

$$A = \begin{bmatrix} 1 & 2 \\ 5 & 10 \end{bmatrix}$$

and recall that the row space of A is

$$\{c_1[1,2] + c_2[5,10] : c_1, c_2 \in \mathbb{R}\}.$$

Is $[3,6]$ in the row space of A? Is $[1,0]$ in the row space of A? We consider

$$\begin{bmatrix} 1 & 2 \\ 5 & 10 \end{bmatrix}^T \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}.$$

For $[3,6]$, we have

$$\begin{bmatrix} 1 & 5 & | & 3 \\ 2 & 10 & | & 6 \end{bmatrix},$$

which reduces to

$$\begin{bmatrix} 1 & 5 & | & 3 \\ 0 & 0 & | & 0 \end{bmatrix}.$$

The system has (infinitely many) solutions, so $[3,6]$ is in the row space of A. For $[1,0]$, we have

$$\begin{bmatrix} 1 & 5 & | & 1 \\ 2 & 10 & | & 0 \end{bmatrix},$$

which reduces to

$$\begin{bmatrix} 1 & 5 & | & 1 \\ 0 & 0 & | & -2 \end{bmatrix}.$$

The system has no solution, so $[1,0]$ is not in the row space of A.

Consider the following theorems.

Theorem 2.5.4: Transitivity of Linear Combinations

Suppose that $\vec{x}$ is a linear combination of $\vec{q}_1, \dots, \vec{q}_k$, and suppose also that each of $\vec{q}_1, \dots, \vec{q}_k$ is itself a linear combination of $\vec{r}_1, \dots, \vec{r}_\ell$. Then, $\vec{x}$ is a linear combination of $\vec{r}_1, \dots, \vec{r}_\ell$.

Proof. Because $\vec{x}$ is a linear combination of $\vec{q}_1, \dots, \vec{q}_k$,

$$\vec{x} = c_1\vec{q}_1 + \cdots + c_k\vec{q}_k$$

for $c_1, \dots, c_k \in \mathbb{R}$. Then, since each of $\vec{q}_1, \dots, \vec{q}_k$ can be written as a linear combination of $\vec{r}_1, \dots, \vec{r}_\ell$, there exist scalars $d_{11}, \dots, d_{k\ell}$ such that

$$\begin{aligned}
\vec{q}_1 &= d_{11}\vec{r}_1 + d_{12}\vec{r}_2 + \cdots + d_{1\ell}\vec{r}_\ell \\
\vec{q}_2 &= d_{21}\vec{r}_1 + d_{22}\vec{r}_2 + \cdots + d_{2\ell}\vec{r}_\ell \\
&\vdots \\
\vec{q}_k &= d_{k1}\vec{r}_1 + d_{k2}\vec{r}_2 + \cdots + d_{k\ell}\vec{r}_\ell
\end{aligned}$$

Then,

$$\begin{aligned}
\vec{x} &= c_1(d_{11}\vec{r}_1 + d_{12}\vec{r}_2 + \cdots + d_{1\ell}\vec{r}_\ell) \\
&\quad + c_2(d_{21}\vec{r}_1 + d_{22}\vec{r}_2 + \cdots + d_{2\ell}\vec{r}_\ell) \\
&\quad \vdots \\
&\quad + c_k(d_{k1}\vec{r}_1 + d_{k2}\vec{r}_2 + \cdots + d_{k\ell}\vec{r}_\ell) \\
&= (c_1d_{11} + c_2d_{21} + \cdots + c_kd_{k1})\vec{r}_1 \\
&\quad + (c_1d_{12} + c_2d_{22} + \cdots + c_kd_{k2})\vec{r}_2 \\
&\quad \vdots \\
&\quad + (c_1d_{1\ell} + c_2d_{1\ell} + \cdots + c_kd_{k\ell})\vec{r}_\ell.
\end{aligned}$$

We have just written $\vec{x}$ as a linear combination of $\vec{r}_1, \dots, \vec{r}_\ell$. □

Note that this theorem may be rephrased as follows: If $\vec{x}$ is in the row space of a matrix Q and each row of Q is in the row space of a matrix R, $\vec{x}$ is in the row space of R.

Theorem 2.5.5: Row Equivalence Implies Equal Row Space

Suppose A and B are row equivalent. Then, the row space of A is equal to the row space of B.

2.6 Lecture 11: September 16, 2022

2.6.1 Linear Maps

Consider the following definition.

Definition 2.6.1: Linear Maps

Given $A \in \mathcal{M}_{mn}$, we define

$$\begin{aligned} T_A &: \mathbb{R}^n \to \mathbb{R}^m \\ &: \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \mapsto A \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}. \end{aligned}$$

Consider the following example.

Example 2.6.1: Some Special Maps in $\mathbb{R}^2$

Consider the following maps, and name them.

1. If $A = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$, $T_A : \mathbb{R}^2 \to \mathbb{R}^2$ is the zero map.
2. If $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $T_A : \mathbb{R}^2 \to \mathbb{R}^2$ is the identity map.
3. If $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, $T_A : \mathbb{R}^2 \to \mathbb{R}^2$ is the projection onto the x axis.
4. If $A = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$, $T_A : \mathbb{R}^2 \to \mathbb{R}^2$ is the projection onto the y axis.

We will revisit linear maps in much greater detail in Chapter 5.

2.7 Lecture 12: September 19, 2022

2.7.1 Inverses of Matrices

Consider the following definitions and theorems.

Definition 2.7.1: Multiplicative Inverse of a Matrix

Let $A \in \mathcal{M}_{nn}$. Then, $B \in \mathcal{M}_{nn}$ is a multiplicative inverse of A if and only if

$$AB = BA = I_n.$$

Consider the following examples.

Example 2.7.1: The Inverse of the Identity

Let $A = I_n$. Find the inverse of A.

Proof. Since $I_n I_n = I_n I_n = I_n$, I_n is the inverse of A. □

Example 2.7.2: The Inverse of the Zero

Let $A = 0_n$. Show that A does not have an inverse.

Proof. For all $B \in \mathcal{M}_{nn}$, $AB = 0_n B = 0_n \neq I_n$. □

Theorem 2.7.1: Inverse Commutativity

Let $A, B \in \mathcal{M}_{nn}$. If either AB or BA equals I_n, the other product also equals I_n, and A and B are inverses of each other.

Definition 2.7.2: Singularity

A matrix is *singular* if and only if it is square and does not have an inverse. A matrix is *nonsingular* if and only if it is square and has an inverse.

Theorem 2.7.2: Uniqueness of the Inverse

If B and C are both inverses of $A \in \mathcal{M}_{nn}$, $B = C$.

Proof. $B = BI_n = B(AC) = (BA)C = I_n C = C$. □

We denote the unique inverse of A as A^{-1}. We can use the inverse to define negative integral powers of a nonsingular matrix A. Consider the following definition.

Definition 2.7.3: Negative Integral Powers of a Nonsingular Matrices

Let A be a nonsingular matrix. Then, the negative integral powers of A are given as follows: A^{-1} is the unique inverse of A. For $k \geq 2$, $A^{-k} = (A^{-1})^k$.

Theorem 2.7.3: Properties of Nonsingular Matrices

Let A and B be nonsingular $n \times n$ matrices. Then,

1. A^{-1} is nonsingular, and $(A^{-1})^{-1} = A$.

   Proof. We have that $A^{-1}A = AA^{-1} = I_n$ since A^{-1} is the inverse of A, hence A^{-1} is nonsingular and $(A^{-1})^{-1} = A$. □

2. A^k is nonsingular, and $(A^k)^{-1} = (A^{-1})^k = A^{-k}$, for $k \in \mathbb{Z}$.

   Proof. We have that $A^k A^{-k} = A^{-k+k} = A^0 = I_n$. Hence, A^k is nonsingular and $(A^k)^{-1} = A^{-k}$. □

3. AB is nonsingular, and $(AB)^{-1} = B^{-1}A^{-1}$.

   Proof. We have that $(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AI_nA^{-1} = AA^{-1} = I_n$. Hence, AB is nonsingular and $(AB)^{-1} = B^{-1}A^{-1}$. □

4. A^T is nonsingular, and $(A^T)^{-1} = (A^{-1})^T$.

   Proof. We have that $(A^T)(A^{-1})^T = (A^{-1}A)^T = I_n^T = I_n$. □

Now, we will fully provide statements of matrix exponent laws.

Theorem 2.7.4: Matrix Exponent Laws

If A is nonsingular and $p, q \in \mathbb{Z}$,

1. $A^{p+q} = (A^p)(A^q)$.
2. $(A^p)^q = A^{pq} = (A^q)^p$.

Consider the following theorem.

Theorem 2.7.5: 2×2 **Inverse**

Suppose $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. Then, A is nonsingular if and only if $ad - bc \neq 0$. In this case,

$$A = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

Proof. Suppose $ad - bc \neq 0$. We consider

$$\frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Thus, A^{-1} exists and the formula holds. Now, suppose that $ad - bc = 0$. We wish to show that A is singular. Consider

$$\begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} ad - bc & 0 \\ 0 & ad - bc \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Suppose A has an inverse. Then,

$$\begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \left(AA^{-1}\right) = \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

But,

$$\left(\begin{bmatrix} d & -b \\ -c & a \end{bmatrix} A\right) A^{-1} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} A^{-1} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

This implies $a = b = c = d = 0$, meaning $A = 0_2$. From Example 2.7.2, we know that the matrix 0_n is singular. $\square$

2.8 Lecture 13: September 21, 2022

2.8.1 Finding Inverses of Matrices

We have provided a lot of theory about the inverse of a matrix, but given $A \in \mathcal{M}_{nn}$, how do we determine if A is invertible. That is, how do we determine if A is nonsingular? If it is, how do we find A^{-1}? Consider the matrix equation

$$A \begin{bmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & \ddots & \vdots \\ x_{n1} & \cdots & x_{nn} \end{bmatrix} = \begin{bmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{bmatrix} = I_n.$$

Our goal is to solve for the elements x_{ij}. We thus have n linear systems to solve. That is,

$$A \begin{bmatrix} x_{11} \\ \vdots \\ x_{n1} \end{bmatrix} = \begin{bmatrix} 1 \\ \vdots \\ 0 \end{bmatrix} \quad , \dots , \quad A \begin{bmatrix} x_{1n} \\ \vdots \\ x_{nn} \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 1 \end{bmatrix}.$$

To solve equations of the sort, consider the following steps.

1. Form $[A|I_n]$.
2. Row reduce $[A|I_n]$ to $[C|D]$ until we have reduced row echelon form.
3. If $C = I_n$, A^{-1} exists and $A^{-1} = D$. Otherwise, A^{-1} does not exist.

Consider the following example.

Example 2.8.1: ☢☢ Finding a 2×2 Inverse

Consider the matrix
$$\begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}.$$
First we form
$$\begin{bmatrix} 2 & 1 & | & 1 & 0 \\ 1 & 1 & | & 0 & 1 \end{bmatrix}.$$
We perform the row operation $\langle 2 \rangle \leftrightarrow \langle 1 \rangle$, yielding
$$\begin{bmatrix} 1 & 1 & | & 0 & 1 \\ 2 & 1 & | & 1 & 0 \end{bmatrix}.$$
Then, we execute $2\langle 1 \rangle - \langle 2 \rangle \rightarrow \langle 2 \rangle$, which yields
$$\begin{bmatrix} 1 & 1 & | & 0 & 1 \\ 0 & 1 & | & -1 & 2 \end{bmatrix}.$$
We then compute $\langle 2 \rangle - \langle 1 \rangle \rightarrow \langle 1 \rangle$, producing
$$\begin{bmatrix} -1 & 0 & | & -1 & 1 \\ 0 & 1 & | & -1 & 2 \end{bmatrix}.$$
Finally, we simply perform $-\langle 1 \rangle \rightarrow \langle 1 \rangle$, which obtains
$$\begin{bmatrix} 1 & 0 & | & 1 & -1 \\ 0 & 1 & | & -1 & 2 \end{bmatrix}.$$
Therefore,
$$\begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix}$$
is the inverse of
$$\begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}.$$
Note that here, we did not apply the formula in Theorem 2.7.5, but instead, the general algorithm.

Example 2.8.2: Finding a 3×3 Inverse

Consider the matrix

$$\begin{bmatrix} 2 & -6 & 5 \\ -4 & 12 & -9 \\ 2 & -9 & 8 \end{bmatrix}.$$

First, we form

$$\left[\begin{array}{ccc|ccc} 2 & -6 & 5 & 1 & 0 & 0 \\ -4 & 12 & -9 & 0 & 1 & 0 \\ 2 & -9 & 8 & 0 & 0 & 1 \end{array}\right].$$

After row reduction, we have

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{5}{2} & \frac{1}{2} & -1 \\ 0 & 1 & 0 & \frac{7}{3} & 1 & -\frac{1}{3} \\ 0 & 0 & 1 & 2 & 1 & 0 \end{array}\right].$$

Thus,

$$\begin{bmatrix} 2 & -6 & 5 \\ -4 & 12 & -9 \\ 2 & -9 & 8 \end{bmatrix} \begin{bmatrix} \frac{5}{2} & \frac{1}{2} & -1 \\ \frac{7}{3} & 1 & -\frac{1}{3} \\ 2 & 1 & 0 \end{bmatrix} = I_3.$$

Finally, we present an important theorem about the existence and uniqueness of solutions to $AX = B$.

Theorem 2.8.1: Uniqueness of Solutions to Linear Systems

Let $A \in \mathcal{M}_{nn}$ and $AX = B$ be a linear system. If A is nonsingular, the system has a unique solution. If A is singular, the system either has no solution or infinitely many solutions. That is, $AX = B$ has a unique solution if and only if A is nonsingular.

Proof. Consider the case where A is nonsingular; by definition, A^{-1} exists. Consider $X = A^{-1}B$ as a prospective solution. To verify it, we have

$$A(A^{-1}B) = (AA^{-1})B = I_nB = B.$$

Thus, $X = A^{-1}B$ is a valid solution. Now, for uniqueness, suppose that $X = Y$ is a solution. That is, suppose $AY = B$. We multiply both sides by A^{-1}, on the left, to obtain

$$A^{-1}(AY) = A^{-1}B \implies (A^{-1}A)Y = A^{-1}B \implies I_nY = A^{-1}B \implies Y = A^{-1}B,$$

showing that $X = A^{-1}B$ is a unique solution. Now, if A is singular, $\operatorname{rank} A < n$, meaning we will have no solution or infinitely many solutions. This is because we will either have a row of all zeroes with a nonzero entry after the augmentation bar yielding an empty solution set, or if the system has at least one solution, there will be at least one free variable, guaranteeing infinitely many solutions. □

2.9 Additional Topics: The LU Factorization

2.9.1 Presented by Olver and Shakiban's Applied Linear Algebra

Consider the following definition.

Definition 2.9.1: Elementary Matrices

The elementary matrix associated with a row operation on a matrix with m rows is the $m \times m$ matrix obtained by applying the row operation to I_m.

Consider the following examples.

Example 2.9.1: Finding an Elementary Matrix

Consider the row operation $-2\langle 1\rangle + \langle 2\rangle \to \langle 2\rangle$ applied to a matrix with 3 rows. Find the elementary matrix, denoted E_1.

We take

$$I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

and perform the row operation to obtain

$$E_1 = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

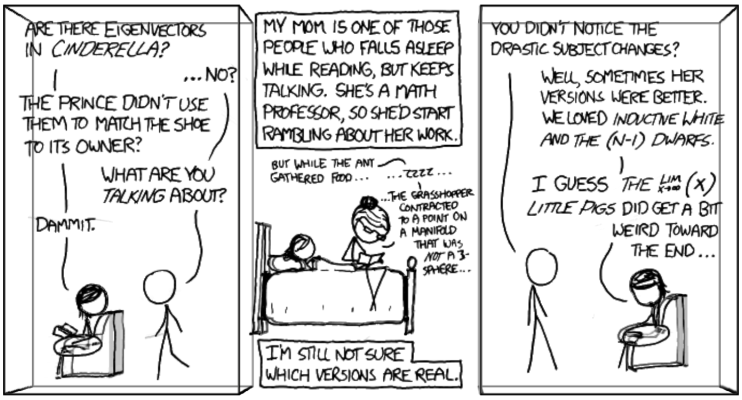
Consider the following theorem.

Theorem 2.9.1: Elementary Matrices are Equivalent to Row Operations

Suppose $A \in \mathcal{M}_{mn}$ and $E_1, \dots, E_k$ are elementary matrices corresponding to row operations $R_1, \dots, R_k$. The multiplication

$$E_k(E_{k-1}(\dots(E_2(E_1))))A$$

produces the same matrix that applying the row operations, in order.



3

Determinants and Eigenvalues

3.1 Lecture 13: September 21, 2022

3.1.1 Defining the Determinant

Consider the following theorems and definitions.

Theorem 3.1.1: The Determinant Determines the Area in $\mathbb{R}^2$

Consider $\vec{x} = [x_1, x_2]$ and $\vec{y} = [y_1, y_2]$. If we form

$$A = \begin{bmatrix} x_1 & x_2 \\ y_1 & y_2 \end{bmatrix},$$

$$|\det A| = ||\vec{x} \times \vec{y}||.$$

That is, $|\det A|$ provides the area of the parallelogram determined by $\vec{x}$ and $\vec{y}$.

Theorem 3.1.2: The Determinant Determines the Volume in $\mathbb{R}^3$

Consider $\vec{x} = [x_1, x_2, x_3]$, $\vec{y} = [y_1, y_2, y_3]$, and $\vec{z} = [z_1, z_2, z_3]$. If we form

$$A = \begin{bmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{bmatrix},$$

$$|\det A| = \vec{x} \cdot (\vec{y} \times \vec{z}).$$

That is, $|\det A|$ provides the volume of the parallelepiped determined by $\vec{x}$, $\vec{y}$, and $\vec{z}$.

Definition 3.1.1: The (i, j) Submatrix

Suppose $A \in \mathcal{M}_{nn}$. The (i, j) submatrix of A is the $(n-1) \times (n-1)$ matrix obtained by removing the ith row and the jth column. We denote this by $A_{(i,j)}$

Definition 3.1.2: The (i, j) Minor

Suppose $A \in \mathcal{M}_{nn}$. The (i, j) minor of A is the determinant of the (i, j) submatrix of A.

Definition 3.1.3: The (i, j) Cofactor

Suppose $A \in \mathcal{M}_{nn}$. The (i, j) cofactor of A is

$$A_{ij} = (-1)^{i+j} \det(A_{(i,j)}).$$

Definition 3.1.4: The Determinant

Suppose $A \in \mathcal{M}_{nn}$. Then,

1. If $n = 1$, and $A = \begin{bmatrix} a_{11} \end{bmatrix}$, $\det A = a_{11}$.
2. If $n = 2$, and $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$, $\det A = a_{11}a_{22} - a_{12}a_{21}$.
3. If $n > 2$, and $A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}$, $\det A = (a_{11}A_{11} + \cdots + a_{1n}A_{1n}) + \cdots + (a_{n1}A_{n1} + \cdots + a_{nn}A_{nn})$.

For fun, consider the following Python 3 implementation of computing the determinant of any $n \times n$ matrix.

```
# Delete nth Column for Cofactor
def delete_nth_column_helper(matrix, n):

    return [m[:n]+m[n+1:] for m in matrix[1:]]

# Calculate Matrix Determinant
def determinant(matrix):

    # Base Cases
    if len(matrix) == 1: return matrix[0][0]
    if len(matrix) == 2: return matrix[0][0]*matrix[1][1]-matrix[0][1]*matrix[1][0]

    # Calculate Determinant Parts
    for _ in range(len(matrix[0])):

        determinant_components = [(-1)**i*matrix[0][i]*determinant(delete_nth_column_helper(matrix, i)) for i in range(len(matrix[0]))]

    # Return Sum
    return sum(determinant_components)

if __name__ == "__main__":

    matrix = []

    # Get Matrix
    for _ in range(int(input())): matrix.append([int(i) for i in input().split()])

    # Output
    print(determinant(matrix))
```

3.2 Lecture 14: September 23, 2022

3.2.1 Determinants of Upper Triangular Matrices

Consider the following theorems.

Theorem 3.2.1: The Determinant of an Upper Triangular Matrix

If $A \in \mathcal{M}_{nn}$ is upper triangular,

$$\det A = a_{11}a_{22}\cdots a_{nn}.$$

Recall that A is upper triangular if and only if all elements below the main diagonal are zero.

Proof. We proceed by induction on n. For $n = 1$, $A = \begin{bmatrix} a_{11} \end{bmatrix}$, so $\det A = a_{11}$. Suppose for all $k \in \mathbb{N}$, and some upper triangular $A \in \mathcal{M}_{kk}$,

$$a_{11}a_{22}\cdots a_{kk}.$$

Consider

$$B = \begin{bmatrix} b_{11} & \cdots & b_{1(k+1)} \\ \vdots & \ddots & \vdots \\ 0 & \cdots & b_{(k+1)(k+1)} \end{bmatrix}.$$

Then, we compute $\det B$ using the last row as our "first row."

$$\begin{aligned} \det B &= (-1)^{1+1} b_{11} \det \begin{bmatrix} b_{22} & b_{23} & \cdots & a_{2(k+1)} \\ 0 & b_{33} & \cdots & b_{3(k+1)} \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & b_{(k+1)(k+1)} \end{bmatrix} \\ &= b_{11} \underbrace{b_{22}\cdots b_{(k+1)(k+1)}}_{\text{By the inductive hypothesis.}}. \end{aligned}$$

This is precisely the stipulation of the theorem when $n = k + 1$. $\square$

Theorem 3.2.2: Determinants and Row Operations

Suppose $A \in \mathcal{M}_{nn}$ and let R be a row operation. Then,

1. If R is $c\langle i\rangle \to \langle i\rangle$ for some $c \in \mathbb{R}$,
$$\det R(A) = c \det A.$$
Note that if $c = 0$, R is not a valid row operation.
2. If R is $c\langle i\rangle + \langle j\rangle \to \langle j\rangle$ for some $c \in \mathbb{R}$,
$$\det R(A) = \det A.$$
Note that if $c = 0$, R is not a valid row operation.
3. If R is $\langle i\rangle \leftrightarrow \langle j\rangle$,
$$\det R(A) = -\det A.$$

Note that if $\det A \neq 0$ and R is a row operation, $\det R(A) \neq 0$.

We can use Theorem 3.2.2 in conjunction with row operations to compute the determinant of a matrix easily. We simply use row operations to create an upper triangular matrix, while keeping track of how the determinant changes. We then apply Theorem 3.2.1. Consider the following examples.

Example 3.2.1: Computing a Determinant by Row Reduction

Compute $\det \begin{bmatrix} 1 & 1 & 1 \\ 2 & 3 & -2 \\ 4 & 9 & 4 \end{bmatrix}$.

Let
$$A_0 = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 3 & -2 \\ 4 & 9 & 4 \end{bmatrix}.$$

Consider the following table.

Row Operation	Resultant Matrix	Effect on the Determinant
$-2\langle 1\rangle + \langle 2\rangle \to \langle 2\rangle$	$A_1 = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & -4 \\ 4 & 9 & 4 \end{bmatrix}$	$\det A_1 = \det A_0$
$-4\langle 1\rangle + \langle 3\rangle \to \langle 3\rangle$	$A_2 = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & -4 \\ 0 & 5 & 0 \end{bmatrix}$	$\det A_2 = \det A_1$
$-5\langle 2\rangle + \langle 3\rangle \to \langle 3\rangle$	$A_3 = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & -4 \\ 0 & 0 & 20 \end{bmatrix}$	$\det A_3 = \det A_2$

Thus, $\det A_0 = 20$.

Example 3.2.2: Computing a Determinant by Row Reduction

Compute $\det \begin{bmatrix} 0 & -14 & -8 \\ 1 & 3 & 2 \\ -2 & 0 & 6 \end{bmatrix}$.

Let

$$A_0 = \begin{bmatrix} 0 & -14 & -8 \\ 1 & 3 & 2 \\ -2 & 0 & 6 \end{bmatrix}.$$

Consider the following table.

Row Operation	Resultant Matrix	Effect on the Determinant
$\langle 2\rangle \leftrightarrow \langle 1\rangle$	$A_1 = \begin{bmatrix} 1 & 3 & 2 \\ 0 & -14 & -8 \\ -2 & 0 & 6 \end{bmatrix}$	$\det A_1 = -\det A_0$
$2\langle 1\rangle + \langle 3\rangle \to \langle 3\rangle$	$A_2 = \begin{bmatrix} 1 & 3 & 2 \\ 0 & -14 & -8 \\ 0 & 6 & 10 \end{bmatrix}$	$\det A_2 = \det A_1$
$-\frac{1}{14}\langle 2\rangle \to \langle 2\rangle$	$A_3 = \begin{bmatrix} 1 & 3 & 2 \\ 0 & 1 & \frac{4}{7} \\ 0 & 6 & 10 \end{bmatrix}$	$\det A_3 = -\frac{1}{14}\det A_2$
$-6\langle 2\rangle + \langle 3\rangle \to \langle 3\rangle$	$A_3 = \begin{bmatrix} 1 & 3 & 2 \\ 0 & 1 & \frac{4}{7} \\ 0 & 0 & \frac{46}{7} \end{bmatrix}$	$\det A_4 = \det A_3$

Thus, $\det A_0 = \frac{46}{7}(-14)(-1) = 92$.

Consider the following theorem.

Theorem 3.2.3: Inverses and Determinants

Suppose that $A \in \mathcal{M}_{nn}$. Then, A is nonsingular if and only if $\det A \neq 0$.

Proof. If A is nonsingular, we can row reduce A to produce I_n. Since $\det I_n \neq 0$, by Theorem 3.2.2, $\det A \neq 0$. If $\det A \neq 0$, we form the system $[A|B]$ and reduce it to $[C|D]$. We know that $\det C \neq 0$. Because C is in reduced row echelon form, and square since A is square, it is upper triangular. That means all main diagonal entries are nonzero, meaning they are all 1. Meaning $C = I_n$. This means we were able to row reduce A to I_n, so A is nonsingular. □

Consider the following table summarizing various results. Statements in each column are equivalent.

$A \in \mathcal{M}_{nn}$ is Nonsingular	$A \in \mathcal{M}_{nn}$ is Singular
$\operatorname{rank} A = n$ $\det A \neq 0$ A is row equivalent to I_n. $AX = 0$ has only the trivial solution for X. $AX = B$ has a unique solution for X, and $X = A^{-1}B$.	$\operatorname{rank} A < n$ $\det A = 0$ A is not row equivalent to I_n. $AX = 0$ has a nontrivial solution for X. $AX = B$ does not have a unique solution.

3.3 Lecture 15: September 26, 2022

3.3.1 Further Properties of Determinants

Theorem 3.3.1: Properties of Determinants

Suppose $A, B \in \mathcal{M}_{nn}$. Then,

1. $\det(AB) = \det A \det B$.

   Proof. If A or B is singular, $\det A \det B = 0$. For now, let B be singular. We don't make any assumptions about A, for now. Since B is singular, there exists some $X \neq 0$ such that $BX = 0$. We consider $(AB)X = A(BX) = A(0) = 0$. Thus, X is a nontrivial solution to the homogeneous equation associated with AB. Thus AB is singular and $\det(AB) = 0 = \det A \det B$. Now, suppose that A is singular and B is nonsingular. There exists a nontrivial solution for Y in the system $AY = 0$. Since B is nonsingular, B^{-1} exists. We define $X = B^{-1}Y$ where $X \neq 0$. Then $ABX = AB(B^{-1}Y) = AY = 0$. Thus, ABX and $X \neq 0$ implies that AB is singular, so $\det(AB) = 0$. Now, we consider the case where A and B are both nonsingular. There exists row operations $R_1, \dots, R_k$ such that $A = R_1(\dots(R_k(I_n)\dots))$ and

   $$\begin{aligned}\det(AB) &= \det(R_1(\dots(R_k(I_n)\dots))B)\\ &= c_1 \cdots c_k \det(I_n B)\\ &= c_1 \cdots c_k \det B\\ &= c_1 \cdots c_k \det I_n \det B\\ &= \det(R_1(\dots(R_k(I_n)\dots))) \det B\\ &= \det A \det B,\end{aligned}$$

   hence proving the proposition. □

2. $\det(A^T) = \det A$.

   Proof. Suppose that A is singular, meaning $\det A = 0$. We seek to show that then, $\det(A^T) = 0$. Suppose, for the sake of contradiction, $\det(A^T) \neq 0$, meaning A^T is nonsingular. Then, $(A^T)^T = A$ is also nonsingular, which contradicts our assumption. Suppose that A is nonsingular, meaning A is row equivalent to I_n. That means

   $$\det A = \det(R_k(\dots R_2(R_1(I_n))\dots)) = \det((R_k(\dots R_2(R_1(I_n))\dots))^T) = \det(A^T),$$

   proving the proposition. □

3. $\underbrace{\det(A^{-1}) = \dfrac{1}{\det A}}_{\text{Suppose that } A \text{ is nonsingular.}}$.

   Proof. We know that A is nonsingular. We have $\det I_n = \det(AA^{-1}) = \det A \det(A^{-1}) = 1$. By simple algebra, $\det(A^{-1}) = \frac{1}{\det A}$. □

3.3.2 Eigenvectors, Eigenvalues, and Diagonalization

Consider the following definitions and theorems.

Definition 3.3.1: Similarity

Suppose $A, B \in \mathcal{M}_{nn}$. The matrix B is similar to a matrix A if and only if there exists some nonsingular matrix P such that

$$B = P^{-1}AP.$$

Definition 3.3.2: Diagonalizability

The matrix $A \in \mathcal{M}_{nn}$ is diagonalizable if and only if a diagonal matrix D is similar to A. That is, A is diagonalizable if and only if, for nonsingular matrix P,

$$D = P^{-1}AP.$$

Definition 3.3.3: Eigenvalues and Eigenvectors

For $A \in \mathcal{M}_{nn}$, $\lambda \in \mathbb{R}$ is an eigenvalue of A if and only if there exists $X \neq \vec{0}$ such that

$$AX = \lambda X.$$

If λ is an eigenvalue of A, X is an eigenvector of A with eigenvalue λ.

3.4 Lecture 16: September 28, 2022

3.4.1 The Process of Diagonalization: Part I

Consider the following definition.

Definition 3.4.1: Eigenspace

Given $A \in \mathcal{M}_{nn}$, the eigenspace of a given eigenvalue λ is

$$E_\lambda = \{X : AX = \lambda X\} \cup \{\vec{0}\}.$$

Note that "$\cup\{\vec{0}\}$" is somewhat redundant, as it will always satisfy the equation. However, the zero vector is never an eigenvector.

Our goal is to find all the eigenvalues and eigenspaces of A. Consider the following theorem.

Theorem 3.4.1: Finding Eigenvectors and Eigenvalues

Consider a matrix $A \in \mathcal{M}_{nn}$. By definition, X is an eigenvector of A with eigenvalue $\lambda \in \mathbb{R}$ when

$$AX = \lambda X = \lambda I_n X.$$

That is, when

$$(A - \lambda I_n)X = \vec{0}.$$

The above equation has a nontrivial solution for X if and only if $(A - \lambda I_n)$ is singular, that is, when

$$\det(A - \lambda I_n) = 0.$$

Therefore, the scalar λ is an eigenvalue of A if and only if λ satisfies the above equation.

Note that, for now, we will only consider real eigenvalues; however, complex eigenvalues are incredibly useful and have numerous applications. Consider the following example.

Example 3.4.1: ☢ Eigenspaces of A Diagonal Matrix

Consider $A = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & 7 \end{bmatrix}$. Find the eigenvalues and eigenspaces of A.

We see that the eigenvalues are $\lambda_1 = 5$ and $\lambda_2 = 7$. The eigenspaces are

$$E_{\lambda_1} = \left\{ c \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} : c \in \mathbb{R} \right\}$$

and

$$E_{\lambda_2} = \left\{ c_1 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + c_2 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} : c_1, c_2 \in \mathbb{R} \right\}.$$

We note that for a diagonal matrix, we can simply read off the eigenvalues and eigenspaces. Generally though, we use Theorem 5.5.5, as seen in the following example.

Example 3.4.2: ☢☢ Eigenspaces of a 2×2 Matrix

Consider $A = \begin{bmatrix} 7 & 1 \\ -3 & 3 \end{bmatrix}$. Find the eigenvalues and eigenspaces of A.

To find the eigenvalues, we consider

$$\begin{aligned} \det(A - \lambda I_2) &= \det \begin{bmatrix} 7-\lambda & 1 \\ -3 & 3-\lambda \end{bmatrix} \\ &= (7-\lambda)(3-\lambda) + 3 \\ &= 21 - 10\lambda + \lambda^2 + 3 \\ &= \lambda^2 - 10\lambda + 24 \\ &= 0. \end{aligned}$$

By factoring, we have $(\lambda - 6)(\lambda - 4) = 0$, meaning $\lambda_1 = 4$ and $\lambda_2 = 6$. Now, we seek to find the eigenspace. We substitute in λ_1 and λ_2 into $[(A - \lambda I_2)|\vec{0}]$ and solve for λ. For λ_1, we have

$$\left[\begin{array}{cc|c} 7-4 & 1 & 0 \\ -3 & 3-4 & 0 \end{array}\right] = \left[\begin{array}{cc|c} 3 & 1 & 0 \\ -3 & -1 & 0 \end{array}\right].$$

By row reduction, we have

$$\left[\begin{array}{cc|c} 1 & \frac{1}{3} & 0 \\ 0 & 0 & 0 \end{array}\right]$$

and

$$E_{\lambda_1} = \left\{ \begin{bmatrix} -\frac{1}{3}c \\ c \end{bmatrix} : c \in \mathbb{R} \right\}.$$

By a similar process, we have

$$E_{\lambda_2} = \left\{ \begin{bmatrix} -c \\ c \end{bmatrix} : c \in \mathbb{R} \right\}.$$

Now, we revisit Definition 3.3.2. How do we find D and P? We take D to be the diagonal matrix with all nonzero elements being the eigenvalues of A. Then, we take P to be the matrix with each column being the eigenvector associated with the eigenvalue in the corresponding column of D. Finally, we check that P^{-1} exists. Why does this work? In the general 2×2 case, we have

$$D = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$$

and $P = \begin{bmatrix} \vec{v}_1 & \vec{v}_2 \end{bmatrix}$ where $\vec{v}_1$ and $\vec{v}_2$ are eigenvectors of A, associated with eigenvalues λ_1 and λ_2, respectively. We have checked that P^{-1} exists and has the form $P^{-1} = \begin{bmatrix} \vec{w}_1 \\ \vec{w}_2 \end{bmatrix}$. Then,

$$P^{-1}P = \begin{bmatrix} \vec{w}_1 \cdot \vec{v}_1 & \vec{w}_1 \cdot \vec{v}_2 \\ \vec{w}_2 \cdot \vec{v}_1 & \vec{w}_2 \cdot \vec{v}_2 \end{bmatrix} = I_2.$$

Then, we have

$$\begin{aligned} P^{-1}AP &= P^{-1}(AP) \\ &= P^{-1} \begin{bmatrix} A\vec{v}_1 & A\vec{v}_2 \end{bmatrix} \\ &= P^{-1} \begin{bmatrix} \lambda_1 \vec{v}_1 & \lambda_2 \vec{v}_2 \end{bmatrix} \\ &= \begin{bmatrix} \vec{w}_1 \\ \vec{w}_2 \end{bmatrix} \begin{bmatrix} \lambda_1 \vec{v}_1 & \lambda_2 \vec{v}_2 \end{bmatrix} \\ &= \begin{bmatrix} \lambda_1 \vec{w}_1 \cdot \vec{v}_1 & \lambda_2 \vec{w}_1 \cdot \vec{v}_2 \\ \lambda_1 \vec{w}_2 \cdot \vec{v}_1 & \lambda_2 \vec{w}_2 \cdot \vec{v}_2 \end{bmatrix} \\ &= \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \\ &= D. \end{aligned}$$

3.5 Lecture 17: September 30, 2022

3.5.1 The Process of Diagonalization: Part II

We summarize the process of diagonalization in a more general sense.

Theorem 3.5.1: The Process of Diagonalization

Let $A \in \mathcal{M}_{nn}$. Consider the following steps.

1. Find the solutions of $\det(A - \lambda I_n) = 0$. The solutions $\lambda_1, \ldots, \lambda_k$ are the eigenvalues of A.
2. For each eigenvalue λ_m, solve the system $[A - \lambda_m I_n | 0]$ by row reduction.
3. If there are less than n fundamental eigenvectors, A cannot be diagonalized.
4. Form $P = \begin{bmatrix} \vec{v}_1 & \cdots & \vec{v}_n \end{bmatrix}$. Note that P is nonsingular.
5. Verify that $D = P^{-1}AP$ is a diagonal matrix where each entry d_{ii} is the eigenvalue for the fundamental eigenvector forming the ith column of P.

Consider the following examples.

Example 3.5.1: Diagonalization 1

Given

$$A = \begin{bmatrix} 9 & 5 \\ -25 & -21 \end{bmatrix},$$

construct the diagonal matrix D and form the nonsingular matrix P such that

$$D = P^{-1}AP.$$

To find the eigenvalues, we have

$$\begin{aligned} 0 &= \det \begin{bmatrix} 9-\lambda & 5 \\ -25 & -21-\lambda \end{bmatrix} \\ &= (9-\lambda)(-21-\lambda)+125 \\ &= (\lambda+16)(\lambda-4). \end{aligned}$$

Therefore, our eigenvalues are $\lambda_1 = -16$ and $\lambda_2 = 4$. To find the associated eigenvectors, for λ_1, we have

$$\left[\begin{array}{cc|c} 25 & 5 & 0 \\ -25 & -5 & 0 \end{array}\right],$$

which, by $\langle 1 \rangle \to \langle 2 \rangle \to \langle 2 \rangle$, becomes

$$\left[\begin{array}{cc|c} 25 & 5 & 0 \\ 0 & 0 & 0 \end{array}\right].$$

We have the eigenspace $E_{-16} = \left\{ c \begin{bmatrix} -1 \\ 5 \end{bmatrix} : c \in \mathbb{R} \right\}$. For λ_2, we have

$$\left[\begin{array}{cc|c} 5 & 5 & 0 \\ -25 & -25 & 0 \end{array}\right]$$

which can be row reduced to

$$\left[\begin{array}{cc|c} -25 & -25 & 0 \\ 0 & 0 & 0 \end{array}\right].$$

This produces the eigenspace $E_4 = \left\{ c \begin{bmatrix} -1 \\ 1 \end{bmatrix} : c \in \mathbb{R} \right\}$. We form

$$D = \begin{bmatrix} -16 & 0 \\ 0 & 4 \end{bmatrix}, \quad P = \begin{bmatrix} -1 & -1 \\ 5 & 1 \end{bmatrix}, \quad P^{-1} = \begin{bmatrix} \frac{1}{4} & \frac{1}{4} \\ -\frac{5}{4} & -\frac{1}{4} \end{bmatrix}.$$

To check our work, we have

$$\begin{aligned} P^{-1}AP &= \begin{bmatrix} \frac{1}{4} & \frac{1}{4} \\ -\frac{5}{4} & -\frac{1}{4} \end{bmatrix} \begin{bmatrix} 9 & 5 \\ -25 & 21 \end{bmatrix} \begin{bmatrix} -1 & -1 \\ 5 & 1 \end{bmatrix} \\ &= \begin{bmatrix} -16 & 0 \\ 0 & 4 \end{bmatrix} \\ &= D, \end{aligned}$$

thus verifying our answer.

Example 3.5.2: ☢☢ Diagonalization 2

Given

$$A = \begin{bmatrix} 0 & -6 & 0 \\ 3 & 9 & 0 \\ 0 & 0 & 6 \end{bmatrix},$$

construct the diagonal matrix D and form the nonsingular matrix P such that

$$D = P^{-1}AP.$$

To find the eigenvalues, we have

$$\begin{aligned} 0 &= \det \begin{bmatrix} -\lambda & -6 & 0 \\ 3 & 9-\lambda & 0 \\ 0 & 0 & 6-\lambda \end{bmatrix} \\ &= -\lambda((9-\lambda)(6-\lambda)) + 6(3(6-\lambda)) \\ &= -(\lambda-3)(\lambda-6)^2. \end{aligned}$$

Therefore, our eigenvalues are $\lambda_1 = 3$ and $\lambda_2 = 6$. For λ_1, we have the linear system

$$\left[\begin{array}{ccc|c} -3 & -6 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 \end{array}\right]$$

which can be row reduced to

$$\left[\begin{array}{ccc|c} 1 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

This gives us $E_3 = \left\{ c \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} : c \in \mathbb{R} \right\}$. For λ_2, we have

$$\left[\begin{array}{ccc|c} -6 & -6 & 0 & 0 \\ 3 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right],$$

which can be row reduced to

$$\left[\begin{array}{ccc|c} 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right],$$

giving

$$\begin{aligned} E_6 &= \left\{ \begin{bmatrix} -c_1 \\ c_1 \\ c_2 \end{bmatrix} : c_1, c_2 \in \mathbb{R} \right\} \\ &= \left\{ c_1 \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + c_2 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} : c_1, c_2 \in \mathbb{R} \right\} \end{aligned}$$

We can then form

$$D = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 6 \end{bmatrix}, \quad P = \begin{bmatrix} -2 & -1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad P^{-1} = \begin{bmatrix} -1 & -1 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

3.6 Lecture 18: October 3, 2022

3.6.1 The Process of Diagonalization: Part III

We start with a question; what can go wrong in the diagonalization process? We will trace each step of Theorem 3.5.1.

What Could go Wrong?	Example
There exist complex roots to $\det(A-\lambda I_n)=0$.	$A=\begin{bmatrix}0 & -1\\ 1 & 0\end{bmatrix}$
There are less than n fundamental eigenvectors.	$A=\begin{bmatrix}7 & 1\\ 0 & 7\end{bmatrix}$

We remark that if there exist complex roots, we can resolve it by considering complex eigenvalues in the diagonalization process, as seen in Example 3.8.1; however, we cannot resolve the issue of having less than n fundamental eigenvectors. Naturally, the question of what constitutes a fundamental eigenvector is brought up. Consider the following definitions and theorems.

Definition 3.6.1: Linear Independence

Suppose $\vec{v}_1,\dots,\vec{v}_k$ are vectors in $\mathbb{R}^n$. The set $\{\vec{v}_1,\dots,\vec{v}_k\}$ is linearly independent if and only if all scalars $c_1,\dots,c_k$ that form

$$c_1\vec{v}_1+\cdots+c_k\vec{v}_k=\vec{0}$$

are zero.

Theorem 3.6.1: Diagonalizability

The matrix $A\in\mathcal{M}_{nn}$ is diagonalizable if and only if there exists a set $S=\{\vec{v}_1,\dots,\vec{v}_n\}$ such that

1. Each $\vec{v}_i$ is an eigenvector of A.
2. The set S is linearly independent.

The statement "There are less than n fundamental eigenvectors" means that the set $S=\{\vec{v}_1,\dots,\vec{v}_n\}$ is not linearly independent, or equivalently, linearly dependent. To conclude if we will have a sufficient number of fundamental eigenvectors, we solve the linear homogeneous system

$$\left[\begin{array}{ccc|c}\vec{v}_1 & \cdots & \vec{v}_n & \vec{0}\end{array}\right]$$

for $c_1,\dots,c_n$. Consider the following theorem.

Theorem 3.6.2: Diagonalizability and Rank

Suppose $A\in\mathcal{M}_{nn}$ with $\det(A-\lambda I_n)=(x-\lambda_1)^{n_1}\cdots(x-\lambda_\ell)^{n_\ell}$ for $\lambda_1,\dots,\lambda_\ell\in\mathbb{R}$. Then, there exists P$=\begin{bmatrix}\vec{v}_1 & \cdots & \vec{v}_n\end{bmatrix}$ where P is nonsingular and each $\vec{v}_i$ is an eigenvector if and only if, for each λ_i,

$$\operatorname{rank}(A-\lambda_i I_n)=n-n_i.$$

Consider the following example.

Example 3.6.1: Is A Diagonalizable?

Suppose $A \in \mathcal{M}_{nn}$ and $\det(A - \lambda I_n) = (x - \lambda_1)(x - \lambda_2)\cdots(x - \lambda_n)$, where $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ and $\lambda_1 < \lambda_2 \cdots < \lambda_n$. Is A diagonalizable?

The matrix A is diagonalizable and

$$D = \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix}.$$

The matrix is diagonalizable if all roots $\lambda_1, \ldots, \lambda_n$ are distinct.

Example 3.6.2: A Non-Diagonalizable Matrix

Given

$$A = \begin{bmatrix} 7 & 1 & -1 \\ -11 & -3 & 2 \\ 18 & 2 & -4 \end{bmatrix},$$

construct the diagonal matrix D and form the nonsingular matrix P such that

$$D = P^{-1}AP.$$

We will skip over the determinant calculation and will, instead, assert that the characteristic polynomial is $(\lambda + 2)^2(\lambda - 4)$, meaning the eigenvalues are $\lambda_1 = -2$ and $\lambda_2 = 4$. The first eigenvalue, λ_1, gives the linear system

$$\left[\begin{array}{ccc|c} 9 & 1 & -1 & 0 \\ -11 & -1 & 2 & 0 \\ 18 & 2 & -2 & 0 \end{array}\right]$$

which reduces to

$$\left[\begin{array}{ccc|c} 1 & 0 & -\frac{1}{2} & 0 \\ 0 & 1 & \frac{7}{2} & 0 \\ 0 & 0 & 0 & 0 \end{array}\right],$$

giving us the eigenspace $E_{-2} = \left\{ c \begin{bmatrix} 1 \\ -7 \\ 2 \end{bmatrix} : c \in \mathbb{R} \right\}$. Notice that for λ_1,

$$\operatorname{rank}(A - \lambda_1 I_3) = 2 \neq 3 - 2 = 1.$$

Therefore, A cannot be diagonalized; we need not find the eigenspace of λ_2, but for the sake of a curious student, we have $E_4 = \left\{ c \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix} : c \in \mathbb{R} \right\}$.

3.6.2 Using Diagonalization to Raise Matrices to Powers

One useful application of the diagonalization process is raising matrices to powers. Consider the following theorem.

Theorem 3.6.3: Raising Matrices to Powers

Given a diagonalizable matrix A,

$$A^k = PD^kP^{-1}.$$

Proof. We proceed by induction. For $k = 2$,

$$A^2 = AA = (PDP^{-1})(PDP^{-1}) = PD(P^{-1}P)DP^{-1} = PDI_nDP^{-1} = PD^2P^{-1}.$$

Suppose that for all $m \in \mathbb{N}$, $A^m = PD^mP^{-1}$. Then,

$$A^{m+1} = A^mA = PD^mP^{-1}(PDP^{-1}) = PD^m(P^{-1}P)DP^{-1} = PD^mI_nDP^{-1} = PD^{m+1}P^{-1},$$

proving the theorem. $\square$

3.7 Lecture 19: October 5, 2022

3.7.1 Complex Numbers

Consider the following definition.

Definition 3.7.1: The Complex Numbers

The set of complex numbers, $\mathbb{C}$, is given by

$$\mathbb{C} = \{a + bi : a, b \in \mathbb{R}, i^2 = -1\}.$$

Definition 3.7.2: The Set $\mathbb{C}^n$

We define $\mathbb{C}^n$ as

$$\mathbb{C}^n = \{[z_1, \dots, z_n] : z_i = \mathbb{C}\}.$$

As a set, $\mathbb{C}$ can be visualized as $\mathbb{R}^2$. Consider the following definitions and properties.

Definition 3.7.3: Definitions of Operations With Complex Numbers

Let $a, b, c, d \in \mathbb{R}$. Then, the following operations are defined as follows.

1. $(a + bi) + (c + di) := (a + c) + (b + d)i$.
2. $\Re(a + bi) := a, \quad \Im(a + bi) := b$.
3. $z = a + bi \in \mathbb{C} \implies ||z|| := \sqrt{a^2 + b^2}$.
4. $(a + bi)(c + di) := ac + adi + bci + bdi^2 = (ac - bd) + (ad + bc)i$.
5. $z = a + bi \in \mathbb{C} \implies \overline{z} := a - bi$.

Theorem 3.7.1: Properties of Complex Numbers

Let $a, b, c, d \in \mathbb{R}$. Then,

1. $z\overline{z} = (a + bi)(a - bi) = a^2 + b^2 = ||z||^2$.
2. $z \in \mathbb{C} - \{0\} \implies \frac{1}{z} = \frac{a}{a^2+b^2} - \left(\frac{b}{a^2+b^2}\right) \in \mathbb{C}$.

Theorem 3.7.2: The Fundamental Theorem of Algebra

If $p(z) = a_n z^n + \dots + a_0 z^0$ is an nth degree polynomial with coefficients $a_0, \dots, a_n \in \mathbb{C}$, $p(z)$ has n complex roots. Note that some roots may be repeated.

We now consider vectors in $\mathbb{C}^n$ and define their operations.

Definition 3.7.4: Vector Addition in $\mathbb{C}^n$

Let $\vec{v} = [v_1, \dots, v_n] \in \mathbb{C}^n$ and $\vec{w} = [w_1, \dots, w_n] \in \mathbb{C}^n$. Then,

$$\vec{v} + \vec{w} = [v_1 + w_1, \dots, v_n + w_n].$$

Definition 3.7.5: Scalar Multiplication in $\mathbb{C}^n$

Let $\vec{v} = [v_1, \dots, v_n] \in \mathbb{C}^n$ and $c \in \mathbb{C}$. Then,

$$c\vec{v} = [cv_1, \dots, cv_n].$$

Definition 3.7.6: The Dot Product in $\mathbb{C}^n$

Let $\vec{v} = [v_1, \dots, v_n] \in \mathbb{C}^n$ and $\vec{w} = [w_1, \dots, w_n] \in \mathbb{C}^n$. Then,

$$\vec{v} \cdot \vec{w} = v_1\overline{w_1} + \cdots + v_n\overline{w_n}.$$

Definition 3.7.7: The Magnitude in $\mathbb{C}^n$

Let $\vec{v} = [v_1, \dots, v_n] \in \mathbb{C}^n$. Then,

$$||\vec{v}|| = \sqrt{\vec{v} \cdot \vec{v}}.$$

Definition 3.7.6 provides a good question: do we consider complex conjugates when performing matrix multiplication with matrices of the sort $\mathcal{M}_{mn}^{\mathbb{C}}$? We do not.

Definition 3.7.8: The Adjoint

Given $A \in \mathcal{M}_{mn}^{\mathbb{C}}$,

$$A = \begin{bmatrix} z_{11} & \cdots & z_{1n} \\ \vdots & \ddots & \vdots \\ z_{m1} & \cdots & z_{mn} \end{bmatrix} \implies A^* = \begin{bmatrix} \overline{z_{11}} & \cdots & \overline{z_{m1}} \\ \vdots & \ddots & \vdots \\ \overline{z_{1n}} & \cdots & \overline{z_{mn}} \end{bmatrix}$$

We will now provide a theorem motivating the transpose operation and apply it to our definition of the adjoint.

Theorem 3.7.3: The Transpose and The Dot Product

If $\vec{v}, \vec{w} \in \mathbb{R}^n$, and $A \in \mathcal{M}_{nn}^{\mathbb{R}}$,

$$(A\vec{v}) \cdot \vec{w} = \vec{v} \cdot (A^T\vec{w})$$

Theorem 3.7.4: The Adjoint and The Dot Product

If $\vec{v}, \vec{w} \in \mathbb{C}^n$, and $A \in \mathcal{M}_{nn}^{\mathbb{C}}$,

$$(A\vec{v}) \cdot \vec{w} = \vec{v} \cdot (A^*\vec{w})$$

The analog of symmetric and skew-symmetric matrices is the notion of hermitian and skew-hermitian matrices. Consider the following definition. It may also be useful to recall Definition 1.3.12.

Definition 3.7.9: Hermitian and Skew-Hermitian Matrices

Suppose $A \in \mathcal{M}_{nn}^{\mathbb{C}}$. Then,

1. A is hermitian if and only if $A = A^*$.
2. A is skew-hermitian if and only if $A = -A^*$.
3. A is normal if and only if $AA^* = A^*A$.

3.8 Lecture 20: October 7, 2022

3.8.1 Diagonalization Using Complex Eigenvalues

We now consider the diagonalization process for matrices with complex eigenvalues. Consider the following motivating example.

Example 3.8.1: ☢☢ Diagonalization With Complex Eigenvalues

Diagonalize the matrix

$$A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}.$$

The roots of the characteristic polynomial, $\det(A - \lambda I_2)$ are $\lambda_1 = i$ and $\lambda_2 = -i$. For λ_1, we have

$$\left[\begin{array}{cc|c} i & -1 & 0 \\ 1 & i & 0 \end{array}\right].$$

We perform $\langle 1 \rangle \leftrightarrow \langle 2 \rangle$ to obtain

$$\left[\begin{array}{cc|c} 1 & i & 0 \\ i & -1 & 0 \end{array}\right].$$

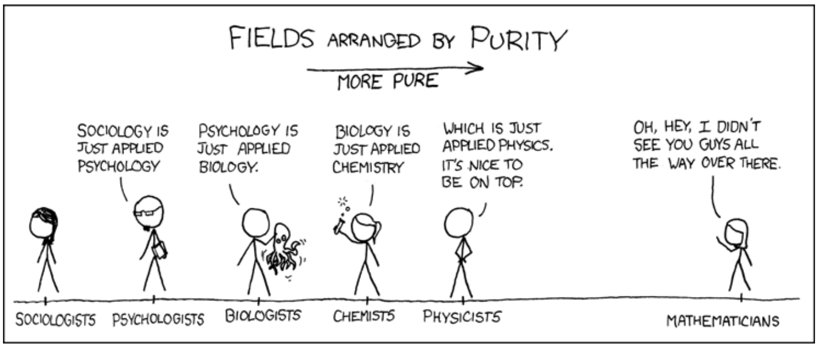
Then, we have $-i\langle 1 \rangle + \langle 2 \rangle \to \langle 2 \rangle$, producing

$$\left[\begin{array}{cc|c} 1 & i & 0 \\ 0 & 0 & 0 \end{array}\right].$$

Thus, $E_i = \left\{ c \begin{bmatrix} -i \\ 1 \end{bmatrix} : c \in \mathbb{C} \right\}$. By a similar process $E_{-i} = \left\{ c \begin{bmatrix} i \\ 1 \end{bmatrix} : c \in \mathbb{C} \right\}$. Now, we form

$$D = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}, \quad P = \begin{bmatrix} -i & i \\ 1 & 1 \end{bmatrix}.$$

Note that we still may not have "enough" eigenvectors. This is when the set of fundamental eigenvectors is linearly dependent. Recall Theorem 3.6.2.



4
Finite Dimensional Vector Spaces

4.1 Lecture 20: October 7, 2022

4.1.1 The Process of Abstraction

Consider the following definition.

Definition 4.1.1: Vector Spaces

Let $\mathbb{F}$ be a field of scalars. For now, $\mathbb{F} = \mathbb{R} \vee \mathbb{C}$. A vector space V over $\mathbb{F}$ is a set with two operations:

1. Vector Addition: $+ : V \times V \to V, (\vec{v}, \vec{w}) \mapsto \vec{v} + \vec{w}$.
2. Scalar Multiplication: $\cdot : \mathbb{F} \times V \to V, (c, \vec{v}) \mapsto c\vec{v}$.

The following axioms must hold for each $\vec{u}, \vec{v}, \vec{w} \in V$ and $c_1, c_2 \in \mathbb{F}$.

1. $\vec{u} + \vec{v} = \vec{v} + \vec{u}$.
2. $\vec{u} + (\vec{v} + \vec{w}) = (\vec{u} + \vec{v}) + \vec{w}$.
3. $\exists \vec{0} \in V, \forall \vec{v} \in V, \vec{v} + \vec{0} = \vec{v}$.
4. $\forall \vec{v} \in V, \exists!(-\vec{v}) \in V, \vec{v} + (-\vec{v}) = \vec{0}$.
5. $c_1(\vec{u} + \vec{v}) = c_1\vec{u} + c_1\vec{v}$.
6. $(c_1 + c_2)\vec{u} = c_1\vec{u} + c_2\vec{u}$.
7. $(c_1 c_2)\vec{u} = c_1(c_2\vec{u})$.
8. $1\vec{u} = \vec{u}$.

We remark that $0\vec{v} = \vec{0}$ is *not* an axiom of a vector space; we must *prove* that it holds. Now, we will justify our use of abstraction.

1. The set $\mathbb{R}^n$ is a vector space.
2. The set $\mathbb{C}^n$ is a vector space.
3. The set $\{\vec{0}\}$, with $\vec{0} + \vec{0} + \vec{0}$ and $c\vec{0} = \vec{0}$, is a vector space.
4. The set $\mathcal{M}_{mn}$ is a vector space.
5. Let S be a nonempty set and $F(S) = \{f : S \to \mathbb{R}\}$ with $(f+g)(s) = f(s)+g(s)$ and $(cf)(s) = cf(s)$. The set $F(S)$ is a vector space.
6. The set $\{a_n x^n + \cdots + a_0 x^0 : a_0, \ldots a_n \in \mathbb{R}, \text{degree is less than or equal to } n\}$ is a vector space.

We note that all the above examples need *proof*. We remark that, for now, we will primarily consider vector spaces over $\mathbb{R}$. We will emphasize when we consider vector spaces over $\mathbb{C}$. Consider the following examples of sets that are not vector spaces.

1. The set $\mathbb{R}^+ = \{x \in \mathbb{R} : x \geq 0\}$, with usual addition and usual multiplication in $\mathbb{R}$, is not a vector space. The set does not satisfy the closure axiom for scalar multiplication; $(-1) \cdot 1 = -1 \notin \mathbb{R}^+$.
2. The set $S = \{f : \mathbb{R} \to \mathbb{R} : f(0) = 1\}$, with $(f+g)(x) = f(x)+g(x)$ and $(cf)(x) = cf(x)$, is not a vector space. The set does not satisfy the closure axiom for vector addition; $(f+g)(0) = f(0)+g(0) = 2$, so $(f+g)(x) \notin S$.

4.2 Lecture 21: October 14, 2022

4.2.1 Derived Properties of Vector Spaces

Consider the following theorems. While they may seem obvious, they require proof by the vector space axioms, provided in Definition 4.1.1.

Theorem 4.2.1: Derived Property of Vector Space 1

Suppose V is a vector space, $\vec{v} \in V$, and $c \in \mathbb{F}$. Then,

$$c\vec{0} = \vec{0}.$$

Proof. We have that

$$\begin{aligned} c\vec{0} &= c\vec{0} + \vec{0} && \text{by axiom 3,} \\ &= c\vec{0} + (c\vec{0} + (-(c\vec{0}))) && \text{by axiom 4,} \\ &= (c\vec{0} + c\vec{0}) + (-(c\vec{0})) && \text{by axiom 2,} \\ &= c(\vec{0} + \vec{0}) + (-(c\vec{0})) && \text{by axiom 5,} \\ &= c\vec{0} + (-(c\vec{0})) && \text{by axiom 3,} \\ &= \vec{0} && \text{by axiom 4,} \end{aligned}$$

as desired. □

Theorem 4.2.2: Derived Property of Vector Space 2

Suppose V is a vector space, $\vec{v} \in V$, and $c \in \mathbb{F}$. Then,

$$0\vec{v} = \vec{0}.$$

Proof. We have that

$$\begin{aligned} 0\vec{v} &= 0\vec{v} + \vec{0} && \text{by axiom 3,} \\ &= 0\vec{v} + (0\vec{v} + (-(0\vec{v}))) && \text{by axiom 4,} \\ &= (0\vec{v} + 0\vec{v}) + (-(0\vec{v})) && \text{by axiom 2,} \\ &= (0 + 0)\vec{v} + (-(0\vec{v})) && \text{by axiom 6,} \\ &= 0\vec{v} + (-(0\vec{v})) && \text{by properties of } \mathbb{R}, \\ &= \vec{0} && \text{by axiom 4,} \end{aligned}$$

as desired. □

Theorem 4.2.3: Derived Property of Vector Space 3

Suppose V is a vector space, $\vec{v} \in V$, and $c \in \mathbb{F}$. Then,

$$(-1)\vec{v} = -\vec{v}.$$

Proof. We have that

$$\begin{aligned}
\vec{v} + (-1)\vec{v} &= 1\vec{v} + (-1)\vec{v} && \text{by axiom 8,} \\
&= (1 + (-1))\vec{v} && \text{by axiom 6,} \\
&= 0\vec{v} && \text{by properties of } \mathbb{R}, \\
&= \vec{0} && \text{by Theorem 4.2.2.}
\end{aligned}$$

Now, let $\vec{v}' = (-1)\vec{v}$ and let $\vec{v}'' \in V$ be another vector such that

$$\vec{v} + \vec{v}'' = \vec{0}.$$

Then,

$$\vec{v} + \vec{v}'' = \vec{v} + \vec{v}'.$$

We can then add $\vec{v}'$ to both sides to obtain

$$\vec{v}' + (\vec{v} + \vec{v}'') = \vec{v}' + (\vec{v} + \vec{v}'),$$

which becomes

$$(\vec{v}' + \vec{v}) + \vec{v}'' = (\vec{v}' + \vec{v}) + \vec{v}'.$$

Then, we have $\vec{v}'' = \vec{v}' = (-1)\vec{v}$. Thus, $(-1)\vec{v} = -\vec{v}$, as desired. $\square$

Theorem 4.2.4: Derived Property of Vector Space 4

Suppose V is a vector space, $\vec{v} \in V$, and $c \in \mathbb{F}$. Then,

$$c\vec{v} = \vec{0} \iff c = 0 \vee \vec{v} = \vec{0}$$

Proof. The statement $c = 0 \vee \vec{v} = \vec{0} \implies c\vec{v} = \vec{0}$ is governed by the first two derived results. Then, to show that $c\vec{v} = \vec{0} \implies c = 0 \vee \vec{v} = \vec{0}$, we suppose that $c \neq 0$ and wish to show that $\vec{v} = \vec{0}$. We have

$$c\vec{v} = \vec{0}$$

with $c \neq 0$. Then,

$$\left(\frac{1}{c}\right)(c\vec{v}) = \frac{1}{c}\vec{0} = \vec{0},$$

by Theorem 4.2.1. But, by axiom 7,

$$\left(\frac{1}{c}\right)(c\vec{v}) = \left(\frac{1}{c}c\vec{v}\right) = \vec{v},$$

which implies $\vec{v} = \vec{0}$. $\square$

4.3 Lecture 22: October 17, 2022

4.3.1 Subspaces

Consider the following definition.

Definition 4.3.1: Subspaces

Suppose V is a vector space and let W be a set with $W \subseteq V$. Then, W is a subspace of V if and only if W is a vector space with the same operations as V.

Consider the following theorem.

Theorem 4.3.1: Showing a Vector Space is a Subspace

Suppose V is a vector space and let W be a set with $W \subseteq V$. Then, W is a subspace of V if and only if

1. $W \neq \emptyset$.
2. $\vec{w}_1, \vec{w}_2 \in W \implies \vec{w}_1 + \vec{w}_2 \in W$.
3. $c \in \mathbb{F}, \vec{w} \in W \implies c\vec{w} \in W$.

Proof. We know that $W \neq \emptyset$, because $\vec{0} \in W$. Next, $\vec{w}_1, \vec{w}_2 \in W \implies \vec{w}_1 + \vec{w}_2 \in W$ by the closure property of addition of a vector space W. Then, $\vec{w} \in W \implies c\vec{w} \in W, c \in \mathbb{F}$ by the closure property of scalar multiplication of a vector space W. Now, we must show that all vector space axioms hold in W. Axioms 1, 2, 5, 6, 7, and 8 all hold true, as they relate to the operations of W. Since W has the same operations as V, a known vector space, these properties hold true in W, a subset of V, as well. We will now prove property 3. Since W is nonempty, we consider some $\vec{w} \in W$. Since W is closed under scalar multiplication, $0\vec{w} \in W$. This is the same operation as the one described in Theorem 4.2.2, and so $0\vec{w} = \vec{0}$. Since $\vec{v} + \vec{0} = \vec{v}$ for all $\vec{v} \in V$, $\vec{w} + \vec{0} = \vec{w}$ for all $\vec{w} \in W$. We will now prove property 4. Since $\vec{w} \in W$, $\vec{w} \in V$. The additive inverse of $\vec{w}$ in V is $(-1)\vec{w}$. We see that $(-1)\vec{w} \in W$, as W is closed under scalar multiplication. ☐

Consider the following examples.

Example 4.3.1: ☢ Is it a Subspace? 1

Let $V = \mathbb{R}^2$ and let

$$W = \{[x, 0] : x \in \mathbb{R}\} \subseteq \mathbb{R}^2.$$

Is W a subspace of V?.

We see that $W \neq \emptyset$, as $[0, 0] \in W$. Then, if $[x_1, 0], [x_2, 0] \in W$,

$$[x_1, 0] + [x_2, 0] = [x_1 + x_2, 0] \in W.$$

Then, if $c \in \mathbb{R}$ and $[x, 0] \in W$,

$$c[x, 0] = [cx, 0] \in W.$$

Because $W \neq \emptyset$ is closed under vector addition and scalar multiplication, W is a subspace of V.

Example 4.3.2: ☢☢ Is it a Subspace? 2

Let $V = \mathcal{M}_{nn}$ and let

$$W = \mathcal{D}_n \subseteq \mathcal{U}_n \subseteq \mathcal{M}_{nn},$$

where $\mathcal{D}_n$ is the set of all $n \times n$ diagonal matrices and $\mathcal{U}_n$ is the set of all $n \times n$ upper triangular matrices. Is W a subspace of V?.

We see that $W \neq \emptyset$, since $\begin{bmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{bmatrix} \in W$. Then, for $A, B \in \mathcal{D}_n$,

$$\begin{aligned} (A + B)_{ij} &= A_{ij} + B_{ij} \\ &= \begin{cases} A_{ii} + B_{ii} & i = j \\ 0 & i \neq j \end{cases} \\ &\in \mathcal{D}_n. \end{aligned}$$

Then, for some $c \in \mathbb{R}$,

$$\begin{aligned} (cA)_{ij} &= cA_{ij} \\ &= \begin{cases} cA_{ii} & i = j \\ 0 & i \neq j \end{cases} \\ &\in \mathcal{D}_n. \end{aligned}$$

Therefore, $\mathcal{D}_n$ is a subspace of $\mathcal{M}_{nn}$.

Example 4.3.3: ☢☢ Is it a Subspace? 3

Let $V = \mathcal{M}_{nn}$ and let

$$W = \mathcal{U}_n \subseteq \mathcal{M}_{nn},$$

where $\mathcal{U}_n$ is the set of all $n \times n$ upper triangular matrices. Is W a subspace of V?.

We see that $W \neq \emptyset$, since $\begin{bmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{bmatrix} \in W$. Then, for $A, B \in \mathcal{U}_n$,

$$\begin{aligned} (A+B)_{ij} &= A_{ij} + B_{ij} \\ &= \begin{cases} A_{ij} + B_{ij} & i \leq j \\ 0 & i > j \end{cases} \\ &\in \mathcal{U}_n. \end{aligned}$$

Then, for some $c \in \mathbb{R}$,

$$\begin{aligned} (cA)_{ij} &= cA_{ij} \\ &= \begin{cases} cA_{ij} & i \leq j \\ 0 & i > j \end{cases} \\ &\in \mathcal{U}_n. \end{aligned}$$

Therefore, $\mathcal{U}_n$ is a subspace of $\mathcal{M}_{nn}$.

Consider the following examples of subsets of $\mathbb{R}^n$ that are not subspaces.

1. The set of n dimensional vectors whose first coordinate is nonnegative.
2. The set of unit n dimensional vectors.
3. The set of n dimensional vectors with a zero in at least one coordinate, where $n \geq 2$.
4. The set of n dimensional vectors having all integer coordinates.
5. The set of all n dimensional vectors whose first two coordinates add up to 3.

Consider the following examples of subsets of $\mathcal{M}_{nn}$ that are not subspaces.

1. The set of nonsingular $n \times n$ matrices.
2. The set of singular $n \times n$ matrices.
3. The set of $n \times n$ matrices in reduced row echelon form.

Consider the following theorem.

Theorem 4.3.2: Eigenspaces are Subspaces

Let $A \in \mathcal{M}_{nn}$ and let λ be an eigenvalue of A with eigenspace E_λ. Then, E_λ is a subspace of $\mathbb{R}^n$.

Proof. By definition,

$$E_\lambda = \{X : AX = \lambda X\}.$$

We see that $E_\lambda \neq \emptyset$, as $\vec{0} \in E_\lambda$. Let $\vec{x}_1, \vec{x}_2 \in E_\lambda$. We must show that $\vec{x}_1 + \vec{x}_2 \in E_\lambda$. That is, we wish to show that

$$A(\vec{x}_1 + \vec{x}_2) = \lambda(\vec{x}_1 + \vec{x}_2).$$

We realize that

$$\begin{aligned} A(\vec{x}_1 + \vec{x}_2) &= A\vec{x}_1 + A\vec{x}_2 \\ &= \lambda\vec{x}_1 + \lambda\vec{x}_2 \\ &= \lambda(\vec{x}_1 + \vec{x}_2), \end{aligned}$$

as desired. Then, we must show that for some scalar c, $c\vec{x}_1 \in E_\lambda$. We wish to show that

$$A(c\vec{x}_1) = \lambda c\vec{x}_1.$$

We see that

$$\begin{aligned} A(c\vec{x}_1) &= cA\vec{x}_1 \\ &= c\lambda\vec{x}_1 \\ &= \lambda c\vec{x}_1, \end{aligned}$$

as desired. Because we have showed that the closure properties hold for E_λ, E_λ is a subspace of $\mathbb{R}^n$. $\square$

4.4 Lecture 23, October 19, 2022

4.4.1 Span

We will now revisit the notion of linear combinations. Consider the following definitions.

Definition 4.4.1: Finite Linear Combinations

Let S be a nonempty, and possibly infinite, subset of a vector space V. Then, a vector $\vec{v} \in V$ is a finite linear combination of the vectors in S if and only if there exists some finite subset $S' = \{\vec{v}_1, \dots \vec{v}_n\}$ of S such that

$$\vec{v} = c_1\vec{v}_1 + \cdots + c_n\vec{v}_n$$

for scalars $c_1, \dots, c_n$.

Theorem 4.4.1: Subspaces are Closed Under Linear Combinations

Let W be a subspace of a vector space V, and let $\vec{v}_1, \dots, \vec{v}_n \in W$. For scalars $c_1, \dots, c_n$, we have

$$c_1\vec{v}_1 + \cdots + c_n\vec{v}_n \in W.$$

Proof. We proceed by induction. For the proposition when $n = 1$, consider the scalar $c_1 \in \mathbb{F}$ and the vector $\vec{v}_1 \in W$. By the closure property of scalar multiplication, $c_1\vec{v}_1 \in W$. Suppose the theorem holds for all $n = k$. That is, for scalars $c_1, \dots, c_k$ and vectors $\vec{v}_1, \dots, \vec{v}_k$, we have

$$c_1\vec{v}_1 + \cdots + c_k\vec{v}_k \in W.$$

Then, for scalar $c_{k+1} \in \mathbb{F}$ and vector $\vec{v}_{k+1} \in W$, $c_{k+1}\vec{v}_{k+1} \in W$ by the closure property of scalar multiplication, and by the closure property of vector addition, since $c_1\vec{v}_1 + \cdots + c_k\vec{v}_k \in W$,

$$c_1\vec{v}_1 + \cdots + c_k\vec{v}_k + c_{k+1}\vec{v}_{k+1} \in W,$$

as desired. □

We now, introduce the notion of "span."

Definition 4.4.2: Span

Let S be a nonempty subset of a vector space V. Then, $\operatorname{span}(S)$ is the set of all possible finite linear combinations of the vectors in S. If $S = \emptyset$, $\operatorname{span}(S) = \{\vec{0}\}$.

Consider the following examples.

Example 4.4.1: ☢ Find Span 1

Let $S = \{[0,1,0],[0,0,1]\} \subseteq \mathbb{R}^3$. Find span (S).

We realize that

$$\begin{aligned}\text{span}\,(S) &= \{c_1[0,1,0] + c_2[0,0,1] : c_1, c_2 \in \mathbb{R}\}\\ &= \{[0, c_1, c_2] : c_1, c_2 \in \mathbb{R}\}.\end{aligned}$$

Example 4.4.2: ☢☢ Find Span 2

Let $S = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\} \subseteq \mathcal{M}_{22}$. Find span (S).

We realize that

$$\begin{aligned}\text{span}\,(S) &= \left\{ c_1 \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + c_2 \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + c_3 \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} : c_1, c_2, c_3 \in \mathbb{R} \right\}\\ &= \left\{ \begin{bmatrix} c_1 & 0 \\ c_2 & c_3 \end{bmatrix} : c_1, c_2, c_3 \in \mathbb{R} \right\}.\end{aligned}$$

Sometimes, for some $S \subseteq V$, span $(S) = V$. Here, we say that S spans V, or equivalently, V is spanned by S. Consider the following example.

Example 4.4.3: ☢☢ Find Span 3

Let $S = \{[1,0,0],[0,1,0],[0,0,1]\} \subseteq \mathbb{R}^3$. Find span (S).

We realize that

$$\begin{aligned}\text{span}\,(S) &= \{c_1[1,0,0] + c_2[0,1,0] + c_3[0,0,1] : c_1, c_2, c_3 \in \mathbb{R}\}\\ &= \{[c_1, c_2, c_3] : c_1, c_2, c_3 \in \mathbb{R}\}\\ &= \mathbb{R}^3,\end{aligned}$$

meaning that S spans $\mathbb{R}^3$.

Consider the following theorems.

Theorem 4.4.2: A Complete Characterization of the Span

Let S be a nonempty subset of a vector space V. Then,

1. $S \subseteq \operatorname{span}(S)$.

   Proof. Suppose that $\vec{v} \in S$. Then, $1\vec{v} = \vec{v} \in \operatorname{span}(S)$. □

2. $\operatorname{span}(S)$ is a subspace of V.

   Proof. We know that $\operatorname{span}(S) \neq \emptyset$, by definition. Suppose that $\vec{w}_1, \vec{w}_2 \in \operatorname{span}(S)$. We know that
   $$\vec{w}_1 = c_1\vec{v}_1 + \cdots + c_n\vec{v}_n$$
   for some $c_1, \ldots, c_n \in \mathbb{F}$ and $\vec{v}_1, \ldots, \vec{v}_n \in S$. We also know that
   $$\vec{w}_2 = d_1\vec{u}_1 + \cdots + d_k\vec{u}_k.$$
   for some $d_1, \ldots, d_k \in \mathbb{F}$ and $\vec{u}_1, \ldots, \vec{u}_k \in S$. Then,
   $$\vec{w}_1 + \vec{w}_2 = (c_1\vec{v}_1 + \cdots + c_k\vec{v}_k) + (d_1\vec{u}_1 + \cdots + d_k\vec{u}_k).$$
   The sum is a linear combination of the elements in S, therefore, $\vec{w}_1 + \vec{w}_2 \in \operatorname{span}(S)$, meaning $\operatorname{span}(S)$ is closed under vector addition. Now, we take $c \in \mathbb{F}$ and $\vec{v} \in \operatorname{span}(S)$ and wish to show that $c\vec{v} \in \operatorname{span}(S)$. We know that
   $$\vec{v} = c_1\vec{v}_1 + \cdots + c_n\vec{v}_n$$
   for some $c_1, \ldots, c_n \in \mathbb{F}$ and $\vec{v}_1, \ldots, \vec{v}_n \in S$. Then,
   $$c\vec{v} = c(c_1\vec{v}_1) + \cdots + c(c_n\vec{v}_n),$$
   which is a linear combination of the elements in S, meaning that it is in $\operatorname{span}(S)$. □

3. If W is a subspace of V with $S \subseteq W$, then, $\operatorname{span}(S) \subseteq W$.

   Proof. We let $\vec{v} \in \operatorname{span}(S)$ and wish to show that $\vec{v} \in W$. We know that
   $$\vec{v} = c_1\vec{v}_1 + \cdots + c_n\vec{v}_n$$
   for some $c_1, \ldots, c_n \in \mathbb{F}$ and $\vec{v}_1, \ldots, \vec{v}_n \in S$. Since $S \subseteq W$, and $\vec{v}$ is a linear combination of elements of S, $\vec{v}$ is a linear combination of elements of W. It follows that $\vec{v} \in W$, by Theorem 4.4.1. □

4. $\operatorname{span}(S)$ is the smallest subspace of V containing S.

   Proof. By part 1 and part 2, $\operatorname{span}(S)$ is a subspace of V containing S. Then, by part 3, $\operatorname{span}(S)$ is a subset of all subspaces W of V, so $\operatorname{span}(S)$ is the smallest subspace of V containing S. □

Theorem 4.4.3: Two Subsets of a Vector Space and Their Span

Let V be a vector space, and let $S_1 \subseteq S_2 \subseteq V$. Then,

$$\text{span}(S_1) \subseteq \text{span}(S_2).$$

Proof. We see that $S_1 \subseteq \text{span}(S_1)$ and $S_2 \subseteq \text{span}(S_2)$. Because $\text{span}(S_2)$ is the smallest subspace of V containing S_2, we write

$$S_1 \subseteq S_2 \subseteq \text{span}(S_2) \subseteq V.$$

Then, because $\text{span}(S_2)$ is a subspace of V with $S_1 \subseteq \text{span}(S_2)$, we have $\text{span}(S_1) \subseteq \text{span}(S_2)$, as desired. □

Theorem 4.4.4: Span as an Intersection of Subspaces

Let V be a vector space, with $S \subseteq V$. Then, $\text{span}(S)$ is the intersection of all the subspaces of V containing S.

Proof. Let W be an arbitrary subspace of V with $S \subseteq W$. We wish to show

$$\text{span}(S) \subseteq \bigcap W, \quad \bigcap W \subseteq \text{span}(S).$$

By the third part of Theorem 4.4.2, $\text{span}(S) \subseteq \bigcap W$ since $\text{span}(S)$ is a subset of all subspaces W. Because $\text{span}(S)$, is itself, a subspace of V containing S, $\bigcap W$ must be contained in $\text{span}(S)$, since if some $\vec{v} \in \bigcap W$, $\vec{v}$ is in all subspaces of V containing S, including $\text{span}(S)$. We have now shown both properties, and conclude that

$$\text{span}(S) = \bigcap W,$$

as desired. □

Now, we turn to the question of how we can determine which vectors lie in span (S) given

$$S = \{\vec{v}_1, \dots, \vec{v}_k\} \subseteq \mathbb{R}^n.$$

Consider the following theorem.

Theorem 4.4.5: Span and Row Space

Let A be the matrix having $S = \{\vec{v}_1, \dots, \vec{v}_k\} \subseteq \mathbb{R}^n$ as its rows. Then, the span of S is the row space of A.

Consider the following examples.

Example 4.4.4: Computing Span as a Row Space 1

Let $\vec{v}_1 = [3, 6, 0]$, $\vec{v}_2 = [0, -1, 1]$, and $S = \{\vec{v}_1, \vec{v}_2\}$. Find span (S).

We have the matrix

$$A = \begin{bmatrix} 3 & 6 & 0 \\ 0 & -1 & 1 \end{bmatrix}.$$

The row space of A, or equivalently, the span of S is

$$\text{span}\,(S) = \{c_1[1, 0, 2] + c_2[0, 1, -1] : c_1, c_2 \in \mathbb{R}\}.$$

Example 4.4.5: Computing Span as a Row Space 2

Let $\vec{v}_1 = [-1, 1, 0]$, $\vec{v}_2 = [1, 0, 1]$, and $\vec{v}_2 = [1, 0, -1]$. Let $S = \{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$. Find span (S).

We have the matrix

$$A = \begin{bmatrix} -1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & -1 \end{bmatrix}.$$

The row space of A, or equivalently, the span of S is

$$\text{span}\,(S) = \{c_1[1, 0, 0] + c_2[0, 1, 0] + c_3[0, 0, 1] : c_1, c_2, c_3 \in \mathbb{R}\}.$$

Theorem 4.4.6: Is a Vector in the Span?

Let A be the matrix having $S = \{\vec{v}_1, \dots, \vec{v}_k\} \subseteq \mathbb{R}^n$ as its rows. Then, $B \in \text{span}\,(S)$ if and only if the linear system $[A^T|B]$ has at least one solution.

A similar algorithm is available for general vector spaces; but we will postpone the justification of it.

4.5 Lecture 24: October 21, 2022

4.5.1 Linear Independence

We have mentioned the notion of linear independence and linear dependence before, but now, we will make these definitions more precise. As such, this section will be heavy on theory.

Definition 4.5.1: Linear Independence and Dependence

Suppose $S = \{\vec{v}_1, \dots, \vec{v}_n\} \subseteq V$. Then,

1. S is linearly dependent if and only if there exist scalars $c_1, \dots, c_n \in \mathbb{F}$ not all zero, such that
$$c_1\vec{v}_1 + \cdots + c_n\vec{v}_n = \vec{0}.$$
2. S is linearly independent if S is not linearly dependent. That is,
$$c_1\vec{v}_1 + \cdots + c_n\vec{v}_n = \vec{0} \iff c_1 = \cdots = c_n = 0.$$

If $S = \emptyset$, S is linearly independent.

Definition 4.5.2: Generalizing Linear Independence and Dependence to Infinite Sets

Suppose an infinite set S with $S \subseteq V$ where V is a vector space. We say S is linearly independent if and only if each finite subset of S is linearly independent.

Consider the following theorems, providing methods of showing linear independence of subsets of $\mathbb{R}^n$. Similar algorithms exist for general vector spaces, but we will postpone the justification of them.

Theorem 4.5.1: Showing Linear Independence 1

Let $A \in \mathcal{M}_{nn}$. If $\det A \neq 0$, the columns of A are linearly independent. If $\det A = 0$, the columns of A are linearly dependent.

Proof. If $\det A \neq 0$, A is nonsingular and has only the trivial solution for $[A|\vec{0}]$. This linear system is precisely
$$c_1\vec{v}_1 + \cdots + c_n\vec{v}_n = \vec{0}$$
for $c_1, \dots, c_n \in \mathbb{R}$. Since the system has only the trivial solution, $X = [c_1, \dots, c_n]^T = [0, \dots, 0]^T$ and $c_1 = \cdots = c_n = 0$, which by definition, means that the set $\{\vec{v}_1, \dots, \vec{v}_n\}$, or equivalently the columns of A, are linearly independent. If $\det A = 0$, A is singular and has infinitely many nontrivial solutions for $[A|\vec{0}]$. Thus, $c_1, \cdots, c_n$ will not all be zero, meaning that the columns of A are linearly dependent. □

Theorem 4.5.2: Showing Linear Independence 2

Suppose $S \subseteq V$ where V is a vector space and S is a finite set having at least two elements. Then, S is linearly independent if and only if

$$\forall \vec{v} \in S, \vec{v} \notin \operatorname{span}(S - \{\vec{v}\}).$$

Equivalently, S is linearly dependent if and only if

$$\exists \vec{v} \in S, \vec{v} \in \operatorname{span}(S - \{\vec{v}\}).$$

Proof. Suppose S is linearly dependent. That is, we have $c_1, \ldots, c_n \in \mathbb{F}$ such that

$$c_1\vec{v}_1 + \cdots + c_{i-1}\vec{v}_{i-1} + c_i\vec{v}_i + c_{i+1}\vec{v}_{i+1} + \cdots + c_n\vec{v}_n = \vec{0}$$

with $c_i \neq 0$ for some i. Then,

$$\vec{v}_i = \left(-\frac{c_1}{c_i}\right)\vec{v}_1 + \cdots + \left(-\frac{c_{i-1}}{c_i}\right)\vec{v}_{i-1} + \left(-\frac{c_{i+1}}{c_i}\right)\vec{v}_{i+1} + \cdots + \left(-\frac{c_n}{c_i}\right)\vec{v}_n.$$

We have constructed a vector in S as a linear combination of the other vectors of S. Now, we will assume that there is a vector $\vec{v}_i \in S$ that is a linear combination of the other vectors in S. Without loss of generality, suppose $\vec{v}_i = \vec{v}_1$, meaning $i = 1$. Then, there exist $c_2, \ldots, c_n \in \mathbb{F}$ such that

$$\vec{v}_1 = c_2\vec{v}_2 + \cdots + c_n\vec{v}_n.$$

Letting $c_1 = -1$, we have

$$\begin{aligned} -\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_n\vec{v}_n &= -(c_2\vec{v}_2 + \cdots + c_n\vec{v}_n) + (c_2\vec{v}_2 + \cdots + c_n\vec{v}_n) \\ &= \vec{0}. \end{aligned}$$

We constructed $c_1 \neq 0$, meaning S is linearly dependent. □

We will now consider the linear dependence of sets with one or two elements.

Theorem 4.5.3: Linear Dependence of Sets With One or Two Elements

Let $S = \{\vec{v}\}$, and we have that if $c\vec{v} = \vec{0}$ for some $c \neq 0$, S is linearly dependent. If S is linearly dependent, we know that $\vec{v} = \vec{0}$ by Theorem 4.2.4. If $\vec{v} = \vec{0}$, then $1\vec{v} = \vec{0}$ and S is linearly dependent. Therefore S is linearly dependent if and only if $\vec{v} = \vec{0}$. Now, suppose that $S = \{\vec{v}_1, \vec{v}_2\}$. If $\vec{v}_1 = \vec{0}$, S is linearly dependent because

$$1\vec{v}_1 + 0\vec{v}_2 = \vec{0}.$$

Then, if $\vec{v}_2 = \vec{0}$, S is linearly dependent because

$$0\vec{v}_1 + 1\vec{v}_2 = \vec{0}.$$

If both $\vec{v}_1$ and $\vec{v}_2$ are nonzero, and there exist $c_1, c_2 \in \mathbb{F}$, not both zero, such that

$$c_1\vec{v}_1 + c_2\vec{v}_2 = \vec{0}.$$

If $c_1 \neq 0$, $\vec{v}_1 = -\frac{c_2}{c_1}\vec{v}_2$. That is, $\vec{v}_1$ is a scalar multiple of $\vec{v}_2$. If $c_2 \neq 0$, $\vec{v}_1 = -\frac{c_1}{c_2}\vec{v}_2$. That is, $\vec{v}_1$ is a scalar multiple of $\vec{v}_2$. That means, a set of two nonzero vectors is linearly dependent if and only if the vectors are scalar multiples of each other.

Now, we will consider the linear dependence of a finite subset of a vector space V that contains the zero vector.

Theorem 4.5.4: Linear Dependence of Finite Subsets of a Vector Space Containing $\vec{0}$

Any finite subset of a vector space that contains the zero vector $\vec{0}$ is linearly dependent.

Proof. Recall that if S is a set with $|S| = 1$ or $|S| = 2$, S is linearly dependent if $\vec{0} \in S$. Now, let $S = \{\vec{v}_1, \dots, \vec{v}_n\}$, a finite subset of a vector space V and $\vec{0} \in S$. If $\vec{v}_k = \vec{0}$, we have

$$0\vec{v}_1 + \cdots + 1\vec{v}_k + \cdots + 0\vec{v}_n = \vec{0},$$

meaning that S is linearly dependent. $\square$

Consider another characterization of linear independence.

Theorem 4.5.5: Linear Independence of Nonempty Sets

A nonempty set $S = \{\vec{v}_1, \dots, \vec{v}_n\}$ is linearly independent if and only if both the following conditions hold:

1. $\vec{v}_1 \neq \vec{0}$.
2. $\forall k, 2 \leq k \leq n, \vec{v}_k \notin \operatorname{span}(\{\vec{v}_1, \dots, \vec{v}_{k-1}\})$.

Consider the following table, describing equivalent characteristics of linear independence of a set S.

S is Linearly Independent
If $\{\vec{v}_1, \dots, \vec{v}_n\} \subseteq S$ and $c_1\vec{v}_1 + \cdots + c_n\vec{v}_n = \vec{0}$, $c_1 = \cdots = c_n = 0$.
No vector in S is a finite linear combination of other vectors in S.
For every $\vec{v} \in S$, $\vec{v} \notin \operatorname{span}(S - \{\vec{v}\})$.
For every $\vec{v} \in S$, $\operatorname{span}(S - \{\vec{v}\})$ does not contain all the vectors of $\operatorname{span}(S)$.
If $S = \{\vec{v}_1, \dots, \vec{v}_n\}$, $\vec{v}_1 \neq \vec{0}$, and, for each $k \geq 2$, $\vec{v}_k \notin \operatorname{span}(\{\vec{v}_1, \dots, \vec{v}_{k-1}\})$
Every finite subset of S is linearly independent.
Every vector in $\operatorname{span}(S)$ can be uniquely expressed as a linear combination of the vectors in S.

Consider the following proofs involving linear independence.

Example 4.5.1: Proving a Property of Linear Independence 1

Suppose $S = \{\vec{v}_1, \dots, \vec{v}_n\}$ is a finite subset of a vector space V and $\vec{v} \in \operatorname{span}(S)$, with $\vec{v} \notin S$. Then, there exists some vector $\vec{w} \in T = S \cup \{\vec{v}\}$ that can be expressed in more than one way as a linear combination of vectors in T.

Proof. We take $\vec{w} = \vec{v}$. Since $\vec{v} \in \operatorname{span}(S)$, we can write

$$\vec{v} = c_1\vec{v}_1 + \cdots + c_n\vec{v}_n,$$

and therefore,

$$\vec{w} = c_1\vec{v}_1 + \cdots + c_n\vec{v}_n + 0\vec{v}.$$

for scalars $c_1, \dots, c_n \in \mathbb{F}$. We can also write

$$\vec{w} = 1\vec{v} + 0\vec{v}_1 + \cdots + 0\vec{v}_n.$$

We have now shown two distinct ways to write $\vec{w}$ as a linear combination of the elements of T. The two ways are indeed distinct, because, since $\vec{v} \notin S$, $\vec{v} \neq \vec{v}_1, \dots, \vec{v} \neq \vec{v}_n$. □

Example 4.5.2: ☢☢ Proving a Property of Linear Independence 2

For each $n \in \mathbb{N}$, the set

$$\{1, x, x^2, x^3, \dots, x^n\}$$

is a linearly independent subset of the vector space of real valued functions.

Proof. We proceed by induction on n. For the base case $n = 0$, since $|\{1\}| = 1$, the set $\{1\}$ is linearly dependent if and only if $1 = 0$, which is not the case. Thus, $\{1\}$ is linearly independent. Now, suppose that for all $k \in \mathbb{N}$, the set

$$\{1, x, x^2, x^3, \dots, x^k\}$$

is a linearly independent subset of the vector space of real valued functions. For $n = k + 1$, we have the set

$$\{1, x, x^2, x^3, \dots, x^k, x^{k+1}\}$$

and must show that this set is linearly independent. Consider the linear combination

$$c_0 + c_1x + c_2x^2 + c_3x^3 + \cdots + c_kx^k + c_{k+1}x^{k+1} = 0.$$

After differentiating both sides, we have

$$c_1 + 2c_2x + 3c_3x^2 + \cdots + kc_kx^{k-1} + (k+1)c_{k+1}x^k = 0.$$

By the inductive hypothesis, $c_1 = 2c_2 = 3c_3 = \cdots = kc_k = (k+1)c_{k+1} = 0$. Since $1, 2, 3, \dots, k, k+1$ are all not zero, this implies $c_1 = \cdots = c_{k+1} = 0$. if we substitute this result into

$$c_0 + c_1x + c_2x^2 + c_3x^3 + \cdots + c_kx^k + c_{k+1}x^{k+1} = 0,$$

we have $c_0 = 0$, so $c_0 = \cdots = c_{k+1} = 0$, meaning

$$\{1, x, x^2, x^3, \dots, x^k, x^{k+1}\}$$

is linearly independent, as desired. Thus, by induction, for all $n \in \mathbb{N}$,

$$\{1, x, x^2, x^3, \dots, x^n\}$$

is a linearly independent subset of the vector space of real valued functions. □

4.6 Lecture 25: October 24, 2022

4.6.1 Determining Linear Independence in $\mathbb{R}^n$

We now consider a method to test for linear independence using row reduction in $\mathbb{R}^n$, and present a related result on linear dependence in $\mathbb{R}^n$.

Theorem 4.6.1: A Test for Linear Independence in $\mathbb{R}^n$

Let S be a finite nonempty set of vectors in $\mathbb{R}^n$. To determine whether S is linearly independent, perform the following steps:

1. Create the matrix A whose columns are the vectors in S.
2. Solve the system $[A|\vec{0}]$ by row reduction.

Then, S is linearly independent if the system has only the trivial solution, and linearly dependent otherwise.

Theorem 4.6.2: A Test for Linear Dependence in $\mathbb{R}^n$

Suppose $S \subseteq \mathbb{R}^n$ and S contains distinct elements $\vec{v}_1, \ldots, \vec{v}_k$ where $k > n$. Then, S is linearly dependent.

Consider the following examples.

Example 4.6.1: Determining Linear Dependence in $\mathbb{R}^n$ 1

Let $S = \{[1, -1, 0, 2], [0, -2, 1, 0], [2, 0, -1, 1]\} \subseteq \mathbb{R}^4$. We build the system

$$\left[\begin{array}{ccc|c} 1 & 0 & 2 & 0 \\ -1 & -2 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 2 & 0 & 1 & 0 \end{array}\right],$$

which, by row reduction, becomes

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

We see that the system only has the trivial solution, so S is linearly independent.

Example 4.6.2: ☢☢ Determining Linear Dependence in $\mathbb{R}^n$ 2

Let $S = \{[3, 1, -1], [-5, -2, 2], [2, 2, -1]\} \subseteq \mathbb{R}^3$. We build the system

$$\left[\begin{array}{ccc|c} 3 & -5 & 2 & 0 \\ 1 & -2 & 2 & 0 \\ -1 & 2 & -1 & 0 \end{array}\right],$$

which, by row reduction, becomes

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right].$$

We see that the system only has the trivial solution, so S is linearly independent.

Example 4.6.3: ☢☢ Determining Linear Dependence in $\mathbb{R}^n$ 3

Let $S = \{[2, 1], [-1, 3], [1, 4]\} \subseteq \mathbb{R}^2$. We build the system

$$\left[\begin{array}{ccc|c} 2 & -1 & 1 & 0 \\ 1 & 3 & 4 & 0 \end{array}\right],$$

which, by row reduction, becomes

$$\left[\begin{array}{ccc|c} 1 & -\frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 1 & 1 & 0 \end{array}\right].$$

We see that the system has infinitely many solutions, since the third column gives a free variable. Thus, S is linearly dependent. Note that the same conclusion is reached by realizing that $k = |S| = 3$, $n = 2$, and $k > n$.

4.6.2 Bases: Part I

We now define the basis of a vector space.

Definition 4.6.1: Basis

Suppose V is a vector space. Then, $B \subseteq V$ is a basis of V if and only if

1. $\text{span}(B) = V$.
2. B is linearly independent.

Consider the following examples.

Example 4.6.4: The Standard Basis of $\mathbb{R}^3$

Verify that $B = \{[1,0,0],[0,1,0],[0,0,1]\}$ is a basis of $\mathbb{R}^3$.

Let A be the matrix with the elements of B as its columns. We form the linear system $[A|\vec{0}]$, or

$$\begin{bmatrix} 1 & 0 & 0 & | & 0 \\ 0 & 1 & 0 & | & 0 \\ 0 & 0 & 1 & | & 0 \end{bmatrix}$$

and realize that it is already in reduced row echelon form. The system has only the trivial solution, and so B is linearly independent. Now, we consider the span of B, or the row space of A^T. We have

$$\text{span}(B) = \{[c_1, c_2, c_3] : c_1, c_2, c_3 \in \mathbb{R}^3\} = \mathbb{R}^3.$$

Thus, B is a basis for $\mathbb{R}^3$.

Example 4.6.5: ☢☢ Another Basis of $\mathbb{R}^3$

Verify that $B = \{[2,2,2],[5,0,0],[0,-3,1]\}$ is a basis of $\mathbb{R}^3$.

Let A be the matrix with the elements of B as its columns. We form the linear system $[A|\vec{0}]$, or

$$\begin{bmatrix} 2 & 5 & 0 & | & 0 \\ 2 & 0 & -3 & | & 0 \\ 2 & 0 & 1 & | & 0 \end{bmatrix}.$$

By row reduction, we obtain

$$\begin{bmatrix} 1 & 0 & 0 & | & 0 \\ 0 & 1 & 0 & | & 0 \\ 0 & 0 & 1 & | & 0 \end{bmatrix}$$

The system has only the trivial solution, and so B is linearly independent. Now, we consider the span of B, or the row space of A^T. We don't need to compute the row reduction in this case as A is square. If A can be row reduced to I_n, $\det A = \det A^T \neq 0$, meaning that A^T can be row reduced to I_n. This means

$$\operatorname{span}(B) = \{[c_1, c_2, c_3] : c_1, c_2, c_3 \in \mathbb{R}^n\}.$$

Thus, B is a basis for $\mathbb{R}^3$.

Example 4.6.6: ☢☢☢ A Basis of $\mathcal{M}_{22}$

Verify that

$$B = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$$

is a basis of $\mathbb{R}^3$.

We start by computing $\operatorname{span}(B)$, to produce

$$\begin{aligned} \operatorname{span}(B) &= \left\{ c_1 \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + c_2 \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + c_3 \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + c_4 \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} : c_1, c_2, c_3, c_4 \in \mathbb{R} \right\} \\ &= \left\{ \begin{bmatrix} c_1 & c_2 \\ c_3 & c_4 \end{bmatrix} : c_1, c_2, c_3, c_4 \in \mathbb{R} \right\} \\ &= \mathcal{M}_{22}. \end{aligned}$$

Then, to verify linear independence, we see that if

$$c_1 \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + c_2 \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + c_3 \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + c_4 \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} c_1 & c_2 \\ c_3 & c_4 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix},$$

$c_1 = c_2 = c_3 = c_4 = 0$, meaning B is linearly independent. Thus, B is a basis for $\mathcal{M}_{22}$.

Example 4.6.7: ☢☢☢ A Basis of $\mathcal{P}_n$

Verify that $B = \{x^0, \dots, x^n\}$ is a basis of $\mathcal{P}_n$, the set of polynomials with degree at most n.

We start by computing span (B), to produce

$$\begin{aligned} \text{span}\,(B) &= \{c_0 x_0 + \cdots + c_n x^n : c_0, \dots, c_n \in \mathbb{R}^n\} \\ &= \mathcal{P}_n. \end{aligned}$$

Then, to verify linear independence, see Example 4.5.2. Thus, B is a basis for $\mathcal{P}_n$.

Example 4.6.8: ☢☢☢ A Basis of $\{\vec{0}\}$

Verify that $B = \emptyset$ is a basis of $\{\vec{0}\}$.

We note that span $(B) = \vec{0}$, by Definition 4.4.2. Then, we recall that B is linearly independent, by Definition 4.5.1. Thus, B is a basis for $\{\vec{0}\}$.

Example 4.6.9: ☢☢☢ A Basis of $\mathcal{D}_3$

Find a basis of $\mathcal{D}_3$, the set of 3×3 diagonal matrices.

Consider the set

$$B = \left\{ \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \right\}.$$

To see that $\operatorname{span}(B) = \mathcal{D}_3$, consider

$$A_1 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

The row space of A_1 is $\{[c_1, 0, 0, 0, c_2, 0, 0, 0, c_3] : c_1, c_2, c_3 \in \mathbb{R}\}$. Thus, $\operatorname{span}(B) = \mathcal{D}_3$. Then, to verify that B is linearly independent, consider

$$A_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

which is easily row reduced to a matrix with pivots in each column, so B is linearly independent and is a basis for $\mathcal{D}_3$.

4.7 Lecture 26: October 26, 2022

4.7.1 Bases: Part II

Consider the following nonexample of a basis.

Example 4.7.1: ☢☢ Not a Basis of $\mathcal{P}_2$

Verify that $B = \{1, x^2\}$ is not a basis for $\mathcal{P}_2$.

We see that $x \in \mathcal{P}_2$ but $x \notin \operatorname{span}(B)$, so B is not a basis for $\mathcal{P}_2$.

Proof. Suppose $x \in \operatorname{span}(B)$. Then,

$$x = c_1(1) + c_2(x^2)$$

for some $c_1, c_2 \in \mathbb{R}$. If $x = 0$, we have

$$0 = c_1 + c_2(0)^2 \implies c_1 = 0.$$

Therefore, we have $x = c_2x^2$. After we take the derivative of both sides, we get

$$1 = 2c_2x,$$

and when $x = 0$, that implies $1 = 0$, and we arrive at a contradiction. Hence, $x \notin \operatorname{span}(B)$. $\square$

Consider the following theorems.

Theorem 4.7.1: A Useful Lemma for Bases

Let S and T be subsets of a vector space V such that $\operatorname{span}(S) = V$ where S is finite and T is linearly independent. Then, T is finite and $|T| \leq |S|$.

Proof. If $S = \emptyset$, then we have that $V = \{\vec{0}\}$. Since $\{\vec{0}\}$ is linearly independent, $T = \emptyset$. Suppose $|S| = n \geq 1$. Suppose, for the sake of contradition, that T is infinite or $|T| > |S| = n$. Since every finite subset of T is linearly independent, there is a linearly independent set $Y \subseteq T$ such that $|Y| = n + 1$. Let $S = \{\vec{v}_1, \ldots, \vec{v}_n\}$ and let $Y = \{\vec{w}_1, \ldots, \vec{w}_n, \ldots, \vec{w}_{n+1}\}$. We will show that Y is linearly dependent. Then, we have $c_{ij} \in \mathbb{F}$ for $1 \leq i \leq n+1$ and $1 \leq j \leq n$ such that

$$\vec{w}_1 = c_{11}\vec{v}_1 + \cdots + c_{1n}\vec{v}_n, \ldots, \vec{w}_{n+1} = c_{n+1,1}\vec{v}_1 + \cdots + c_{n+1,n}\vec{v}_n.$$

Let C be the $(n+1) \times n$ matrix whose (i, j) entry is c_{ij}. Then, $[C^T|\vec{0}]$ has $n + 1$ variables with only n equations, and has a nontrivial solution $\vec{u} = [u_1, \ldots, u_{n+1}]$. Then,

$$\begin{aligned} u_1\vec{w}_1 + \cdots + u_{n+1}\vec{w}_{n+1} &= u_1(c_{11}\vec{v}_1 + \cdots + c_{1n}\vec{v}_n) + \cdots + u_{n+1}(c_{n+1,1}\vec{v}_1 + \cdots + c_{n+1,n}\vec{v}_n) \\ &= (c_{11}u_1 + \cdots + c_{n+1,1}u_{n+1})\vec{v}_1 + \cdots + (c_{n1}u_1 + \cdots + c_{n+1,n}u_{n+1})\vec{v}_n. \end{aligned}$$

But, the coefficient of each $\vec{v}_i$ in the last expression is the ith entry of $C^T\vec{u}$. Since $C^T\vec{u} = \vec{0}$, the coefficient of each $\vec{v}_i$ is 0. Then, for

$$u_1\vec{w}_1 + \cdots + u_{n+1}\vec{w}_{n+1} = \vec{0},$$

we have $u_1 = \cdots = u_{n+1} = 0$. But, since $\vec{u}$ is a nontrivial solution, at least one $u_i \neq 0$ and Y is linearly dependent, providing a contradiction. □

Theorem 4.7.2: Bases Have Equivalent Cardinality

Suppose V is a vector space and B_1 is a basis for V with finitely many elements. If B_2 is a basis for V with finitely many elements,

$$|B_1| = |B_2|.$$

Proof. Since $\operatorname{span}(B_1) = V$ and B_2 is finite and linearly independent, $|B_2| \leq |B_1|$. Since $\operatorname{span}(B_2) = V$ and B_1 is finite and linearly independent, $|B_1| \leq |B_2|$. Thus, $|B_1| = |B_2|$. □

Now, we come to the definition of the dimension of a vector space.

Definition 4.7.1: Dimension

If V is a vector space with a finite basis B,

$$\dim V = |B|.$$

If $V = \{\vec{0}\}$, $\dim V = 0$. Otherwise, $\dim V$ is infinite.

Consider the following theorem.

Theorem 4.7.3: General Statements About Span and Dimension 1

Suppose V is a vector space. If $S \subseteq V$ is finite and $\operatorname{span}(S) = V$, then, $\dim V \leq |S|$.

Proof. See Theorem 4.7.1 and Definition 4.7.1. □

Theorem 4.7.4: General Statements About Span and Dimension 2

Suppose V is a vector space. If $S \subseteq V$ is finite and $\operatorname{span}(S) = V$, then, S is a basis if and only if $\dim V = |S|$.

Proof. Suppose S is a basis. Then, by Definition 4.7.1, $\dim V = |S|$. If $\dim V = |S|$, let $n = |S| = \dim V$ so $S = \{\vec{v}_1, \dots, \vec{v}_n\}$. Since $\operatorname{span}(S) = V$, we need only show that S is linearly independent. Suppose S is linearly dependent, then, there exists $\vec{v}_i$ such that

$$\operatorname{span}(\{\vec{v}_1, \dots, \vec{v}_{i-1}, \vec{v}_{i+1}, \vec{v}_n\}) = V.$$

This contradicts $\dim V = n$. □

Theorem 4.7.5: General Statements About Linear Independence and Dimension 1

Suppose V is a vector space. If $T \subseteq V$ is linearly independent, then, $|T| \leq \dim V$.

Proof. Let S be a basis of V. Thus, $\operatorname{span}(S) = V$ and $\dim V = |S|$. By Theorem 4.7.1, if T is linearly independent, $|T| \leq |S|$, so $|T| \leq \dim V$. □

Theorem 4.7.6: General Statements About Linear Independence and Dimension 2

Suppose V is a vector space. If $T \subseteq V$ is linearly independent and finite, T is a basis for V if and only if $\dim V = |T|$.

Proof. Suppose T is a basis. Then, by Definition 4.7.1, $\dim V = |T|$. If $\dim V = |T|$, Let $n = |T| = \dim V$ so $T = \{\vec{v}_1, \dots, \vec{v}_n\}$. Since T is linearly independent, we need only show that $\operatorname{span}(T) = V$. Suppose not, then, there exists some $\vec{v} \in V$ where $\vec{v} \notin \operatorname{span}(T)$. Consider the set $T \cup \{\vec{v}\}$. We see that

$$|T \cup \{\vec{v}\}| = |T| + 1 = \dim V + 1$$

since $\vec{v} \notin \operatorname{span}(T)$, and because $T \subseteq \operatorname{span}(T)$, $\vec{v} \notin T$. We also see that $T \cup \{\vec{v}\}$ is linearly independent by considering

$$\vec{0} = c_1\vec{v}_1 + \cdots + c_n\vec{v}_n + c_{n+1}\vec{v}$$

for $c_1, \dots, c_{n+1} \in \mathbb{F}$ and $\vec{v}_1, \dots, \vec{v}_k \in T$. We have that $c_{n+1} = 0$ since, otherwise,

$$\vec{v} = -\frac{1}{c_{n+1}}(c_1\vec{v}_1 + \cdots + c_n\vec{v}_n).$$

Since we've written $\vec{v}$ as a linear combination of the elements of T, $\vec{v} \in \operatorname{span}(T)$, which is contradictory. We also have

$$\vec{0} = c_1\vec{v}_1 + \cdots + c_n\vec{v}_n$$

since T is linearly independent and $\vec{v}_1, \dots, \vec{v}_n \in T$. Thus, $c_1 = \cdots = c_n = c_{n+1} = 0$. By the previous part, we have $|T \cup \{\vec{v}\}| \leq \dim V$. However, the left hand side is $\dim V + 1$, which is contradictory. Thus, $\operatorname{span}(T) = V$, so T is a basis for V. $\square$

4.8 Lecture 27: October 28, 2022

4.8.1 Bases: Part III

Consider the following theorems.

Theorem 4.8.1: Dimensions of Subspaces

If W is a subspace of a finite dimensional vector space V,

$$\dim W \leq \dim V.$$

Moreover, if $\dim W = \dim V$, then, $W = V$.

Proof. Suppose that W is the trivial subspace, that is, $W = \{\vec{0}\}$. Then, $\dim W = 0 \leq \dim V$. Othersie, we take some $\vec{w}_1 \in W$ where $\vec{w}_1 \neq \vec{0}$. Then, $\{\vec{w}_1\}$ is linearly independent. If $\operatorname{span}(\{\vec{w}_1\}) \neq W$, we can take some $\vec{w}_2 \in (W - \operatorname{span}(\{\vec{w}_1\}))$. Then, $\{\vec{w}_1, \vec{w}_2\}$ is linearly independent. Until our set spans W, we can keep adding $\vec{w}_{k+1} \in (W - \operatorname{span}(\{\vec{w}_1, \dots, \vec{w}_k\}))$ to eventually obtain a linearly independent spanning set

$$\{\vec{w}_1, \dots, \vec{w}_{k+1}\} \in W.$$

Since this set is linearly independent in W, it is also linearly independent in V, meaning $|\{\vec{w}_1, \dots, \vec{w}_{k+1}\}| \leq \dim V$, as desired. For the next part of the theorem, suppose $\dim W = \dim V = n$. Then, let $B_W = \{\vec{w}_1, \dots, \vec{w}_n\}$ and $B_V = \{\vec{v}_1, \dots, \vec{v}_n\}$. Suppose, for the sake of contradiction, $W \neq V$. That is, suppose that $\operatorname{span}(B_W) = W \neq V$. Then, there exists some $\vec{v} \in V$ with $\vec{v} \notin \operatorname{span}(B_W)$. Then, $B_W \cup \{\vec{v}\}$ is linearly independent in V, and by construction, $\operatorname{span}(B_W \cup \{\vec{v}\}) = V$. But, $|B_W \cup \{\vec{v}\}| = n + 1$, and an n-dimensional space cannot have a basis with $n + 1$ elements, providing our contradiction. Thus, $W = V$, as desired. □

Theorem 4.8.2: Diagonalizability, Revisited: Part I

Suppose $A \in \mathcal{M}_{nn}$. Then, A is diagonalizable if and only if there exists a basis of $\mathbb{R}^n$ that consists of eigenvectors of A.

Proof. Recall that by Theorem 3.6.1, $A \in \mathcal{M}_{nn}$ is diagonalizable if and only if there exists a set of n linearly indpendent eigenvectors $\vec{v}_1, \dots, \vec{v}_n$. Let $V = \mathbb{R}^n$ and $W = \{\vec{v}_1, \dots, \vec{v}_n\}$. Then, $\dim V = n$. We also see that $W \subseteq V$ is linearly independent and finite with $|W| = \dim V$. By Theorem 4.7.6, W is a basis for V. Our proof is reversible, as all implications are bidirectional. □

We now posit two questions. Given $S \subseteq V$, where V is a vector space, how would one construct a basis for $\text{span}(S)$? Similarly, given a linearly independent set S where $S \subseteq V$ and V is a vector space, how can we expand S to form a basis of V? Consider the following theorems.

Theorem 4.8.3: Finding a Basis for $\text{span}(S)$ by Contraction

If $\text{span}(S) = V$, where V is a finite dimensional vector space, there exists some $B \subseteq S$ where B is a basis for V. Given $S = \{\vec{v}_1, \dots, \vec{v}_k\} \subseteq \mathbb{R}^n$. Consider the following steps.

1. We form the matrix A with the elements of S as columns.
2. We row reduce A into reduced row echelon form to obtain a matrix C.
3. The basis of $\text{span}(S)$, B, is formed by removing the vectors in A associated with the columns containing free variables.

Note that the vectors in B are in S.

Note that we can also construct a basis for $\text{span}(S)$ by forming a matrix with the elements of S as rows and finding the row space. In general though, the basis will not contain vectors in S.

Theorem 4.8.4: Finding a Basis by Expansion

Let T be a linearly independent subset of a finite dimensional vector space V. Then, V has a basis B with $T \subseteq B$. If $T = \{\vec{t}_1, \dots, \vec{t}_k\} \subseteq V$, perform the following steps to find B:

1. Find a spanning set $A = \{\vec{a}_1, \dots, \vec{a}_n\}$. For $\mathbb{R}^n$, this will often be $\{\vec{e}_1, \dots, \vec{e}_n\}$.
2. Form the ordered spanning set $S = \{\vec{t}_1, \dots, \vec{t}_k, \vec{a}_1, \dots, \vec{a}_n\}$ for V.
3. Use Theorem 4.8.3 to find a basis B for V, containing T.

Consider the following examples.

Example 4.8.1: ✣✣ Finding a Basis 1

Let $S = \{[1,2,0,1],[3,1,5,-7],[-2,4,-8,14]\}$. We form the matrix A with the elements of S as columns, giving

$$A = \begin{bmatrix} 1 & 3 & -2 \\ 2 & 1 & 4 \\ 0 & 5 & -8 \\ 1 & -7 & 14 \end{bmatrix}.$$

By row reduction, we get

$$C = \begin{bmatrix} 1 & 0 & \frac{14}{5} \\ 0 & 1 & -\frac{8}{5} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Now, we remove the vectors associated with the columns containing free variables to get the basis. Our basis is then

$$B = \{[1,2,0,1],[3,1,5,-7]\}.$$

Example 4.8.2: ✣✣ Finding a Basis 2

Let $S = \{x^3 - 3x^2 + 2, 2x^3 - 7x^2 + x - 3, 4x^3 - 13x^2 + x + 5\}$. We instead equivalently consider $S' = \{[2,0,-3,1],[-3,1,-7,2],[5,1,-13,4]\}$. We form the matrix A with the elements of S' as columns, giving

$$A = \begin{bmatrix} 1 & -3 & 5 \\ 0 & 1 & 1 \\ -3 & -7 & -13 \\ 1 & 2 & 4 \end{bmatrix}.$$

By row reduction, we get

$$C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

There are no vectors to remove and we have $B = S$.

Example 4.8.3: ☢☢ Finding a Basis 3

Let $S = \{x^3-8x^2+1, 3x^3-2x^2+x, 4x^3+2x-10, x^3-20x^2-x+12, x^3+24x^2+2x-13\}$. We instead equivalently consider $S' = \{[1,0,-8,1],[0,1,-2,3],[-10,2,0,4],[12,-1,-20,1],[-13,2,24,1]\}$. We form the matrix A with the elements of S' as columns, giving

$$A = \begin{bmatrix} 1 & 0 & -10 & 12 & -13 \\ 0 & 1 & 2 & -1 & 2 \\ -8 & -2 & 0 & -20 & 24 \\ 1 & 3 & 4 & 1 & 1 \end{bmatrix}.$$

By row reduction, we get

$$C = \begin{bmatrix} 1 & 0 & 0 & 0 & -3 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}.$$

We remove the last vector to get $B = \{x^3-8x^2+1, 3x^3-2x^2+x, 4x^3+2x-10, x^3-20x^2-x+12\}$.

Example 4.8.4: ☢☢ Finding a Basis 4

Let $T = \{[1,2,0,1],[3,1,5,-7]\}$. We consider

$$S = \{[1,2,0,1],[3,1,5,-7],[1,0,0,0],[0,1,0,0],[0,0,1,0],[0,0,0,1]\}$$

Note that S is the union of the standard basis of $\mathbb{R}^4$ and T. Since

$$\{[1,0,0,0],[0,1,0,0],[0,0,1,0],[0,0,0,1]\}$$

is a basis of $\mathbb{R}^4$, $\operatorname{span}(S) = \mathbb{R}^4$. Now, we form

$$A = \begin{bmatrix} 1 & 3 & 1 & 0 & 0 & 0 \\ 2 & 3 & 0 & 1 & 0 & 0 \\ 0 & 5 & 0 & 0 & 1 & 0 \\ 1 & -7 & 0 & 0 & 0 & 1 \end{bmatrix}$$

and row reduce to

$$C = \begin{bmatrix} 1 & 0 & 0 & 0 & \frac{7}{5} & 1 \\ 0 & 1 & 0 & 0 & \frac{1}{5} & 0 \\ 0 & 0 & 1 & 0 & -2 & -1 \\ 0 & 0 & 0 & 1 & -3 & 2 \end{bmatrix}.$$

Hence, our basis for $\mathbb{R}^4$ is

$$B = \{[1,2,0,1],[3,1,5,-7],[1,0,0,0],[0,1,0,0]\}.$$

Example 4.8.5: ☢☢ Finding a Basis 5

Let $T = \{[2, 0, 4, -12], [0, -1, -3, 9]\}$. We consider

$$S = \{[2, 0, 4, -12], [0, -1, -3, 9], [1, 0, 0, 0], [0, 1, 0, 0], [0, 0, 1, 0], [0, 0, 0, 1]\}$$

Note that S is the union of the standard basis of $\mathbb{R}^4$ and T. Since

$$\{[1, 0, 0, 0], [0, 1, 0, 0], [0, 0, 1, 0], [0, 0, 0, 1]\}$$

is a basis of $\mathbb{R}^4$, span $(T) = \mathbb{R}^4$. Now, we form

$$A = \begin{bmatrix} 2 & 0 & 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 1 & 0 & 0 \\ 4 & -3 & 0 & 0 & 1 & 0 \\ -12 & 9 & 0 & 0 & 0 & 1 \end{bmatrix}$$

and row reduce to

$$C = \begin{bmatrix} 1 & 0 & 0 & -\frac{3}{4} & 0 & -\frac{1}{12} \\ 0 & 1 & 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & \frac{3}{2} & 0 & \frac{1}{6} \\ 0 & 0 & 0 & 0 & 1 & \frac{1}{3} \end{bmatrix}.$$

Hence, our basis for $\mathbb{R}^4$ is

$$B = \{[2, 0, 4, -12], [0, -1, -3, 9], [1, 0, 0, 0], [0, 0, 1, 0]\}.$$

We end this section with two examples on finding bases of “weird” sets.

Example 4.8.6: ☢☢ Finding a Weird Basis 1

Find the basis and the dimension of $V = \{[x, y, z, w] : x, y, z, w \in \mathbb{R}, x - y - w + z = 0\}$.

The membership condition of the set indicates that $w = x - y + z$. We then have

$$V = \{[x, y, z, x - y + z] : x, y, z \in \mathbb{R}\}.$$

We hypothesize that the dimension of V will be 3, because there are only three variables making up the set; however, we will confirm this after we have found the basis. We form, somewhat arbitrarily, the set

$$S = \{[1, 1, 1, 1], [1, 1, 2, 2], [1, 2, 1, 0]\}.$$

To find $\operatorname{span}(S)$, we consider

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 2 & 2 \\ 1 & 2 & 1 & 0 \end{bmatrix} \underbrace{\rightarrow}_{\text{RREF}} \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 1 \end{bmatrix}.$$

Thus,

$$\operatorname{span}(S) = \{[c_1, c_2, c_3, c_1 - c_2 + c_3] : c_1, c_2, c_3 \in \mathbb{R}\} = V.$$

Since $\operatorname{span}(S) = V$, we check if S is linearly independent by considering

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 2 \\ 1 & 2 & 1 \\ 1 & 2 & 0 \end{bmatrix} \underbrace{\rightarrow}_{\text{RREF}} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

All columns have pivots, so S is linearly independent, spans V, and is therefore a basis for V. Because there are 3 elements in S, $\dim V = 3$.

Example 4.8.7: ☢☢ Finding a Weird Basis 2

Find the basis and the dimension of

$$V = \{a_0 + a_1x + a_2x^2 + a_3x^3 : a_0, a_1, a_2, a_3 \in \mathbb{R}, 2a_0 + a_1 = 0, a_2 - 2a_3 = 0\}$$

The membership condition of the set indicates that $a_1 = -2a_0$ and $a_3 = \frac{1}{2}a_2$. We then have

$$V = \{a_0 - 2a_0x + a_2x^2 + \frac{1}{2}a_2x^3 : a_0, a_2 \in \mathbb{R}\}.$$

We hypothesize that the dimension of V will be 2, because there are only two variables making up the set; however, we will confirm this after we have found the basis. We form, somewhat arbitrarily, the set

$$S = \{1 - 2x + 2x^2 + x^3, 1 - 2x + 4x^2 + 2x^3\}.$$

Consider $S' = \{[1, -2, 2, 1], [1, -2, 4, 2]\}$. To find span (S'), we consider

$$\begin{bmatrix} 1 & -2 & 2 & 1 \\ 1 & -2 & 4 & 2 \end{bmatrix} \underbrace{\rightarrow}_{\text{RREF}} \begin{bmatrix} 1 & -2 & 0 & 0 \\ 0 & 0 & 1 & \frac{1}{2} \end{bmatrix}$$

so span $(S') = \{[c_1, -2c_1, c_2, \frac{1}{2}c_2] : c_1, c_2 \in \mathbb{R}\}$ and

$$\text{span}\,(S) = \left\{c_1 - 2c_1x + c_2x^2 + \frac{1}{2}c_2x^3\right\} = V.$$

Since span $(S) = V$, we check if S is linearly independent by considering

$$\begin{bmatrix} 1 & 1 \\ -2 & -2 \\ 2 & 4 \\ 1 & 2 \end{bmatrix} \underbrace{\rightarrow}_{\text{RREF}} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

All columns have pivots, so S is linearly independent, spans V, and is therefore a basis for V. Since there are 2 elements in S, dim $V = 2$.

4.9 Lecture 28: October 31, 2022

4.9.1 Coordinatization: Part I

Consider the following definition.

Definition 4.9.1: Ordered Bases

An ordered basis of an n dimensional vector space V is an n-tuple of vectors $(\vec{v}_1, \dots, \vec{v}_n)$ such that the set $\{\vec{v}_1, \dots, \vec{v}_n\}$ is a basis for V.

Consider the following example.

Example 4.9.1: Ordered Bases of $\mathbb{R}^3$

The tuples

$$B_1 = ([1, 0, 0], [0, 1, 0], [0, 0, 1])$$

and

$$B_2 = ([0, 1, 0], [1, 0, 0], [0, 0, 1])$$

are two distinct ordered bases of $\mathbb{R}^3$.

The concept of ordering allows us to define coordinates with respect to a basis.

Definition 4.9.2: Coordinates With Respect to a Basis

Let $B = (\vec{v}_1, \dots, \vec{v}_n)$ be an ordered basis for a vector space V. Suppose that

$$\vec{w} = c_1\vec{v}_1 + \cdots + c_n\vec{v}_n \in V$$

for $c_1, \dots, c_n \in \mathbb{F}$. Then,

$$[\vec{w}]_B = [c_1, \dots, c_n].$$

The quantity $[\vec{w}]_B$ is the coordinatization of $\vec{w}$ with respect to B, or equivalently, "$\vec{w}$ expressed in B coordinates."

Consider the following examples.

Example 4.9.2: Coordinatization 1

Let $\vec{w} = [1, 2, 3] \in \mathbb{R}^3$. Then,

$$[\vec{w}]_{B_1=([1,0,0],[0,1,0],[0,0,1])} = [1, 2, 3].$$

Example 4.9.3: ☢ Coordinatization 2

Let $\vec{w} = [1, 2, 3] \in \mathbb{R}^3$. Then,

$$[\vec{w}]_{B_2=([0,1,0],[1,0,0],[0,0,1])} = [2, 1, 3].$$

Example 4.9.4: ☢☢ Coordinatization 3

Find the matrix associated to

$$[-2, 6, -1, -11]_B$$

where

$$B = \left(\begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \right).$$

We have the matrix

$$-2 \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix} + 6 \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} - 1 \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} - 11 \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 4 & -9 \\ -5 & -1 \end{bmatrix}.$$

We now present the general algorithm for find the coordinatization of a vector with respect to an ordered basis.

Theorem 4.9.1: ⛔ Coordinatization

Let V be a nontrivial subspace of $\mathbb{R}^n$ and let $B = (\vec{v}_1, \ldots, \vec{v}_k)$ be an ordered basis for V, and let $\vec{v} \in \mathbb{R}^n$. The following steps will find $[\vec{v}]_B$, if it exists:

1. Form the system $[A|\vec{v}]$ by using the vectors in B as the columns for A, in order. Use $\vec{v}$ as a column as well.
2. Row reduce to obtain the system $[C|\vec{w}]$ in reduced row echelon form.
3. If the system has no solutions, $\vec{v} \notin \operatorname{span}(B)$ and coordinatization is not possible. We note that B is not a basis if this happens.
4. Otherwise, $\vec{v} \in \operatorname{span}(B) = V$. Eliminate all zeroes consisting solely of zeroes in $[C|\vec{w}]$ to obtain $[I_k|\vec{y}]$. Then, $[\vec{v}]_B = \vec{y}$.

Theorem 4.9.2: ⛔ Properties of Coordinatization

Let $B = (\vec{v}_1, \ldots, \vec{v}_n)$ be an ordered basis of V and $\vec{w}_1, \ldots, \vec{w}_k \in V$ and $c_1, \ldots, c_k \in \mathbb{F}$. Then,

1. $[\vec{w}_1 + \vec{w}_2]_B = [\vec{w}_1]_B + [\vec{w}_2]_B$.
2. $[c_1\vec{w}_1]_B = c_1[\vec{w}_1]_B$.
3. $[c_1\vec{w}_1 + \cdots + c_k\vec{w}_k]_B = c_1[\vec{w}_1]_B + \cdots + c_k[\vec{w}_k]_B$.

Example 4.9.5: ☢☢☢ Coordinatization 4

Let B be an ordered basis of a subspace of $\mathbb{R}^n$ where

$$B = ([2, 6, -4, 2], [-2, -4, -8, 1], [0, 1, -6, -2]).$$

Let $\vec{v} = [-2, -4, -8, -2]$. Find $[\vec{v}]_B$.

To find $[\vec{v}]_B$, we form

$$\begin{bmatrix} 2 & -2 & 0 & | & -2 \\ 6 & -4 & 1 & | & -4 \\ -4 & -8 & -6 & | & -8 \\ 2 & 1 & -2 & | & -2 \end{bmatrix}$$

and row reduce to get

$$\begin{bmatrix} 1 & 0 & 0 & | & -\frac{3}{7} \\ 0 & 1 & 0 & | & \frac{4}{7} \\ 0 & 0 & 1 & | & \frac{6}{7} \\ 0 & 0 & 0 & | & 0 \end{bmatrix}.$$

Thus, $[\vec{v}]_B = \left[-\frac{3}{7}, \frac{4}{7}, \frac{6}{7}\right]$.

Example 4.9.6: ☢☢☢ Coordinatization 5

Let B be an ordered basis of a subspace of $\mathbb{R}^n$ where

$$B = ([1, 0, 16, 0], [0, 1, -6, 0], [0, 0, 0, 1]).$$

Let $\vec{v} = [-2, -4, -8, -2]$. Find $[\vec{v}]_B$.

To find $[\vec{v}]_B$, we form

$$\begin{bmatrix} 1 & 0 & 0 & | & -2 \\ 0 & 1 & 0 & | & -4 \\ 16 & -6 & 0 & | & -8 \\ 0 & 0 & 1 & | & -2 \end{bmatrix}$$

and row reduce to get

$$\begin{bmatrix} 1 & 0 & 0 & | & -2 \\ 0 & 1 & 0 & | & -4 \\ 0 & 0 & 1 & | & -2 \\ 0 & 0 & 0 & | & 0 \end{bmatrix}.$$

Thus, $[\vec{v}]_B = [-2, -4, -2]$.

4.10 Lecture 29: November 2, 2022

4.10.1 Coordinatization: Part II

We now posit a question: Given two ordered bases B_1 and B_2, how are $[\vec{v}]_{B_1}$ and $[\vec{v}]_{B_2}$ related? Consider the following definition and theorem.

Definition 4.10.1: Transition Matrices

Suppose that V is a nontrivial vector space with $\dim V = n$. Let B_1 and B_2 be ordered bases of V. Let P be the $n \times n$ matrix whose ith column, for $1 \leq i \leq n$, equals $[\vec{b}_{1i}]_{B_2}$, where $\vec{b}_{1i}$ is the ith basis vector in B_1. Then, P is called the transition matrix from B_1 to B_2.

Theorem 4.10.1: Finding a Transition Matrix

Given $B_1 = (\vec{v}_1, \dots, \vec{v}_n)$ and $B_2 = (\vec{w}_1, \dots, \vec{w}_n)$. To go from B_1 to B_2, we form

$$\begin{bmatrix} \vec{w}_1 & \cdots & \vec{w}_n & | & \vec{v}_1 & \cdots & \vec{v}_n \end{bmatrix}$$

and row reduce to get $[I_n|P]$. The matrix P is the transition matrix from the coordinatization with respect to B_1 to the coordinatization with respect to B_2. Then,

$$P[\vec{v}]_{B_1} = [\vec{v}]_{B_2}.$$

Consider the following example.

Example 4.10.1: Finding a Transition Matrix

Let

$$B_1 = ([1,1,0,1],[1,0,0,1],[0,1,0,1],[1,1,1,0]), \quad B_2 = ([1,0,0,0],[1,1,0,0],[1,1,0,1],[1,1,1,1]).$$

Find the associated transition matrix P.

We form

$$\begin{bmatrix} 1 & 1 & 1 & 1 & | & 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & | & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & | & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & | & 1 & 1 & 1 & 0 \end{bmatrix}$$

and row reduce to get

$$\begin{bmatrix} 1 & 0 & 0 & 0 & | & 0 & 1 & -1 & 0 \\ 0 & 1 & 0 & 0 & | & 0 & -1 & 0 & 1 \\ 0 & 0 & 1 & 0 & | & 1 & 1 & 1 & -1 \\ 0 & 0 & 0 & 1 & | & 0 & 0 & 0 & 1 \end{bmatrix}.$$

Thus,

$$P = \begin{bmatrix} 0 & 1 & -1 & 0 \\ 0 & -1 & 0 & 1 \\ 1 & 1 & 1 & -1 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Example 4.10.2: ☢☢ Checking a Transition Matrix

Check that the P found in Example 4.10.1 yields the correct result for

$$\vec{v}_1 = [1, 1, 0, 1].$$

We see that $[\vec{v}_1]_{B_1} = [1, 0, 0, 0]$. Then, we have

$$\begin{aligned} P[\vec{v}_1]_{B_1} &= \begin{bmatrix} 0 & 1 & -1 & 0 \\ 0 & -1 & 0 & 1 \\ 1 & 1 & 1 & -1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \\ &= \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \\ &= [\vec{v}_1]_{B_2}, \end{aligned}$$

as desired.

Consider the following theorem.

Theorem 4.10.2: Properties of Transition Matrices

Suppose B_1, B_2, and B_3 are ordered bases of $\mathbb{R}^n$. Then,

1. The transition matrix from B_1 to B_2 is unique.

   Proof. Suppose we have $P[\vec{v}]_{B_1} = Q[\vec{v}]_{B_1} = [\vec{v}]_{B_2}$. Then, we have $P[\vec{v}_i]_{B_1} = Q[\vec{v}_i]_{B_1}$ for each $\vec{v}_i \in B_1$. Also, note that for each $\vec{v}_i \in B_1$, $[\vec{v}_i]_{B_1} = \vec{e}_i$, so we have

   $$P\vec{e}_i = Q\vec{e}_i.$$

   For any matrix A, we have that $A\vec{e}_i$ is the ith column of A. Therefore, each column of P is equal to the respective corresponding column of Q, so $P = Q$, as desired. □

2. If P is the transition matrix from B_1 to B_2 and Q is the transition matrix from B_2 to B_3, QP is the transition matrix from B_1 to B_3.

   Proof. We consider

   $$\begin{aligned} QP([\vec{v}_{B_1}]) &= Q(P[\vec{v}]_{B_1}) \\ &= Q[\vec{v}]_{B_2} \\ &= [\vec{v}]_{B_3}. \end{aligned}$$

   Then, uniqueness implies we have the correct transition matrix QP. □

3. If P is the transition matrix from B_1 to B_2, P^{-1} is the transition matrix from B_2 to B_1.

   Proof. Uniqueness implies that the transition matrix from B_1 to B_1 is I_n. By existence, we have P, the transition matrix from B_1 to B_2 and Q, the transition matrix from B_2 to B_1. We know that

   $$QP = I_n,$$

   which only occurs if $Q = P^{-1}$. □

We now revisit diagonalization, yet again.

Theorem 4.10.3: Diagonalizability, Revisited: Part II

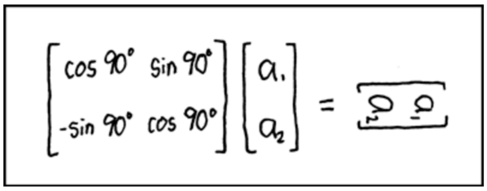
Suppose $A \in \mathcal{M}_{nn}$ and we have found nonsingular P with

$$D = P^{-1}AP.$$

To find P, we have

$$P = \begin{bmatrix} \vec{v}_1 & \cdots & \vec{v}_n \end{bmatrix}.$$

By Theorem 4.8.2, $B = (\vec{v}_1, \dots, \vec{v}_n)$ is an ordered basis of $\mathbb{R}^n$. Then, P is the transition matrix from B to the standard basis.



5 Linear Transformations

5.1 Lecture 30: November 4, 2022

5.1.1 An Introduction to Linear Transformations

Before proceeding into linear transformations, for a review of functions and associated terminology, consult Appendix B. Consider the following definition.

Definition 5.1.1: Linear Transformations

Let V and W be vector spaces. Let $F : V \to W$ be a function. Then, F is a linear transformation if and only if both the following conditions hold:

1. $\forall \vec{v}_1, \vec{v}_2 \in V, F(\vec{v}_1 + \vec{v}_2) = F(\vec{v}_1) + F(\vec{v}_2)$.
2. $\forall c \in \mathbb{F}, \forall \vec{v} \in V, F(c\vec{v}) = cF(\vec{v})$.

We remark that a linear transformation "preserves" the operations that give structure to the vector spaces involved: vector addition and scalar multiplication.

Consider the following examples.

Example 5.1.1: ☢☢ Is it a Linear Transformation? 1

Let $F : \mathcal{M}_{mn} \to \mathcal{M}_{nm}$ where $F(A) = A^T$. Is F a linear transformation?

For matrices $A_1, A_2 \in \mathcal{M}_{mn}$ and scalar $c \in \mathbb{R}$, we have

$$\begin{aligned} F(A_1 + A_2) &= (A_1 + A_2)^T \\ &= A_1^T + A_2^T \\ &= F(A_1) + F(A_2) \end{aligned}$$

and

$$\begin{aligned} F(cA_1) &= (cA_1)^T \\ &= cA_1^T \\ &= cF(A_1). \end{aligned}$$

Thus, F is a linear transformation.

Example 5.1.2: ☢☢ Is it a Linear Transformation? 2

Let $F : \mathcal{P}_n \to \mathcal{P}_{n-1}$ where $F(\vec{p}) = \vec{p}\,'$, the derivative of $\vec{p}$. Is F a linear transformation?

For $\vec{p}_1, \vec{p}_2 \in \mathcal{P}_n$, we know, from Calculus, that the derivative of a sum is the sum of the derivatives, so

$$\begin{aligned} F(\vec{p}_1 + \vec{p}_2) &= (\vec{p}_1 + \vec{p}_2)' \\ &= \vec{p}_1' + \vec{p}_2' \\ &= F(\vec{p}_1) + F(\vec{p}_2). \end{aligned}$$

For $c \in \mathbb{R}$, the constant multiple rule, from Calculus, tells us that

$$\begin{aligned} F(c\vec{p}_1) &= (c\vec{p}_1)' \\ &= c\vec{p}_1\,' \\ &= cF(\vec{p}_1). \end{aligned}$$

Thus, F is a linear transformation.

Example 5.1.3: ☢☢ Is it a Linear Transformation? 3

Let $F : \mathcal{P}_n \to W$ where $W = \text{span}\left(\left\{\frac{1}{s}, \dots, \frac{1}{s^{n+1}}\right\}\right)$ and

$$\begin{aligned} F(\vec{p}) &= \mathcal{L}\left\{\vec{p}(t)\right\}(s) \\ &= \int_0^\infty e^{-st}\vec{p}(t)\,\mathrm{d}t. \end{aligned}$$

Is F a linear transformation?

For $\vec{p}_1(t), \vec{p}_2(t) \in \mathcal{P}_n$, we have

$$\begin{aligned} F(\vec{p}_1(t) + \vec{p}_2(t)) &= \int_0^\infty e^{-st}(\vec{p}_1(t) + \vec{p}_2(t))\,\mathrm{d}t \\ &= \int_0^\infty e^{-st}\vec{p}_1(t) + e^{-st}\vec{p}_2(t)\,\mathrm{d}t \\ &= \int_0^\infty e^{-st}\vec{p}_1(t)\,\mathrm{d}t + \int_0^\infty e^{-st}\vec{p}_2(t)\,\mathrm{d}t \\ &= F(\vec{p}_1(t)) + F(\vec{p}_2(t)). \end{aligned}$$

For $c \in \mathbb{R}$, we have

$$\begin{aligned} F(c\vec{p}_1(t)) &= \int_0^\infty ce^{-st}\vec{p}_1(t)\,\mathrm{d}t \\ &= c\int_0^\infty e^{-st}\vec{p}_1(t)\,\mathrm{d}t \\ &= cF(\vec{p}_1(t)). \end{aligned}$$

Thus, F is a linear transformation.

Example 5.1.4: ☢☢ Is it a Linear Transformation? 4

Let V be a vector space with $\dim V = n$. Let B be an ordered basis for V. Then, every $\vec{v} \in V$ has coordinatization $[\vec{v}]_B$ with respect to B. Consider the function $F : V \to \mathbb{R}^n$ given by

$$F(\vec{v}) = [\vec{v}]_B.$$

Is F a linear transformation?

For $\vec{v}_1, \vec{v}_2 \in V$, we have

$$\begin{aligned} F(\vec{v}_1 + \vec{v}_2) &= [\vec{v}_1 + \vec{v}_2]_B \\ &= [\vec{v}_1]_B + [\vec{v}_2]_B \\ &= F(\vec{v}_1) + F(\vec{v}_2), \end{aligned}$$

by Theorem 4.9.2. Then, for $c \in \mathbb{R}$, we have

$$\begin{aligned} F(c\vec{v}_1) &= [c\vec{v}_1]_B \\ &= c[\vec{v}_1]_B \\ &= cF(\vec{v}), \end{aligned}$$

also by Theorem 4.9.2. Thus, F is a linear transformation.

We now state some properties of linear transformations.

Theorem 5.1.1: Properties of Linear Transformations

Let V and W be vector spaces, and let $L : V \to W$ be a linear transformation. Let $\vec{0}_V$ be the zero vector in V and $\vec{0}_W$ be the zero vector in W. Then,

1. $L(\vec{0}_V) = L(\vec{0}_W)$.

   Proof. Consider $L(\vec{0}_V) = L(0\vec{0}_V) = 0L(\vec{0}_V) = \vec{0}_W$, as desired. □

2. $L(-\vec{v}) = -L(\vec{v})$.

   Proof. Consider $L(-\vec{v}) = L(-1\vec{v}) = -L(\vec{v})$, as desired. □

3. $L(c_1\vec{v}_1 + \cdots + c_n\vec{v}_n) = c_1L(\vec{v}_1) + \cdots + c_nL(\vec{v}_n)$ for $c_1, \dots, c_n \in \mathbb{F}$ and $\vec{v}_1, \dots, \vec{v}_n \in V$ with $n \geq 2$.

   Proof. We proceed by induction. For the base case when $n = 2$, we have

   $$\begin{aligned} L(c_1\vec{v}_1 + c_2\vec{v}_2) &= L(c_1\vec{v}_1) + L(c_2\vec{v}_2) \\ &= c_1L(\vec{v}_1) + c_2L(\vec{v}_2). \end{aligned}$$

   Then, suppose that for all $n = k$,

   $$L(c_1\vec{v}_1 + \cdots + c_k\vec{v}_k) = c_1L(\vec{v}_1) + \cdots + c_kL(\vec{v}_k).$$

   Then, we have

   $$\begin{aligned} L(c_1\vec{v}_1 + \cdots + c_k\vec{v}_k + c_{k+1}v_{k+1}) &= c_1L(\vec{v}_1) + \cdots + c_kL(\vec{v}_k) + L(c_{k+1}\vec{v}_{k+1}) \\ &= c_1L(\vec{v}_1) + \cdots + c_kL(\vec{v}_k) + c_{k+1}L(\vec{v}_{k+1}), \end{aligned}$$

   as desired. □

We remark that not every function between vector spaces is a linear transformation. To show that some function between vector spaces is not a linear transformation, we must show a counterexample of the conditions in Definition 5.1.1.

We now turn to compositions of linear transformations.

Theorem 5.1.2: Compositions of Linear Transformations

Let V_1, V_2, and V_3 be vector spaces and $L_1 : V_1 \to V_2$ and $L_2 : V_2 \to V_3$ be linear transformations. Then, $(L_2 \circ L_1) : V_1 \to V_3$ with $(L_2 \circ L_1)(\vec{v}) = L_2(L_1(\vec{v}))$ is a linear transformation for $\vec{v} \in V_1$.

Proof. For $\vec{v}_1, \vec{v}_2 \in V_1$, we have

$$\begin{aligned}(L_2 \circ L_1)(\vec{v}_1 + \vec{v}_2) &= L_2(L_1(\vec{v}_1 + \vec{v}_2)) \\ &= L_2(L_1(\vec{v}_1) + L_1(\vec{v}_2)) \\ &= L_2(L_1(\vec{v}_1)) + L_2(L_1(\vec{v}_2)) \\ &= (L_2 \circ L_1)(\vec{v}_1) + (L_2 \circ L_1)(\vec{v}_2).\end{aligned}$$

Then, for $c \in \mathbb{F}$, we have

$$\begin{aligned}(L_2 \circ L_1)(c\vec{v}_1) &= L_2(L_1(c\vec{v}_1)) \\ &= L_2(cL_1(\vec{v}_1)) \\ &= cL_2(L_1(\vec{v}_1)) \\ &= c(L_2 \circ L_1)(\vec{v}_1),\end{aligned}$$

as desired. □

We now define a special case of linear transformations: linear operators.

Definition 5.1.2: Linear Operators

Let V be a vector space. A linear operator on V is a linear transformation whose domain and codomain are both V.

Two special linear operators are the identity linear operator and the zero linear operator. Consider the following definitions.

Definition 5.1.3: The Identity Linear Operator

Let V be a vector space. Then, the function $i : V \to V, \vec{v} \mapsto \vec{v}$ is the identity linear operator.

Definition 5.1.4: The Zero Linear Operator

Let V be a vector space. Then, the function $z : V \to V, \vec{v} \mapsto \vec{0}_V$ is the zero linear operator.

We end this section with a result about subspaces and linear transformations.

Theorem 5.1.3: Linear Transformations and Subspaces

Let $L : V \to W$ be a linear transformation. Then,

1. If V' is a subspace of V, $L(V') = \{L(\vec{v}) : \vec{v} \in V'\}$, the image of V' in W, is a subspace of W. That is, the range of L is a subspace of W.

   Proof. We know $\vec{0}_V \in V'$ since V' is a subspace. Then, $\vec{0}_W \in L(V')$ since $L(\vec{0}_V) = \vec{0}_W$. Next, we take $\vec{w}_1, \vec{w}_2 \in L(V')$. By definition, $\vec{w}_1 = L(\vec{v}_1)$ and $\vec{w}_2 = L(\vec{v}_2)$ for some $\vec{v}_1, \vec{v}_2 \in V'$. Then,

   $$\begin{aligned} \vec{w}_1 + \vec{w}_2 &= L(\vec{v}_1) + L(\vec{v}_2) \\ &= L(\vec{v}_1 + \vec{v}_2). \end{aligned}$$

   Since V' is a subspace, $\vec{v}_1 + \vec{v}_2 \in V'$. Then, $\vec{w}_1 + \vec{w}_2$ is the image of $\vec{v}_1 + \vec{v}_2$, so $(\vec{w}_1 + \vec{w}_2) \in L(V')$. Now, for $c \in \mathbb{F}$, we have $c\vec{w}_1 = cL(\vec{v}_1) = L(c\vec{v}_1)$. Since V' is a subspace, $c\vec{v}_1 \in V'$, so $c\vec{w}_1$ is the image of $c\vec{v}_1$, so $c\vec{w}_1 \in L(V')$. $\square$

2. If W' is subspace of W, then $L^{-1}(W') = \{\vec{v} \in V : L(\vec{v}) \in W'\}$, the pre-image of W' in V, is a subspace of V.

   Proof. We know $\vec{0}_W \in W'$ since W' is a subspace. Then, $\vec{0}_V \in L^{-1}(W')$ since $L(\vec{0}_V) = \vec{0}_W \in W'$. Next, we take $\vec{v}_1, \vec{v}_2 \in L^{-1}(W')$, hence $L(\vec{v}_1), L(\vec{v}_2) \in W'$. Since W' is a subspace,

   $$L(\vec{v}_1) + L(\vec{v}_2) \in W'.$$

   Because L is linear, $L(\vec{v}_1) + L(\vec{v}_2) = L(\vec{v}_1 + \vec{v}_2)$. Thus, $L(\vec{v}_1 + \vec{v}_2) \in W'$. That is, $(\vec{v}_1 + \vec{v}_2) \in L^{-1}(W')$. Finally, for $c \in \mathbb{F}$, we have $L(c\vec{v}_1) = cL(\vec{v}_1)$. Since W' is a subspace and $L(\vec{v}_1) \in W'$, we have that $cL(\vec{v}_1) \in W'$. Thus, $L(c\vec{v}_1) \in W'$, so $c\vec{v}_1 \in L^{-1}(W')$. $\square$

5.2 Lecture 31: November 7, 2022

5.2.1 Linear Transformations and Bases

We begin with an important theorem.

Theorem 5.2.1: Linear Transformations and Bases

Let $B = \{\vec{v}_1, \dots, \vec{v}_n\}$ be a basis for a vector space V. Let W be a vector space with arbitrary $\vec{w}_1, \dots, \vec{w}_n \in W$. Then, there exists a unique linear transformation $L : V \to W$ such that

$$L(\vec{v}_1) = \vec{w}_1, \dots, L(\vec{v}_n) = \vec{w}_n.$$

Proof. Let $L : V \to W$ be a linear transformation with

$$(c_1\vec{v}_1 + \cdots + c_n\vec{v}_n) \mapsto (c_1\vec{w}_1 + \cdots + c_n\vec{w}_n)$$

for scalars $c_1, \dots, c_n$. We note that L is well-defined because $c_1, \dots, c_n$ are unique. We will show L is linear by considering

$$\begin{aligned}
L(\vec{v} + \vec{v}') &= L(c_1\vec{v}_1 + \cdots + c_n\vec{v}_n + c_1'\vec{v}_1 + \cdots + c_n'\vec{v}_n) \\
&= L((c_1 + c_1')\vec{v}_1 + \cdots + (c_n + c_n')\vec{v}_n) \\
&= (c_1 + c_1')\vec{w}_1 + \cdots + (c_n + c_n')\vec{w}_n \\
&= c_1\vec{w}_1 + c_1'\vec{w}_1 + \cdots + c_n\vec{w}_n + c_2'\vec{w}_n \\
&= c_1\vec{w}_1 + \cdots + c_n\vec{w}_n + c_1'\vec{w}_1 + \cdots + c_n'\vec{w}_n \\
&= L(\vec{v}) + L(\vec{v}').
\end{aligned}$$

We now consider

$$\begin{aligned}
L(c\vec{v}) &= L(cc_1\vec{v}_1 + \cdots + cc_n\vec{v}_n) \\
&= cc_1\vec{w}_1 + \cdots + cc_n\vec{w}_n \\
&= c(c_1\vec{w}_1 + \cdots + c_n\vec{w}_n) \\
&= cL(\vec{v}).
\end{aligned}$$

Now, we will show that $L(\vec{v}_i) = \vec{w}_i$. We have that

$$\begin{aligned}
L(\vec{v}_i) &= L(0\vec{v}_1 + \cdots + 1\vec{v}_i + \cdots + 0\vec{v}_n) \\
&= \vec{w}_i.
\end{aligned}$$

We have shown existence, and now, will show uniqueness. Suppose $R : V \to W$ is a linear transformation and $R(\vec{v}_i) = \vec{w}_i$ for $1 \le i \le n, i \in \mathbb{N}$. We will show that R and L are equal. Let $\vec{v} \in V$. Then,

$$\begin{aligned}
R(\vec{v}) &= R(c_1\vec{v}_1 + \cdots + c_n\vec{v}_n) \\
&= c_1R(\vec{v}_1) + \cdots + c_nR(\vec{v}_n) \\
&= c_1\vec{w}_1 + \cdots + c_n\vec{w}_n \\
&= L(\vec{v}).
\end{aligned}$$

Thus, L and R are the same transformation. □

5.3 Lecture 32: November 9, 2022

5.3.1 The Matrix of a Linear Transformation

Consider the following theorem.

Theorem 5.3.1: Matrices and Linear Transformations

Let V and W be nontrivial vector spaces. Let $B = (\vec{v}_1, \dots, \vec{v}_n)$ and $C = (\vec{w}_1, \dots, \vec{w}_m)$ be ordered bases for V and W, respectively. Let $L : V \to W$ be a linear transformation. Then, there exists a unique $A_{BC} \in \mathcal{M}_{mn}$ such that

$$A_{BC}[\vec{v}]_B = [L(\vec{v})]_C.$$

For $1 \le i \le n$, the ith column of A_{BC} is $[L(\vec{v}_i)]_C$.

Proof. Consider $A_{BC} \in \mathcal{M}_{mn}$ with ith column $[L(\vec{v}_i)]_C$, for $1 \le i \le n$. We will first show that $A_{BC}[\vec{v}]_B = [L(\vec{v})]_C$. Suppose that $[\vec{v}]_B = [c_1, \dots, c_n]$. Then,

$$\vec{v} = c_1\vec{v}_1 + \cdots + c_n\vec{v}_n.$$

Then, we have

$$L(\vec{v}) = c_1 L(\vec{v}_1) + \cdots + c_n L(\vec{v}_n).$$

Next,

$$\begin{aligned}
[L(\vec{v})]_C &= [c_1 L(\vec{v}_1) + \cdots + c_n L(\vec{v}_n)]_C \\
&= c_1[L(\vec{v}_1)]_C + \cdots + c_n[L(\vec{v}_n)]_C \\
&= A_{BC}\begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} \\
&= A_{BC}[\vec{v}]_B.
\end{aligned}$$

Note that the third step in the above transitive chain comes from the fact that the ith column of A_{BC} is $[L(\vec{v}_i)]_C$. For uniqueness, suppose $H \in \mathcal{M}_{nn}$ such that $H[\vec{v}]_B = [L(\vec{v})]_C$. We can show that $H = A_{BC}$ if we can show that the ith column of H is the ith column of A_{BC}, or equivalently, $[L(\vec{v}_i)]_C$. Consider $\vec{v}_i \in B$. We know $[\vec{v}_i]_B = \vec{e}_i$. Then, the ith column of H is $H\vec{e}_i = H[\vec{v}_i]_B = [L(\vec{v}_i)]_C$, which is also the ith column of A_{BC}. □

As a remark, Theorem 5.3.1 shows that once we have picked ordered bases for V and W, each linear transformation $L : V \to W$ is equivalent to multiplication by a unique corresponding matrix. This matrix, A_{BC} is called the matrix of the linear transformation L with respect to the ordered bases B and C. To compute A_{BC}, we simply apply the linear transformation on each basis element $\vec{v}_i$, and then express the result with respect to C to get the respective columns of A_{BC}.

Consider the following example.

Example 5.3.1: ☢☢ Finding a Matrix for a Linear Transformation 1

Consider
$$L : \mathbb{R}^2 \to \mathbb{R}^2, \begin{bmatrix} x \\ y \end{bmatrix} \mapsto \begin{bmatrix} x+y \\ x \end{bmatrix}.$$

Find the matrix for the linear transformation L with respect to the ordered bases $B = ([1,0],[0,1])$ and $C = ([1,0],[0,1])$.

Consider

$$\begin{aligned} A_{BC} &= \begin{bmatrix} L\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) & L\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) \end{bmatrix} \\ &= \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}. \end{aligned}$$

Note that we did not have to explicitly coordinatize after finding the image of each element in B under L since we are using the standard basis for the codomain as well.

Example 5.3.2: ☢☢ Finding a Matrix for a Linear Transformation 2

Consider
$$L : \mathbb{R}^2 \to \mathbb{R}^2, \begin{bmatrix} x \\ y \end{bmatrix} \mapsto \begin{bmatrix} x+y \\ x \end{bmatrix}.$$

Find the matrix for the linear transformation L with respect to the ordered bases $B = ([1,0],[0,1])$ and $C = ([0,1],[1,0])$.

Consider

$$\begin{aligned} A_{BC} &= \begin{bmatrix} \left[L\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right)\right]_C & \left[L\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right)\right]_C \end{bmatrix} \\ &= \begin{bmatrix} \left[\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right]_C & \left[\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right]_C \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}. \end{aligned}$$

Note that L is the same linear transformation as the one given in Example 5.3.1; however, we did need to explicitly coordinatize, since we had a nonstandard basis.

Example 5.3.3: ☢☢ Finding a Matrix for a Linear Transformation 3

Consider

$$L : \mathcal{P}_3 \to \mathbb{R}^3, c_0 + c_1x + c_2x^2 + c_3x^3 \mapsto [c_0 + c_1, 2c_2, c_3 - c_0].$$

Find the matrix for the linear transformation L with respect to the ordered bases $B = (x^3, x^2, x, 1)$ for $\mathcal{P}_3$ and $C = ([1, 0, 0], [0, 1, 0], [0, 0, 1])$ for $\mathbb{R}^3$.

By the definition of L, we see that $L(x^3) = [0, 0, 1]$, $L(x^2) = [0, 2, 0]$, $L(x) = [1, 0, 0]$, and $L(1) = [1, 0, -1]$. We need not perform any explicit coordinatization since we are using the standard basis for $\mathbb{R}^3$, so,

$$A_{BC} = \begin{bmatrix} 0 & 0 & 1 & 1 \\ 0 & 2 & 0 & 0 \\ 1 & 0 & 0 & -1 \end{bmatrix}.$$

Example 5.3.4: ☢☢ Finding a Matrix for a Linear Transformation 4

Consider

$$L : \mathcal{P}_3 \to \mathbb{R}^3, c_0 + c_1x + c_2x^2 + c_3x^3 \mapsto [c_0 + c_1, 2c_2, c_3 - c_0].$$

Find the matrix for the linear transformation L with respect to the ordered bases $B = (x^3 + x^2, x^2 + x, x + 1, 1)$ for $\mathcal{P}_3$ and $C = ([-2, 1, -3], [1, -3, 0], [3, -6, 2])$ for $\mathbb{R}^3$.

By the definition of L, we see that $L(x^3 + x^2) = [0, 2, 0]$, $L(x^2 + x) = [1, 2, 0]$, $L(x + 1) = [2, 0, -1]$, and $L(1) = [1, 0, -1]$. Now, we must find $[0, 2, 1]_C$, $[1, 2, 0]_C$, $[2, 0, -1]_C$, and $[1, 0, -1]_C$. We consider

$$\left[\begin{array}{ccc|cccc} -2 & 1 & 3 & 0 & 1 & 2 & 1 \\ 1 & -3 & -6 & 2 & 2 & 0 & 0 \\ -3 & 0 & 2 & 1 & 0 & -1 & -1 \end{array}\right] \underbrace{\to}_{\text{RREF}} \left[\begin{array}{ccc|cccc} 1 & 0 & 0 & -1 & -10 & -15 & -9 \\ 0 & 1 & 0 & 1 & 26 & 41 & 25 \\ 0 & 0 & 1 & -1 & -15 & -23 & -14 \end{array}\right]$$

Thus,

$$A_{BC} = \begin{bmatrix} -1 & -10 & -15 & -9 \\ 1 & 26 & 41 & 25 \\ -1 & -15 & -23 & -14 \end{bmatrix}.$$

Consider the following theorem.

Theorem 5.3.2: Matrices for Linear Transformation, Considering Different Bases

Let V and W be nontrivial vector spaces with B and D be distinct ordered bases for V and C and E be distinct ordered bases for W. Suppose $L : V \to W$ is a linear transformation with matrix A_{BC}. Then,

$$A_{DE} = QA_{BC}P^{-1}$$

where P is the transition matrix from B to D and Q is the transition matrix from C to E.

Proof. For $\vec{v} \in V$, consider $A_{BC}[\vec{v}]_B = [L(\vec{v})]_C$. First, we see that $P^{-1}[\vec{v}]_D = [\vec{v}]_B$, so we may substitute to obtain $A_{BC}P^{-1}[\vec{v}]_D = [L(\vec{v})]_C$. If we multiply by Q on both sides, on the left, we have

$$\begin{aligned} QA_{BC}P^{-1}[\vec{v}]_D &= Q[L(\vec{v})]_C \\ &= [L(\vec{v})]_E. \end{aligned}$$

Since A_{DE} is the unique matrix such that $A_{DE}[\vec{v}]_D = [L(\vec{v})]_E$, $A_{DE} = QA_{BC}P^{-1}$. $\square$

Consider the following example.

Example 5.3.5: ☢☢☢ Finding a Matrix for a Linear Transformation 5

Consider
$$L : \mathcal{P}_3 \to \mathbb{R}^3, c_0 + c_1x + c_2x^2 + c_3x^3 \mapsto [c_0 + c_1, 2c_2, c_3 - c_0].$$
As seen in Example 5.3.3, the matrix for L using the standard bases for $\mathcal{P}_3$ and $\mathbb{R}^3$ was
$$A_{BC} = \begin{bmatrix} 0 & 0 & 1 & 1 \\ 0 & 2 & 0 & 0 \\ 1 & 0 & 0 & -1 \end{bmatrix},$$
where, again $B = (x^3, x^2, x, 1)$ and $C = ([1, 0, 0], [0, 1, 0], [0, 0, 1])$. We will now check our work in Example 5.3.4, where we saw the matrix for L, with respect to the bases $D = (x^3+x^2, x^2+x, x+1, 1)$ and $E = ([-2, 1, -3], [1, 3, 0], [3, -6, 2])$ was
$$A_{DE} = \begin{bmatrix} -1 & -10 & -15 & -9 \\ 1 & 26 & 41 & 25 \\ -1 & -15 & -23 & -14 \end{bmatrix}.$$
To calculate the transition matrix P^{-1} from D to B, we have
$$\left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 \end{array}\right],$$
so
$$P^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}.$$
To calculate the transition matrix from C to E, we have
$$\left[\begin{array}{ccc|ccc} -2 & 1 & 3 & 1 & 0 & 0 \\ 1 & 3 & -6 & 0 & 1 & 0 \\ -3 & 0 & 2 & 0 & 0 & 1 \end{array}\right] \underbrace{\to}_{\text{RREF}} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -6 & -2 & 3 \\ 0 & 1 & 0 & 16 & 5 & -9 \\ 0 & 0 & 1 & -9 & -3 & 5 \end{array}\right].$$
so
$$Q = \begin{bmatrix} -6 & -2 & 3 \\ 16 & 5 & -9 \\ -9 & -3 & 5 \end{bmatrix}.$$
Then,
$$\begin{aligned} A_{DE} = QA_{BC}P^{-1} &= \begin{bmatrix} -6 & -2 & 3 \\ 16 & 5 & -9 \\ -9 & -3 & 5 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 & 1 \\ 0 & 2 & 0 & 0 \\ 1 & 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix} \\ &= \begin{bmatrix} -1 & -10 & -15 & -9 \\ 1 & 26 & 41 & 25 \\ -1 & -15 & -23 & -14 \end{bmatrix}, \end{aligned}$$
as desired.

We now revisit and consider similar matrices.

Theorem 5.3.3: Similar Matrices and Linear Operators

Let V be a vector space with bases C and D. Let $L : V \to V$ be a linear operator, so there exists some A_{CC} and A_{DD}. Let P be the transition matrix from D to C. Then, by Theorem 5.3.2,

$$A_{CC} = PA_{DD}P^{-1}$$

and

$$A_{DD} = P^{-1}A_{CC}P.$$

Thus, A_{CC} and A_{DD} are similar. Generally, any two matrices for the same linear operator, with respect to different bases, are similar, by Definition 3.3.1.

Finally, we present an important result about compositions of linear transformations and matrix multiplication.

Theorem 5.3.4: The Matrix of a Composition of Linear Transformations

Let V_1, V_2 and V_3 be nontrivial finite dimensional vector spaces with ordered bases B, C, and D, respectively. Let $L_1 : V_1 \to V_2$ be a linear transformation with matrix A_{BC}, and let $L_2 : V_2 \to V_3$ be a linear transformation with matrix A_{CD}. Then, the matrix, A_{BD}, for the composite linear transformation $L_2 \circ L_1 : V_1 \to V_3$, with respect to bases B and D, is $A_{CD}A_{BC}$.

5.4 Lecture 33: November 11, 2022

5.4.1 Kernel and Range

We now define some important concepts.

Definition 5.4.1: Kernel

The kernel of a linear transformation $L : V \to W$, is given by

$$\ker(L) = \{\vec{v} \in V : L(\vec{v}) = \vec{0}_W\}.$$

Definition 5.4.2: Range

The range of a linear transformation $L : V \to W$, is given by

$$\text{range}\,(L) = \{\vec{w} \in W : \exists \vec{v} \in V, L(\vec{v}) = \vec{w}\}.$$

Consider the following theorem.

Theorem 5.4.1: Kernel and Range are Subspaces

Let $L : V \to W$ be a linear transformation. Then, $\ker(L)$ is a subspace of V and $\operatorname{range}(L)$ is a subspace of W.

Proof. For $\ker(L)$, we have $\vec{0}_V \in \ker(L)$ since $L(\vec{0}_V) = \vec{0}_W$. Then, for $\vec{v}_1, \vec{v}_2 \in \ker(L)$, we have

$$\begin{aligned} L(\vec{v}_1 + \vec{v}_2) &= L(\vec{v}_1) + L(\vec{v}_2) \\ &= \vec{0}_W + \vec{0}_W \\ &= \vec{0}_W, \end{aligned}$$

so $\vec{v}_1 + \vec{v}_2 \in \ker(L)$. Then, we also have

$$\begin{aligned} L(c\vec{v}_1) &= cL(\vec{v}_1) \\ &= c\vec{0}_W \\ &= \vec{0}_W, \end{aligned}$$

for some $c \in \mathbb{F}$. Thus, $c\vec{v}_1 \in \ker(L)$, as desired. For $\operatorname{range}(L)$, we have $\vec{0}_W \in \operatorname{range}(L)$ since $L(\vec{0}_V) = \vec{0}_W$. Then, if $\vec{w}_1, \vec{w}_2 \in \operatorname{range}(L)$, $L(\vec{v}_1) = \vec{w}_1$ and $L(\vec{v}_2) = \vec{w}_2$ for some $\vec{v}_1, \vec{v}_2 \in V$. Then,

$$\begin{aligned} L(\vec{v}_1 + \vec{v}_2) &= L(\vec{v}_1) + L(\vec{v}_2) \\ &= \vec{w}_1 + \vec{w}_2, \end{aligned}$$

meaning $\vec{w}_1 + \vec{w}_2 \in \operatorname{range}(L)$. Then, we also have

$$\begin{aligned} L(c\vec{v}_1) &= cL(\vec{v}_1) \\ &= c\vec{w} \end{aligned}$$

for some $c \in \mathbb{F}$. Thus, $c\vec{w} \in \operatorname{range}(L)$, as desired. $\square$

Consider the linear transformation

$$L_A : \mathbb{R}^n \to \mathbb{R}^m, \vec{v} \mapsto A\vec{v}$$

for $A \in \mathcal{A}_{mn}$. Now, we want to find $\ker(L_A)$ and $\operatorname{range}(L_A)$. Consider the following theorems.

Theorem 5.4.2: Finding the Kernel of a Linear Transformation

Let $A \in \mathcal{M}_{mn}$. Let L_A be a linear transformation with

$$L_A : \mathbb{R}^n \to \mathbb{R}^m, \vec{v} \mapsto A\vec{v}.$$

We find a basis of $\ker(L)$ by finding particular solutions to $[A|\vec{0}]$. We find each particular solution $\vec{v}_i$ by setting the ith free variable in the system to 1 and the other free variables to 0. We end up with the set $\{\vec{v}_1, \dots, \vec{v}_k\}$ as a basis for $\ker(L)$. Then, $\ker(L) = \operatorname{span}(\{\vec{v}_1, \dots, \vec{v}_k\})$. As a remark, $\dim(\ker(L))$ is the number of free variables in the homogeneous solution set.

Theorem 5.4.3: Finding the Range of a Linear Transformation

Let $A \in \mathcal{M}_{mn}$. Let L_A be a linear transformation with

$$L_A : \mathbb{R}^n \to \mathbb{R}^m, \vec{v} \mapsto A\vec{v}.$$

To find $\operatorname{range}(L)$, we need to row reduce A. The columns in A corresponding to the pivot columns after row reduction is complete form a basis of $\operatorname{range}(L)$. This is because the span of the columns of A is $\operatorname{range}(L)$, and we must then form a basis by keeping only the pivot columns. As a remark, $\dim(\operatorname{range}(L)) = \operatorname{rank} A$, the number of pivot columns.

Consider the following examples.

Example 5.4.1: ☢☢ Find Kernel

Let $L : \mathbb{R}^5 \to \mathbb{R}^4, \vec{v} \mapsto A\vec{v}$, where

$$A = \begin{bmatrix} 8 & 4 & 16 & 32 & 0 \\ 4 & 2 & 10 & 22 & -4 \\ -2 & -1 & -5 & -11 & 7 \\ 6 & 3 & 15 & 33 & -7 \end{bmatrix}.$$

Find $\ker(L)$.

To find $\ker(L)$, we solve $[A|\vec{0}]$ by row reduction to obtain

$$\begin{bmatrix} 1 & \frac{1}{2} & 0 & -2 & 0 & | & 0 \\ 0 & 0 & 1 & 3 & 0 & | & 0 \\ 0 & 0 & 0 & 0 & 1 & | & 0 \\ 0 & 0 & 0 & 0 & 0 & | & 0 \end{bmatrix}.$$

We see that there are free variables in the second and fourth columns. The solution set to the system is

$$\left\{ \left[-\frac{1}{2}c_1 + 2c_2, c_1, -3c_2, c_2, 0 \right] : c_1, c_2 \in \mathbb{R} \right\}.$$

If $c_1 = 1$ and $c_2 = 0$, we have the particular solution $\vec{v}_1 = [-\frac{1}{2}, 1, 0, 0, 0]$. If $c_1 = 0$ and $c_2 = 1$, we have $\vec{v}_2 = [2, 0, -3, 1, 0]$. Thus,

$$\ker(L) = \left\{ c_1 \left[-\frac{1}{2}, 1, 0, 0, 0 \right] + c_2[2, 0, -3, 1, 0] : c_1, c_2 \in \mathbb{R} \right\}.$$

It is worth noting that the initial solution set was indeed also $\ker(L)$, but it is nice to see $\ker(L)$ as the span of a basis of $\ker(L)$. We will further simplify to obtain

$$\ker(L) = \{c_1 [-1, 2, 0, 0, 0] + c_2[2, 0, -3, 1, 0] : c_1, c_2 \in \mathbb{R}\}.$$

Example 5.4.2: ☢☢ Find Range

Let $L : \mathbb{R}^5 \to \mathbb{R}^4$, $\vec{v} \mapsto A\vec{v}$, where

$$A = \begin{bmatrix} 8 & 4 & 16 & 32 & 0 \\ 4 & 2 & 10 & 22 & -4 \\ -2 & -1 & -5 & -11 & 7 \\ 6 & 3 & 15 & 33 & -7 \end{bmatrix}.$$

Find range (L).

To find range (L), row reduce A to obtain

$$\begin{bmatrix} 1 & \frac{1}{2} & 0 & -2 & 0 \\ 0 & 0 & 1 & 3 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

We see that there are pivots in the first, third, and fifth columns. Thus,

$$\text{range}\,(L) = \{c_1[8, 4, -2, 6] + c_2[16, 10, -5, 15] + c_3[0, -4, 7, -7] : c_1, c_2, c_3 \in \mathbb{R}\}.$$

The next theorems combine the results of Theorem 5.4.2 and Theorem 5.4.3 to find

$$\dim(\ker(L)) + \dim(\operatorname{range}(L)).$$

But first, consider the following definitions.

Definition 5.4.3: Nullity of a Linear Transformation

Suppose V and W are finite dimensional vector spaces and $L : V \to W$ is a linear transformation. Then,

$$\operatorname{nullity}(L) = \dim(\ker(L)).$$

Definition 5.4.4: Rank of a Linear Transformation

Suppose V and W are finite dimensional vector spaces and $L : V \to W$ is a linear transformation. Then,

$$\operatorname{rank}(L) = \dim(\operatorname{range}(L)).$$

Now, consider the following theorems.

Theorem 5.4.4: The Dimension Theorem (The Rank-Nullity Theorem), in $\mathbb{R}^n$

Let $A \in \mathcal{M}_{mn}$. Let L_A be a linear transformation with

$$L_A : \mathbb{R}^n \to \mathbb{R}^m, \vec{v} \mapsto A\vec{v}.$$

Then,

1. $\dim(\operatorname{range}(L)) = \operatorname{rank} A$.
2. $\dim(\ker(L)) = n - \operatorname{rank} A$.
3. $\dim(\ker(L)) + \dim(\operatorname{range}(L)) = n$.

These results are verified by Theorems 5.4.2 and 5.4.3.

Theorem: The Dimension Theorem (The Rank-Nullity Theorem)

Suppose V and W are finite dimensional vector spaces and $L : V \to W$ is a linear transformation. Then,

$$\dim(\ker(L)) + \dim(\operatorname{range}(L)) = \dim V.$$

We will postpone the proof of the above theorem to a later section; hence, we have omitted the reference number.

5.4.2 Injections, Surjections, Bijections, and Isomorphisms

We state two theorems about if linear transformations are injective or surjective.

Theorem 5.4.5: Determining Injectivity and Surjectivity

Suppose V and W are finite dimensional vector spaces and $L : V \to W$ is a linear transformation. Then,

1. The linear transformation L is injective if and only if $\ker(L) = \{\vec{0}_V\}$.

   Proof. Suppose L is injective and let $\vec{v} \in \ker(L)$. Now, $L(\vec{v}) = \vec{0}_W$. Similarly, $L(\vec{0}_V) = \vec{0}_W$, and since L is injective, $\vec{v} = \vec{0}_V$. Now, we suppose $\ker(L) = \{\vec{0}_V\}$. We must show L is injective. Let $\vec{v}_1, \vec{v}_2 \in V$ with $L(\vec{v}_1) = L(\vec{v}_2)$. We wish to show $\vec{v}_1 = \vec{v}_2$. Now, we have $L(\vec{v}_1) - L(\vec{v}_2) = \vec{0}_W$, implying that $L(\vec{v}_1 - \vec{v}_2) = \vec{0}_W$. Thus, $\vec{v}_1 - \vec{v}_2 \in \ker(L)$. Since $\ker(L) = \{\vec{0}_V\}$, $\vec{v}_1 - \vec{v}_2 = \vec{0}_V$, and so, $\vec{v}_1 = \vec{v}_2$, as desired. $\square$

2. The linear transformation L is surjective if and only if $\dim(\operatorname{range}(L)) = \dim W$.

   Proof. By definition, L is surjective if and only if $\operatorname{range}(L) = W$. Then, since $\operatorname{range}(L)$ is a subspace of W, $\operatorname{range}(L) = W$ if and only if $\dim(\operatorname{range}(L)) = \dim W$ by Theorem 4.8.1. $\square$

Theorem 5.4.6: Determining Injectivity and Surjectivity With Equivalent Dimensions

Suppose V and W are finite dimensional vector spaces with $\dim V = \dim W$. Let $L : V \to W$ be a linear transformation. Then, L is injective if and only if L is surjective.

Proof. We know L is injective if and only if $\ker(L) = \{\vec{0}_V\}$, meaning $\dim(\ker(L)) = 0$. By the Rank-Nullity Theorem, $\dim V = \dim(\operatorname{range}(L)) + \dim(\ker(L)) = \dim(\operatorname{range}(L))$. Since $\dim V = \dim W$, $\dim W = \dim(\operatorname{range}(L))$, meaning L is surjective, by definition. Conversely, if L is surjective, $\dim(\ker(L)) = 0$, meaning that $\ker(L) = \{\vec{0}_V\}$, which is equivalent to L being injective. $\square$

Consider the following examples.

Example 5.4.3: ☢☢ Determining Injectivity and Surjectivity 1

Consider the linear transformation

$$L : \mathbb{R}^2 \to \mathbb{R}, \begin{bmatrix} x \\ y \end{bmatrix} \mapsto x + 2y.$$

Is L injective? Is L surjective?

With respect to the standard basis for $\mathbb{R}^2$, the matrix for L is

$$A = \begin{bmatrix} 1 & 2 \end{bmatrix}.$$

Consider the linear system

$$\left[\begin{array}{cc|c} 1 & 2 & 0 \end{array}\right],$$

with solution set $\{(-2c, c) : c \in \mathbb{R}\}$. Therefore, $\ker(L) = \{c[-2, 1] : c \in \mathbb{R}\}$. Since $\ker(L) \neq \{\vec{0}\}$, L is not injective. By the Rank-Nullity Theorem, we have

$$\begin{aligned} \dim(\text{range}\,(L)) &= \dim \mathbb{R}^2 - \dim(\ker(L)) \\ &= 2 - 1 = \dim \mathbb{R}, \end{aligned}$$

meaning that L is surjective.

Example 5.4.4: ☢☢ Determining Injectivity and Surjectivity 2

Consider the linear transformation

$$L : \mathcal{M}_{22} \to \mathbb{R}, \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \mapsto a_{11} + a_{22}.$$

Is L injective? Is L surjective?

With respect to the standard basis for $\mathcal{M}_{22}$, the matrix for L is

$$A = \begin{bmatrix} 1 & 0 & 0 & 1 \end{bmatrix}.$$

Consider the linear system

$$\left[\begin{array}{cccc|c} 1 & 0 & 0 & 1 & 0 \end{array}\right],$$

with solution set $\{(-c_3, c_1, c_2, c_3) : c_1, c_2, c_3 \in \mathbb{R}\}$. Therefore,

$$\ker(L) = \{c_1[0, 1, 0, 0] + c_2[0, 0, 1, 0] + c_3[-1, 0, 0, 1] : c_1, c_2, c_3 \in \mathbb{R}\}.$$

Since $\ker(L) \neq \{\vec{0}\}$, L is not injective. By the Rank-Nullity Theorem, we have

$$\begin{aligned} \dim(\text{range}\,(L)) &= \dim \mathcal{M}_{22} - \dim(\ker(L)) \\ &= 4 - 3 = \dim \mathbb{R}, \end{aligned}$$

meaning that L is surjective.

Example 5.4.5: ☢☢ Determining Injectivity and Surjectivity 3

Consider the linear transformation

$$L : \mathbb{R}^2 \to \mathbb{R}^2, \begin{bmatrix} x \\ y \end{bmatrix} \mapsto \begin{bmatrix} x - y \\ 3y \end{bmatrix}.$$

Is L injective? Is L surjective?

With respect to the standard basis for $\mathbb{R}^2$, the matrix for L is

$$A = \begin{bmatrix} 1 & -1 \\ 0 & 3 \end{bmatrix}.$$

Consider the linear system

$$\left[\begin{array}{cc|c} 1 & -1 & 0 \\ 0 & 3 & 0 \end{array}\right].$$

Since A is nonsingular, there exists only a trivial solution for the above system, so $\ker(L) = \{\vec{0}\}$ and L is injective. Since the dimensions of the domain and codomain of L are equal, and L is injective, L is surjective.

Example 5.4.6: ☢☢☢ Determining Injectivity and Surjectivity 4

Consider the linear transformation

$$L : \mathcal{P} \to \mathcal{P}, p(x) \mapsto x^2 p(x) + xp'(x).$$

Is L injective? Is L surjective?

Consider the polynomial

$$p(x) = c_0 + c_1 x + c_2 x^2 + \cdots + c_n x^n$$

with $c_0, \ldots, c_n \in \mathbb{R}$. Then,

$$\begin{aligned} L(p(x)) &= x^2(c_0 + c_1 x + c_2 x^2 + \cdots + c_n x^n) + x(c_1 + 2c_2 x + \cdots + nc_n x^{n-1}) \\ &= c_0 x^2 + c_1 x^3 + c_2 x^4 + \cdots + c_n x^{n+2} + c_1 x + 2c_2 x^2 + 3c_3 x^3 + \cdots + nc_n x^n \\ &= c_1 x + (c_0 + 2c_2)x^2 + (c_1 + 3c_3)x^3 + \cdots + (c_{n-3} + (n-1)c_{n-1}) \\ &\quad + (c_{n-2} + nc_n)x^n + c_{n-1}x^{n+1} + c_n x^{n+2}. \end{aligned}$$

We have expressed $L(p(x))$ as a finite linear combination of a basis of $\mathcal{P}$. Therefore, if $L(p(x)) = 0$, $c_1 = c_{n-1} = c_n = 0$ and $c_0 + 2c_2 = c_1 + 3c_3 = \cdots = c_{n-2} + nc_n = 0$. Since $c_n = 0$, $c_{n-2} = 0$. Similarly, since $c_{n-1} = 0$, $c_{n-3} = 0$. We see that all previous terms $c_0 = \cdots = c_{n-4} = 0$ because we can substitute terms we know to be zero into the summation. For example, next, we will have that since $c_{n-2} = 0$, and $c_{n-4} + (n-2)c_{n-2}$ is a term in the summation, $c_{n-4} = 0$. We will proceed accordingly to find that $c_0 = \cdots = c_n = 0$, meaning that $p(x) = 0$. Thus, $\ker(L) = \{\vec{0}\}$ and L is injective. Now, we will note that the degree of $L(p(x))$ is always greater than or equal to 2, unless $p(x) = 0$, in which case the degree of $L(p(x))$ is zero. Thus, there does not exist $p(x)$ such that $L(p(x)) = 1$, and L is not surjective.

Consider the following theorem about linear independence and spanning, with regards to linear transformations.

Theorem 5.4.7: Injectivity Implies Linear Independence, Surjectivity Implies Spanning

Suppose V and W are vector spaces and $L : V \to W$ is a linear transformation. Then,

1. If L is injective, and T is a linearly independent subset of V, $L(T)$ is linearly independent in W.

   Proof. Suppose that L is injective, and T is a linearly independent subset of V. We wish to show that all finite subsets of $L(T)$ are linearly independent. Suppose $\{L(\vec{v}_1), \dots, L(\vec{v}_n)\} \subseteq L(T)$ for $\vec{v}_1, \dots, \vec{v}_n \in T$. Suppose
   $$c_1 L(\vec{v}_1) + \cdots + c_n L(\vec{v}_n) = \vec{0}_W,$$
   which implies
   $$L(c_1\vec{v}_1 + \cdots + c_n\vec{v}_n) = \vec{0}_W$$
   for scalars $c_1, \dots, c_n$. Thus, $(c_1\vec{v}_1 + \cdots + c_n\vec{v}_n) \in \ker(L)$. Since $\ker(L) = \{\vec{0}_V\}$ because L is injective,
   $$c_1\vec{v}_1 + \cdots + c_n\vec{v}_n = \vec{0}_V.$$
   Since $\{\vec{v}_1, \dots, \vec{v}_n\} \subseteq T$ are linearly independent, $c_1 = \cdots = c_n = 0$. Thus, $\{L(\vec{v}_1), \dots, L(\vec{v}_n)\}$ is linearly independent as well, meaning that $L(T)$ is linearly independent, as desired. □

2. If L is surjective and S spans V, $L(S)$ spans W.

   Proof. Suppose that L is surjective and S spans V. We wish to show that all $\vec{w} \in W$ can be written as a linear combination of vectors in $L(S)$. Since L is surjective, there exists $\vec{v} \in V$ with $L(\vec{v}) = \vec{w}$. Since S spans V, we have
   $$\vec{v} = c_1\vec{v}_1 + \cdots + c_n\vec{v}_n$$
   for $\vec{v}_1, \dots, \vec{v}_n \in S$ and scalars $c_1, \dots, c_n$. Then,
   $$\begin{aligned} \vec{w} = L(\vec{v}) &= L(c_1\vec{v}_1 + \cdots + c_n\vec{v}_n) \\ &= c_1 L(\vec{v}_1) + \cdots + c_n L(\vec{v}_n). \end{aligned}$$
   We have written arbitrary $\vec{w} \in W$ as a linear combination of elements in $L(S)$, so $L(S)$ spans W, as desired. □

Consider the following definition.

Definition 5.4.5: Isomorphisms

Suppose V and W are finite dimensional vector spaces and $L : V \to W$ is a linear transformation; L an isomorphism from V to W if and only if L is both injective and surjective, or bijective.

Definition 5.4.6: Invertible Linear Transformations

Suppose V and W are finite dimensional vector spaces and $L : V \to W$ is a linear transformation. Then, L is an invertible linear transformation if and only if there exists some function $M : W \to V$ such that

$$(M \circ L)(\vec{v}) = \vec{v}$$

for all $\vec{v} \in V$ and

$$(L \circ M)(\vec{w}) = \vec{w}$$

for all $\vec{w} \in W$.

Theorem 5.4.8: Isomorphism If And Only If Invertible

Let $L : V \to W$ be a linear transformation. Then, L is an isomorphism if and only if L is an invertible linear transformation. If L is invertible, L^{-1} is also a linear transformation.

Proof. The first part of this theorem follows from Theorem B.3.2, as by definition, L is an isomorphism if and only if L is bijective. Now, we just seek to show that L^{-1} is a linear transformation. First, consider $\vec{w}_1, \vec{w}_2 \in W$. Since L is surjective, we have $\vec{w}_1 = L(\vec{v}_1)$ and $\vec{w}_2 = L(\vec{v}_2)$ for $\vec{v}_1, \vec{v}_2 \in V$. We have

$$\begin{aligned} L^{-1}(\vec{w}_1 + \vec{w}_2) &= L^{-1}(L(\vec{v}_1) + L(\vec{v}_2)) \\ &= L^{-1}(L(\vec{v}_1 + \vec{v}_2)) \\ &= \vec{v}_1 + \vec{v}_2 \\ &= L^{-1}(\vec{w}_1) + L^{-1}(\vec{w}_2). \end{aligned}$$

Now, for some $c \in \mathbb{F}$, we have

$$\begin{aligned} L^{-1}(c\vec{w}_1) &= L^{-1}(cL(\vec{v}_1)) \\ &= cL^{-1}(L(\vec{v}_1)) \\ &= c\vec{v}_1 \\ &= cL^{-1}(\vec{w}_1), \end{aligned}$$

as desired. Note that for the last step of both transitive chains, we used the fact that L is injective. $\square$

The following theorem allows us to determine whether a linear transformation between finite dimensional vector spaces is an isomorphism, and if so, how to find the inverse.

Theorem 5.4.9: Finding an Inverse Matrix, if it Exists

Suppose V and W are nontrivial finite dimensional vector spaces with ordered bases B and C, respectively. Let $L : V \to W$ be a linear transformation. Then, L is an isomorphism if and only if the matrix representation A_{BC} associated to L is nonsingular. If L is indeed an isomorphism, the matrix A_{CB} for L^{-1} is A_{BC}^{-1}.

Consider the following examples.

Example 5.4.7: Find Inverse 1

Consider

$$L : \mathbb{R}^2 \to \mathbb{R}^2, \begin{bmatrix} x \\ y \end{bmatrix} \mapsto \begin{bmatrix} 3x + y \\ x + y \end{bmatrix}.$$

Is L invertible? If so, find its inverse.

Let B be an ordered basis of $\mathbb{R}^2$ with

$$B = \left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right).$$

Then,

$$A_{BB} = \begin{bmatrix} 3 & 1 \\ 1 & 1 \end{bmatrix}.$$

Since this matrix is invertible, as $\det A_{BB} \neq 0$, the inverse is given by

$$L^{-1} : \mathbb{R}^2 \to \mathbb{R}^2, \begin{bmatrix} x \\ y \end{bmatrix} \mapsto A_{BB}^{-1} \begin{bmatrix} x \\ y \end{bmatrix}$$

$$, \begin{bmatrix} x \\ y \end{bmatrix} \mapsto \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{3}{2} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}.$$

Example 5.4.8: ☢☢ Find Inverse 2

Consider

$$L : \mathbb{R}^3 \to \mathbb{R}^3, \begin{bmatrix} x \\ y \\ z \end{bmatrix} \mapsto A \begin{bmatrix} x \\ y \\ z \end{bmatrix}.$$

where

$$A = \begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{bmatrix}.$$

Is L invertible? If so, find its inverse.

Since this matrix is invertible, as $\det A_{BB} \neq 0$, the inverse is given by

$$L^{-1} : \mathbb{R}^2 \to \mathbb{R}^2, \begin{bmatrix} x \\ y \\ z \end{bmatrix} \mapsto A^{-1} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

$$, \begin{bmatrix} x \\ y \\ z \end{bmatrix} \mapsto \begin{bmatrix} 1 & 0 & -3 \\ 0 & 1 & -3 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}.$$

We end with an important, yet unsurprising, theorem.

Theorem 5.4.10: Isomorphisms Preserve Linear Independence and Span

Let V and W be vector spaces, and let $L : V \to W$ be an isomorphism.

1. If T is a linearly independent subset of V, $L(T)$ is linearly independent in W.

   Proof. Because L is an isomorphism, L is injective. Therefore, $L(T)$ is linearly independent in W by the first part of Theorem 5.4.7. $\square$

2. If S spans V, $L(S)$ spans W.

   Proof. Because L is an isomorphism, L is surjective. Therefore, $L(S)$ spans W by the second part of Theorem 5.4.7. $\square$

3. If B is a basis of V, $L(B)$ is a basis for W.

   Proof. If B is a basis for V, B is linearly independent, so $L(B)$ is linearly independent. We also have that $\operatorname{span}(B) = V$, so $\operatorname{span}(L(B)) = W$. Since $L(B)$ is linearly independent and spans W, $L(B)$ is a basis for W. $\square$

5.5 Lecture 34, November 18, 2022

5.5.1 Isomorphic Vector Spaces

We will now define the notion of equivalence of vector spaces.

Definition 5.5.1: Isomorphic Vector Spaces

Suppose V and W are vector spaces. Then, V is isomorphic to W, that is, $V \cong W$, if and only if there exists some linear transformation $L : V \to W$ that is an isomorphism.

Theorem 5.5.1: $\cong$ is an Equivalence Relation

Suppose V and W are vector spaces. Then, $\cong$ is an equivalence relation.

Proof. We must show that $\cong$ is reflexive, symmetric and transitive. Consider the following.

1. We wish to show $V \cong V$. Consider the linear transformation $i : V \to V, \vec{v} \mapsto \vec{v}$, the identity linear operator, defined by Definition 5.1.3. We wish to show that $i : V \to V$ is an isomorphism. Suppose we have $\vec{v}_1, \vec{v}_2 \in V$ such that $\vec{v}_1 \neq \vec{v}_2$. By definition, $i(\vec{v}_1) = \vec{v}_1$ and $i(\vec{v}_2) = \vec{v}_2$ meaning $i(\vec{v}_1) \neq i(\vec{v}_2)$, so i is injective. Suppose we have $\vec{w} \in V$. We wish to show there exists some $\vec{v} \in V$ such that $i(\vec{v}) = \vec{w}$. By definition, $\vec{v} = \vec{w}$, so i is surjective. We have found that i is an isomorphism from V to V, so $V \cong V$.

2. Suppose $V \cong W$ via $L : V \to W$. Because L is an isomorphism, $L^{-1} : W \to V$ exists. Since $(L^{-1})^{-1} = L$, L^{-1} is an isomorphism, and $W \cong V$.

3. Suppose $V_1 \cong V_2$ via $L_1 : V_1 \to V_2$ and $V_2 \cong V_3$ via $L_2 : V_2 \to V_3$. We wish to show that $L_2 \circ L_1 : V_1 \to V_3$ is an isomorphism. Suppose

$$(L_2 \circ L_1)(\vec{v_1}) = L_2(L_1(\vec{v_1})) = (L_2 \circ L_1)(\vec{v}_2) = L_2(L_1(\vec{v}_2))$$

for $\vec{v}_1, \vec{v}_2 \in V_1$. We wish to show $\vec{v}_1 = \vec{v}_2$. Since L_2 is injective, $L_1(\vec{v}_1) = L_1(\vec{v}_2)$. Then, since L_1 is injective, $\vec{v}_1 = \vec{v}_2$, as desired. Thus, $L_2 \circ L_1$ is injective. Now, suppose we have $\vec{w} \in V_3$. We wish to show there exists some $\vec{v} \in V_1$ such that $(L_2 \circ L_1)(\vec{v}) = L_2(L_1(\vec{v})) = \vec{w}$. Since L_2 is surjective, there exists some $\hat{v} \in V_2$ such that $L_2(\hat{v}) = \vec{w}$. Since L_1 is surjective, there exists some $\vec{v} \in V_1$ such that $L_1(\vec{v}) = \hat{v}$. Thus, $(L_2 \circ L_1)(\vec{v}) = L_2(L_1(\vec{v})) = L_2(\hat{v}) = \vec{w}$, so $L_2 \circ L_1$ is surjective. We have shown that $L_2 \circ L_1$ is both injective and surjective, so it is an isomorphism, as desired.

We have now shown that $\cong$ satisfies the necessary properties. □

We will now restate the Dimension Theorem, or the Rank-Nullity Theorem, with an accompanying proof.

Theorem 5.5.2: The Dimension Theorem (The Rank-Nullity Theorem)

Suppose V and W are finite dimensional vector spaces and $L : V \to W$ is a linear transformation. Then,

$$\dim(\ker(L)) + \dim(\text{range}\,(L)) = \dim V.$$

Proof. Let $B_{\ker(L)} = \{\vec{u}_1, \dots, \vec{u}_m\}$ be a basis for $\ker(L)$. We can extend $B_{\ker(L)}$ to form a basis for V, B_V, with $B_V = \{\vec{u}_1, \dots, \vec{u}_m, \vec{v}_1, \dots, \vec{v}_n\}$. Note that $\dim(\ker(L)) = m$ and $\dim V = m + n$. Then, for some $\vec{v} \in V$, we have

$$\vec{v} = c_1\vec{u}_1 + \cdots + c_m\vec{u}_m + c_{m+1}\vec{v}_1 + \cdots + c_n\vec{v}_n$$

for $c_1, \dots, c_m, c_{m+1}, \dots, c_n \in \mathbb{F}$. If we apply L to both sides, we have

$$\begin{aligned} L(\vec{v}) &= L(c_1\vec{u}_1 + \cdots + c_m\vec{u}_m + c_{m+1}\vec{v}_1 + \cdots + c_n\vec{v}_n) \\ &= c_1L(\vec{u}_1) + \cdots + c_mL(\vec{u}_m) + c_{m+1}L(\vec{v}_1) + \cdots + c_nL(\vec{v}_n). \end{aligned}$$

Since $B_{\ker(L)}$ is a basis of $\ker(L)$, $L(\vec{u}_1) = \cdots = L(\vec{u_m}) = \vec{0}_W$ and we have

$$L(\vec{v}) = c_{m+1}L(\vec{v}_1) + \cdots + c_nL(\vec{v}_n).$$

Therefore, since we have written arbitrary $L(\vec{v})$ as a linear combination of $L(\vec{v}_1), \dots, L(\vec{v}_n)$, we have

$$\text{range}\,(L) = \text{span}\,(\{L(\vec{v}_1), \dots, L(\vec{v}_n)\}).$$

Consider

$$\begin{aligned} \vec{0}_W &= b_1L(\vec{v}_1) + \cdots + b_nL(\vec{v}_n) \\ &= L(b_1\vec{v}_1 + \cdots + b_n\vec{v}_n) \end{aligned}$$

for $b_1, \dots, b_n \in \mathbb{F}$. We see that $b_1\vec{v}_1 + \cdots + b_n\vec{v}_n \in \ker(L)$, so

$$b_1\vec{v}_1 + \cdots + b_n\vec{v}_n = d_1\vec{u}_1 + \cdots + d_m\vec{u}_m.$$

If we subtract the right hand side from both sides, we have

$$b_1\vec{v}_1 + \cdots + b_n\vec{v}_n + (-d_1)\vec{u}_1 + \cdots + (-d_m)\vec{u}_m = \vec{0}_V$$

Since B_V is a basis for V, B_V is linearly independent, so $b_1 = \cdots = b_n = d_1 = \cdots = d_m = 0$. Since $b_1 = \cdots = b_n = 0$, $B_{\text{range}\,(L)} := \{L(\vec{v}_1), \dots, L(\vec{v}_n)\}$ is linearly independent. Since $B_{\text{range}\,(L)}$ both spans range (L) and is linearly independent, $B_{\text{range}\,(L)}$ is a basis for range (L). We see that $\dim(\text{range}\,(L)) = n$, so

$$\begin{aligned} \dim V &= m + n \\ &= \dim(\ker(L)) + \dim(\text{range}\,(L)), \end{aligned}$$

as desired. □

Consider the following important theorems.

Theorem 5.5.3: Isomorphism Implies Equivalent Dimension

Suppose V and W are finite dimensional vector spaces. Then, $V \cong W$ if and only if $\dim V = \dim W$.

Proof. Suppose $V \cong W$. Then, there exists some linear transformation $L : V \to W$ where L is an isomorphism. Therefore, we have that $\dim(\text{range}\,(L)) = \dim W$. We also have $\ker(L) = \{\vec{0}_V\}$, so $\dim(\ker(L)) = 0$. Therefore, by Theorem 5.5.2,

$$\begin{aligned} \dim V &= \dim(\ker(L)) + \dim(\text{range}\,(L)) \\ &= 0 + \dim W \\ &= \dim W. \end{aligned}$$

Now, suppose $\dim V = \dim W$. Let $B_V = \{\vec{v}_1, \dots, \vec{v}_n\}$ be a basis for V and $B_W = \{\vec{w}_1, \dots, \vec{w}_n\}$ be a basis for W. Let

$$L : V \to W, c_1\vec{v}_1 + \cdots + c_n\vec{v}_n \mapsto c_1\vec{w}_1 + \cdots + c_n\vec{w}_n$$

for $c_1, \dots, c_n \in \mathbb{F}$. Consider arbitrary $\vec{w} \in W$. We wish to find some $\vec{v} \in V$ such that $L(\vec{v}) = \vec{w}$ to show that L is surjective. Under the supposition $\dim V = \dim W$, L is an isomorphism if and only if L is surjective by Theorem 5.4.6. We have

$$\vec{w} = c_1\vec{w}_1 + \cdots + c_n\vec{w}_n.$$

By the definition of L, we know

$$L(c_1\vec{v}_1 + \cdots + c_n\vec{v}_n) = c_1\vec{w}_1 + \cdots + c_n\vec{w}_n,$$

so L is surjective and is an isomorphism, so $V \cong W$, as desired. □

Theorem 5.5.4: All n-Dimensional Vector Spaces are Isomorphic to $\mathbb{R}^n$

Suppose V is a finite dimensional vector space with $\dim V = n$. Then, $V \cong \mathbb{R}^n$.

Proof. We have that $\dim V = n$, and we know that $\dim \mathbb{R}^n = n$, so by Theorem 5.5.3, $V \cong \mathbb{R}^n$. □

5.5.2 Diagonalization of Linear Operators

Consider the following definitions, similar to the notions discussed in Chapter 3, but in the context of linear transformations.

Definition 5.5.2: Eigenvalues and Eigenvectors

Suppose V is a vector space. Let $L : V \to V$ be a linear operator. A scalar λ is an eigenvalue of L if and only if there exists $\vec{v} \in V$, where $\vec{v} \neq \vec{0}_V$, such that $L(\vec{v}) = \lambda\vec{v}$. If λ is an eigenvalue of L, $\vec{v}$ is an eigenvector of L with eigenvalue λ.

Definition 5.5.3: Eigenspace

Suppose V is a vector space. Let $L : V \to V$ be a linear operator. The eigenspace of a given eigenvalue λ is

$$E_\lambda = \{\vec{v} \in V : L(\vec{v}) = \lambda\vec{v}\} \cup \{\vec{0}_V\}.$$

Note that "$\cup\{\vec{0}_V\}$" is somewhat redundant, as it will always satisfy the equation. However, the zero vector is never an eigenvector.

Our goal is to find all the eigenvalues and eigenspaces of L. Consider the following theorem.

Theorem 5.5.5: Finding Eigenvectors and Eigenvalues

Let L be a linear operator on a nontrivial finite dimensional vector space V. Suppose $A \in \mathcal{M}_{nn}$ is the matrix representation of L with respect to some ordered basis of V. The scalar λ is an eigenvalue of L if and only if λ satisfies

$$\det(A - \lambda I_n) = 0.$$

We will now define what it means for a linear operator to be diagonalizable, while also providing a theorem to provide an equivalent condition.

Definition 5.5.4: Diagonalizability of a Linear Operator

A linear operator L on a finite dimensional vector space is diagonalizable if and only if the matrix representation of L with respect to some ordered basis for V is a diagonal matrix.

Theorem 5.5.6: Diagonalizability of a Linear Operator

Suppose L is a linear operator on an n-dimensional vector space V. Then, L is diagonalizable if and only if there exists a set of n linearly independent eigenvectors for L.

Proof. Suppose L is diagonalizable. Then, there exists some ordered basis $B = (\vec{v}_1, \ldots, \vec{v}_n)$ for V such that the matrix representation for L is a diagonal matrix D. Because B is a basis, B is linearly independent. We wish to show that each $\vec{v}_i \in B$, with $1 \leq i \leq n$, is an eigenvector corresponding to some eigenvalue for L. Let d_{ii} be the (i, i) element of D. For each $\vec{v}_i$, we have

$$[L(\vec{v_i})]_B = D[\vec{v_i}]_B = D\vec{e}_i = d_{ii}\vec{e}_i = d_{ii}[\vec{v_i}]_B = [d_{ii}\vec{v}_i]_B.$$

We have shown that $L(\vec{v}_i) = d_{ii}\vec{v}_i$. That is, we have shown that each $\vec{v}_i \in B$ is an eigenvector of L with eigenvalue d_{ii}. Thus, B is a set of n linearly independent eigenvectors for L. Conversely, suppose $B = \{\vec{w}_1, \ldots, \vec{w}_n\}$ is a set of n linearly independent eigenvectors for L, corresponding to eigenvalues $\lambda_1, \ldots, \lambda_n$. These eigenvalues need not be distinct. We also note that B is a basis for V, by Theorem 4.7.6. We wish to show that the matrix representation for L, with respect to B is diagonal. The ith column for A is given by

$$[L(\vec{w}_i)]_B = [\lambda_i\vec{w}_i]_B = \lambda_i[\vec{w}_i]_B = \lambda_i\vec{e}_i.$$

Thus, A is diagonal, and L is diagonalizable, as desired. $\square$

Note that Theorem 5.5.6 requires that we find "enough" linearly independent eigenvectors. We now provide a theorem guaranteeing the linear independence of eigenvectors in certain conditions.

Theorem 5.5.7: Eigenvectors With Distinct Eigenvalues are Linearly Independent

Suppose L is a linear operator on V. Let $\lambda_1, \dots, \lambda_n$ be distinct eigenvalues for L. If $\vec{v}_1, \dots, \vec{v}_n$ are eigenvectors for L corresponding to $\lambda_1, \dots, \lambda_n$, respectively, the set $\{\vec{v}_1, \dots, \vec{v}_n\}$ is linearly independent.

Proof. We proceed by induction on n. For $n = 1$, any eigenvector $\vec{v}_1$, for any eigenvalue, by definition is nonzero, so $\{\vec{v}_1\}$ is linearly independent. Suppose that for all $k \in \mathbb{N}$, the proposition holds. Now, for distinct eigenvalues $\lambda_1, \dots, \lambda_{k+1}$. We wish to show that $\{\vec{v}_1, \dots, \vec{v}_{k+1}\}$ is linearly independent. Consider

$$c_1\vec{v}_1 + \cdots + c_{k+1}\vec{v}_{k+1} = \vec{0}_V$$

for scalars $c_1, \dots, c_n$. If we apply L to both sides, we have

$$\begin{aligned} L(\vec{0}_V) &= L(c_1\vec{v}_1 + \cdots + c_{k+1}\vec{v}_{k+1}) \\ &= c_1L(\vec{v}_1) + \cdots + c_{k+1}L(\vec{v}_{k+1}) \end{aligned}$$

which implies

$$c_1\lambda_1\vec{v}_1 + \cdots + c_{k+1}\lambda_{k+1}\vec{v}_{k+1} = \vec{0}_V.$$

If we multiply $c_1\vec{v}_1 + \cdots + c_{k+1}\vec{v}_{k+1} = \vec{0}_V$ by λ_{k+1}, we have

$$c_1\lambda_{k+1}\vec{v}_1 + \cdots + c_{k+1}\lambda_{k+1}\vec{v}_{k+1} = \vec{0}_V = c_1\lambda_1\vec{v}_1 + \cdots + c_{k+1}\lambda_{k+1}\vec{v}_{k+1} = \vec{0}_V.$$

which we can rewrite as

$$c_1(\lambda_1 - \lambda_{k+1})\vec{v}_1 + \cdots + c_k(\lambda_k - \lambda_{k+1})\vec{v}_k = \vec{0}_V$$

By the inductive hypothesis,

$$c_1(\lambda_1 - \lambda_{k+1}) = \cdots = c_k(\lambda_k - \lambda_{k+1}) = 0.$$

Since $\lambda_1, \dots, \lambda_{k+1}$ are distinct, none of the differences in the above equation can be zero, so $c_1 = \cdots = c_k = 0$. Thus, for

$$c_1\vec{v}_1 + \cdots + c_{k+1}\vec{v}_{k+1} = \vec{0}_V,$$

we have $c_{k+1}\vec{v}_{k+1} = \vec{0}_V$. Since $\vec{v}_{k+1} \neq \vec{0}_V$, we have $c_{k+1} = 0$, as desired. $\square$

Note that Theorem 5.5.7 provides that if L is a linear operator on an n-dimensional vector space and L has n distinct eigenvalues, L is diagonalizable. The converse is false.

Consider the following theorems and definitions.

Definition 5.5.5: Algebraic Multiplicity

Let L be a linear operator on a finite dimensional vector space. Let λ be an eigenvalue for L. Suppose $(x-\lambda)^k$ is the highest power of $x-\lambda$ that divides the characteristic polynomial of L. Then, k is the algebraic multiplicity of λ.

Definition 5.5.6: Geometric Multiplicity

Let L be a linear operator on a finite dimensional vector space. Let λ be an eigenvalue for L. Then, $\dim E_\lambda$ is the geometric multiplicity of λ.

Theorem 5.5.8: Algebraic and Geometric Multiplicities and Diagonalizability

Suppose V is a finite dimensional vector space. Let $L : V \to V$ be a linear operator. Then, L is diagonalizable if and only if

1. The sum of the algebraic multiplicities of all eigenvalues of L is $\dim V$.
2. The geometric multiplicity of each eigenvalue is equal to its algebraic multiplicity.

Both conditions must be true.

Theorem 5.5.9: Union and Intersection of Bases for Eigenspaces

Suppose V is a finite dimensional vector space. Let $L : V \to V$ be a linear operator, and let $B_1, \ldots, B_k$ be bases for eigenspaces $E_{\lambda_1}, \ldots, E_{\lambda_k}$ for L, where $\lambda_1, \ldots, \lambda_k$ are distinct eigenvalues for L. Then, $B_i \cap B_j = \emptyset$ for $1 \le i < j \le k$, and $B_1 \cup \cdots \cup B_k$ is a linearly independent subset of V.

Theorem 5.5.10: The Process of Diagonalization for Linear Operators

Let V be an n-dimensional vector space and let $L : V \to V$ be a linear operator. Consider the following steps.

1. Find a basis, C, for V. Then, find the matrix representation A of L with respect to C.
2. Apply the steps of Theorem 3.5.1 on A to find the eigenvalues $\lambda_1, \ldots, \lambda_k$ and a basis in $\mathbb{R}^n$ for each eigenspace E_λ. If $|\bigcup_i E_{\lambda_i}| < n$, L is not diagonalizable. Otherwise, let $Z = (\vec{w}_1, \ldots, \vec{w}_n) = \bigcup_i E_{\lambda_i}$ be an ordered basis for $\mathbb{R}^n$.
3. Find an ordered basis $B = (\vec{v}_1, \ldots, \vec{v}_n)$ of V such that $[\vec{v}_i]_C = \vec{w}_i$.
4. Form D by finding the matrix representation for L with respect to B.
5. If needed, form $P = \begin{bmatrix} [\vec{v}_1]_C & \cdots & [\vec{v}_n]_C \end{bmatrix} = \begin{bmatrix} \vec{w}_1 & \cdots & \vec{w}_n \end{bmatrix}$. Recall $D = P^{-1}AP$.

Richard Friedman: I think that issue is entirely orthogonal to the issue here because the Commonwealth is acknowledging–

Chief Justice Roberts: I'm sorry. Entirely what?

Richard Friedman: Orthogonal. Right angle. Unrelated. Irrelevant.

Chief Justice Roberts: Oh.

Justice Scalia: What was that adjective? I liked that.

Richard Friedman: Orthogonal.

Chief Justice Roberts: Orthogonal.

Richard Friedman: Right, right.

Justice Scalia: Orthogonal, ooh.

Justice Kennedy: I knew this case presented us a problem.

6 Orthogonality

6.1 Lecture 35: November 28, 2022

6.1.1 Inner Product Spaces

Consider the following definition.

Definition 6.1.1: Inner Products and Inner Product Spaces

An $\mathbb{F}$-valued inner product on a vector space V is a function

$$\langle \cdot, \cdot \rangle : V \times V \to \mathbb{F}$$

such that

1. $\forall \vec{v} \in V, \langle \vec{v}, \vec{v} \rangle \geq 0$.
2. $\vec{v} = \vec{0}_V \iff \langle \vec{v}, \vec{v} \rangle = 0$.
3. $\forall \vec{u}, \vec{v} \in V, \langle \vec{u}, \vec{v} \rangle = \overline{\langle \vec{v}, \vec{u} \rangle}$.
4. $\forall \vec{u}, \vec{v}, \vec{w} \in V, \langle \vec{u} + \vec{v}, \vec{w} \rangle = \langle \vec{u}, \vec{w} \rangle + \langle \vec{v}, \vec{w} \rangle$.
5. $\forall c \in \mathbb{F}, \forall \vec{u}, \vec{v} \in V, \langle c\vec{u}, \vec{v} \rangle = c \langle \vec{u}, \vec{v} \rangle$.

The pair $(V, \langle \cdot, \cdot \rangle)$ is called an inner product space.

Consider the following inner product spaces.

1. The pair $(\mathbb{R}^n, \left\langle \vec{u}, \vec{v} \right\rangle = \vec{u} \cdot \vec{v} = u_1 v_1 + \cdots + u_n v_n)$ is a real inner product space.
2. The pair $(\mathbb{C}^n, \left\langle \vec{u}, \vec{v} \right\rangle = \vec{u} \cdot \vec{v} = u_1 \overline{v_1} + \cdots + u_n \overline{v_n})$ is a complex inner product space.
3. The pair
$$\left(V, \left\langle \vec{f}, \vec{g} \right\rangle = \int_0^1 f(x) g(x) \, \mathrm{d}x \right)$$
is an real inner product space, for
$$V = \left\{ f : [0,1] \to \mathbb{R} : \forall c, \lim_{x \to c} = f(c) \right\}.$$
4. The pair
$$\left(\mathcal{P}_n, \left\langle \vec{p}, \vec{q} \right\rangle = \int_{-1}^1 p(x) \overline{q(x)} \, \mathrm{d}x \right)$$
is a real inner product space.

Consider the following theorems and definitions.

Theorem 6.1.1: Properties of Inner Products

Let V be an inner product space. Suppose $\vec{u}, \vec{v}, \vec{w} \in V$ and $c \in \mathbb{F}$. Then,

1. $F : V \to \mathbb{F}, \vec{v} \mapsto \left\langle \vec{v}, \vec{u} \right\rangle$ is a linear transformation.

   Proof. Consider $\vec{v}_1, \vec{v}_2 \in V$. Then,

$$\begin{aligned} F(\vec{v}_1 + \vec{v}_2) &= \left\langle \vec{v}_1 + \vec{v}_2, \vec{u} \right\rangle \\ &= \left\langle \vec{v}_1, u \right\rangle + \left\langle \vec{v}_2, \vec{u} \right\rangle \\ &= F(\vec{v}_1) + F(\vec{v}_2). \end{aligned}$$

   For $c \in \mathbb{F}$, we have

$$\begin{aligned} F(c\vec{v}_1) &= \left\langle c\vec{v}_1, \vec{u} \right\rangle \\ &= c \left\langle \vec{v}_1, \vec{u} \right\rangle, \end{aligned}$$

   as desired. □

2. $\left\langle \vec{0}_V, \vec{v} \right\rangle = \vec{0}_V = \left\langle \vec{v}, \vec{0}_V \right\rangle$.

   Proof. The first equality is derived by the first part since, for all linear transformations L, $\vec{0}_V \in \ker(L)$. The second equality is derived from conjugate symmetry. □

3. $\left\langle \vec{u}, \vec{v} + \vec{w} \right\rangle = \left\langle \vec{u}, \vec{v} \right\rangle + \left\langle \vec{u}, \vec{w} \right\rangle$.

   Proof. Consider $\left\langle \vec{u}, \vec{v} + \vec{w} \right\rangle = \overline{\left\langle \vec{v} + \vec{w}, \vec{u} \right\rangle} = \overline{\left\langle \vec{v}, \vec{u} \right\rangle} + \overline{\left\langle \vec{w}, \vec{u} \right\rangle} = \left\langle \vec{u}, \vec{v} \right\rangle + \left\langle \vec{u}, \vec{w} \right\rangle$. □

4. $\left\langle \vec{u}, c\vec{v} \right\rangle = \overline{c} \left\langle \vec{u}, \vec{v} \right\rangle$.

   Proof. Consider $\left\langle \vec{u}, c\vec{v} \right\rangle = \overline{\left\langle c\vec{v}, \vec{u} \right\rangle} = \overline{c}\overline{\left\langle \vec{v}, \vec{w} \right\rangle} = \overline{c} \left\langle \vec{u}, \vec{v} \right\rangle$. □

Definition 6.1.2: Norms

The norm associated to the inner product $\langle \cdot, \cdot \rangle$ is

$$|| \cdot || : V \to [0, \infty), \vec{v} \mapsto \sqrt{\langle \vec{v}, \vec{v} \rangle}.$$

Theorem 6.1.2: Properties of Norms

Suppose V is an inner product space and $\vec{v} \in V$. Then,

1. $||\vec{v}|| = 0 \iff \vec{v} = \vec{0}_V$.

   Proof. If $\vec{v} = \vec{0}_V$, $\langle \vec{v}, \vec{v} \rangle = \sqrt{\langle \vec{v}, \vec{v} \rangle} = ||\vec{v}|| = 0$. If $||\vec{v}|| = 0$, $\sqrt{\langle \vec{v}, \vec{v} \rangle} = \langle \vec{v}, \vec{v} \rangle = 0$, meaning that $\vec{v} = \vec{0}_V$. □

2. $\forall c \in \mathbb{F}, ||c\vec{v}|| = |c|||\vec{v}||$.

   Proof. Consider

$$\begin{aligned} ||c\vec{v}|| &= \sqrt{\langle c\vec{v}, c\vec{v} \rangle} \\ &= \sqrt{c\bar{c} \langle \vec{v}, \vec{v} \rangle} \\ &= \sqrt{|c|^2 \langle \vec{v}, \vec{v} \rangle} \\ &= |c|\sqrt{\langle \vec{v}, \vec{v} \rangle} \\ &= |c|||\vec{v}||, \end{aligned}$$

   as desired. □

We will now present some generalized proofs of theorems we covered special cases of in Chapter 1.

Theorem 6.1.3: The Cauchy-Schwarz Inequality

Let V be an inner product space with $\vec{v}, \vec{w} \in V$. Then,

$$\left|\langle \vec{v}, \vec{w} \rangle\right| \leq ||\vec{v}||||\vec{w}||.$$

Proof. If $\vec{w} = \vec{0}_V$, the statement is trivial since both sides are zero. We have that for all $t \in \mathbb{F}$,

$$\begin{aligned} 0 &\leq ||\vec{v} - t\vec{w}||^2 \\ &= \langle \vec{v} - t\vec{w}, \vec{v} - t\vec{w} \rangle \\ &= \langle \vec{v}, \vec{v} - t\vec{w} \rangle - t \langle \vec{w}, \vec{v} - t\vec{w} \rangle \\ &= ||\vec{v}||^2 - t \langle \vec{w}, \vec{v} \rangle - \bar{t} \langle \vec{v}, \vec{w} \rangle + |t|^2 ||\vec{w}||^2. \end{aligned}$$

The quadratic has a minimum at $t = \frac{\langle \vec{v}, \vec{w} \rangle}{||\vec{w}||^2} = \frac{\overline{\langle \vec{w}, \vec{v} \rangle}}{||\vec{w}||^2}$ and the inequality should hold. The inequality for this t is

$$0 \leq ||\vec{v}||^2 - \frac{\langle \vec{v}, \vec{w} \rangle^2}{||\vec{w}||^2}.$$

Thus, $\left|\langle \vec{v}, \vec{w} \rangle\right|^2 \leq ||\vec{v}||||\vec{w}||^2 \implies \left|\langle \vec{v}, \vec{w} \rangle\right| \leq ||\vec{v}||||\vec{w}||$. □

Theorem 6.1.4: The Triangle Inequality

Let V be an inner product space with $\vec{v}, \vec{w} \in V$. Then,

$$||\vec{v} + \vec{w}|| \leq ||\vec{v}|| + ||\vec{w}||.$$

Proof. We have

$$\begin{aligned} ||\vec{v} + \vec{w}||^2 &= \left\langle \vec{v} + \vec{w}, \vec{v} + \vec{w} \right\rangle \\ &= \left\langle \vec{v}, \vec{v} \right\rangle + \left\langle \vec{v}, \vec{w} \right\rangle + \left\langle \vec{w}, \vec{v} \right\rangle + \left\langle \vec{w}, \vec{w} \right\rangle \\ &\leq \left\langle \vec{v}, \vec{v} \right\rangle + \left\langle \vec{w}, \vec{w} \right\rangle + 2\left|\left\langle \vec{v}, \vec{w} \right\rangle\right| \\ &\leq \left\langle \vec{v}, \vec{v} \right\rangle + \left\langle \vec{w}, \vec{w} \right\rangle + 2||\vec{v}||||\vec{w}|| \\ &= (||\vec{v}|| + ||\vec{w}||)^2, \end{aligned}$$

as desired. □

We now provide notions of distance and angle.

Definition 6.1.3: Distance

Let V be an inner product space. The distance betwen $\vec{v}, \vec{w} \in V$ is $||\vec{v} - \vec{w}||$.

Definition 6.1.4: Angle

Let V be a real inner product space. That is, $\mathbb{F} = \mathbb{R}$. The angle between $\vec{v} \neq \vec{0}_V$ and $\vec{w} \neq \vec{0}_V$ is given by

$$\theta = \arccos\left(\frac{\langle \vec{v}, \vec{w} \rangle}{||\vec{v}||||\vec{w}||}\right)$$

Definition 6.1.5: Orthogonality

Let V be an inner product space. Then,

1. $\vec{v}, \vec{w} \in V$ are orthogonal if and only if $\langle \vec{v}, \vec{w} \rangle = 0$.
2. A set $S \subseteq V$ is orthogonal if and only if for each $\vec{v}, \vec{w} \in S$, $\langle \vec{v}, \vec{w} \rangle = 0$.
3. A set $S \subseteq V$ is orthonormal if and only if it is orthogonal and each $\vec{v} \in S$ has $||\vec{v}|| = 1$.

Consider the following example.

Example 6.1.1: The Standard Basis of $\mathbb{R}^n$

The set $\{\vec{e}_1, \dots, \vec{e}_n\}$ is orthonormal when considered as a subset of $\mathbb{R}^n$ or $\mathbb{C}^n$.

6.1.2 Orthonormal Bases and the Gram-Schmidt Process

Consider the following theorems and definitions.

Theorem 6.1.5: Orthonormal Implies Linearly Independent

If $\{\vec{v}_1, \dots, \vec{v}_n\}$ is orthonormal in an inner product space V, $\{\vec{v}_1, \dots, \vec{v}_n\}$ is linearly independent.

Proof. Suppose

$$c_1\vec{v}_1 + \cdots + c_n\vec{v}_n = \vec{0}_V$$

for some $c \in \mathbb{F}$. We must show $c_1 = \cdots = c_n = 0$. For each i, $1 \leq i \leq n$, consider

$$\begin{aligned} 0 &= \langle c_1\vec{v}_1 + \cdots + c_n\vec{v}_n, \vec{v}_i\rangle \\ &= \langle c_1\vec{v}_1, \vec{v}_i\rangle + \cdots + \langle c_n\vec{v}_n, \vec{v}_i\rangle \\ &= c_1\langle \vec{v}_1, \vec{v}_i\rangle + \cdots + c_n\langle \vec{v}_n, \vec{v}_i\rangle \\ &= c_i\langle \vec{v}_i, \vec{v}_i\rangle \end{aligned}$$

We see that $\langle \vec{v}_i, \vec{v}_i\rangle > 0$, since $\vec{v}_i \neq \vec{0}_V$. We have $0 = c_i||\vec{v}_i||^2$, implying that $c_i = 0$ for each i. □

The converse is false.

Definition 6.1.6: Orthogonal Bases

A set $B \subseteq V$ is an orthogonal basis if and only if B is a basis of V and B is orthogonal.

Definition 6.1.7: Orthonormal Bases

A set $B \subseteq V$ is an orthonormal basis if and only if B is a basis of V and B is orthonormal.

We would like to form an orthogonal set of n vectors from any linearly independent set of n vectors such that both sets span the same subspace. We present the Gram-Schmidt process.

Theorem 6.1.6: The Gram-Schmidt Process

Let $S_1 = \{\vec{w}_1, \dots, \vec{w}_n\}$ such that S_1 is linearly independent. We will create $S_2 = \{\vec{v}_1, \dots, \vec{v}_n\}$ such that $\text{span}(S_1) = \text{span}(S_2)$.

- Let $\vec{v}_1 = \vec{w}_1$.
- Let $\vec{v}_2 = \vec{w}_2 - \left(\frac{\langle \vec{w}_2, \vec{v}_1\rangle}{\langle \vec{v}_1, \vec{v}_1\rangle}\right)\vec{v}_1$.
- Let $\vec{v}_3 = \vec{w}_3 - \left(\frac{\langle \vec{w}_3, \vec{v}_1\rangle}{\langle \vec{v}_1, \vec{v}_1\rangle}\right)\vec{v}_1 - \left(\frac{\langle \vec{w}_3, \vec{v}_2\rangle}{\langle \vec{v}_2, \vec{v}_2\rangle}\right)\vec{v}_2$.
- Continue.
- Let $\vec{v}_n = \vec{w}_n - \left(\frac{\langle \vec{w}_n, \vec{v}_1\rangle}{\langle \vec{v}_1, \vec{v}_1\rangle}\right)\vec{v}_1 - \left(\frac{\langle \vec{w}_n, \vec{v}_2\rangle}{\langle \vec{v}_2, \vec{v}_2\rangle}\right)\vec{v}_2 - \cdots - \left(\frac{\langle \vec{w}_n, \vec{v}_{n-1}\rangle}{\langle \vec{v}_{n-1}, \vec{v}_{n-1}\rangle}\right)\vec{v}_{n-1}$.

Consider the following examples.

Example 6.1.2: Gram-Schmidt 1

Apply the Gram-Schmidt process to $\{[1,1,0,0],[1,1,0,-1]\} \subseteq \mathbb{R}^4$.

Let $\vec{v}_1 = [1,1,0,0]$. Then,

$$\begin{aligned} v_2 &= [1,1,0,-1] - \frac{\langle [1,1,0,-1],[1,1,0,0]\rangle}{\langle [1,1,0,0],[1,1,0,0]\rangle}[1,1,0,0] \\ &= [1,1,0,-1] - \frac{2}{2}[1,1,0,0] \\ &= [0,0,0,-1]. \end{aligned}$$

Let $S = \{[1,1,0,0],[0,0,0,-1]\}$. Then, $\text{span}(S) = \text{span}(\{[1,1,0,0],[1,1,0,-1]\})$.

Example 6.1.3: ☢☢ Gram-Schmidt 2

Apply the Gram-Schmidt process to $\{[i,-i,1,1],[i,0,0,1],[0,1,-i,0]\} \subseteq \mathbb{C}^4$.

Let $\vec{v}_1 = [i,-i,1,1]$. Then,

$$\begin{aligned}\vec{v}_2 &= [i,0,0,1] - \frac{\langle [i,0,0,1],[i,-i,1,1]\rangle}{\langle [i,-i,1,1],[i,-i,1,1]\rangle}[i,-i,1,1] \\ &= [i,0,0,1] - \frac{1}{2}[i,-i,1,1] \\ &= \left[\frac{1}{2}i, \frac{1}{2}i, -\frac{1}{2}, \frac{1}{2}\right].\end{aligned}$$

Then,

$$\begin{aligned}\vec{v}_3 = [0,1,-i,0] &- \frac{\langle [0,1,-i,0],[i,-i,1,1]\rangle}{\langle [i,-i,1,1],[i,-i,1,1]\rangle}[i,-i,1,1] \\ &- \frac{\langle [0,1,-i,0],[i,-i,1,1]\rangle}{\left\langle \left[\frac{1}{2}i,\frac{1}{2}i,-\frac{1}{2},\frac{1}{2}\right],\left[\frac{1}{2}i,\frac{1}{2}i,-\frac{1}{2},\frac{1}{2}\right]\right\rangle}\left[\frac{1}{2}i,\frac{1}{2}i,-\frac{1}{2},\frac{1}{2}\right] \\ = [0,1,-i,0].\end{aligned}$$

Let $S = \left\{[i,-i,1,1], \left[\frac{1}{2}i,\frac{1}{2}i,-\frac{1}{2},\frac{1}{2}\right], [0,1,-i,0]\right\}$. Then,

$$\operatorname{span}(S) = \operatorname{span}(\{[i,-i,1,1],[i,0,0,1],[0,1,-i,0]\}).$$

Example 6.1.4: ☢☢ Gram-Schmidt 3

Apply the Gram-Schmidt process to $\{1, x, x^2\} \subseteq \mathcal{P}$ with inner product given by

$$\langle p(x), q(x)\rangle = \int_0^1 p(x)q(x)\,\mathrm{d}x.$$

Let $\vec{v}_1 = 1$. Then,

$$\begin{aligned}\vec{v}_2 &= x - \frac{\langle x, 1\rangle}{\langle 1, 1\rangle}1 \\ &= x - \frac{1}{2}.\end{aligned}$$

Then,

$$\begin{aligned}\vec{v}_3 &= x^2 - \frac{\langle x^2, 1\rangle}{\langle 1, 1\rangle}1 - \frac{\langle x^2, x - \frac{1}{2}\rangle}{\langle x - \frac{1}{2}, x - \frac{1}{2}\rangle}\left(x - \frac{1}{2}\right) \\ &= x^2 - \frac{1}{3} - \frac{\frac{1}{4} - \frac{1}{6}}{\frac{\left(1-\frac{1}{2}\right)^3 - \left(0-\frac{1}{2}\right)^3}{3}}\left(x - \frac{1}{2}\right) \\ &= x^2 - \frac{1}{3} - \frac{\frac{1}{4}}{\frac{1}{4}}\left(x - \frac{1}{2}\right) \\ &= x^2 - x + \frac{1}{6}.\end{aligned}$$

Let $S = \left\{1, x - \frac{1}{2}, x^2 - x + \frac{1}{6}\right\}$. Then,

$$\operatorname{span}(S) = \operatorname{span}(\{1, x, x^2\}).$$

We can also normalize an orthogonal set into an orthonormal one by just dividing each $\vec{v}_i$ by $||\vec{v}_i||$.

Theorem 6.1.7: 🛑 Every Inner Product Space Has an Orthonormal Basis

Every inner product space has an orthonormal basis.

Proof. Suppose V is a finite dimensional inner product space. Let $B = \{\vec{w}_1, \dots, \vec{w}_n\}$ be a basis for V. Apply the Gram-Schmidt process to find orthogonal $C = \{\vec{v}_1, \dots, \vec{v}_n\}$ and normalize it. This set is orthonormal and, therefore, linearly independent, and $\operatorname{span}(B) = V = \operatorname{span}(C)$. $\square$

Example 6.1.5: ☢☢ Finding an Orthonormal Basis

Let $W = \{[x, y, z, w] : x + y - z - w = 0; x, y, z, w \in \mathbb{C}\}$. Find an orthonormal basis for W.

By the membership condition of W, we can rewrite it as

$$W = \{[x, y, z, x + y - z] : x, y, z, w \in \mathbb{C}\}.$$

We note that $[1, -1, 0, 0], [1, 0, 1, 0], [1, 1, 1, 1] \in W$. Let $B_\neg = \{[1, -1, 0, 0], [1, 0, 1, 0], [1, 1, 1, 1]\}$. Consider

$$\begin{bmatrix} 1 & -1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix} \underbrace{\rightarrow}_{\text{RREF}} \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & -1 \end{bmatrix}$$

implying that $\operatorname{span}(B_\neg) = \{[c_1, c_2, c_3, c_1 + c_2 - c_3] : c_1, c_2, c_3 \in \mathbb{C}\} = W$. Now, consider

$$\begin{bmatrix} 1 & 1 & 1 \\ -1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \underbrace{\rightarrow}_{\text{RREF}} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix},$$

implying that $B_\neg$ is linearly independent and therefore a basis for W. Now, we apply the Gram-Schmidt process to obtain an orthogonal basis. Let $\vec{v}_1 = [1, -1, 0, 0]$. Then,

$$\begin{aligned} \vec{v}_2 &= [1, 0, 1, 0] - \frac{\langle 1, 0, 1, 0, [1, -1, 0, 0] \rangle}{\langle [1, -1, 0, 0], [1, -1, 0, 0] \rangle}[1, -1, 0, 0] \\ &= [1, 0, 1, 0] - \frac{1}{2}[1, -1, 0, 0] \\ &= \left[\frac{1}{2}, \frac{1}{2}, 1, 0\right]. \end{aligned}$$

Then,

$$\begin{aligned} \vec{v}_3 &= [1, 1, 1, 1] - \frac{\langle [1, 1, 1, 1], [1, -1, 0, 0] \rangle}{\langle 1, -1, 0, 0, 1, -1, 0, 0 \rangle}[1, -1, 0, 0] - \frac{\left\langle [1, 1, 1, 1], \left[\frac{1}{2}, \frac{1}{2}, 1, 0\right] \right\rangle}{\left\langle \left[\frac{1}{2}, \frac{1}{2}, 1, 0\right], \left[\frac{1}{2}, \frac{1}{2}, 1, 0\right] \right\rangle} \left[\frac{1}{2}, \frac{1}{2}, 1, 0\right] \\ &= [1, 1, 1, 1] - \frac{4}{3}\left[\frac{1}{2}, \frac{1}{2}, 1, 0\right] \\ &= \left[\frac{1}{3}, \frac{1}{3}, -\frac{1}{3}, 1\right]. \end{aligned}$$

Our orthogonal basis for W is then $B_\perp = \left\{[1, -1, 0, 0], \left[\frac{1}{2}, \frac{1}{2}, 1, 0\right], \left[\frac{1}{3}, \frac{1}{3}, -\frac{1}{3}, 1\right]\right\}$. To find an orthonormal basis, we simply normalize each vector to obtain

$$B_{\hat{\perp}} = \left\{ \left[\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}, 0, 0\right], \left[\frac{\sqrt{6}}{6}, \frac{\sqrt{6}}{6}, \frac{\sqrt{6}}{3}, 0\right], \left[\frac{\sqrt{3}}{6}, \frac{\sqrt{3}}{6}, -\frac{\sqrt{3}}{6}, \frac{\sqrt{3}}{2}\right] \right\}.$$

We have mentioned many methods of finding orthogonal and orthonormal bases and spanning sets. The following theorems illustrate the utility of doing so.

Theorem 6.1.8: Coordinatization With Respect to an Orthogonal Basis

Let V be a finite dimensional inner product space. Let W be a subspace of V. Suppose $B = (\vec{v}_1, \dots, \vec{v}_n)$ is a nonempty ordered orthogonal basis for W. Then, for all $\vec{w} \in W$,

$$[\vec{w}]_B = \left[\frac{\left\langle \vec{w}, \vec{v}_1 \right\rangle}{\left\langle \vec{v}_1, \vec{v}_1 \right\rangle}, \dots, \frac{\left\langle \vec{w}, \vec{v}_n \right\rangle}{\left\langle \vec{v}_n, \vec{v}_n \right\rangle}\right].$$

Proof. We know that $[\vec{w}]_B = [c_1, \dots, c_n]$ for $c_1, \dots, c_n \in \mathbb{F}$. We wish to show that for each c_i, with $1 \leq i \leq n$,

$$c_i = \frac{\left\langle \vec{v}_i, \vec{v}_i \right\rangle}{\left\langle \vec{v}_i, \vec{v}_i \right\rangle}.$$

We have that

$$\vec{w} = c_1 \vec{v}_1 + \cdots + c_n \vec{v}_n.$$

Then,

$$\begin{aligned}\left\langle \vec{w}, \vec{v}_i \right\rangle &= \left\langle c_1 \vec{v}_1 + \cdots + c_n \vec{v}_n, \vec{v}_i \right\rangle \\ &= c_1 \left\langle \vec{v}_1, \vec{v}_i \right\rangle + \cdots + c_n \left\langle \vec{v}_n, \vec{v}_i \right\rangle \\ &= c_i \left\langle \vec{v}_i, \vec{v}_i \right\rangle.\end{aligned}$$

Therefore,

$$c_i = \frac{\left\langle \vec{w}, \vec{v}_i \right\rangle}{\left\langle \vec{v}_i, \vec{v}_i \right\rangle},$$

as desired. □

Theorem 6.1.9: Coordinatization With Respect to an Orthonormal Basis

Let V be a finite dimensional inner product space. Let W be a subspace of V. Suppose $B = (\vec{v}_1, \dots, \vec{v}_n)$ is a nonempty ordered orthonormal basis for W. Then, for all $\vec{w} \in W$,

$$[\vec{w}]_B = \left[\left\langle \vec{w}, \vec{v}_1 \right\rangle, \dots, \left\langle \vec{w}, \vec{v}_n \right\rangle\right].$$

Proof. If B is orthonormal, $\left\langle \vec{v}_1, \vec{v}_1 \right\rangle = \cdots = \left\langle \vec{v}_n, \vec{v}_n \right\rangle = 1$, so Theorem 6.1.8 simplifies to the above. □

6.2 Lecture 36: November 30, 2022

6.2.1 Orthogonal Complements

Consider the following theorems and definitions.

Definition 6.2.1: Orthogonal Complements

Let V be an inner product space. Let $S \subseteq V$. Then,

$$S^\perp = \{\vec{v} \in V : \forall \vec{w} \in S, \langle \vec{v}, \vec{w} \rangle = 0\}.$$

Theorem 6.2.1: A Useful Lemma for Finding Orthogonal Complements

Let V be an inner product space, and let W be a subspace of V. Let $B = \{\vec{w}_1, \dots, \vec{w}_n\}$ be a basis of W. Then, $\vec{v} \in W^\perp$ if and only if $\langle \vec{v}, \vec{w}_1 \rangle = \cdots = \langle \vec{v}, \vec{w}_n \rangle = 0$.

Proof. If $\vec{v} \in W^\perp$, we easily have that $\langle \vec{v}, \vec{w}_1 \rangle = \cdots = \langle \vec{v}, \vec{w}_n \rangle = 0$ since $B \subseteq W$. Now, suppose $\langle \vec{v}, \vec{w}_1 \rangle = \cdots = \langle \vec{v}, \vec{w}_n \rangle = 0$. Since B is a basis for W, for $\vec{w} \in W$, we have

$$\vec{w} = c_1\vec{w}_1 + \cdots + c_n\vec{w}_n$$

for $c_1, \dots, c_n \in \mathbb{F}$. Then,

$$\begin{aligned}
\langle \vec{v}, \vec{w} \rangle &= \langle \vec{v}, c_1\vec{w}_1 + \cdots + c_n\vec{w}_n \rangle \\
&= \langle \vec{v}, c_1\vec{w}_1 \rangle + \cdots + \langle \vec{v}, c_n\vec{w}_n \rangle \\
&= \overline{c}_1 \langle \vec{v}, \vec{w}_1 \rangle + \cdots + \overline{c}_n \langle \vec{v}, \vec{w}_n \rangle \\
&= c_1(0) + \cdots + c_n(0) \\
&= 0,
\end{aligned}$$

so $\vec{v} \in W^\perp$, as desired. □

Theorem 6.2.1 allows us to only consider the basis vectors of a subspace when finding basis vectors for the orthogonal complement of that subspace. In general, we will form a linear system by setting the inner product of an arbitrary vector and each basis vector to zero and solving.

Consider the following examples.

Example 6.2.1: ☢☢ Find Orthogonal Complement 1

Let $W = \text{span}(\{[1,0,0],[0,1,0]\}) \subseteq \mathbb{R}^3$. Find $W^\perp$.

We, without any degree of formality, see that

$$W^\perp = \text{span}(\{[0,0,1]\}).$$

Geometrically, W is the xy plane and $W^\perp$ is the z axis.

Example 6.2.2: ☢☢ Find Orthogonal Complement 2

Let $W = \text{span}(\{[1,-1,0,0],[1,0,1,0]\})$. Find $W^\perp$.

We see that W is a subspace of $\mathbb{C}^4$. We must find linearly independent $\vec{v}_1, \vec{v}_2 \in \mathbb{C}^4$ such that

$$\left\langle \vec{v}_1, [1,-1,0,0] \right\rangle = \left\langle \vec{v}_1, [1,0,1,0] \right\rangle = 0$$

and

$$\left\langle \vec{v}_2, [1,-1,0,0] \right\rangle = \left\langle \vec{v}_2, [1,0,1,0] \right\rangle = 0.$$

Let $\vec{v}_1 = [x_1, y_1, z_1, w_1]$ and $\vec{v}_2 = [x_2, y_2, z_2, w_2]$. We can then form the system

$$\begin{bmatrix} 1 & -1 & 0 & 0 & | & 0 \\ 1 & 0 & 1 & 0 & | & 0 \end{bmatrix} \underbrace{\to}_{\text{RREF}} \begin{bmatrix} 1 & 0 & 1 & 0 & | & 0 \\ 0 & 1 & 1 & 0 & | & 0 \end{bmatrix},$$

implying that $\vec{v}_1 = [0,0,0,1]$ and $\vec{v}_2 = [-1,-1,1,0]$. Note that both vectors are particular solutions to the linear system. Then, we have

$$W^\perp = \text{span}(\{[0,0,0,1],[-1,-1,1,0]\}).$$

Consider the following theorems.

Theorem 6.2.2: Subsets and Subspaces

Let V be an inner product space. Let $S \subseteq V$. Then,

1. $S^{\perp}$ is a subspace of V.

   Proof. We know that $\vec{0}_V \in S^{\perp}$ since $\left\langle \vec{0}_V, \vec{w} \right\rangle = 0$ for all $\vec{w} \in S$. If we take $\vec{v}_1, \vec{v}_2 \in S^{\perp}$, we consider

$$\begin{aligned} \langle \vec{v}_1 + \vec{v}_2, \vec{w} \rangle &= \langle \vec{v}_1, \vec{w} \rangle + \langle \vec{v}_2, \vec{w} \rangle \\ &= 0 + 0 \\ &= 0, \end{aligned}$$

   so $\vec{v}_1 + \vec{v}_2 \in S^{\perp}$. For $c \in \mathbb{F}$, we consider

$$\begin{aligned} \langle c\vec{v}_1, \vec{w} \rangle &= c \langle \vec{v}_1, \vec{w} \rangle \\ &= c(0) \\ &= 0, \end{aligned}$$

   so $c\vec{v}_1 \in S^{\perp}$, as desired. □

2. If W is a subspace of V, $W \cap W^{\perp} = \{\vec{0}_V\}$.

   Proof. Since W is a subspace of V, $\vec{0}_V \in W$. Since $W^{\perp}$ is a subspace of V because $W \subseteq V$, $\vec{0}_V \in W^{\perp}$. Then, suppose $\vec{w} \in W \cap W^{\perp}$. By Definition 6.2.1, $\langle \vec{w}, \vec{w} \rangle = 0$, so $\vec{w} = \vec{0}_V$. □

3. $S \subseteq (S^{\perp})^{\perp}$.

   Proof. Suppose we have $\vec{v} \in S$. We wish to show that $\vec{v} \in (S^{\perp})^{\perp}$. That is, we must show that $\vec{v}$ is orthogonal to all vectors in $S^{\perp}$, which is true since $\vec{v} \in S$. □

Theorem 6.2.3: Finite Dimensional Inner Product Spaces and Subspaces

Suppose V is a finite dimensional inner product space and W is a subspace of V. Then,

1. If $B_W = \{\vec{v}_1, \dots, \vec{v}_k\}$ is an orthonormal basis of W and $B_V = \{\vec{v}_1, \dots, \vec{v}_k, \vec{w}_1, \dots, \vec{w}_\ell\}$ is an orthonormal basis for V, $\{\vec{w}_1, \dots, \vec{w}_\ell\}$ is an orthonormal basis for $W^\perp$.

   Proof. We have that $\{\vec{w}_1, \dots, \vec{w}_\ell\}$ is linearly independent because it is a subset of a linearly independent set. Let $S = \operatorname{span}(\{\vec{w}_1, \dots, \vec{w}_\ell\})$. We wish to show that $S \subseteq W^\perp$ and $W^\perp \subseteq S$. Suppose we have some $\vec{s} \in S$. Then,

   $$\vec{s} = c_1\vec{w}_1 + \cdots + c_\ell\vec{w}_\ell$$

   for $c_1, \dots, c_\ell \in \mathbb{F}$. Since B_V is orthonormal, each of $\vec{w}_1, \dots, \vec{w}_\ell$ is orthogonal to each of $\vec{v}_1, \dots, \vec{v}_k$. As a result, $\vec{s}$ is orthogonal to each $\vec{v}_i \in B_W$ for $1 \le i \le k$. Therefore, $\vec{s} \in W^\perp$, so $S \subseteq W$. Now, suppose $\vec{s} \in W^\perp$. By Theorem 6.1.8,

   $$\begin{aligned}[\vec{s}]_{B_V} &= \left[\frac{\langle \vec{s}, \vec{v}_1\rangle}{\langle \vec{v}_1, \vec{v}_1\rangle}, \dots, \frac{\langle \vec{s}, \vec{v}_k\rangle}{\langle \vec{v}_k, \vec{v}_k\rangle}, \frac{\langle \vec{s}, \vec{w}_1\rangle}{\langle \vec{w}_1, \vec{w}_1\rangle}, \dots, \frac{\langle \vec{s}, \vec{w}_\ell\rangle}{\langle \vec{w}_\ell, \vec{w}_\ell\rangle}\right] \\ &= \left[0, \dots, 0, \frac{\langle \vec{s}, \vec{w}_1\rangle}{\langle \vec{w}_1, \vec{w}_1\rangle}, \dots, \frac{\langle \vec{s}, \vec{w}_\ell\rangle}{\langle \vec{w}_\ell, \vec{w}_\ell\rangle}\right],\end{aligned}$$

   so

   $$\vec{s} = \frac{\langle \vec{s}, \vec{w}_1\rangle}{\langle \vec{w}_1, \vec{w}_1\rangle}\vec{w}_1 + \cdots + \frac{\langle \vec{s}, \vec{w}_\ell\rangle}{\langle \vec{w}_\ell, \vec{w}_\ell\rangle}\vec{w}_\ell,$$

   meaning that $\vec{s} \in \operatorname{span}(\{\vec{w}_1, \dots, \vec{w}_\ell\})$, or equivalently, $\vec{s} \in S$, so $W^\perp \subseteq S$, as desired. □

2. $\dim V = \dim W + \dim W^\perp$.

   Proof. By the previous part, B_W is a basis for W, B_V is a basis for V, and $B_{W^\perp} = \{\vec{w}_1, \dots, \vec{w}_\ell\}$ is a basis of $W^\perp$. We see that $|B_W| = k$, $|B_{W^\perp}| = \ell$, and $|B_V| = k + \ell$. Then,

   $$\dim W + \dim W^\perp = |B_W| + |B_{W^\perp}| = n + \ell = |B_V| = \dim V,$$

   as desired. □

3. $W = (W^\perp)^\perp$.

   Proof. By Theorem 6.2.2, $W \subseteq (W^\perp)^\perp$. Let $\dim V = n$. Then, by the previous parts, $\dim((W^\perp)^\perp) = n - (n - \dim W) = \dim W$. Therefore, $W = (W^\perp)^\perp$, as desired. □

Theorem 6.2.4: The Orthogonal Complement of the Ambient Space

Let V be an inner product space. Then, $V^{\perp} = \{\vec{0}_V\}$.

Proof. Let $B_V = \{\vec{v}_1, \dots, \vec{v}_n\}$ be a basis for V. Then, by Theorem 6.2.3, the basis for $V^{\perp}$, or $B_{V^{\perp}}$ is $\emptyset$. Therefore,

$$V^{\perp} = \operatorname{span}(B_{V^{\perp}}) = \vec{0}_V,$$

as desired. $\square$

Theorem 6.2.5: The Orthogonal Complement of the Zero Vector

Let V be an inner product space. Then, $\{\vec{0}_V\}^{\perp} = V$.

Proof. By definition, $\emptyset$ is a basis for $\vec{0}_V$. Then, let $B_V = \{\vec{v}_1, \dots, \vec{v}_n\}$ be a basis for V. By Theorem 6.2.3,

$$\{\vec{0}_V\}^{\perp} = \operatorname{span}(B_V) = V,$$

as desired. $\square$

6.3 Lecture 37: December 2, 2022

6.3.1 Projections Onto a Subspace

Consider the following definition.

Definition 6.3.1: Projections Onto a Subspace With Orthogonal Bases

Let V be an inner product space. Let W be a subspace of V. Let $\{\vec{w}_1, \dots, \vec{w}_n\}$ be an orthogonal basis for W. Let $\vec{v} \in V$. Then,

$$\text{proj}_W \vec{v} = \frac{\langle \vec{v}, \vec{w}_1 \rangle}{\langle \vec{w}_1, \vec{w}_1 \rangle}\vec{w}_1 + \cdots + \frac{\langle \vec{v}, \vec{w}_n \rangle}{\langle \vec{w}_n, \vec{w}_n \rangle}\vec{w}_n.$$

If W is the trivial subspace, $\text{proj}_W \vec{v} = \vec{0}_V$.

Theorem 6.3.1: Projections Onto a Subspace With Orthonormal Bases

Let V be an inner product space. Let W be a subspace of V. Let $\{\vec{w}_1, \dots, \vec{w}_n\}$ be an orthonormal basis for W. Let $\vec{v} \in V$. Then,

$$\text{proj}_W \vec{v} = \langle \vec{v}, \vec{w}_1 \rangle \vec{w}_1 + \cdots + \langle \vec{v}, \vec{w}_n \rangle \vec{w}_n.$$

If W is the trivial subspace, $\text{proj}_W \vec{v} = \vec{0}_V$.

Proof. If $\{\vec{w}_1, \dots, \vec{w}_n\}$ is orthonormal, $\langle \vec{w}_1, \vec{w}_1 \rangle = \cdots = \langle \vec{w}_n, \vec{w}_n \rangle = 1$, so Definition 6.3.1 simplifies to the above. □

Consider the following example.

Example 6.3.1: Projection 1

Let $V = \mathbb{R}^3$. Let W be the xy plane with basis $B_W = \{[1, 0.0], [0, 1, 0]\}$. Find $\text{proj}_W \vec{v}$.

Note that B_W is an orthonormal basis of W. If $\vec{v} = [v_1, v_2, v_3]$,

$$\begin{aligned}\text{proj}_W \vec{v} &= \langle v, [1, 0, 0] \rangle [1, 0, 0] + \langle v, [0, 1, 0] \rangle [0, 1, 0] \\ &= [v_1, v_2, 0],\end{aligned}$$

Consider the following theorem.

Theorem 6.3.2: Projection Theorem

If W is a finite dimensional subspace of an inner product space V, with $\vec{v} \in V$, there exist unique $\vec{w} \in W$ and $\hat{w} \in W^{\perp}$ such that $\vec{v} = \vec{w} + \hat{w}$.

Proof. Let $B_W = \{\vec{w}_1, \dots, \vec{w}_k\}$ be an orthonormal basis for W. We extend this basis to be an orthonormal basis of $B_V = \{\vec{w}_1, \dots, \vec{w}_k, \vec{v}_1, \dots, \vec{v}_\ell\}$. Then, $B_{W^\perp} = \{\vec{v}_1, \dots, \vec{v}_\ell\}$ is a basis of $W^{\perp}$. For some $\vec{v} \in V$, we can write

$$\vec{v} = c_1\vec{w}_1 + \cdots + c_k\vec{w}_k + d_1\vec{v}_1 + \cdots + d_\ell\vec{v}_\ell$$

for some $c_1, \dots, c_k, d_1, \dots, d_\ell \in \mathbb{F}$. We know

$$c_1\vec{w}_1 + \cdots + c_k\vec{w}_k \in W, \quad d_1\vec{v}_1 + \cdots + d_\ell\vec{v}_\ell \in W^{\perp}.$$

Thus, $\vec{w} = c_1\vec{w}_1 + \cdots + c_k\vec{w}_k$ and $\hat{w} = d_1\vec{v}_1 + \cdots + d_\ell\vec{v}_\ell$ so $\vec{v} = \vec{w} + \hat{w}$. We have shown existence. For uniqueness, suppose $\vec{v} = \vec{q} + \hat{q}$ where $\vec{q} \in W$ and $\hat{q} \in W^{\perp}$. We will show $\vec{w} = \vec{q}$ and $\hat{w} = \hat{q}$. We have

$$\vec{v} = \vec{q} + \hat{q} = \vec{w} + \hat{w},$$

which implies

$$\vec{q} - \vec{w} = \hat{w} - \hat{q}.$$

We see that $\vec{q} - \vec{w} \in W$, meaning that $\hat{w} - \hat{q} \in W$. But, we also know $\hat{w} - \hat{q} \in W^{\perp}$, so $\vec{q} - \vec{w} \in W^{\perp}$. Thus, $\vec{q} - \vec{w}, \hat{w} - \hat{q} \in W \cap W^{\perp}$. Since $W \cap W^{\perp} = \{\vec{0}_V\}$, we have $\vec{q} - \vec{w} = \vec{0}_V$ and $\hat{w} - \hat{q} = \vec{0}_V$, as desired. $\square$

Example 6.3.2: Projection 2

Let $V = \mathbb{R}^2$. Let $W = \{[x, y] : x - y = 0; x, y \in \mathbb{R}\}$. For $\vec{v} \in \mathbb{R}^2$, we want to find $\vec{w}_1 + \vec{w}_2$ such that $\vec{v} = \vec{w}_1 + \vec{w}_2$.

Notice that $\operatorname{span}(\{[1, 1]\}) = W$. Let $\vec{v} = [v_1, v_2]$. Then,

$$\begin{aligned}
\vec{w}_1 &= \frac{\langle [v_1, v_2], [1, 1] \rangle}{\langle [1, 1], [1, 1] \rangle}[1, 1] \\
&= \frac{v_1 + v_2}{2}[1, 1] \\
&= \left[\frac{v_1 + v_2}{2}, \frac{v_1 + v_2}{2}\right].
\end{aligned}$$

Then,

$$\begin{aligned}
\vec{w}_2 &= [v_1, v_2] - \left[\frac{v_1 + v_2}{2}, \frac{v_1 + v_2}{2}\right] \\
&= \left[\frac{2v_1 - v_1 - v_2}{2}, \frac{2v_2 - v_1 - v_2}{2}\right] \\
&= \left[\frac{v_1 - v_2}{2}, \frac{v_1 - v_2}{2}\right].
\end{aligned}$$

Example 6.3.3: Projection 3

Let V be the vector space of real continuous functions. Define the inner product in V with

$$\left\langle \vec{f}, \vec{g} \right\rangle = \int_{-\pi}^{\pi} f(t)g(t)\,\mathrm{d}t.$$

Write every continuous function $\vec{f} \in V$ as $\vec{f}_1 + \vec{f}_2$.

6.4 Lecture 38: December 5, 2022

6.4.1 Orthogonal and Unitary Matrices

Consider the following definitions and theorems.

Definition 6.4.1: Orthogonal Matrices

The nonsingular matrix $A \in \mathcal{M}_{nn}^{\mathbb{R}}$ is orthogonal if and only if $A^{-1} = A^T$.

Definition 6.4.2: Unitary Matrices

The nonsingular matrix $A \in \mathcal{M}_{nn}^{\mathbb{C}}$ is unitary if and only if $A^{-1} = A^*$.

Theorem 6.4.1: Orthogonal Matrices and Orthonormal Bases

The matrix $A \in \mathcal{M}_{nn}^{\mathbb{R}}$ is orthogonal if and only if the columns of A form an orthonormal basis for $\mathbb{R}^n$.

Theorem 6.4.2: Unitary Matrices and Orthonormal Bases

The matrix $A \in \mathcal{M}_{nn}^{\mathbb{C}}$ is unitary if and only if the columns of A form an orthonormal basis for $\mathbb{C}^n$.

Definition 6.4.3: Orthogonal Diagonalizability

The matrix $A \in \mathcal{M}_{nn}^{\mathbb{R}}$ is orthogonally diagonalizable if and only if there exists some orthogonal P such that

$$D = P^{-1}AP = P^TAP$$

for some diagonal matrix D.

Definition 6.4.4: Unitary Diagonalizability

The matrix $A \in \mathcal{M}_{nn}^{\mathbb{C}}$ is unitarily diagonalizable if and only if there exists some unitary P such that

$$D = P^{-1}AP = P^*AP$$

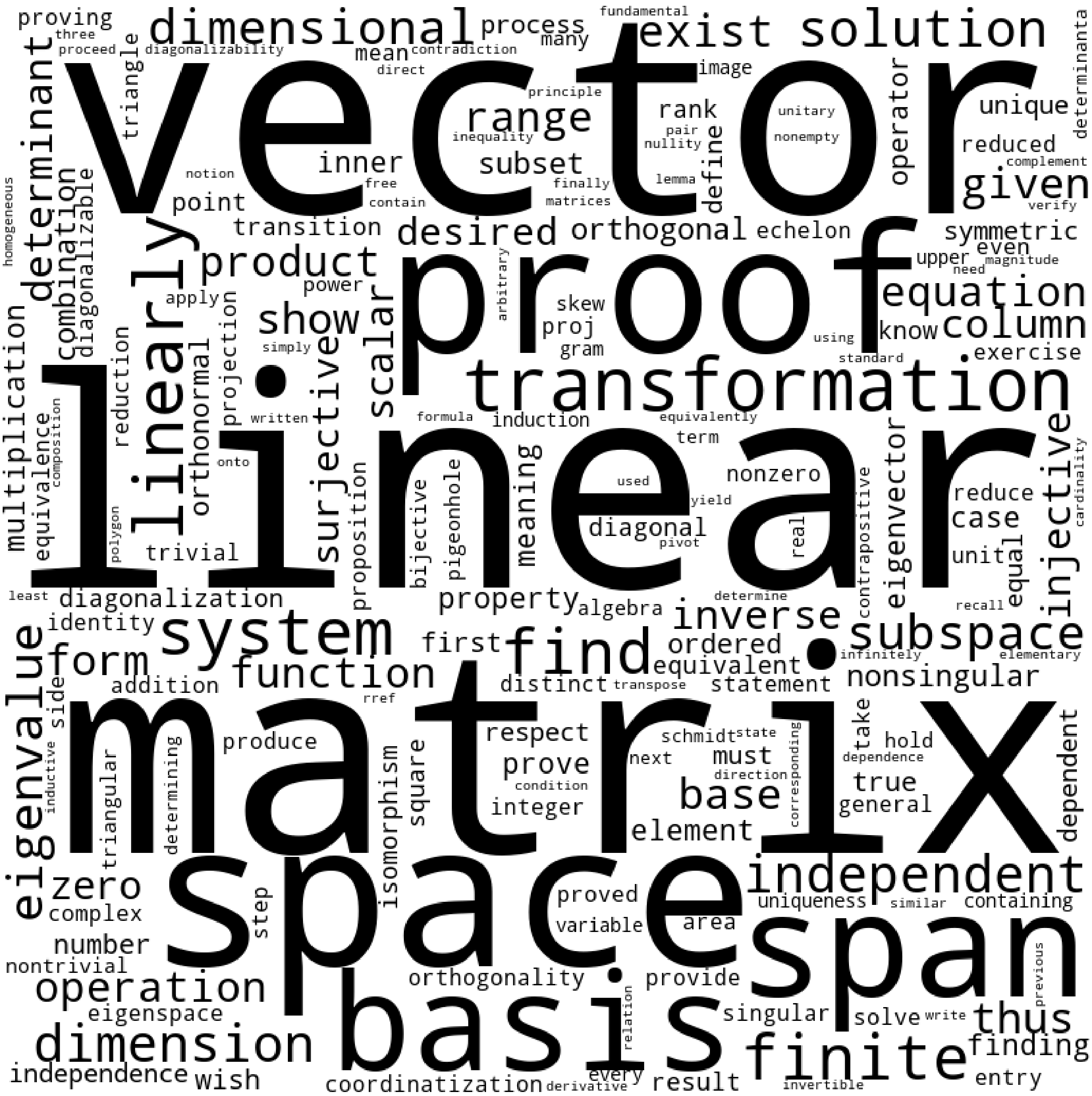
for some diagonal matrix D.

Theorem 6.4.3: Orthogonal Diagonalizability

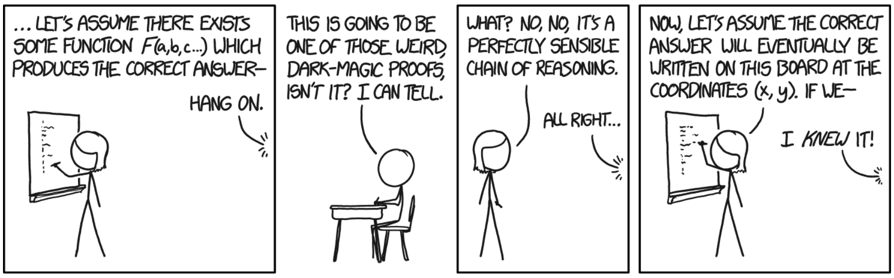
The matrix $A \in \mathcal{M}_{nn}^{\mathbb{R}}$ is orthogonally diagonalizable if and only if A is symmetric. That is, $A = A^T$.

Theorem 6.4.4: Unitary Diagonalizability

The matrix $A \in \mathcal{M}_{nn}^{\mathbb{C}}$ is unitarily diagonalizable if and only if A is normal. That is, $A^*A = AA^*$.


vector
linear
matrix
space
proof
basis
span
finite
transformation
linearly
system
find
subspace
dimensional
exist
solution
dimension
independent
eigenvalue
determinant
product
range
given
equation
column
show
scalar
surjective
injective
function
form
inverse
operation
base
desired
orthogonal
multiplication
combination
diagonalizable
orthonormal
projection
reduction
equivalence
eigenvector
diagonalization
identity
nonsingular
ordered
equivalent
statement
addition
first
property
meaning
proposition
bijective
pigeonhole
diagonal
trivial
nonzero
reduce
case
unit
equal
contrapositive
real
respect
prove
must
true
general
integer
element
schmidt
dependent
isomorphism
square
triangular
determining
produce
zero
complex
number
nontrivial
eigenspace
independence
wish
coordinatization
orthogonality
provide
singular
solve
thus
finding
entry
result
proved
variable
area
uniqueness
containing
step
proving
three
proceed
diagonalizability
process
many
mean
contradiction
direct
fundamental
image
principle
rank
unitary
pair
nullity
inequality
nonempty
unique
reduced
determinanta
complement
verify
subset
inner
operator
define
notion
free
contain
finally
matrices
lemma
point
transition
echelon
symmetric
upper
even
magnitude
need
power
apply
skew
proj
gram
arbitrary
simply
using
standard
know
exercise
homogeneous
triangle
written
formula
induction
equivalently
term
onto
used
yield
composition
cardinality
polygon
pivot
least
determine
algebra
recall
infinitely
elementary
side
rref
transpose
distinct
take
state
hold
next
corresponding
dependence
direction
condition
inductive
similar
previous
relation
every
derivative
invertible
write

Appendices



A

Introduction to Proofs

A.1 Introduction to Proofs

Before we delve into techniques to write proofs, let us first define what a proof is.

Definition A.1.1: Proofs

Mathematical proofs are logical arguments to show that stated premises guarantee that a mathematical statement must be true.

There are multiple techniques to write proofs, but here, we will explore the Proof by Induction, the Direct Proof, the Proof by Contrapositive, the Proof by Contradiction, the Proof by Cases, and the Combinatorial Proof. In Appendix B, proofs based on the Pigeonhole Principle will be discussed.

A.2 Proof by Induction

We will use quantifiers to state induction.

Let $P(n)$ be a statement with $n \in \mathbb{N}$. Consider the following Rule of Inference.

$$\begin{array}{c}
\hline
P(0) \\
P(0) \implies P(1) \\
P(1) \implies P(2) \\
\vdots \\
P(n) \implies P(n+1) \\
\hline
\therefore \forall n \in \mathbb{N}, P(n). \\
\hline
\end{array}.$$

This may be further collapsed into

$$\frac{\begin{array}{c} P(0) \\ \forall k \in \mathbb{N}, P(k) \implies P(k+1) \end{array}}{\therefore \forall n \in \mathbb{N}, P(n).}.$$

Generally, in Proofs by Induction, we follow the following steps.

- Start with an iterative propostition that depends on some $n \in \mathbb{N}$, or $P(n)$.
- Prove that the proposition is true for some base case $n = n_0$. That is, show that the proposition is true for the smallest fixed number that the proposition makes sense for.
- Prove the inductive step. Suppose that the proposition holds true for $n = k$, and then prove that the proposition holds for $n = k + 1$. Essentially, suppose $P(k)$ is true, and prove that $P(k + 1)$ is true.
- Then, the proposition is proved $\forall n \in \mathbb{N}$, where $n \geq n_0$.

Consider the following examples and exercises.

Example A.2.1: ☢☢ Gauss' Formula

Prove that the sum of consecutive integers starting at 1 can be found by Gauss' formula. That is,

$$1 + 2 + 3 \cdots + n = \frac{n(n+1)}{2}.$$

Proof. Consider the base case $n = 1$. Then the left hand side is 1, and the right hand side is

$$\frac{1(1+1)}{2} = 1.$$

Therefore, the left hand side is equal to the right hand side, proving the base case.

We suppose that the relationship is true for $n = k$ where $k \in \mathbb{N}$. That is, we suppose that

$$1 + 2 + 3 + \cdots + k = \frac{k(k+1)}{2}.$$

If we add $k + 1$ to both sides, we obtain

$$\begin{aligned} 1 + 2 + 3 + \cdots + k + k + 1 &= \frac{k(k+1)}{2} + k + 1 \\ &= \frac{k(k+1) + 2k + 2}{2} \\ &= \frac{k^2 + 3k + 2}{2} \\ &= \frac{(k+2)(k+1)}{2} \\ &= \frac{(k+1)((k+1)+1)}{2}. \end{aligned}$$

This result is the proposition where $n = k+1$. Therefore, the inductive step is true. Therefore, Gauss' formula is true for all $n \in \mathbb{N}$ where $n \geq 1$. □

Example A.2.2: ☢☢ Sum of Consequtive Squares

Prove that for all $n \in \mathbb{N}, n \geq 1$,

$$1^2 + 2^2 + 3^2 \cdots + n^2 = \frac{n(n+1)(2n+1)}{6}.$$

Proof. Consider the base case $n = 1$. Then the left hand side is 1, and the right hand side is

$$\frac{1(1+1)(2+1)}{6} = 1.$$

Therefore, the left hand side is equal to the right hand side, proving the base case.

We suppose that the relationship is true for $n = k$ where $k \in \mathbb{N}$. That is, we suppose that

$$1^2 + 2^2 + 3^2 \cdots + k^2 = \frac{k(k+1)(2k+1)}{6}.$$

If we add $k+1$ to both sides, we obtain

$$\begin{aligned}
1^2 + 2^2 + 3^2 \cdots + k^2 + (k+1)^2 &= \frac{k(k+1)(2k+1)}{6} + (k+1)^2 \\
&= \frac{k(k+1)(2k+1)}{6} + k^2 + 2k + 1 \\
&= \frac{k(k+1)(2k+1) + 6k^2 + 12k + 6}{6} \\
&= \frac{2k^3 + 9k^2 + 13k + 6}{6} \\
&= \frac{(k+1)(k+2)(2k+3)}{6} \\
&= \frac{(k+1)((k+1)+1)(2(k+1)+1)}{6}.
\end{aligned}$$

This result is the proposition where $n = k+1$. Therefore, the inductive step is true. Therefore, the above formula is true for all $n \in \mathbb{N}$ where $n \geq 1$. □

Exercise A.2.1: ☢☢ The Power Rule for Derivatives

Prove that for all $n \in \mathbb{N}, n \geq 0$,

$$\frac{\mathrm{d}}{\mathrm{d}x}x^n = nx^{n-1}.$$

Proof. Consider the base case $n = 0$. Then the left hand side is 0, as the derivative of any constant is zero, and the right hand side is

$$0x^{0-1} = 0.$$

Therefore, the left hand side is equal to the right hand side, proving the base case.

We suppose that the relationship is true for $n = k$ where $k \in \mathbb{N}$. That is, we suppose that

$$\frac{\mathrm{d}}{\mathrm{d}x}x^k = kx^{k-1}.$$

Consider $\frac{\mathrm{d}}{\mathrm{d}x}[x^{k+1}]$, or $\frac{\mathrm{d}}{\mathrm{d}x}[xx^k]$. Then we have

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}x}[x^{k+1}] &= \frac{\mathrm{d}}{\mathrm{d}x}[xx^k] \\ &= x^k + x(kx^{k-1}) \\ &= x^k + kx^k \\ &= (k+1)x^k.\end{aligned}$$

This result is the proposition where $n = k + 1$. Therefore, the inductive step is true. Therefore, the power rule for derivatives is true for all $n \in \mathbb{N}$ where $n \geq 0$. □

Exercise A.2.2: ☢☢ *n*th Derivative

Prove that the nth Derivative of $f(x) = \frac{1}{x}$ is

$$f^{(n)}(x) = \frac{(-1)^n n!}{x^{n+1}}.$$

Proof. Consider the base case $n = 0$. The zeroth derivative of $f(x)$ is $f(x)$ itself. Using the formula, we have

$$\begin{aligned} f^{(0)}(x) &= \frac{(-1)^0 0!}{x^{0+1}} \\ &= \frac{1}{x} \\ &= f(x). \end{aligned}$$

Therefore, the base case is true. We suppose that the relationship is true for $n = k$ where $k \in \mathbb{N}$. That is, we suppose that

$$f^{(k)}(x) = \frac{(-1)^k k!}{x^{k+1}}.$$

To find the $(k+1)$th derivative, we differentiate $f^{(k)}$, producing

$$\begin{aligned} f^{(k+1)}(x) &= \frac{\mathrm{d}}{\mathrm{d}x} \frac{(-1)^k k!}{x^{k+1}} \\ &= \frac{(-1)^k k!}{x^{k+1+1}}(-(k+1)) \\ &= \frac{(-1)^{(k+1)}(k+1)!}{x^{(k+1)+1}}. \end{aligned}$$

This result is the proposition where $n = k + 1$. Therefore, the inductive step is true. Therefore, the proposition is proved for all $n \in \mathbb{N}$ where $n \geq 0$. □

Exercise A.2.3: ☢☢☢ Reduction

Prove that for all $n \in \mathbb{N}, n \geq 0$,

$$\int x^n e^{-x}\,\mathrm{d}x = -e^{-x}\left(x^n + nx^{n-1} + n(n-1)x^{n-2} + n(n-1)(n-2)x^{n-3} + \cdots + n!\right) + C.$$

Proof. Consider the base case $n = 0$. Then, the left hand side is equal to

$$\int e^{-x}\,\mathrm{d}x = -e^{-x} + C.$$

The right hand side is equal to $-e^{-x} + C$. Therefore, the left hand side is equal to the right hand side, proving the base case.

We assume that the relationship is true for $n = k$. That is, we assume that

$$\int x^k e^{-x}\,\mathrm{d}x = -e^{-x}(x^k + kx^{k-1} + k(k-1)x^{k-2} + k(k-1)(k-2)x^{k-3} + \cdots + k!) + C.$$

Let

$$u = e^{-x}(x^k + kx^{k-1} + k(k-1)x^{k-2} + k(k-1)(k-2)x^{k-3} + \cdots + k!).$$

Then,

$$\begin{aligned}
\int x^{k+1}e^{-x}\,\mathrm{d}x &= -x^{k+1}e^{-x} - \int -e^{-x}x^k(k+1)\,\mathrm{d}x \\
&= -x^{k+1}e^{-x} - (k+1)\int -x^k e^{-x}\,\mathrm{d}x \\
&= -x^{k+1}e^{-x} + (k+1)\int x^k e^{-x}\,\mathrm{d}x \\
&= -x^{k+1}e^{-x} - e^{-x}(k+1)\frac{u}{e^{-x}} + C \\
&= -e^{-x}\left(x^{k+1} + \frac{u(k+1)}{e^{-x}}\right) + C \\
&= -e^{-x}\left(x^{k+1} + (k+1)x^k + k(k+1)x^{k-1} + \cdots + (k+1)!\right) + C.
\end{aligned}$$

This result is the proposition where $n = k + 1$. Therefore, the inductive step is true. Therefore, the above formula is true for all $n \in \mathbb{N}$ where $n \geq 0$. □

Exercise A.2.4: ☢☢☢ The Shoelace Lemma

The following is a statement of the Shoelace Lemma.

> Consider a simple polygon with vertices $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$, oriented clockwise. Let $(x_{n+1}, y_{n+1}) = (x_1, y_1)$. The area of the polygon is given by
> $$A_n = \frac{1}{2}\left[\sum_{i=1}^{n} x_i y_{i+1} - x_{i+1} y_i\right].$$

Prove the above proposition.

Proof. Consider a polygon with three vertices: (x_1, y_1), (x_2, y_2), and (x_3, y_3). The area, given by the Shoelace Lemma, is

$$A_3 = \frac{1}{2}\left[\sum_{i=1}^{3} x_i y_{i+1} - x_{i+1} y_i\right] = \frac{1}{2}\left[x_1 y_2 - x_2 y_1 + x_2 y_3 - x_3 y_2 + x_3 y_1 - y_3 x_1\right].$$

Then, if we define two vectors $(\vec{v}, \vec{w}) \in \mathbb{R}^3$ such that

$$\vec{v} = (x_2 - x_1, y_2 - y_1, 0), \quad \vec{w} = (x_3 - x_1, y_3 - y_1, 0),$$

we may see that the area of the parallelogram formed by the two vectors is given by

$$A_{||GRAM} = ||\vec{v} \times \vec{w}||$$

Either of the two triangles formed by the parallelogram's diagonals correspond to our polygon. The area is then given by

$$\begin{aligned} A_3 &= \frac{1}{2}||\vec{v} \times \vec{w}|| \\ &= \frac{1}{2}||(0, 0, x_1 y_2 - x_2 y_1 + x_2 y_3 - x_3 y_2 + x_3 y_1 - y_3 x_1)|| \\ &= \frac{1}{2}\left[x_1 y_2 - x_2 y_1 + x_2 y_3 - x_3 y_2 + x_3 y_1 - y_3 x_1\right]. \end{aligned}$$

Therefore, we have proved the Shoelace Lemma in the case of a polygon with three vertices. By induction, we suppose that for a polygon with k vertices $(x_1, y_1), (x_2, y_2), \dots, (x_k, y_k)$, the area is

$$A_k = \frac{1}{2}\left[\sum_{i=1}^{k} x_i y_{i+1} - x_{i+1} y_i\right].$$

The area of a polygon with vertices $(x_1, y_1), (x_2, y_2), \dots, (x_{k+1}, y_{k+1})$ is given by the sum of the area of the polygon with vertices $(x_1, y_1), (x_2, y_2), \dots, (x_k, y_k)$ and the area of the polygon with vertices (x_1, y_1), (x_k, y_k), and (x_{k+1}, y_{k+1}). That is,

$$\begin{aligned} A_{k+1} &= \frac{1}{2}\left[\sum_{i=1}^{k} x_i y_{i+1} - x_{i+1} y_i\right] + \frac{1}{2}\left[x_1 y_k - x_k y_1 + x_k y_{k+1} - x_{k+1} y_k + x_{k+1} y_1 - y_{k+1} x_1\right] \\ &= \frac{1}{2}\left[\sum_{i=1}^{k+1} x_i y_{i+1} - x_{i+1} y_i\right]. \end{aligned}$$

The above result is the consequence of the Shoelace Lemma in the case of a polygon with $k + 1$ vertices. Therefore, the Shoelace Lemma is proved. □

Exercise A.2.5: ☢☢☢ Fibonacci

Let f_n represent the sequence of Fibonacci numbers, which is defined recursively as

$$f_0 = 1,\ f_1 = 1,\ f_n = f_{n-1} + f_{n-2}.$$

Prove that

$$\sum_{i=0}^{n} (f_i)^2 = f_n f_{n+1}.$$

Proof. Consider the base case $n = 0$. Then the left hand side is equal to 1, and the right hand side is

$$(1)f_1 = 1.$$

Therefore, the left hand side is equal to the right hand side, proving the first base case. Then, consider the base case $n = 1$. The left hand side is equal to 2, and the right hand side is $1 \cdot f_2$ where $f_2 = f_1 + f_0 = 2$. Therefore the right hand side is 2 and is equal to the left hand side proving the second base case.

We suppose that the relationship is true for $n = k$ where $k \in \mathbb{N}$. That is, we suppose that

$$\sum_{i=0}^{k} (f_i)^2 = f_k f_{k+1}.$$

If we add $(f_{k+1})^2$ to both sides, we have

$$\begin{aligned}(f_{k+1})^2 + \sum_{i=0}^{k} (f_i)^2 &= f_k f_{k+1} + (f_{k+1})^2 \\ &= f_{k+1}(f_k + f_{k+1}) \\ &= f_{k+1} f_{k+2}.\end{aligned}$$

This result is the proposition where $n = k + 1$. Therefore, the inductive step is true. Therefore, the above formula is true for all $n \in \mathbb{N}$ where $n \geq 0$. □

A.3 Direct Proofs

Direct Proofs are the simplest style of proofs, and are especially useful when proving implications. Consider the following examples.

Example A.3.1: ☢☢ Direct Proof 1

Prove that for all integers n, if n is even, then n^2 is even.

Proof. Let $n \in \mathbb{Z}$ and suppose that n is even. Let $m \in \mathbb{Z}$. Thus, $n = 2m$. Then, $n^2 = (2m)^2 = 4m^2 = 2(2m^2)$. Because $2m^2 \in \mathbb{Z}$, n^2 is even. $\square$

Example A.3.2: ☢☢ Direct Proof 2

Prove that for all integers a, b, and c, if $a|b$ and $b|c$, then $a|c$.

Proof. Let $(a, b, c, p, q, r) \in \mathbb{Z}$ and suppose that $a|b$ and$b|c$. Because $a|b$, $b = pa$. Because $b|c$, $c = qb = pqa$. Because c is an integer multiple of a, $a|c$. $\square$

Consider the following exercises.

Exercise A.3.1: ☢☢ Direct Proof 1

Prove that for any two odd integers, their sum is even.

Proof. Let $(m, n) \in \mathbb{Z} : m \bmod 2 \neq 0 : n \bmod 2 \neq 0$ and let $(p, q) \in \mathbb{Z}$. Because m and n are odd, $m = 2p + 1$ and $n = 2q + 1$. Therefore,

$$\begin{aligned} m + n &= (2p + 1) + (2q + 1) \\ &= 2p + 2q + 2 \\ &= 2(p + q + 1). \end{aligned}$$

Because $(p + q + 1) \in \mathbb{Z}$, $m + n$ is even. $\square$

Exercise A.3.2: ☢☢ Direct Proof 2

Prove that for all integers n, if n is odd, then n^2 is odd.

Proof. Let $n \in \mathbb{Z} : n \bmod 2 \neq 0$ and let $p \in \mathbb{Z}$. Because n is odd, $n = 2p + 1$. Therefore,

$$\begin{aligned} n^2 &= (2p + 1)^2 \\ &= 4p^2 + 4p + 1 \\ &= 2(2p^2 + 2p) + 1. \end{aligned}$$

Because $(2p^2 + 2p) \in \mathbb{Z}$, n^2 is odd. $\square$

A.4 Proof by Contrapositive

Recall that for two statements P and Q, $(P \implies Q) \iff (\neg Q \implies \neg P)$. In a Proof by Contrapositive, we produce a direct proof of the contrapositive of the implication. This is equivalent to proving the implication, because the implication is logically equivalent to the contrapositive. Consider the following examples.

Example A.4.1: ☢☢ Proof by Contrapositive 1

Prove that for all integers n, if n^2 is even, then n is even.

Proof. Let $n \in \mathbb{Z} : n \bmod 2 \neq 0$. By Exercise A.3.2, n^2 is odd. □

Example A.4.2: ☢☢ Proof by Contrapositive 2

Prove that for all integers a and b, if $a + b$ is odd, then a is odd or b is odd.

Proof. Let $(a, b, p, q) \in \mathbb{Z}$. Suppose that a is even and b is even. Then, $a = 2p$ and $b = 2q$. We see that

$$\begin{aligned} a + b &= 2p + 2q \\ &= 2(p + q). \end{aligned}$$

Because $(p + q) \in \mathbb{Z}$, $a + b$ is even. □

Consider the following exercises.

Exercise A.4.1: ☢☢ Proof by Contrapositive 1

Prove that for real numbers a and b, if ab is irrational, then a or b must be an irrational number.

Proof. Let $(p, q, r, s) \in \mathbb{Z}$. Suppose that $(a, b) \in \mathbb{Q}$. Therefore, $a = \frac{p}{q}$ and $b = \frac{r}{s}$. We see that

$$ab = \frac{pr}{qs} \in \mathbb{Q}$$

Therefore ab is rational. □

Exercise A.4.2: ☢☢ Proof by Contrapositive 2

Prove that for integers a and b, if ab is even, then a or b must be even.

Proof. Let $(a, b, p, q) \in \mathbb{Z}$. Suppose that $a = 2p + 1$ and $b = 2q + 1$. We see that

$$\begin{aligned} ab &= (2p+1)(2q+1) \\ &= 4pq + 2p + 2q + 1 \\ &= 2(2pq + p + q) + 1 \end{aligned}$$

Because $(2pq + p + q) \in \mathbb{Z}$, ab is odd. □

Exercise A.4.3: ☢☢ Proof by Contrapositive 3

Prove that for any integer a, if a^2 is not divisible by 4, then a is odd.

Proof. Let $a, p \in \mathbb{Z}$. Suppose that a is even, and $a = 2p$. Then, $a^2 = 4p^2$, and $4|4p^2$, so $4|a^2$. □

A.5 Proof by Contradiction

Sometimes, a statement, P cannot be rephrased as an implication. In these cases, it may be useful to prove that $P \implies Q$, and also prove that $P \implies \neg Q$. Then, we conclude $\neg P$. Consider the following example.

Example A.5.1: ☢☢ Proof by Contradiction 1

Prove that $\sqrt{2}$ is irrational.

Proof. Suppose that $\sqrt{2}$ is rational. Then,

$$\sqrt{2} = \frac{p}{q}$$

where $(p, q) \in \mathbb{Z}$ and $\frac{p}{q}$ is in lowest terms. By squaring both sides of the equation, we have

$$2 = \frac{p^2}{q^2}.$$

This means that

$$2q^2 = p^2,$$

and as $q^2 \in \mathbb{Z}$, p^2 is even, which means that by Example A.4.1, p is even. We see that $p = 2k$ for some $k \in \mathbb{Z}$. Then, we have

$$2q^2 = (2k)^2 = 4k^2$$

meaning that

$$q^2 = 2k^2.$$

Therefore, q is even. If p and q are both even, $\frac{p}{q}$ is not in lowest terms. Therefore, $\sqrt{2}$ is irrational. □

A.6 Combinatorial Proofs

We start with a visual exercise.

Exercise A.6.1: ☢ Visual Analysis of Pascal's Triangle

Consider the first six rows of Pascal's Triangle, provided below.

$$
\begin{array}{lccccccccccccc}
n=0 & & & & & & & 1 & & & & & & \\
n=1 & & & & & & 1 & & 1 & & & & & \\
n=2 & & & & & 1 & & 2 & & 1 & & & & \\
n=3 & & & & 1 & & 3 & & 3 & & 1 & & & \\
n=4 & & & 1 & & 4 & & 6 & & 4 & & 1 & & \\
n=5 & & 1 & & 5 & & 10 & & 10 & & 5 & & 1 & \\
n=6 & 1 & & 6 & & 15 & & 20 & & 15 & & 6 & & 1
\end{array}
$$

Name, without proof, some patterns that are visible.

- Border entries are all 1.
- Non-border entries are the sum of the two entries above.
- The triangle is symmetric.
- The sum of all entries on any row is a power of 2.

Recall that each entry in Pascal's Triangle can be written as a binomial coefficient. The kth entry in the nth row, where both k and n are zero-indexed, is $\binom{n}{k}$. That is, the first six rows are

$$
\begin{array}{lccccccccccccc}
n=0 & & & & & & & \binom{0}{0} & & & & & & \\
n=1 & & & & & & \binom{1}{0} & & \binom{1}{1} & & & & & \\
n=2 & & & & & \binom{2}{0} & & \binom{2}{1} & & \binom{2}{2} & & & & \\
n=3 & & & & \binom{3}{0} & & \binom{3}{1} & & \binom{3}{2} & & \binom{3}{3} & & & \\
n=4 & & & \binom{4}{0} & & \binom{4}{1} & & \binom{4}{2} & & \binom{4}{3} & & \binom{4}{4} & & \\
n=5 & & \binom{5}{0} & & \binom{5}{1} & & \binom{5}{2} & & \binom{5}{3} & & \binom{5}{4} & & \binom{5}{5} & \\
n=6 & \binom{6}{0} & & \binom{6}{1} & & \binom{6}{2} & & \binom{6}{3} & & \binom{6}{4} & & \binom{6}{5} & & \binom{6}{6}.
\end{array}
$$

Consider the following exercise.

Exercise A.6.2: ☢☢ Rewriting Hypotheses in Pascal's Triangle

Rewrite the patterns discovered in Exercise A.6.1 in terms of binomial coefficients.

- $\binom{n}{0} = \binom{n}{n} = 1$.
- $\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}$.
- $\binom{n}{k} = \binom{n}{n-k}$.
- $\binom{n}{0} + \binom{n}{1} + \binom{n}{2} + \cdots + \binom{n}{n} = 2^n$.

The above identities were given without proof. To prove the identities, methods of either algebraic or combinatorial proof may be used. While algebraic proof certainly can show that an identity is true, it would not show *why*. Instead, combinatorial proof is based precisely on *why*. In general, combinatorial proofs for binomial identities are in the following format.

1. Find a counting problem that can be answered in two ways.
2. Explain why one answer to the counting problem is the left hand side of the identity.
3. Explain why one answer to the counting problem is the right hand side of the identity.
4. Conclude that the left hand side is equal to the right hand side because they are both answers to the same question.

Consider the following example.

Example A.6.1: ☢☢☢ Pascal Proof

Prove the binomial identity

$$\binom{n}{0} + \binom{n}{1} + \binom{n}{2} + \cdots + \binom{n}{n} = 2^n.$$

Proof. Consider a set S with cardinality n. We note that $|\mathcal{P}(S)| = 2^n$, which is precisely the right hand side of the statement. We wish to show that the left hand side also produces the cardinality of $\mathcal{P}(S)$. The term $\binom{n}{0}$ corresponds to the fact that $\emptyset \in \mathcal{P}(S)$, and the term $\binom{n}{n}$ corresponds to the fact that $S \in \mathcal{P}(S)$. The quantities $\binom{n}{0}$ and $\binom{n}{n}$ together provide 2 sets in $\mathcal{P}(S)$. The terms $\binom{n}{1}, \binom{n}{2}, \ldots, \binom{n}{n-1}$ correspond to the sets of cardinality $1, 2, \ldots, n-1$ in $\mathcal{P}(S)$. In $\mathcal{P}(S)$ there are $\binom{n}{1} = \frac{n!}{1!(n-1)!} = n$ sets of cardinality 1, $\binom{n}{2} = \frac{n!}{2!(n-2)!} = \frac{n(n-1)}{2!}$ sets of cardinality 2, $\binom{n}{3} = \frac{n!}{3!(n-3)!} = \frac{n(n-1)(n-2)}{3!}$ sets of cardinality 3, etc. until we get to $\binom{n}{n-1} = \frac{n!}{(n-1)!(n-n+1)!} = n!$ sets of cardinality $n-1$. It is now clear that the left hand side also proves the cardinality of $\mathcal{P}(S)$. The left hand side sums the number of sets with cardinality $0, 1, \ldots, n$ in $\mathcal{P}(S)$. In contrast, the right hand side is the result of allowing each element of the set to "choose" to be in an n-length subset of S. □

Example A.6.2: ☢☢☢ Team Captain

Prove the binomial identity

$$\binom{n}{1} + 2\binom{n}{2} + 3\binom{n}{3} + \cdots + n\binom{n}{n} = n2^{n-1}.$$

Proof. If k escape room enthusiasts out of a group of n people wanted to go to an escape room, and out of those k enthusiasts, one needed to be the team captain, how many ways are there to accomplish this? To select the escape room enthusiasts, we have $\binom{n}{k}$. From the group of k escape room enthusiasts, there are always $\binom{k}{1} = k$ ways to select the team captain. We multiply the two outcomes to yield $k\binom{n}{k}$. Also, since there may be anywhere from 1 to n enthusiasts, inclusive, we sum the outcomes in the following manner:

$$\sum_{k=1}^{n} k\binom{n}{k} = \binom{n}{1} + 2\binom{n}{2} + 3\binom{n}{3} + \cdots + n\binom{n}{n}.$$

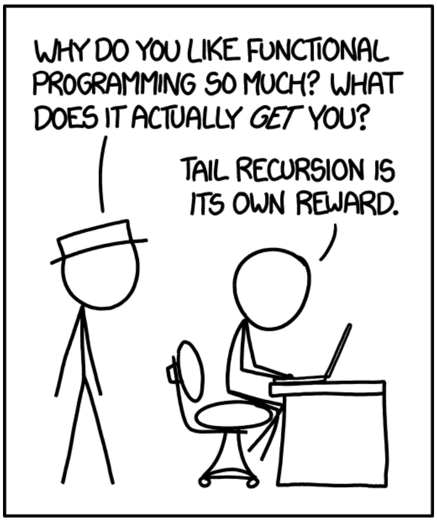
Alternatively, we may select one person from the group of n people that also want to go to an escape room to be the team captain; this gives n outcomes. Then, each remaining person from the group of $n-1$ has the choice of going to the escape room or not; there are 2^{n-1} ways to select the escape room enthusiasts from the group of $n-1$. We multiply the outcomes to yield

$$n2^{n-1}.$$

We now see that since the solution involving binomial coefficients and the solution involving powers of two were the answer to the same question, the equality

$$\binom{n}{1} + 2\binom{n}{2} + 3\binom{n}{3} + \cdots + n\binom{n}{n} = n2^{n-1}.$$

is hence proved. □



B
Functions

B.1 An Introduction to the Terminology of Functions

Consider the following definitions.

Definition B.1.1: Functions, Domains, and Codomains

A function F, from a domain A to a codomain B, that is, $F : A \to B$ is a map from the elements of a set A to a set B such that for all $a \in A$, there exists a unique $b \in B$ such that a is mapped to b by F.

Definition B.1.2: Images and Pre-Images

Let $F : A \to B$ be a function. For $a \in A$, the image of a is written as $F(a)$ and is the unique element of B to which a is mapped to by F. For $b \in B$, the pre-images of b are the elements of A that map to b by F.

Definition B.1.3: Range

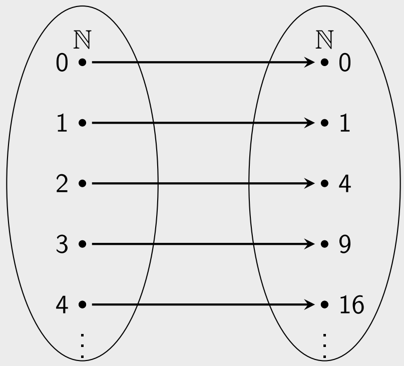
Let $F : A \to B$ be a function. The image of the domain, A, is the range of F.

Consider the following example.

Example B.1.1: ☢☢ Is it a Function?

Define $R : \mathbb{N} \to \mathbb{N}$ where for $a, b \in \mathbb{N}$, $a \sim_R b \iff b = a^2$. Determine if R is a function. State the domain, codomain, and range of R.

Consider the following diagram.



We recognize that R is a function, as each output only has one input. Notice that if R were a relation on $\mathbb{Z}$, this would not be true. Furthermore, the domain and codomain of R is $\mathbb{N}$. The range of R is the set of all perfect squares.

B.2 Injections, Surjections, and Bijections

Consider the following definitions.

Definition B.2.1: Injective Functions

Given a function $F : A \to B$, F is injective, or one-to-one, if and only if

$$\forall a_1, a_2 \in A, a_1 \neq a_2 \implies F(a_1) \neq F(a_2).$$

That is, F is injective if and only if

$$\forall a_1, a_2 \in A, F(a_1) = F(a_2) \implies a_1 = a_2.$$

Definition B.2.2: Surjective Functions

Given a function $F : A \to B$, F is surjective, or onto, if and only if

$$\operatorname{range} F = B.$$

That is, F is surjective if and only if

$$\forall b \in B, \exists a \in A, F(a) = b.$$

Definition B.2.3: Bijective Functions

Given a function $F : A \to B$, F is bijective, if and only if F is both injective and surjective. That is, F is bijective if and only if

$$(\forall a_1, a_2 \in A, F(a_1) = F(a_2) \implies a_1 = a_2) \wedge (\forall b \in B, \exists a \in A, F(a) = b).$$

Consider the following examples.

Example B.2.1: The Arctangent: Part I

Consider $F : \mathbb{R}^2 \to \mathbb{R}^2$ given by $F(x) = \arctan x$. Determine if F is injective, surjective, or bijective.

- Injective: F is injective.
  - Suppose $\arctan(a_1) = \arctan(a_2)$. If we take the tangent of both sides, we see that $a_1 = a_2$.
  - Alternatively, if $a_1 \neq a_2$, we observe that $\arctan(a_1) \neq \arctan(a_2)$ because $\arctan x$ is monotonically increasing.
- Surjective: F is not surjective.
  - We see that $\operatorname{range} F = \left\{x : -\frac{\pi}{2} < x < \frac{\pi}{2}\right\}$.
- Bijective: F is not bijective.

Example B.2.2: ☢☢ The Arctangent: Part II

Consider $F : \mathbb{R}^2 \to \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ given by $F(x) = \arctan x$. Determine if F is injective, surjective, or bijective.

- Injective: F is injective.
- Surjective: F is surjective.
- Bijective: F is bijective.

We note that to make a function injective, we can often restrict the domain. Similarly, to make a function surjective, we can modify the codomain.

B.3 Composition and Inverses

Consider the following definitions and theorems.

Definition B.3.1: Compositions

For functions $F : A \to B$ and $G : B \to C$, the composition of F and G is $G \circ F : A \to C$, which is given by

$$(G \circ F)(a) = G(F(a)).$$

Theorem B.3.1: Compositions, Injectivity, and Surjectivity

Let $F : A \to B$ and $G : B \to C$ be functions. Then,

1. If both F and G are injective, $G \circ F : A \to C$ is injective.

   Proof. Suppose both F and G are injective. Now, suppose

   $$(G \circ F)(a_1) = G(F(a_1)) = (G \circ F)(a_2) = G(F(a_2))$$

   for $a_1, a_2 \in A$. We wish to show that $a_1 = a_2$. We see that $F(a_1) = F(a_2)$ since G is injective. Then, since F is injective, $a_1 = a_2$. □

2. If both F and G are surjective, $G \circ F : A \to C$ is surjective.

   Proof. Suppose both F and G are surjective. Consider some arbitrary $c \in C$. We wish to find some $a \in A$ such that $(G \circ F)(a) = G(F(a)) = c$. Since G is surjective, there exists some $b \in B$ such that $G(b) = c$. Since F is surjective, there exists some $a \in A$ such that $F(a) = b$. Thus, $G(F(a)) = G(b) = c$. □

Definition B.3.2: Inverse Functions

The functions $F : A \to B$ and $G : B \to A$ are inverses of each other if and only if, for all $a \in A$ and $b \in B$,

$$(G \circ F)(a) = a$$

and

$$(F \circ G)(b) = b.$$

Theorem B.3.2: Existence of Inverse Functions

The function $F : A \to B$ has an inverse $G : B \to A$ if and only if F is bijective.

Proof. Suppose $F : A \to B$ has an inverse $G : B \to A$. Suppose $F(a_1) = F(a_2)$ for $a_1, a_2 \in A$. Since $F(a_1) = F(a_2)$, $G(F(a_1)) = G(F(a_2))$, but since G is an inverse of F,

$$G(F(a_1)) = a_1 = G(F(a_2)) = a_2.$$

Thus, F is injective. Consider some arbitrary $b \in B$. Then, we have $G(b) = a$ since G will map all $b \in B$ to some $a \in A$. Since F and G are inverses, we have $F(a) = F(G(b)) = b$, so F is surjective. We have, at this point, shown that F is bijective. Now, suppose F is bijective. Consider some arbitrary $b \in B$. Then, since F is surjective, there exists some $a \in A$ such that $F(a) = b$. Since F is surjective, a is unique. Now, consider the map $G : Y \to X$ which maps each $b \in B$ to its unique pre-image $a \in A$ under F. Then, $(F \circ G)(b) = F(G(b)) = F(a) = b$. We also have $(G \circ F)(a)$ to be the unique pre-image of $F(a)$ under F. We have that a is the unique pre-image, so $(G \circ F)(a) = a$, so F and G are inverses. $\square$

Theorem B.3.3: Uniqueness of Inverse Functions

If $F : A \to B$ has an inverse $G : B \to A$, G is the only inverse of X.

Proof. Suppose $G_1 : B \to A$ and $G_2 : B \to A$ are both inverses of F. We wish to show that for all $b \in B$, $G_1(b) = G_2(b)$. We have $(G_2 \circ F)(a) = a$ for all $a \in A$, since F and G_2 are inverses. Similarly, we have $(F \circ G_1)(b) = b$. We know $G_1(b) \in A$, so

$$G_1(b) = (G_2 \circ F)(G_1(b)) = G_2(F(G_1(b))) = G_2((F \circ G_1)(b)) = G_2(b),$$

as desired. $\square$

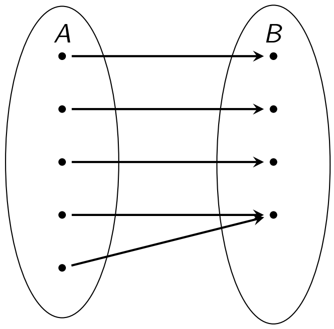
B.4 The Pigeonhole Principle

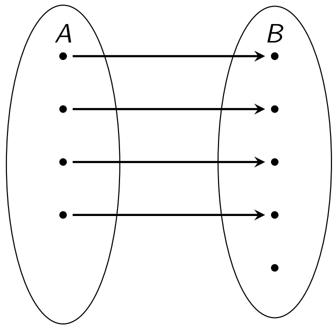
Consider the following theorem.

Theorem B.4.1: The Pigeonhole Principle

Let $F : A \to B$ be a function with finite sets A and B. If $|A| > |B|$, F is not injective.

The above result gets its name from the conceptual problem of a function that maps pigeons to holes. If there are more pigeons than holes, there exists a hole with more than one pigeon. To visualize this, consider the following diagrams.





While the Pigeonhole Principle may seem trivial, it may be used to construct various proofs. There are three parts to every Pigeonhole argument.

1. Define A, the set of pigeons.
2. Define B, the set of pigeonholes, such that $|A| > |B|$.
3. Define $F : A \to B$, the method of assigning pigeons to pigeonholes.

We may then conclude that

$$\exists a_1, a_2 \in A, a_1 \neq a_2 \wedge F(a_1) = F(a_2).$$

Consider the following examples.

Example B.4.1: ☢☢ Hairs

Prove that two people from the state of Colorado have the same number of hairs on their head.

Proof. The state of Colorado, at the time of writing, has roughly 5.8×10^6 people, and it is safe to assume that the number of human hairs is less than 5×10^5. Let A be the set of people in Colorado, with $|A| = 5.8 \times 10^6$ and let B be the set of all integers from 0 to 5×10^5, inclusive, noninclusive. We note that $|B| = 5 \times 10^5$. Let the function $F : A \to B$ be the function that maps a given person to the number of hairs on their head. We see that F is non-injective, therefore, at least two people from the state of Colorado must have the same number of hairs on their head. $\square$

Example B.4.2: ☢☢ Sphere

Prove that given 5 points on the surface of a sphere, there exists a hemisphere containing at least four of them. Any point on the boundary between the hemispheres is simultaneously in both hemispheres.

Proof. Pick two points, and cut the sphere in half such that the two points lie on the cut. Let A be the set of the three remaining points, and let B be the set of the two pieces of the sphere–the hemispheres. Let $F : A \to B$ map the points to their corresponding hemisphere. As $|A| = 3$ and $|B| = 2$, we see that F is non-injective, meaning that at least two remaining points will fall on the same hemisphere. These two points add to the two points that lie on the cut, giving four points in the hemisphere. $\square$

Example B.4.3: ☢☢☢ 1978 Putnam

Prove that any 20 distinct integers chosen from the set $S = \{1, 4, 7, 10, \dots, 100\}$ will contain a pair that sums to 104.

Before delving into the proof itself, we will proceed with some informal experimentation. Consider the following pairs S that sum to 104.

$$\begin{aligned} 104 &= \underbrace{4}_{1+1(3)} + 100 \\ &= \underbrace{7}_{1+2(3)} + 97 \\ &= \underbrace{10}_{1+3(3)} + 94 \\ &= \underbrace{13}_{1+4(3)} + 91 \\ &\vdots \\ &= \underbrace{49}_{1+16(3)} + 55. \end{aligned}$$

Note that there are 16 pairs of numbers that add to 104. Also, 1 and 52 are not able to be used in a pair that sums to 104. Therefore, any choice of 20 distinct integers from S will contain at least 18 distinct integers selected from $S - \{1, 52\}$. We are now ready to begin our proof.

Proof. Let A be 18 distinct integers chosen from $S - \{1, 52\}$. Let B be the set of pairs of integers that sum to 104. That is,

$$\begin{aligned} B &= \{\{4, 100\}, \{7, 97\}, \{10, 94\}, \dots, \{49, 55\}\} \\ &= \{\{1 + 3n, 103 - 3n\} : n \in \{1, 2, 3, \dots, 16\}\}. \end{aligned}$$

Note that $|A| = 18$ and $|B| = 16$. Let $F : A \to B$ be given by

$$F(a) = \{a, 104 - a\}.$$

for $a \in A$. The function F is non-injective by the Pigeonhole Principle, so there are two distinct elements of A that are mapped to the same element of B. These two elements are the pair that will sum to 104. $\square$

Example B.4.4: ☢☢☢ Subset Sum

Let S be a set of 10 positive 2-digit numbers. Prove that there exist two disjoint subsets of S that have the same sum.

Proof. A set with n elements has 2^n subsets. Therefore, S has 1024 subsets. The upper bound for the sum of a subset of S would be the sum of the largest subset of S when $S = \{90, 91, 92, 93, 94, 95, 96, 97, 98, 99\}$. This sum is 945. Therefore, the sum of a subset of S will be in the interval $[0, 945]$. Let A be the power set of S and let B be the set of all integers in the closed interval $[0, 945]$. Let $F : A \to B$ be the function that maps a subset to its sum. As $|A| > |B|$, at least two subsets must have the same sum. If these two subsets are not disjoint, we just remove the elements in the intersection from both subsets. The sum of the subsets will still be equal. □

The Extended Pigeonhole Principle is, well, an extended form of the Pigeonhole Principle. Consider the following statement.

Theorem B.4.2: The Extended Pigeonhole Principle

If n "pigeons" land into k "pigeonholes," there exists at least one pigeonhole with at least $\lfloor \frac{n-1}{k} \rfloor = \lceil \frac{n}{k} \rceil$ pigeons.

We may use Theorem B.4.2 to better quantify the "population" of the holes. Consider the following example.

Example B.4.5: ☢☢☢ Equilateral Triangle

Prove that given 9 points in an equilateral triangle with unit sides, there exist 3 that define a triangle of area less than or equal to $\frac{\sqrt{3}}{8}$.

Proof. Let A be the set of the nine points in the triangle. Let B be the set of four equilateral triangles given by the first iteration of Sierpinski's Triangle. Let $F : A \to B$ map each point to the triangle that contains the point. There must exist a triangle containing $\lceil \frac{9}{4} \rceil = 3$ points. The four equilateral triangles have area $\frac{\sqrt{3}}{4 \cdot 2}$, and the triangle formed by the three points is therefore less than or equal to $\frac{\sqrt{3}}{4 \cdot 2}$. □

Example B.4.6: ☢☢☢ Square

Prove that among any nine points placed in a unit square, there exist three that form a triangle of area less than or equal to $\frac{1}{8}$.

Proof. Let A be the set of all nine points in the unit square. Let B be the set of four squares of equivalent area given by the corners of the unit square. This can informally be thought of as "folding" the square in half twice. Each of these smaller squares has area $\frac{1}{4}$. Let $F : A \to B$ map each point in the unit square to the smaller square that contains the point. There must exist three points that make up a triangle in one of the small squares because $\lceil \frac{9}{4} \rceil = 3$. The largest triangle that can be enclosed in one of the smaller squares has area $\frac{1}{8}$, so the three points mentioned earlier must form a triangle with area $A \leq \frac{1}{8}$. □