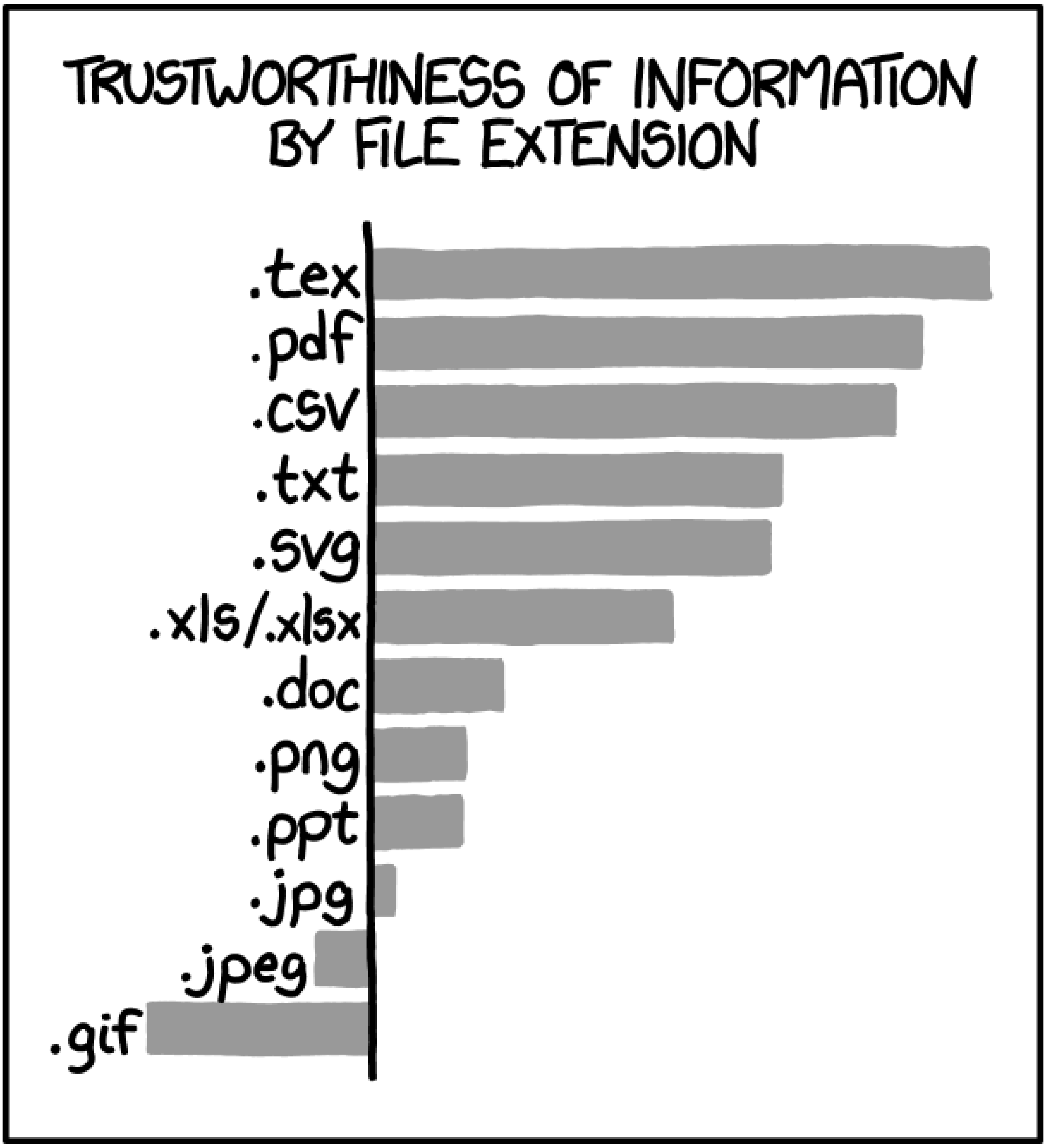
MATH2135: LINEAR ALGEBRA

ADITHYA BHASKARA
PROFESSOR: ROBIN DEELEY

TEXTBOOK: STEPHEN ANDRILLI & DAVID HECKER

UNIVERSITY OF COLORADO BOULDER

Draft: October 17, 2022


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

Contents

1 Vectors and Matrices **1**
1.1 Lecture 1: August 22, 2022 . . . 1
1.1.1 Notations, Definitions, and Conventions . . . 1
1.2 Lecture 2: August 24, 2022 . . . 5
1.2.1 Vector Operations . . . 5
1.3 Lecture 3: August 26, 2022 . . . 8
1.3.1 Matrices . . . 8
1.4 Lecture 4: August 29, 2022 . . . 14
1.4.1 The Dot Product . . . 14
1.5 Lecture 5: August 31, 2022 . . . 19
1.5.1 Matrix Multiplication . . . 19

2 Systems of Linear Equations **22**
2.1 Lecture 6: September 2, 2022 . . . 22
2.1.1 Systems of Linear Equations . . . 22
2.2 Lecture 7: September 7, 2022 . . . 23
2.2.1 Systems of Linear Equations as Matrices . . . 23
2.2.2 Matrix Row Operations . . . 23
2.3 Lecture 8: September 9, 2022 . . . 25
2.3.1 Solving Linear Systems . . . 25
2.4 Lecture 9: September 12, 2022 . . . 30
2.4.1 Formalizing Previous Notions: Part I . . . 30
2.5 Lecture 10: September 14, 2022 . . . 34
2.5.1 Formalizing Previous Notions: Part II . . . 34
2.6 Lecture 11: September 16, 2022 . . . 41
2.6.1 Linear Maps . . . 41
2.7 Lecture 12: September 19, 2022 . . . 42
2.7.1 Inverses of Matrices . . . 42
2.8 Lecture 13: September 21, 2022 . . . 45
2.8.1 Finding Inverses of Matrices . . . 45

3 Determinants and Eigenvalues **48**
3.1 Lecture 13: September 21, 2022 . . . 48
3.1.1 Defining the Determinant . . . 48
3.2 Lecture 14: September 23, 2022 . . . 51
3.2.1 Determinants of Upper Triangular Matrices . . . 51
3.3 Lecture 15: September 26, 2022 . . . 54
3.3.1 Further Properties of Determinants . . . 54
3.3.2 Eigenvectors, Eigenvalues, and Diagonalization . . . 55
3.4 Lecture 16: September 28, 2022 . . . 56

3.4.1 The Process of Diagonalization: Part I . . . 56
3.5 Lecture 17: September 30, 2022 . . . 60
3.5.1 The Process of Diagonalization: Part II . . . 60
3.6 Lecture 18: October 3, 2022 . . . 63
3.6.1 The Process of Diagonalization: Part III . . . 63
3.6.2 Using Diagonalization to Raise Matrices to Powers . . . 65
3.7 Lecture 19: October 5, 2022 . . . 66
3.7.1 Complex Numbers . . . 66
3.8 Lecture 20: October 7, 2022 . . . 69
3.8.1 Diagonalization Using Complex Eigenvalues . . . 69

4 Finite Dimensional Vector Spaces 70
4.1 Lecture 20: October 7, 2022 . . . 70
4.1.1 The Process of Abstraction . . . 70
4.2 Lecture 21: October 14, 2022 . . . 72
4.2.1 Derived Properties of Vector Spaces . . . 72
4.3 Lecture 22: October 17, 2022 . . . 74
4.3.1 Subspaces . . . 74

5 Linear Transformations 80

6 Orthogonality 81

Appendices

A Introduction to Proofs 83
A.1 Introduction to Proofs . . . 83
A.2 Proof by Induction . . . 83
A.3 Direct Proofs . . . 91
A.4 Proof by Contrapositive . . . 92
A.5 Proof by Contradiction . . . 93
A.6 Combinatorial Proofs . . . 94

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

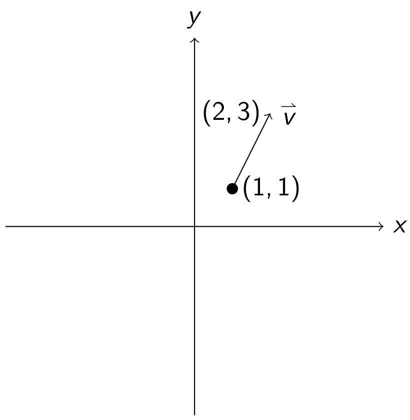
1.1.1 Notations, Definitions, and Conventions

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
$\mathbb{R}^n$	$\{[v_1, v_2, \dots, v_n] : v_1, \dots, v_n \in \mathbb{R}\}$

.

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

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

Definition 1.1.2: Vector Equality

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

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

Definition 1.1.3: Vector Magnitude

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

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

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

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

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 $v_1, \ldots, v_n$ are zero, which, by definition, means that $\vec{v} = \vec{0}$. $\square$

1.2 Lecture 2: August 24, 2022

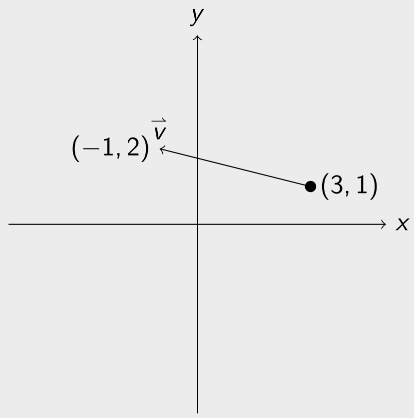
1.2.1 Vector Operations

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 $\frac{[3,-1,-\pi]}{\sqrt{3^2+(-1)^2+(-\pi)^2}}$.

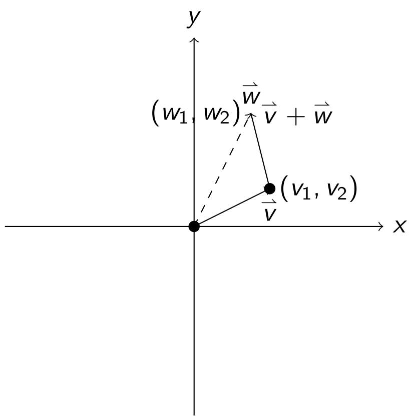
We will now define addition and subtraction with vectors and provide a few properties.

Definition 1.2.1: Addition and Subtraction With Vectors

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

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

Geometrically, we may visualize vector addition as



.

Theorem 1.2.1: Properties of Vector Addition and Scalar Multiplication

Let $\vec{u} = [u_1, \ldots, u_n]$, $\vec{v} = [v_1, \ldots, v_n]$, and $\vec{w} = [w_1, \ldots, w_n]$ 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 $\vec{v} = [v_1, \ldots, v_n]$ and $\vec{w} = [w_1, \ldots, w_n]$ be vectors in $\mathbb{R}^n$. Prove that

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

Proof. Consider the following.

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

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,

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

Proof. First, we wish to show that

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

Suppose $c = 0$. We have

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

We wish to show that all of $cv_1, \ldots, cv_n$ 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

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

and wish to show that all of $cv_1, \ldots, cv_n$ 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, $\left(c = 0 \vee \vec{v} = \vec{0}\right) \implies c\vec{v} = \vec{0}$. Now, we wish to show that

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

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

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

There must be some nonzero $v_1, \ldots, v_n$. 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 $\vec{v} = [v_1, \ldots, v_n]$ and $\vec{w} = [w_1, \ldots, w_n]$ 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. $\square$

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

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

Consider the following example.

Example 1.4.4: Projections

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

We see that

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

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. □

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. □

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

$$\left[\begin{array}{cccc|c} 3 & -6 & 0 & 3 & 9 \\ -2 & 4 & 2 & -1 & -11 \\ 4 & -8 & 6 & 7 & -5 \end{array}\right].$$

By row operations, we obtain

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

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

$$\left[\begin{array}{cccc|c} 3 & 1 & 7 & 2 & 13 \\ 2 & -4 & 14 & -1 & -10 \\ 5 & 11 & -7 & 8 & 59 \\ 2 & 5 & -4 & -3 & 39 \end{array}\right].$$

By row operations, we obtain

$$\left[\begin{array}{cccc|c} 1 & \frac{1}{3} & \frac{7}{3} & \frac{2}{3} & \frac{13}{3} \\ 0 & 1 & -2 & \frac{1}{2} & 4 \\ 0 & 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right].$$

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:

$$\begin{bmatrix} 2 & -1 & 1 & \mid & 0 \\ 1 & 3 & 4 & \mid & 0 \end{bmatrix}$$

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

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

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

$$\begin{bmatrix} 1 & -\frac{1}{2} & \frac{1}{2} & \mid & 0 \\ 0 & -\frac{7}{2} & -\frac{7}{2} & \mid & 0 \end{bmatrix}.$$

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

$$\begin{bmatrix} 1 & -\frac{1}{2} & \frac{1}{2} & \mid & 0 \\ 0 & 1 & 1 & \mid & 0 \end{bmatrix}.$$

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,

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

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 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 $[E|F]$. $\square$

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 1

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

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

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

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

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

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

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

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

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

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

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

$$\begin{bmatrix} 2 & -4 & | & 14 \\ -3 & 6 & | & -21 \\ 1 & 2 & | & 7 \end{bmatrix}.$$

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

$$\begin{bmatrix} 1 & -2 & | & 7 \\ -3 & 6 & | & -21 \\ 1 & 2 & | & 7 \end{bmatrix}.$$

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

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

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

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

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

$$\begin{bmatrix} 1 & -2 & | & 7 \\ 2 & 0 & | & 14 \\ 0 & 0 & | & 0 \end{bmatrix}.$$

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

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

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

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

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

$$\begin{bmatrix} 1 & 0 & | & 7 \\ 0 & 1 & | & 0 \\ 0 & 0 & | & 0 \end{bmatrix}.$$

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

$$\left[\begin{array}{cc|c} 1 & 5 & 3 \\ 2 & 10 & 6 \end{array}\right],$$

which reduces to

$$\left[\begin{array}{cc|c} 1 & 5 & 3 \\ 0 & 0 & 0 \end{array}\right].$$

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

$$\left[\begin{array}{cc|c} 1 & 5 & 1 \\ 2 & 10 & 0 \end{array}\right],$$

which reduces to

$$\left[\begin{array}{cc|c} 1 & 5 & 1 \\ 0 & 0 & -2 \end{array}\right].$$

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.

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

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

We perform the row operation $\langle 2 \rangle \leftrightarrow \langle 1 \rangle$, yielding

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

Then, we execute $2\langle 1 \rangle - \langle 2 \rangle \to \langle 2 \rangle$, which yields

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

We then compute $\langle 2 \rangle - \langle 1 \rangle \to \langle 1 \rangle$, producing

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

Finally, we simply perform $-\langle 1 \rangle \to \langle 1 \rangle$, which obtains

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

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.

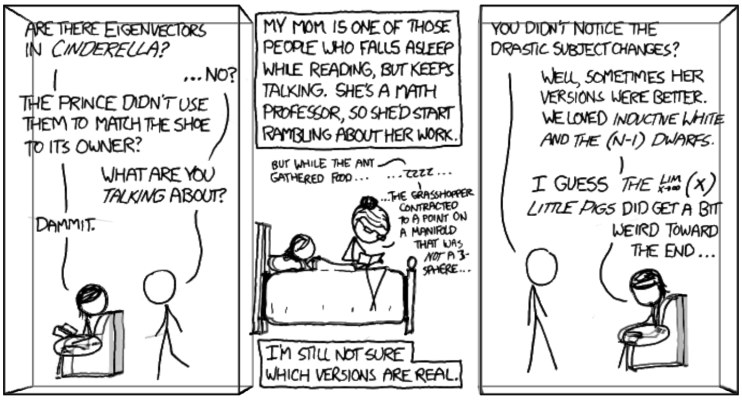
Proof. Consider the case where A is nonsingular; by definition, A^{-1} exists. Consider $X = A^{-1}B$ as a prospective solution. To verify it, we have

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

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. □



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.

```python
# 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 $R_1, \ldots, R_k$ such that $A = R_1(\ldots(R_k(I_n)\ldots))$ and

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

   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

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

   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 $X \neq 0$ 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 3.4.1, 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 $\lambda 2$ into $(A - \lambda I_2) = 0$ and solve for λ. For λ_1, we have

$$\begin{aligned} \begin{bmatrix} 0 \\ 0 \end{bmatrix} &= \begin{bmatrix} 7-4 & 1 \\ -3 & 3-4 \end{bmatrix} \\ &= \begin{bmatrix} 3 & 1 \\ -3 & -1 \end{bmatrix}. \end{aligned}$$

By row reduction, we have

$$\begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 & \frac{1}{3} \\ 0 & 0 \end{bmatrix}$$

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 3.3.1. 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 $\lambda_1, \dots, \lambda_k$ 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,

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

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

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

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. $z = a + bi \in \mathbb{C} \implies \bar{z} := a - bi$.

Theorem 3.7.1: Properties of Complex Numbers

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

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

Theorem 3.7.2: The Fundamental Theorem of Algebra

If $p(z) = a_n z^n + \cdots + a_0 z^0$ is an nth degree polynomial with coefficients $a_0, \ldots, a_n \in \mathbb{C}$, $p(z)$ 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,

$$\vec{v} \cdot \vec{w} = v_1\bar{w}_1 + \cdots + v_n\bar{w}_n.$$

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 $\mathcal{M}^{\mathbb{C}}_{mn}$? We do not.

Definition 3.7.8: The Adjoint

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

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

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 $A \in \mathcal{M}^{\mathbb{R}}_{nn}$,

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

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 of the sort $A \in \mathcal{M}_{nn}^{\mathbb{C}}$. 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

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

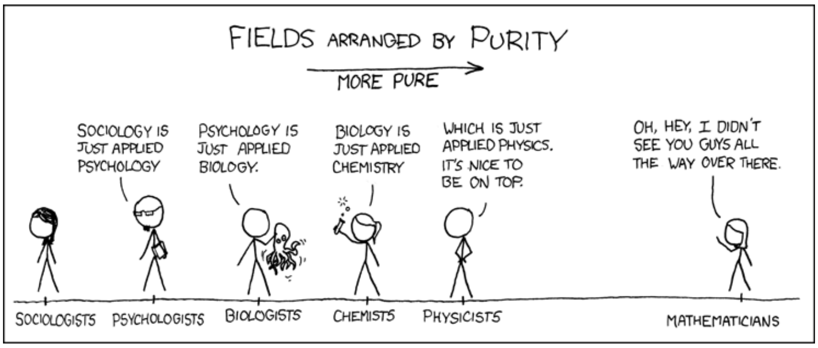
Then, we have $-i\langle 1 \rangle + \langle 2 \rangle \to \langle 2 \rangle$, producing

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

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: $\mathbb{F} \times V \to V, (c, \vec{v}) \mapsto c\vec{v}$.

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. $\exists \vec{0} \in V, \forall \vec{v} \in V, \vec{0} + \vec{v} = \vec{v}$
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 non-empty 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 $\{a_n x^n + \cdots + a_0 x^0 : a_0, \ldots a_n \in \mathbb{R}\}$ is a vector space.
7. 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.

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 note that by axiom 3 of Definition 4.1.1,

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

By axiom 4,

$$c\vec{0} = c\vec{0} + c\vec{0} + (-c\vec{0}).$$

Then, by axiom 5,

$$c\vec{0} = (\vec{0} + \vec{0}) + (-c\vec{0}).$$

By axiom 3, we have

$$c\vec{0} + (-c\vec{0}).$$

Then, by axiom 4, we have

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

and we have proved the proposition. □

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. Consider, by axiom 3,

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

Then, by axiom 4, we have

$$0\vec{v} = 0\vec{v} + 0\vec{v} + (-0\vec{v}).$$

By axiom 6,

$$0\vec{v} = (0 + 0)\vec{v} + (-0\vec{v}) = 0\vec{v} + (-0\vec{v}).$$

Finally, by axiom 4,

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

and we have proved the proposition. □

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. Consider, by axiom 8,

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

Then, by axiom 6,

$$\vec{v} + (-1)\vec{v} = (1 + (-1))\vec{v} = 0\vec{v}.$$

By Theorem 4.2.2,

$$\vec{v} + (-1)\vec{v} = \vec{0},$$

which implies $(-1)\vec{v} = -\vec{v}$, the additive inverse of $\vec{v}$. □

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,

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

which implies $\vec{v} = \vec{0}$. □

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 the set $W \subseteq V$. 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 the set W is a subset of 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. $\vec{w} \in W \implies c\vec{w} \in W, c \in \mathbb{F}$.

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 for W. The closure properties are given by conditions 2 and 3. For vector space axiom 1, we take $\vec{w}_1, \vec{w}_2 \in W$. Then,

$$\begin{aligned}\vec{w}_1 +_W \vec{w}_2 &= \vec{w}_1 +_V \vec{w}_2 \\ &= \vec{w}_2 +_V \vec{w}_1 \\ &= \vec{w}_2 +_W \vec{w}_1.\end{aligned}$$

For vector space axiom 3, we take $\vec{w} \in W$ and by the closure property of scalar multiplication, $(-1)\vec{w} \in W$. By the closure property of vector addition $\vec{w} + (-1)\vec{w} = \vec{0}$, meaning $\vec{0} \in W$. Moreover, for any $\vec{w} \in W$, $\vec{w} + \vec{0} = \vec{0} + \vec{w} = \vec{w}$. For vector space axiom 4, $-1\vec{w}$ is the additive inverse of $\vec{w}$, and by the closure property, $(-1)\vec{w} \in W$. We must show the rest.

□

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 $[x, 0] \in W$,
$$c[x, 0] = [cx, 0].$$
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 $V = \mathcal{M}_{nn}$.
$$W = \mathcal{D}_n \subseteq \mathcal{U}_n \subseteq \mathcal{M}_{nn},$$
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 $V = \mathcal{M}_{nn}$.

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

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 $\mathcal{M}_{nn}$.

To show that W is not a subspace of V,

1. If $\vec{0} \notin S$, S is not a subspace.
2. Find $\vec{v}_1, \vec{v}_2 \in S$ such that $\vec{v}_1 + \vec{v}_2 \notin S$.
3. Find $\vec{v} \in S$ such that $(-1)\vec{v} \notin S$.
4. Find $\vec{v} \in S$ and $c \in \mathbb{F}$ such that $c\vec{v} \notin S$.

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.

We now revisit the notion of linear combinations.

Theorem 4.3.2: Linear Combinations Remain in a Subspace

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, $c_1\vec{v}_1 \in W$. For the proposition when $n = 2$, consider scalars $c_1, c_2 \in \mathbb{F}$ and vectors $\vec{v}_1, \vec{v}_2 \in W$. Then, by the closure property of scalar multiplication, $c_1\vec{v}_1, c_2\vec{v}_2 \in W$. Then, since $c_1\vec{v}_1, c_2\vec{v}_2 \in W$, $c_1\vec{v}_1 + c_2\vec{v}_2 \in W$ by the closure property of vector addition. Suppose the theorem holds for all $n = k$. That is, for scalars $c_1, \dots, c_n$ 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 $c_1\vec{v}_1 + \cdots + c_k\vec{v}_k \in W$,

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

proving the proposition. □

Essentially, Theorem 4.3.2 provides that subspaces are closed under linear combinations. Consider the following theorem.

Theorem 4.3.3: 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 $\mathbb{R}^n$. $\square$

As we did in Theorem 4.3.2, we will one again revisit the notion of linear combinations. Consider the following definitions.

Definition 4.3.2: 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 $c_1, \dots, c_n$.

Definition 4.3.3: Span

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

Consider the following examples.

Example 4.3.4: Find Span 1

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

We realize that

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

Example 4.3.5: 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 $\operatorname{span}(S)$.

We realize that

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

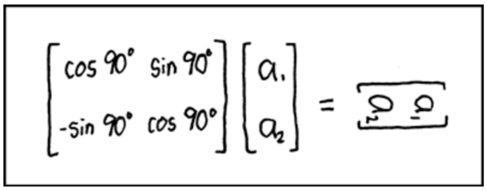
Sometimes, for some $S \subseteq V$, $\operatorname{span}(S) = V$. Here, we say that S spans V, or equivalently, V is spanned by S. Consider the following example.

Example 4.3.6: Find Span 3

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

We realize that

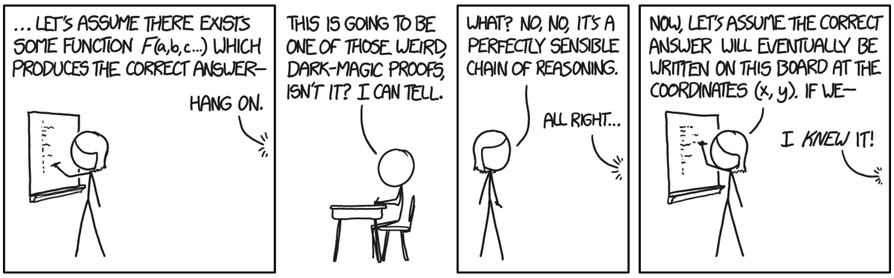
$$\begin{aligned}\operatorname{span}(S) &= \{c_1[1,0,0] + c_2[0,1,0] + c_3[0,0,1] : c_1, c_2, c_3 \in \mathbb{R}\}\\ &= \{[c_1, c_2, c_3] : c_1, c_2, c_3 \in \mathbb{R}\}\\ &= \mathbb{R}^3.\end{aligned}$$



5 Linear Transformations

6
Orthogonality

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

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

$$\begin{array}{c} \hline P(0) \\ \forall k \in \mathbb{N}, P(k) \implies P(k+1) \\ \hline \therefore \forall n \in \mathbb{N}, P(n). \\ \hline \end{array} .$$

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 case base.

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 $n \geq 1$. $\square$

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: ☢☢ nth 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 $n \geq 0$. $\square$

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. □