MATH3430: DIFFERENTIAL EQUATIONS

ADITHYA BHASKARA PROFESSOR: ELIZABETH GRULKE

TEXTBOOK: MARTIN BRAUN

UNIVERSITY OF COLORADO BOULDER





Contents

Pr	Preface iv						
1	First	t Order Differential Equations	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		ond Order Differential 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				
	2.9	Additional Topics: The <i>LU</i> Factorization	48				
		2.9.1 Presented by Olver and Shakiban's Applied Linear Algebra	48				
_			••				
3		tems of Differential Equations	49				
	3.1	Lecture 13: September 21, 2022					
	0.0	3.1.1 Defining the Determinant					
	3.2	Lecture 14: September 23, 2022					
		3.2.1 Determinants of Upper Triangular Matrices	52				

CONTENTS iii

3.3						
	3.3.1 Further Properties of Determinants					
_	3.3.2 Eigenvectors, Eigenvalues, and Diagonalization					
3.						
_	3.4.1 The Process of Diagonalization: Part I					
3						
_	3.5.1 The Process of Diagonalization: Part II					
3.						
	3.6.1 The Process of Diagonalization: Part III					
	3.6.2 Using Diagonalization to Raise Matrices to Powers					
3.						
	3.7.1 Complex Numbers					
3.	,					
	3.8.1 Diagonalization Using Complex Eigenvalues					
4 Q	ualitative Theory of Differential Equations 71					
4.	·					
	4.1.1 The Process of Abstraction					
4.:						
	4.2.1 Derived Properties of Vector Spaces					
4.	·					
	4.3.1 Subspaces					
4.4	•					
	4.4.1 Span					
4.						
	4.5.1 Linear Independence					
4.	·					
	4.6.1 Determining Linear Independence in \mathbb{R}^n					
	4.6.2 Bases: Part I					
4.						
	4.7.1 Bases: Part II					
4.8						
	4.8.1 Bases: Part III					
4.	9 Lecture 28: October 31, 2022					
	4.9.1 Coordinatization: Part I					
4.	10 Lecture 29: November 2, 2022					
	4.10.1 Coordinatization: Part II					
Diffe	Differential Equations as a Word Cloud 112					

Preface

To the interested reader.

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

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

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

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

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

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

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

Best Regards, Adithya Bhaskara

REVISED: December 18, 2022

¹Andrilli, S., & Hecker, D. (2016). *Elementary Linear Algebra* (5th ed.). Academic Press.

²Levin, O. (2019). Discrete Mathematics: An Open Introduction (3rd ed.). Oscar Levin.

³Hammack, R. (2018). Book of Proof (3rd ed.). Richard Hammack.



First Order Differential Equations

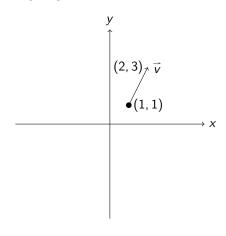
1.1 Lecture 1: August 22, 2022

1.1.1 Notations, Definitions, and Conventions

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

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

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



Consider the following definitions and statements.

Definition 1.1.1: The Zero Vector

We define

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

Definition 1.1.2: Vector Equality

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

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

Definition 1.1.3: • Vector Magnitude

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

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

Definition 1.1.4: Scalar Multiplication

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

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

Consider the following theorem.

Theorem 1.1.1: Scalar Multiplication and Magnitude

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

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

Proof.

$$||c\vec{v}|| = \sqrt{(cv_1)^2 + \dots + (cv_n)^2}$$

$$= \sqrt{c^2v_1^2 + \dots + c^2v_n^2}$$

$$= \sqrt{c^2}\sqrt{v_1^2 + \dots + v_n^2}$$

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

The theorem is hence proved.

Consider the following definitions.

Definition 1.1.5: Wector 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

$$||\vec{u}|| = \left| \left| \frac{1}{||\vec{v}||} \vec{v} \right| \right|$$

$$= \left| \frac{1}{||\vec{v}||} \right| ||\vec{v}||$$

$$= \frac{1}{||\vec{v}||} ||\vec{v}||$$

This means that \vec{u} is a unit vector and, because $\frac{1}{||\vec{v}||}$ is a positive 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 $|c|=rac{1}{||\overrightarrow{v}||}.$ Because c>0, |c|=c, so $\overrightarrow{w}=rac{1}{||\overrightarrow{v}||}\overrightarrow{v}=\overrightarrow{u}.$

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

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

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

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

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

1.2 Lecture 2: August 24, 2022

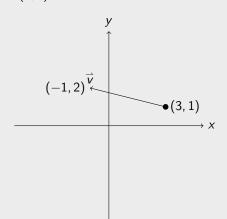
1.2.1 Vector Operations

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: • Vector Addition

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

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

Geometrically, we may visualize vector addition as



Theorem 1.2.1: Properties of Vector Addition and Scalar Multiplication

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

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

Proof. Consider the following.

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

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

The proposition is hence proved.

Theorem 1.2.2: Scalar Multiplication Producing the Zero Vector

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

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

Proof. First, we wish to show that

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

Suppose c = 0. We have

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

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

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

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

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

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

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

There must be some nonzero v_1, \dots, v_n . Let this nonzero number be v_i . The only way $cv_i = 0$ 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 $A=(a_1,\ldots,a_n)$ to $B=(b_1,\ldots,b_n)$. 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,

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

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} ,

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 $a_{ii} = 0$ for all i > j.

Definition 1.3.7: Definition 1.3.7: Definition

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

Definition 1.3.8: \blacksquare **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: Properties of the Transpose

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

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

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,

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

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,

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

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,

$$(A + A^{T})^{T} = A^{T} + (A^{T})^{T}$$
$$= A^{T} + A$$
$$= A + A^{T}.$$

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

$$(A - A^{T})^{T} = A^{T} - (A^{T})^{T}$$
$$= A^{T} - A$$
$$= -(A - A^{T}).$$

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

- 1. $\vec{u} \cdot \vec{v} = \vec{v} \cdot \vec{u}$.
- 2. $\vec{u} \cdot \vec{u} = ||\vec{u}||^2 \ge 0$.
- 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 $\vec{v} = [v_1, ..., v_n]$ and $\vec{w} = [w_1, ..., w_n]$ 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:

$$c(\vec{v} \cdot \vec{w}) = c(v_1 w_1 + \dots + v_n w_n)$$

$$= (cv_1 w_1 + \dots + cv_n w_n)$$

$$= ((cv_1)w_1 + \dots + (cv_n)w_n)$$

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

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} , $\vec{w} \in \mathbb{R}^2$,

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

where θ is the angle between the two vectors.

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

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

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

$$\hat{\mathbf{v}} \cdot \hat{\mathbf{w}} = \cos \alpha \cos \beta + \sin \beta \sin \alpha$$
$$= \cos(\beta - \alpha)$$
$$= \cos \theta.$$

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. \Box

We will now state the famous Cauchy-Schwarz Inequality.

Theorem 1.4.3: The Cauchy-Schwarz Inequality

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

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

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

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

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

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

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

$$||\vec{v} - \vec{w}||^2 = (\vec{v} - \vec{w}) \cdot (\vec{v} - \vec{w})$$

$$= ||\vec{v}||^2 - 2\vec{v} \cdot \vec{w} + ||\vec{w}||^2$$

$$= 2 - 2\vec{v} \cdot \vec{w}$$

$$\geq 0.$$

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

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

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

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

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

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

We have

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

The theorem is hence proved.

We will now define the projection onto a vector.

Definition 1.4.2: The Projection Onto a Vector

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

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

Consider the following example.

Example 1.4.4: * Projections

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

We see that

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

Consider the following theorems.

Theorem 1.4.5: Sum of Parallel and Perpendicular Projections

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

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

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

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

$$\operatorname{proj}_{\overrightarrow{w}} \overrightarrow{v} + (\overrightarrow{v} - \operatorname{proj}_{\overrightarrow{w}} \overrightarrow{v}) = \frac{\overrightarrow{v} \cdot \overrightarrow{w}}{\overrightarrow{w} \cdot \overrightarrow{w}} \overrightarrow{w} + \left(\overrightarrow{v} - \frac{\overrightarrow{v} \cdot \overrightarrow{w}}{\overrightarrow{w} \cdot \overrightarrow{w}} \overrightarrow{w} \right)$$
$$= \overrightarrow{v}.$$

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 $c \neq 0$,

$$\operatorname{proj}_{\overrightarrow{w}}\overrightarrow{v} = \operatorname{proj}_{\overrightarrow{cw}}\overrightarrow{v}.$$

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

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

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 *i*th row of A and the *j*th 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 $k \ge 2$,

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

Second Order Differential 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

$$a_{11}x_1+\cdots+a_{1n}x_n=b_1$$

:

$$a_{m1}x_1+\cdots+a_{mn}x_n=b_m.$$

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 $c\langle r_1 \rangle o \langle r_1 \rangle$.
- 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
$$\begin{bmatrix} 3 & 1 & -1 \\ 1 & 0 & 1 \\ -1 & 1 & 5 \end{bmatrix}. \text{ Find } 4\langle 3 \rangle \rightarrow \langle 3 \rangle.$$

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 $\langle 1 \rangle + (-3)\langle 2 \rangle \rightarrow \langle 1 \rangle$.

We obtain

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

Example 2.2.3: ** Row Operation 3

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

We obtain

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

2.3 Lecture 8: September 9, 2022

2.3.1 Solving Linear Systems

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

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

Consider the following examples.

Example 2.3.1: * No Solution

Consider the matrix

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

By row operations, we obtain

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

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

Example 2.3.2: * Infinitely Many Solutions

Consider the matrix

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

By row operations, we obtain

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

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

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

For x_1 , we have

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

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

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

Example 2.3.3: * No More Back Substitution

Note that the matrix

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

reduces into

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

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

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

Then, we will perform $\langle 1 \rangle + \langle 2 \rangle \rightarrow \langle 1 \rangle$, 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 $\frac{2}{3}\langle 3\rangle + \langle 1\rangle \rightarrow 1$, 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 $R_1, ..., R_n$ are row operations, $R_n(...(R_2(R_1(AB)))...) = (R_n(...(R_2(R_1(A)))...))B$.

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:

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

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

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

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

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

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

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

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 $\langle 1 \rangle - \langle 2 \rangle \rightarrow \langle 2 \rangle$, yielding

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

Then, we perform $\langle 1 \rangle - \langle 3 \rangle \rightarrow \langle 3 \rangle.$ 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 $\langle 3 \rangle + 9 \langle 2 \rangle \rightarrow \langle 3 \rangle.$ 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:

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

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

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

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

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

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

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

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

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

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

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

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

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,

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

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,

$$cX + X_{\circ}$$

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

$$A(cX + \hat{X}) = A(cX) + A\hat{X}$$

$$= cAX + A\hat{X}$$

$$= \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} + \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}$$

$$= B,$$

as desired.

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

Definition 2.4.4: • Equivalence of Linear Systems

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

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

That is, if they have the same solution sets.

Definition 2.4.5: Row Equivalence

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

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

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

Definition 2.4.6: © **Equivalence Relations**

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

- 1. \forall *a* ∈ *S*, *a* ∼ *a*.
- 2. $\forall a, b \in S, a \sim b \implies b \sim a$
- 3. $\forall a, b, c \in S, a \sim b \land b \sim c \implies a \sim c$.

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
$c\langle i \rangle o \langle i \rangle$	$\frac{1}{c}\langle i\rangle \rightarrow \langle i\rangle$
$c\langle i\rangle + \langle j\rangle \to \langle j\rangle$	$-c\langle i\rangle + \langle j\rangle \rightarrow \langle j\rangle$
$\langle i \rangle \leftrightarrow \langle j \rangle$	$\langle i \rangle \leftrightarrow \langle j \rangle$

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

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

2.5 Lecture 10: September 14, 2022

2.5.1 Formalizing Previous Notions: Part II

Consider the following formalization of our last discoveries.

Theorem 2.5.1: ■ Row Equivalence Implies System Equivalence

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

Theorem 2.5.2: Uniqueness of Reduced Row Echelon Form

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

Definition 2.5.1: Rank

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

Consider the following example.

Example 2.5.1: * Rank

Consider

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

By row reduction, we have the matrix

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

and see that rank A = 2.

Consider the following theorem.

Theorem 2.5.3: ■ Number of Solutions to Homogeneous Systems

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

- 1. If rank A < n, AX = 0 has an infinite solution set.
- 2. If 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,

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

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 $2\langle 1\rangle - \langle 2\rangle \rightarrow \langle 2\rangle$, 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 $-\frac{1}{2}\langle 2\rangle \rightarrow \langle 2\rangle$ 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 $\langle 2 \rangle - \langle 3 \rangle \rightarrow \langle 3 \rangle$; 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 $\frac{1}{2}\langle 1 \rangle \rightarrow \langle 1 \rangle$ to obtain

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

Then, we have $3\langle 1 \rangle + \langle 2 \rangle \rightarrow \langle 2 \rangle$, 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 $\langle 1 \rangle + \langle 2 \rangle \rightarrow \langle 2 \rangle$; we have

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

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

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

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

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

Finally, we have the operation $\langle 1 \rangle + 2 \langle 2 \rangle \rightarrow \langle 1 \rangle$, 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

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

By row reduction, we finally obtain

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

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

Consider the following definition.

Definition 2.5.3: Row Space

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

Consider the following examples.

Example 2.5.6: * Row Space 1

Consider

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

The row space of A is

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

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

Example 2.5.7: * Row Space 2

Consider

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

The row space of A is

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

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

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

Example 2.5.8: * Are Vectors in the Row Space?

Consider

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

and recall that the row space of A is

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

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

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

For [3, 6], we have

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

which reduces to

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

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

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

which reduces to

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

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

Consider the following theorems.

Theorem 2.5.4: Transitivity of Linear Combinations

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

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

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

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

$$\vec{q}_{1} = d_{11}\vec{r}_{1} + d_{12}\vec{r}_{2} + \dots + d_{1\ell}\vec{r}_{\ell}$$

$$\vec{q}_{2} = d_{21}\vec{r}_{1} + d_{22}\vec{r}_{2} + \dots + d_{2\ell}\vec{r}_{\ell}$$

$$\vdots$$

$$\vec{q}_{k} = d_{k1}\vec{r}_{1} + d_{k2}\vec{r}_{2} + \dots + d_{k\ell}\vec{r}_{\ell}$$

Then,

$$\vec{x} = c_1(d_{11}\vec{r}_1 + d_{12}\vec{r}_2 + \dots + d_{1\ell}\vec{r}_{\ell}) + c_2(d_{21}\vec{r}_1 + d_{22}\vec{r}_2 + \dots + d_{2\ell}\vec{r}_{\ell})$$

$$\vdots$$

$$+ c_k(d_{k1}\vec{r}_1 + d_{k2}\vec{r}_2 + \dots + d_{k\ell}\vec{r}_{\ell})$$

$$= (c_1d_{11} + c_2d_{21} + \dots + c_kd_{k1})\vec{r}_1 + (c_1d_{12} + c_2d_{22} + \dots + c_kd_{k2})\vec{r}_2$$

$$\vdots$$

$$+ (c_1d_{1\ell} + c_2d_{1\ell} + \dots + c_kd_{k\ell})\vec{r}_{\ell}.$$

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: Definition **Definition 2.6.1:** Definition

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

$$T_A: \mathbb{R}^n \to \mathbb{R}^m$$

$$: \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \mapsto A \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}.$$

Consider the following example.

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

Consider the following maps, and name them.

1. If
$$A=\begin{bmatrix}0&0\\0&0\end{bmatrix}$$
, $T_A:\mathbb{R}^2 o \mathbb{R}^2$ is the zero map.

2. If
$$A=\begin{bmatrix}1&0\\0&1\end{bmatrix}$$
, $T_A:\mathbb{R}^2\to\mathbb{R}^2$ is the identity map.

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

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

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

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 $I_nI_n=I_nI_n=I_n$, 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. $B = BI_n = B(AC) = (BA)C = I_nC = C$.

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 $k \ge 2$, $A^{-k} = (A^{-1})^k$.

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 $A^kA^{-k}=A^{-k+k}=A^0=I_n$. 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,

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

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.

2.8 Lecture 13: September 21, 2022

2.8.1 Finding Inverses of Matrices

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

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

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

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

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

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

Consider the following example.

Example 2.8.1: * Finding a 2×2 Inverse

Consider the matrix

$$\begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$$

First we form

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

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

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

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

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

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

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

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

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

Therefore,

$$\begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix}$$

is the inverse of

$$\begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$$

Note that here, we did not apply the formula in Theorem 2.7.5, but instead, the general algorithm.

Example 2.8.2: * Finding a 3×3 Inverse

Consider the matrix

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

First, we form

$$\begin{bmatrix} 2 & -6 & 5 & | & 1 & 0 & 0 \\ -4 & 12 & -9 & | & 0 & 1 & 0 \\ 2 & -9 & 8 & | & 0 & 0 & 1 \end{bmatrix}.$$

After row reduction, we have

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

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, rank A < n, meaning we will have no solution or infinitely many solutions. This is because we will either have a row of all zeroes with a nonzero entry after the augmentation bar yielding an empty solution set, or if the system has at least one solution, there will be at least one free variable, guaranteeing infinitely many solutions.

2.9 Additional Topics: The LU Factorization

2.9.1 Presented by Olver and Shakiban's Applied Linear Algebra

Consider the following definition.

Definition 2.9.1: © **Elementary Matrices**

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

Consider the following examples.

Example 2.9.1: * Finding an Elementary Matrix

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

We take

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

and perform the row operation to obtain

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

Consider the following theorem.

Theorem 2.9.1: Elementary Matrices are Equivalent to Row Operations

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

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

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

3

Systems of Differential Equations

3.1 Lecture 13: September 21, 2022

3.1.1 Defining the Determinant

Consider the following theorems and definitions.

Theorem 3.1.1: \blacksquare 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 $\vec{x} = [x_1, x_2, x_3], \ \vec{y} = [y_1, y_2, y_3], \ \text{and} \ \vec{z} = [z_1, z_2, z_3].$ 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: \bigcirc **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: \blacksquare **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: \bigcirc **The** (i, j) **Cofactor**

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

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

Definition 3.1.4: The Determinant

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

1. If
$$n = 1$$
, and $A = [a_{11}]$, 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.

```
1 # Delete nth Column for Cofactor
def delete_nth_column_helper(matrix, n):
      return [m[:n]+m[n+1:] for m in matrix[1:]]
6 # Calculate Matrix Determinant
7 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]
11
12
    # Calculate Determinant Parts
     for _ in range(len(matrix[0])):
14
         determinant_components = [(-1)**i*matrix[0][i]*determinant(
16
      delete_nth_column_helper(matrix, i)) for i in range(len(matrix[0]))]
     # Return Sum
18
     return sum(determinant_components)
19
21 if __name__ == "__main__":
23
      matrix = []
24
      # Get Matrix
     for _ in range(int(input())): matrix.append([int(i) for i in input().split()])
26
27
    # 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, $A=\left[a_{11}\right]$, 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."

$$\det B = (-1)^{1+1}b_{11} \det \begin{bmatrix} b_{22} & b_{23} & \cdots & a_{2(k+1)} \\ 0 & b_{33} & \cdots & b_{3(k+1)} \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & b_{(k+1)(k+1)} \end{bmatrix}$$

$$= b_{11} \underbrace{b_{22} \cdots b_{(k+1)(k+1)}}_{\text{By the inductive hypothesis.}}.$$

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

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 $c\langle i\rangle \rightarrow \langle i\rangle$ 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 $c\langle i \rangle + \langle j \rangle \rightarrow \langle j \rangle$ 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			atrix	Effect on the Determinant
	$A_1 =$	1 0 4	1 1 9	1 -4 4	$\det A_1 = \det A_0$
$-4\langle 1\rangle + \langle 3\rangle \rightarrow \langle 3\rangle$	$A_2 =$	1 0 0	1 1 5	1 -4 0	$\det A_2 = \det A_1$
$-5\langle 2\rangle + \langle 3\rangle \rightarrow \langle 3\rangle$	$A_3 =$	1 0 0	1 1 0	1 -4 20	$\det A_3 = \det A_2$

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	
$\langle 2 \rangle \leftrightarrow \langle 1 \rangle$	$A_1 = \begin{bmatrix} 1 & 3 & 2 \\ 0 & -14 & -8 \\ -2 & 0 & 6 \end{bmatrix}$	$\det A_1 = -\det A_0$	
$2\langle 1\rangle + \langle 3\rangle \rightarrow \langle 3\rangle$	$A_2 = egin{bmatrix} 1 & 3 & 2 \ 0 & -14 & -8 \ 0 & 6 & 10 \end{bmatrix}$	$\det A_2 = \det A_1$	
$-\frac{1}{14}\langle 2\rangle \to \langle 2\rangle$	$A_3 = \begin{bmatrix} 1 & 3 & 2 \\ 0 & 1 & \frac{4}{7} \\ 0 & 6 & 10 \end{bmatrix}$	$\det A_3 = -\tfrac{1}{14} \det A_2$	
	$A_3 = \begin{bmatrix} 1 & 3 & 2 \\ 0 & 1 & \frac{4}{7} \\ 0 & 0 & \frac{46}{7} \end{bmatrix}$	$\det A_4 = \det A_3$	

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

$$det(AB) = det(R_1(...(R_k(I_n)...))B)$$

$$= c_1 \cdots c_k \det(I_nB)$$

$$= c_1 \cdots c_k \det B$$

$$= c_1 \cdots c_k \det I_n \det B$$

$$= det(R_1(...(R_k(I_n)...))) \det B$$

$$= det A \det B,$$

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

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

proving the proposition.

3. $\det(A^{-1}) = \frac{1}{\det A}$ Suppose that A 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 $X
eq \vec{0}$ such that

$$AX = \lambda X$$
.

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

3.4 Lecture 16: September 28, 2022

3.4.1 The Process of Diagonalization: Part I

Consider the following definition.

Definition 3.4.1: © **Eigenspace**

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

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

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

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

Theorem 3.4.1: Finding Eigenvectors and Eigenvalues

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

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

That is, when

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

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

$$\mathcal{E}_{\lambda_1} = \left\{ c egin{bmatrix} 1 \ 0 \ 0 \end{bmatrix} : c \in \mathbb{R}
ight\}$$

and

$$\mathcal{E}_{\lambda_2} = \left\{ c_1 egin{bmatrix} 0 \ 1 \ 0 \end{bmatrix} + c_2 egin{bmatrix} 0 \ 0 \ 1 \end{bmatrix} : c_1, c_2 \in \mathbb{R}
ight\}.$$

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

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

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

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

By row reduction, we have

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

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

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

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

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

Then, we have

$$P^{-1}AP = P^{-1}(AP)$$

$$= P^{-1} \begin{bmatrix} A\vec{v}_1 & A\vec{v}_2 \end{bmatrix}$$

$$= P^{-1} \begin{bmatrix} \lambda_1 \vec{v}_1 & \lambda_2 \vec{v}_2 \end{bmatrix}$$

$$= \begin{bmatrix} \vec{w}_1 \\ \vec{w}_2 \end{bmatrix} \begin{bmatrix} \lambda_1 \vec{v}_1 & \lambda_2 \vec{v}_2 \end{bmatrix}$$

$$= \begin{bmatrix} \lambda_1 \vec{w}_1 \cdot \vec{v}_1 & \lambda_2 \vec{w}_1 \cdot \vec{v}_2 \\ \lambda_1 \vec{w}_2 \cdot \vec{v}_1 & \lambda_2 \vec{w}_2 \cdot \vec{v}_2 \end{bmatrix}$$

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

$$= D.$$

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 *i*th 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

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

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

$$\begin{bmatrix} 25 & 5 & | & 0 \\ -25 & -5 & | & 0 \end{bmatrix},$$

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

$$\begin{bmatrix} 25 & 5 & | & 0 \\ 0 & 0 & | & 0 \end{bmatrix}.$$

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

$$\begin{bmatrix} 5 & 5 & | & 0 \\ -25 & -25 & | & 0 \end{bmatrix}$$

which can be row reduced to

$$\begin{bmatrix} -25 & -25 & | & 0 \\ 0 & 0 & | & 0 \end{bmatrix}.$$

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

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

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

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

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

$$\begin{bmatrix} -3 & -6 & 0 & | & 0 \\ 0 & 0 & 0 & | & 0 \\ 0 & 0 & 3 & | & 0 \end{bmatrix}$$

which can be row reduced to

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

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

$$\begin{bmatrix} -6 & -6 & 0 & | & 0 \\ 3 & 3 & 0 & | & 0 \\ 0 & 0 & 0 & | & 0 \end{bmatrix},$$

which can be row reduced to

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

giving

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

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: Definition 3.6.1: Definition 3.6.1:

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

$$\begin{bmatrix} \vec{v}_1 & \cdots & \vec{v}_n | \vec{0} \end{bmatrix}$$

for c_1, \ldots, 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 $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ 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 $\lambda_1, \dots, \lambda_n$ 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

$$\begin{bmatrix} 9 & 1 & -1 & | & 0 \\ -11 & -1 & 2 & | & 0 \\ 18 & 2 & -2 & | & 0 \end{bmatrix}$$

which reduces to

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

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

$${\sf rank}\, ({\it A}-\lambda_1{\it I}_3)=2\neq 3-2=1.$$

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,

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

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,

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

proving the theorem.

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: \blacksquare **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.
$$\Re(a + bi) := a$$
, $\Im(a + bi) := b$.

3.
$$z = a + bi \in \mathbb{C} \implies ||z|| := \sqrt{a^2 + b^2}$$
.

4.
$$(a + bi)(c + di) := ac + adi + bci + bdi^2 = (ac - bd) + (ad + bc)i$$
.

5.
$$z = a + bi \in \mathbb{C} \implies \overline{z} := a - bi$$
.

Theorem 3.7.1: Properties of Complex Numbers

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

1.
$$z\overline{z} = (a + bi)(a - bi) = a^2 + b^2 = ||z||^2$$
.

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

Theorem 3.7.2: The Fundamental Theorem of Algebra

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

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

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

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

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

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

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

$$\overrightarrow{cv} = [cv_1, ..., cv_n].$$

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

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

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

Definition 3.7.7: $\ \ \ \ \$ The Magnitude in $\ \ \ \ \ \ \$

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

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

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

Definition 3.7.8: The Adjoint

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

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

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

Theorem 3.7.3: The Transpose and The Dot Product

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

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

Theorem 3.7.4: The Adjoint and The Dot Product

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

$$(\overrightarrow{Av}) \cdot \overrightarrow{w} = \overrightarrow{v} \cdot (\overrightarrow{A} \cdot \overrightarrow{w})$$

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

Definition 3.7.9: ● Hermitian and Skew-Hermitian Matrices

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

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

3.8 Lecture 20: October 7, 2022

3.8.1 Diagonalization Using Complex Eigenvalues

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

Example 3.8.1: * Diagonalization With Complex Eigenvalues

Diagonalize the matrix

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

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

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

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

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

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

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

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.

Qualitative Theory of Differential Equations

4.1 Lecture 20: October 7, 2022

4.1.1 The Process of Abstraction

Consider the following definition.

Definition 4.1.1: • Vector Spaces

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

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

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

- 1. $\vec{u} + \vec{v} = \vec{v} + \vec{u}$.
- 2. $\vec{u} + (\vec{v} + \vec{w}) = (\vec{u} + \vec{v}) + \vec{w}$.
- 3. $\exists \vec{0} \in V, \forall \vec{v} \in V, \vec{v} + \vec{0} = \vec{v}$.
- 4. $\forall \vec{v} \in V, \exists ! (-\vec{v}) \in V, \vec{v} + (-\vec{v}) = \vec{0}.$
- 5. $c_1(\vec{u} + \vec{v}) = c_1\vec{u} + c_1\vec{v}$.
- 6. $(c_1 + c_2)\vec{u} = c_1\vec{u} + c_2\vec{u}$.
- 7. $(c_1c_2)\vec{u} = c_1(c_2\vec{u}).$
- 8. $1\vec{u} = \vec{u}$.

We remark that $0\vec{v} = 0$ 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 $\vec{c0} = \vec{0}$, is a vector space.
- 4. The set \mathcal{M}_{mn} is a vector space.
- 5. Let S be a nonempty set and $F(S) = \{f : S \to \mathbb{R}\}$ with (f+g)(s) = f(s) + g(s) and (cf)(s) = cf(s). The set F(S) is a vector space.
- 6. The set $\{a_nx^n + \cdots + a_0x^0 : a_0, \dots a_n \in \mathbb{R}, \text{ degree is less than or equal to } n\}$ 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 $\mathbb{R}^+ = \{x \in \mathbb{R} : x \ge 0\}$, with usual addition and usual multiplication in \mathbb{R} , is not a vector space. The set does not satisfy the closure axiom for scalar multiplication; $(-1) \cdot 1 = -1 \notin \mathbb{R}^+$.
- 2. The set $S = \{f : \mathbb{R} \to \mathbb{R} : f(0) = 1\}$, with (f+g)(x) = f(x) + g(x) and (cf)(x) = cf(x), is not a vector space. The set does not satisfy the closure axiom for vector addition; (f+g)(0) = f(0) + g(0) = 2, so $(f+g)(x) \notin S$.

4.2 Lecture 21: October 14, 2022

4.2.1 Derived Properties of Vector Spaces

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

Theorem 4.2.1: Derived Property of Vector Space 1

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

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

Proof. We have that

$$c\vec{0} = c\vec{0} + \vec{0}$$
 by axiom 3,

$$= c\vec{0} + (c\vec{0} + (-(c\vec{0})))$$
 by axiom 4,

$$= (c\vec{0} + c\vec{0}) + (-(c\vec{0}))$$
 by axiom 2,

$$= c(\vec{0} + \vec{0}) + (-(c\vec{0}))$$
 by axiom 5,

$$= c\vec{0} + (-(c\vec{0}))$$
 by axiom 3,

$$= \vec{0}$$
 by axiom 4,

as desired.

Theorem 4.2.2: ■ Derived Property of Vector Space 2

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

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

Proof. We have that

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

as desired.

Theorem 4.2.3: ■ Derived Property of Vector Space 3

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

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

Proof. We have that

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

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

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

Then,

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

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

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

which becomes

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

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

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,

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

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

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

with $c \neq 0$. Then,

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

by Theorem 4.2.1. But, by axiom 7,

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

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

4.3 Lecture 22: October 17, 2022

4.3.1 Subspaces

Consider the following definition.

Definition 4.3.1: Subspaces

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

Consider the following theorem.

Theorem 4.3.1: Showing a Vector Space is a Subspace

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

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

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

Consider the following examples.

Example 4.3.1: * Is it a Subspace? 1

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

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

Is W a subspace of V?.

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

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

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

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

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

Example 4.3.2: * Is it a Subspace? 2

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

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

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

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

$$(A+B)_{ij} = A_{ij} + B_{ij}$$

$$= \begin{cases} A_{ii} + B_{ii} & i = j \\ 0 & i \neq j \end{cases}$$

$$\in \mathcal{D}_n.$$

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

$$(cA)_{ij} = cA_{ij}$$

$$= \begin{cases} cA_{ii} & i = j \\ 0 & i \neq j \end{cases}$$

$$\in \mathcal{D}_{n}.$$

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

Example 4.3.3: * Is it a Subspace? 3

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

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

where 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 $A, B \in \mathcal{U}_n$,

$$(A+B)_{ij} = A_{ij} + B_{ij}$$

$$= \begin{cases} A_{ij} + B_{ij} & i \leq j \\ 0 & i > j \end{cases}$$

$$\in \mathcal{U}_{P}$$

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

$$(cA)_{ij} = cA_{ij}$$

$$= \begin{cases} cA_{ij} & i \leq j \\ 0 & i > j \end{cases}$$

$$\in \mathcal{U}_{n}.$$

Therefore, U_n is a subspace of \mathcal{M}_{nn} .

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

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

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

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

Consider the following theorem.

Theorem 4.3.2: © **Eigenspaces are Subspaces**

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

Proof. By definition,

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

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

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

We realize that

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

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

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

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

4.4 Lecture 23, October 19, 2022

4.4.1 Span

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

Definition 4.4.1: Finite Linear Combinations

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

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

for scalars c_1, \ldots, c_n .

Theorem 4.4.1: Subspaces are Closed Under Linear Combinations

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

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

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

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

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

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

as desired.

We now, introduce the notion of "span."

Definition 4.4.2: Span

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

Consider the following examples.

Example 4.4.1: * Find Span 1

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

We realize that

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

Example 4.4.2: * Find Span 2

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

We realize that

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

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

Example 4.4.3: ** Find Span 3

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

We realize that

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

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

Consider the following theorems.

Theorem 4.4.2: A Complete Characterization of the Span

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

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

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

2. span(S) is a subspace of V.

Proof. We know that span $(S) \neq \emptyset$, by definition. Suppose that $\vec{w}_1, \vec{w}_2 \in \text{span}(S)$. We know that

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

for some $c_1, \ldots, c_n \in \mathbb{F}$ and $\vec{v}_1, \ldots, \vec{v}_n \in S$. We also know that

$$\vec{w}_2 = d_1 \vec{u}_1 + \cdots + d_k \vec{u}_k.$$

for some $d_1, \ldots, d_k \in \mathbb{F}$ and $\vec{u}_1, \ldots, \vec{u}_k \in S$. Then,

$$\vec{w}_1 + \vec{w}_2 = (c_1 \vec{v}_1 + \dots + c_k \vec{v}_k) + (d_1 \vec{u}_1 + \dots + d_k \vec{u}_k).$$

The sum is a linear combination of the elements in S, therefore, $\vec{w}_1 + \vec{w}_2 \in \text{span}(S)$, meaning span (S) is closed under vector addition. Now, we take $c \in \mathbb{F}$ and $\vec{v} \in \text{span}(S)$ and wish to show that $c\vec{v} \in \text{span}(S)$. We know that

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

for some $c_1, \ldots, c_n \in \mathbb{F}$ and $\vec{v}_1, \ldots, \vec{v}_n \in S$. Then,

$$c\overrightarrow{v} = c(c_1\overrightarrow{v}_1) + \cdots + c(c_n\overrightarrow{v}_n),$$

which is a linear combination of the elements in S, meaning that it is in span (S).

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

Proof. We let $\vec{v} \in \text{span}(S)$ and wish to show that $\vec{v} \in W$. We know that

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

for some $c_1, \ldots, c_n \in \mathbb{F}$ and $\vec{v}_1, \ldots, \vec{v}_n \in S$. Since $S \subseteq W$, and \vec{v} is a linear combination of elements of S, \vec{v} is a linear combination of elements of W. It follows that $\vec{v} \in W$, by Theorem 4.4.1.

4. span(S) is the smallest subspace of V containing S.

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

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

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

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

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

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

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

Theorem 4.4.4: Span as an Intersection of Subspaces

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

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

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

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

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

as desired.

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

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

Consider the following theorem.

Theorem 4.4.5: Span and Row Space

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

Consider the following examples.

Example 4.4.4: * Computing Span as a Row Space 1

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

We have the matrix

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

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

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

Example 4.4.5: * Computing Span as a Row Space 2

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

We have the matrix

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

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

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

Theorem 4.4.6: Is a Vector in the Span?

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

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

4.5 Lecture 24: October 21, 2022

4.5.1 Linear Independence

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

Definition 4.5.1: Dependence and Dependence

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

1. S is linearly dependent if and only if there exist scalars $c_1, \ldots, c_n \in \mathbb{F}$ not all zero, such that

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

2. S is linearly independent if S is not linearly dependent. That is,

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

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

Definition 4.5.2: © Generalizing Linear Independence and Dependence to Infinite Sets

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

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

Theorem 4.5.1: Showing Linear Independence 1

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

Proof. If det $A \neq 0$, A is nonsingular and has only the trivial solution for $[A|\vec{0}]$. This linear system is precisely

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

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

Theorem 4.5.2: Showing Linear Independence 2

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

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

Equivalently, S is linearly dependent if and only if

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

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

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

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

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

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

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

Letting $c_1 = -1$, we have

$$-\vec{v}_1 + c_2\vec{v}_2 + \dots + c_n\vec{v}_n = -(c_2\vec{v}_2 + \dots + c_n\vec{v}_n) + (c_2\vec{v}_2 + \dots + c_n\vec{v}_n)$$

= $\vec{0}$.

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

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

Theorem 4.5.3: Dependence of Sets With One or Two Elements

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

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

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

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

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

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

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

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

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

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

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

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

meaning that S is linearly dependent.

Consider another characterization of linear independence.

Theorem 4.5.5: ■ Linear Independence of Nonempty Sets

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

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

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

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

Consider the following proofs involving linear independence.

Example 4.5.1: ** Proving a Property of Linear Independence 1

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

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

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

and therefore,

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

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

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

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

Example 4.5.2: * Proving a Property of Linear Independence 2

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

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

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

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

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

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

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

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

$$c_0 + c_1 x + c_2 x^2 + c_3 x^3 + \dots + c_k x^k + c_{k+1} x^{k+1} = 0.$$

After differentiating both sides, we have

$$c_1 + 2c_2x + 3c_3x^2 + \dots + kc_kx^{k-1} + (k+1)c_{k+1}x^k = 0.$$

By the inductive hypothesis, $c_1=2c_2=3c_3=\cdots=kc_k=(k+1)c_{k+1}=0$. Since 1, 2, 3, ..., k, k+1 are all not zero, this implies $c_1=\cdots=c_{k+1}=0$. if we substitute this result into

$$c_0 + c_1 x + c_2 x^2 + c_3 x^3 + \dots + c_k x^k + c_{k+1} x^{k+1} = 0$$

we have $c_0=0$, so $c_0=\cdots=c_{k+1}=0$, meaning

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

is linearly independent, as desired. Thus, by induction, for all $n \in \mathbb{N}$,

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

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

4.6 Lecture 25: October 24, 2022

4.6.1 Determining Linear Independence in \mathbb{R}^n

We now consider a method to test for linear independence using row reduction in \mathbb{R}^n , and present a related result on linear dependence in \mathbb{R}^n .

Theorem 4.6.1: \bullet A Test for Linear Independence in \mathbb{R}^n

Let S be a finite nonempty set of vectors in \mathbb{R}^n . To determine whether S is linearly independent, perform the following steps:

- 1. Create the matrix A whose columns are the vectors in S.
- 2. Solve the system $[A|\overline{0}]$ by row reduction.

Then, S is linearly independent if the system has only the trivial solution, and linearly dependent otherwise.

Theorem 4.6.2: \odot A Test for Linear Dependence in \mathbb{R}^n

Suppose $S \subseteq \mathbb{R}^n$ and S contains distinct elements $\vec{v}_1, ..., \vec{v}_k$ where k > n. Then, S is linearly dependent.

Consider the following examples.

Example 4.6.1: * Determining Linear Dependence in \mathbb{R}^n 1

Let $S = \{[1, -1, 0, 2], [0, -2, 1, 0], [2, 0, -1, 1]\} \subseteq \mathbb{R}^4$. We build the system

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

which, by row reduction, becomes

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

We see that the system only has the trivial solution, so S is linearly independent.

Example 4.6.2: * Determining Linear Dependence in \mathbb{R}^n 2

Let $S = \{[3,1,-1], [-5,-2,2], [2,2,-1]\} \subseteq \mathbb{R}^3.$ We build the system

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

which, by row reduction, becomes

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

We see that the system only has the trivial solution, so S is linearly independent.

Example 4.6.3: * Determining Linear Dependence in \mathbb{R}^n 3

Let $S = \{[2, 1], [-1, 3], [1, 4]\} \subseteq \mathbb{R}^2$. We build the system

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

which, by row reduction, becomes

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

We see that the system has infinitely many solutions, since the third column gives a free variable. Thus, S is linearly dependent. Note that the same conclusion is reached by realizing that k = |S| = 3, n = 2, and k > n.

4.6.2 Bases: Part I

We now define the basis of a vector space.

Definition 4.6.1: Basis

Suppose V is a vector space. Then, $B \subseteq V$ is a basis of V if and only if

- 1. span(B) = V.
- 2. B is linearly independent.

Consider the following examples.

Example 4.6.4: * The Standard Basis of \mathbb{R}^3

Verify that $B = \{[1, 0, 0], [0, 1, 0], [0, 0, 1]\}$ is a basis of \mathbb{R}^3 .

Let A be the matrix with the elements of B as its columns. We form the linear system $[A|\vec{0}]$, or

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

and realize that it is already in reduced row echelon form. The system has only the trivial solution, and so B is linearly independent. Now, we consider the span of B, or the row space of A^T . We have

$$\mathsf{span}\,(B)=\{[c_1,c_2,c_3]:c_1,c_2,c_3\in\mathbb{R}^3\}=\mathbb{R}^3.$$

Thus, B is a basis for \mathbb{R}^3 .

Example 4.6.5: $\ ^{\bullet}$ **Another Basis of** \mathbb{R}^3

Verify that $B = \{[2, 2, 2], [5, 0, 0], [0, -3, 1]\}$ is a basis of \mathbb{R}^3 .

Let A be the matrix with the elements of B as its columns. We form the linear system $[A|\vec{0}]$, or

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

By row reduction, we obtain

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

The system has only the trivial solution, and so B is linearly independent. Now, we consider the span of B, or the row space of A^T . We don't need to compute the row reduction in this case as A is square. If A can be row reduced to I_n , $\det A = \det A^T \neq 0$, meaning that A^T can be row reduced to I_n . This means

$$\mathrm{span}(B) = \{ [c_1, c_2, c_3] : c_1, c_2, c_3 \in \mathbb{R}^n \}.$$

Thus, B is a basis for \mathbb{R}^3 .

Example 4.6.6: * * A Basis of \mathcal{M}_{22}

Verify that

$$B = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$$

is a basis of \mathbb{R}^3 .

We start by computing span (B), to produce

$$\begin{split} \mathsf{span}\,(B) &= \left\{ c_1 \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + c_2 \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + c_3 \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + c_4 \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} : c_1, c_2, c_3, c_4 \in \mathbb{R} \right\} \\ &= \left\{ \begin{bmatrix} c_1 & c_2 \\ c_3 & c_4 \end{bmatrix} : c_1, c_2, c_3, c_4 \in \mathbb{R} \right\} \\ &= \mathcal{M}_{22}. \end{split}$$

Then, to verify linear independence, we see that if

$$c_1\begin{bmatrix}1&0\\0&0\end{bmatrix}+c_2\begin{bmatrix}0&1\\0&0\end{bmatrix}+c_3\begin{bmatrix}0&0\\1&0\end{bmatrix}+c_4\begin{bmatrix}0&0\\0&1\end{bmatrix}=\begin{bmatrix}c_1&c_2\\c_3&c_4\end{bmatrix}=\begin{bmatrix}0&0\\0&0\end{bmatrix},$$

 $c_1 = c_2 = c_3 = c_4 = 0$, meaning B is linearly independent. Thus, B is a basis for \mathcal{M}_{22} .

Example 4.6.7: ** A Basis of \mathcal{P}_n

Verify that $B = \{x^0, ..., x^n\}$ is a basis of $\overline{\mathcal{P}}_n$, the set of polynomials with degree at most n.

We start by computing span (B), to produce

$$span(B) = \{c_0x_0 + \dots + c_nx^n : c_0, \dots, c_n \in \mathbb{R}^n\}$$
$$= \mathcal{P}_n.$$

Then, to verify linear independence, see Example 4.5.2. Thus, B is a basis for \mathcal{P}_n .

Example 4.6.8: * * A Basis of $\{\vec{0}\}$

Verify that $B = \emptyset$ is a basis of $\{\vec{0}\}$.

We note that span $(B) = \vec{0}$, by Definition 4.4.2. Then, we recall that B is linearly independent, by Definition 4.5.1. Thus, B is a basis for $\{\vec{0}\}$.

Example 4.6.9: * * A Basis of \mathcal{D}_3

Find a basis of \mathcal{D}_3 , the set of 3×3 diagonal matrices.

Consider the set

$$B = \left\{ egin{bmatrix} 1 & 0 & 0 \ 0 & 0 & 0 \ 0 & 0 & 0 \end{bmatrix}, egin{bmatrix} 0 & 0 & 0 \ 0 & 1 & 0 \ 0 & 0 & 0 \end{bmatrix}, egin{bmatrix} 0 & 0 & 0 \ 0 & 0 & 0 \ 0 & 0 & 1 \end{bmatrix}
ight\}.$$

To see that span $(B) = \mathcal{D}_3$, consider

The row space of A_1 is $\{[c_1, 0, 0, 0, c_2, 0, 0, 0, c_3] : c_1, c_2, c_3 \in \mathbb{R}\}$. Thus, span $(B) = \mathcal{D}_3$. Then, to verify that B is linearly independent, consider

which is easily row reduced to a matrix with pivots in each column, so B is linearly independent and is a basis for \mathcal{D}_3 .

4.7 Lecture 26: October 26, 2022

4.7.1 Bases: Part II

Consider the following nonexample of a basis.

Example 4.7.1: * Not a Basis of \mathcal{P}_2

Verify that $B = \{1, x^2\}$ is not a basis for \mathcal{P}_2 .

We see that $x \in \mathcal{P}_2$ but $x \notin \text{span}(B)$, so B is not a basis for \mathcal{P}_2 .

Proof. Suppose $x \in \text{span}(B)$. Then,

$$x = c_1(1) + c_2(x^2)$$

for some c_1 , $c_2 \in \mathbb{R}$. If x = 0, we have

$$0 = c_1 + c_2(0)^2 \implies c_1 = 0.$$

Therefore, we have $x = c_2 x^2$. After we take the derivative of both sides, we get

$$1 = 2c_2x$$
,

and when x = 0, that implies 1 = 0, and we arrive at a contradiction. Hence, $x \notin \text{span}(B)$.

Consider the following theorems.

Theorem 4.7.1: A Useful Lemma for Bases

Let S and T be subsets of a vector space V such that span (S) = V where S is finite and T is linearly independent. Then, T is finite and $|T| \le |S|$.

Proof. If $S=\emptyset$, then we have that $V=\{\vec{0}\}$. Since $\{\vec{0}\}$ is linearly independent, $T=\emptyset$. Suppose $|S|=n\geq 1$. Suppose, for the sake of contradition, that T is infinite or |T|>|S|=n. Since every finite subset of T is linearly independent, there is a linearly independent set $Y\subseteq T$ such that |Y|=n+1. Let $S=\{\vec{v}_1,\ldots,\vec{v}_n\}$ and let $Y=\{\vec{w}_1,\ldots,\vec{w}_n,\ldots,\vec{w}_{n+1}\}$. We will show that Y is linearly dependent. Then, we have $c_{ij}\in\mathbb{F}$ for $1\leq i\leq n+1$ and $1\leq j\leq n$ such that

$$\vec{w}_1 = c_{11}\vec{v}_1 + \dots + c_{1n}\vec{v}_n, \dots, \vec{w}_{n+1} = c_{n+1,1}\vec{v}_1 + \dots + c_{n+1,n}\vec{v}_n.$$

Let C be the $(n+1) \times n$ matrix whose (i,j) entry is c_{ij} . Then, $[C^T|\vec{0}]$ has n+1 variables with only n equations, and has a nontrivial solution $\vec{u} = [u_1, \dots, u_{n+1}]$. Then,

$$u_1\vec{w}_1 + \dots + u_{n+1}\vec{w}_{n+1} = u_1(c_{11}\vec{v}_1 + \dots + c_{1n}\vec{v}_n) + \dots + u_{n+1}(c_{n+1,1}\vec{v}_1 + \dots + c_{n+1,n}\vec{v}_n)$$

$$= (c_{11}u_1 + \dots + c_{n+1,1}u_{n+1})\vec{v}_1 + \dots + (c_{n1}u_1 + \dots + c_{n+1,n}u_{n+1})\vec{v}_n.$$

But, the coefficient of each \vec{v}_i in the last expression is the *i*th entry of $C^T\vec{u}$. Since $C^T\vec{u}=\vec{0}$, the coefficient of each \vec{v}_i is 0. Then, for

$$u_1\vec{w}_1 + \cdots + u_{n+1}\vec{w}_{n+1} = \vec{0}$$
,

we have $u_1 = \cdots = u_{n+1} = 0$. But, since \vec{u} is a nontrivial solution, at least one $u_i \neq 0$ and Y is linearly dependent, providing a contradiction.

Theorem 4.7.2: Bases Have Equivalent Cardinality

Suppose V is a vector space and B_1 is a basis for V with finitely many elements. If B_2 is a basis for V with finitely many elements,

$$|B_1|=|B_2|.$$

Proof. Since span $(B_1) = V$ and B_2 is finite and linearly independent, $|B_2| \le |B_1|$. Since span $(B_2) = V$ and B_1 is finite and linearly independent, $|B_1| \le |B_2|$. Thus, $|B_1| = |B_2|$.

Now, we come to the definition of the dimension of a vector space.

Definition 4.7.1: Dimension

If V is a vector space with a finite basis B,

$$\dim V = |B|$$
.

If $V = \{\vec{0}\}\$, dim V = 0. Otherwise, dim V is infinite.

Consider the following theorems.

Theorem 4.7.3: General Statements About Span and Dimension 1

Suppose V is a vector space. If $S \subseteq V$ is finite and span (S) = V, then, dim $V \leq |S|$.

Proof. See Theorem 4.7.1 and Definition 4.7.1.

Theorem 4.7.4: General Statements About Span and Dimension 2

Suppose V is a vector space. If $S \subseteq V$ is finite and span (S) = V, then, S is a basis if and only if $\dim V = |S|$.

Proof. Suppose S is a basis. Then, by Definition 4.7.1, dim V = |S|. If dim V = |S|, let $n = |S| = \dim V$ so $S = \{\vec{v}_1, \dots, \vec{v}_n\}$. Since span (S) = V, we need only show that S is linearly independent. Suppose S is linearly dependent, then, there exists \vec{v}_i such that

$$\mathrm{span}\left(\{\overrightarrow{v}_1,\ldots,\overrightarrow{v}_{i-1},\overrightarrow{v}_{i+1},\overrightarrow{v}_n\}\right)=V.$$

This contradicts dim V = n.

Theorem 4.7.5: General Statements About Linear Independence and Dimension 1

Suppose V is a vector space. If $T \subseteq V$ is linearly independent, then, $|T| \leq \dim V$.

Proof. Let S be a basis of V. Thus, span (S) = V and dim V = |S|. By Theorem 4.7.1, if T is linearly independent, $|T| \le |S|$, so $|T| \le \dim V$.

Theorem 4.7.6: General Statements About Linear Independence and Dimension 2

Suppose V is a vector space. If $T \subseteq V$ is linearly independent and finite, T is a basis for V if and only if dim V = |T|.

Proof. Suppose T is a basis. Then, by Definition 4.7.1, dim V = |T|. If dim V = |T|, Let $n = |T| = \dim V$ so $T = \{\vec{v}_1, \dots, \vec{v}_n\}$. Since T is linearly independent, we need only show that span (T) = V. Suppose not, then, there exists some $\vec{v} \in V$ where $\vec{v} \notin \operatorname{span}(T)$. Consider the set $T \cup \{\vec{v}\}$. We see that

$$|T \cup \{\vec{v}\}| = |T| + 1 = \dim V + 1$$

since $\vec{v} \notin \text{span}(T)$, and because $T \subseteq \text{span}(T)$, $\vec{v} \notin T$. We also see that $T \cup \{\vec{v}\}$ is linearly independent by considering

$$\vec{0} = c_1 \vec{v}_1 + \cdots + c_n \vec{v}_n + c_{n+1} \vec{v}$$

for $c_1, \ldots, c_{n+1} \in \mathbb{F}$ and $\vec{v}_1, \ldots, \vec{v}_k \in \mathcal{T}$. We have that $c_{n+1} = 0$ since, otherwise,

$$\vec{v} = -\frac{1}{c_{n+1}}(c_1\vec{v}_1 + \cdots + c_n\vec{v}_n).$$

Since we've written \vec{v} as a linear combination of the elements of T, $\vec{v} \in \text{span}(T)$, which is contradictory. We also have

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

since T is linearly independent and $\vec{v}_1, ..., \vec{v}_n \in T$. Thus, $c_1 = \cdots = c_n = c_{n+1} = 0$. By the previous part, we have $|T \cup \{\vec{v}\}| \leq \dim V$. However, the left hand side is dim V + 1, which is contradictory. Thus, span (T) = V, so T is a basis for V.

4.8 Lecture 27: October 28, 2022

4.8.1 Bases: Part III

Consider the following theorems.

Theorem 4.8.1: Dimensions of Subspaces

If W is a subspace of a finite dimensional vector space V,

 $\dim W < \dim V$.

Moreover, if dim $W = \dim V$, then, W = V.

Proof. Let dim V=n. If some $S\subseteq W$ is linearly independent in W, S is linearly independent in V, and therefore, $|S|\leq n$. Let $m\in \mathbb{N}$ be the maximum number of elements a linearly independent set in W can have. Let $B_W=\{\vec{w}_1,\dots,\vec{w}_m\}$ be a linearly independent subset of W. Then, for the sake of contradiction, suppose that span $(B_W)\neq W$, meaning that there exists some $\vec{w}\in W$ such that $\vec{w}\notin \text{span}\,(B_W)$. Then, $B_W\cup\{\vec{w}\}$ is linearly independent in W, but $|B_W\cup\{\vec{w}\}|=m+1$, contradicting our definition of m. Therefore, B_W is a basis for W, and since $|B_W|=m\leq n=\dim V$, dim $W\leq \dim V$. For the next part of the theorem, suppose dim $W=\dim V=n$. Then, let $B_W=\{\vec{w}_1,\dots,\vec{w}_n\}$ and $B_V=\{\vec{v}_1,\dots,\vec{v}_n\}$ be bases of W and V, respectively. Suppose, for the sake of contradiction, $W\neq V$. That is, suppose that span $(B_W)=W\neq V$. Then, there exists some $\vec{v}\in V$ with $\vec{v}\notin \text{span}\,(B_W)$. Then, $B_W\cup\{\vec{v}\}$ is linearly independent in V, and by construction, span $(B_W\cup\{\vec{v}\})=V$. But, $|B_W\cup\{\vec{v}\}|=n+1$, and an n-dimensional space cannot have a basis with n+1 elements, providing our contradiction. Thus, W=V, as desired.

Theorem 4.8.2: Diagonalizability, Revisited: Part I

Suppose $A \in \mathcal{M}_{nn}$. Then, A is diagonalizable if and only if there exists a basis of \mathbb{R}^n that consists of eigenvectors of A.

Proof. Recall that by Theorem 3.6.1, $A \in \mathcal{M}_{nn}$ is diagonalizable if and only if there exists a set of n linearly indpendent eigenvectors $\vec{v}_1, \dots, \vec{v}_n$. Let $V = \mathbb{R}^n$ and $W = \{\vec{v}_1, \dots, \vec{v}_n\}$. Then, dim V = n. We also see that $W \subseteq V$ is linearly independent and finite with $|W| = \dim V$. By Theorem 4.7.6, W is a basis for V. Our proof is reversible, as all implications are bidirectional.

We now posit two questions. Given $S \subseteq V$, where V is a vector space, how would one construct a basis for span (S)? Similarly, given a linearly independent set S where $S \subseteq V$ and V is a vector space, how can we expand S to form a basis of V? Consider the following theorems.

Theorem 4.8.3: \odot Finding a Basis for span (S) by Contraction

If span (S) = V, where V is a finite dimensional vector space, there exists some $B \subseteq S$ where B is a basis for V. Given $S = \{\vec{v}_1, \dots, \vec{v}_k\} \subseteq \mathbb{R}^n$. Consider the following steps.

- 1. We form the matrix A with the elements of S as columns.
- 2. We row reduce A into reduced row echelon form to obtain a matrix C.
- 3. The basis of span (S), B, is formed by removing the vectors in A associated with the columns containing free variables.

Note that the vectors in B are in S.

Note that we can also construct a basis for span (S) by forming a matrix with the elements of S as rows and finding the row space. In general though, the basis will not contain vectors in S.

Theorem 4.8.4: Finding a Basis by Expansion

Let T be a linearly independent subset of a finite dimensional vector space V. Then, V has a basis B with $T \subseteq B$. If $T = \{\vec{t}_1, \dots, \vec{t}_k\} \subseteq V$, perform the following steps to find B:

- 1. Find a spanning set $A = \{\vec{a}_1, \dots, \vec{a}_n\}$. For \mathbb{R}^n , this will often be $\{\vec{e}_1, \dots, \vec{e}_n\}$.
- 2. Form the ordered spanning set $S = \{\vec{t}_1, ..., \vec{t}_k, \vec{a}_1, ..., \vec{a}_n\}$ for V.
- 3. Use Theorem 4.8.3 to find a basis B for V, containing T.

Consider the following examples.

Example 4.8.1: * Finding a Basis 1

Let $S = \{[1, 2, 0, 1], [3, 1, 5, -7], [-2, 4, -8, 14]\}$. We form the matrix A with the elements of S as columns, giving

$$A = \begin{bmatrix} 1 & 3 & -2 \\ 2 & 1 & 4 \\ 0 & 5 & -8 \\ 1 & -7 & 14 \end{bmatrix}.$$

By row reduction, we get

$$C = \begin{bmatrix} 1 & 0 & \frac{14}{5} \\ 0 & 1 & -\frac{8}{5} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Now, we remove the vectors associated with the columns containing free variables to get the basis. Our basis is then

$$B = \{[1, 2, 0, 1], [3, 1, 5, -7]\}.$$

Example 4.8.2: * Finding a Basis 2

Let $S = \{x^3 - 3x^2 + 2, 2x^3 - 7x^2 + x - 3, 4x^3 - 13x^2 + x + 5\}$. We instead equivalently consider $S' = \{[2, 0, -3, 1], [-3, 1, -7, 2], [5, 1, -13, 4]\}$. We form the matrix A with the elements of S' as columns, giving

$$A = \begin{bmatrix} 1 & -3 & 5 \\ 0 & 1 & 1 \\ -3 & -7 & -13 \\ 1 & 2 & 4 \end{bmatrix}.$$

By row reduction, we get

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

There are no vectors to remove and we have B = S.

Example 4.8.3: * Finding a Basis 3

Let $S = \{x^3 - 8x^2 + 1, 3x^3 - 2x^2 + x, 4x^3 + 2x - 10, x^3 - 20x^2 - x + 12, x^3 + 24x^2 + 2x - 13\}$. We instead equivalently consider $S' = \{[1, 0, -8, 1], [0, 1, -2, 3], [-10, 2, 0, 4], [12, -1, -20, 1], [-13, 2, 24, 1]\}$. We form the matrix A with the elements of S' as columns, giving

$$A = \begin{bmatrix} 1 & 0 & -10 & 12 & -13 \\ 0 & 1 & 2 & -1 & 2 \\ -8 & -2 & 0 & -20 & 24 \\ 1 & 3 & 4 & 1 & 1 \end{bmatrix}.$$

By row reduction, we get

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

We remove the last vector to get $B = \{x^3 - 8x^2 + 1, 3x^3 - 2x^2 + x, 4x^3 + 2x - 10, x^3 - 20x^2 - x + 12\}.$

Example 4.8.4: * Finding a Basis 4

Let $T = \{[1, 2, 0, 1], [3, 1, 5, -7]\}$. We consider

$$S = \{[1, 2, 0, 1], [3, 1, 5, -7], [1, 0, 0, 0], [0, 1, 0, 0], [0, 0, 1, 0], [0, 0, 0, 1]\}$$

Note that S is the union of the standard basis of \mathbb{R}^4 and T. Since

$$\{[1, 0, 0, 0], [0, 1, 0, 0], [0, 0, 1, 0], [0, 0, 0, 1]\}$$

is a basis of \mathbb{R}^4 , span $(S) = \mathbb{R}^4$. Now, we form

$$A = \begin{bmatrix} 1 & 3 & 1 & 0 & 0 & 0 \\ 2 & 3 & 0 & 1 & 0 & 0 \\ 0 & 5 & 0 & 0 & 1 & 0 \\ 1 & -7 & 0 & 0 & 0 & 1 \end{bmatrix}$$

and row reduce to

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

Hence, our basis for \mathbb{R}^4 is

$$B = \{[1, 2, 0, 1], [3, 1, 5, -7], [1, 0, 0, 0], [0, 1, 0, 0]\}.$$

Example 4.8.5: * Finding a Basis 5

Let $T = \{[2, 0, 4, -12], [0, -1, -3, 9]\}$. We consider

$$S = \{[2, 0, 4, -12], [0, -1, -3, 9], [1, 0, 0, 0], [0, 1, 0, 0], [0, 0, 1, 0], [0, 0, 0, 1]\}$$

Note that S is the union of the standard basis of \mathbb{R}^4 and T. Since

$$\{[1, 0, 0, 0], [0, 1, 0, 0], [0, 0, 1, 0], [0, 0, 0, 1]\}$$

is a basis of \mathbb{R}^4 , span (T) = \mathbb{R}^4 . Now, we form

$$A = \begin{bmatrix} 2 & 0 & 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 1 & 0 & 0 \\ 4 & -3 & 0 & 0 & 1 & 0 \\ -12 & 9 & 0 & 0 & 0 & 1 \end{bmatrix}$$

and row reduce to

$$C = \begin{bmatrix} 1 & 0 & 0 & -\frac{3}{4} & 0 & -\frac{1}{12} \\ 0 & 1 & 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & \frac{3}{2} & 0 & \frac{1}{6} \\ 0 & 0 & 0 & 0 & 1 & \frac{1}{3} \end{bmatrix}.$$

Hence, our basis for \mathbb{R}^4 is

$$B = \{[2, 0, 4, -12], [0, -1, -3, 9], [1, 0, 0, 0], [0, 0, 1, 0]\}.$$

We end this section with two examples on finding bases of "weird" sets.

Example 4.8.6: * Finding a Weird Basis 1

Find the basis and the dimension of $V = \{[x, y, z, w] : x, y, z, w \in \mathbb{R}, x - y - w + z = 0\}$.

The membership condition of the set indicates that w = x - y + z. We then have

$$V = \{ [x, y, z, x - y + z] : x, y, z \in \mathbb{R} \}.$$

We hypothesize that the dimension of V will be 3, because there are only three variables making up the set; however, we will confirm this after we have found the basis. We form, somewhat arbitrarily, the set

$$S = \{[1, 1, 1, 1], [1, 1, 2, 2], [1, 2, 1, 0]\}.$$

To find span (S), we consider

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 2 & 2 \\ 1 & 2 & 1 & 0 \end{bmatrix} \xrightarrow[\mathsf{RREF}]{} \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 1 \end{bmatrix}.$$

Thus,

$$\mathsf{span}\left(S\right) = \{[c_1, c_2, c_3, c_1 - c_2 + c_3] : c_1, c_2, c_3 \in \mathbb{R}\} = V.$$

Since span (S) = V, we check if S is linearly independent by considering

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 2 \\ 1 & 2 & 1 \\ 1 & 2 & 0 \end{bmatrix} \xrightarrow[\mathsf{RREF}]{} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

All columns have pivots, so S is linearly independent, spans V, and is therefore a basis for V. Because there are 3 elements in S, dim V=3.

Example 4.8.7: * Finding a Weird Basis 2

Find the basis and the dimension of

$$V = \{a_0 + a_1x + a_2x^2 + a_3x^3 : a_0, a_1, a_2, a_3 \in \mathbb{R}, 2a_0 + a_1 = 0, a_2 - 2a_3 = 0\}$$

The membership condition of the set indicates that $a_1 = -2a_0$ and $a_3 = \frac{1}{2}a_2$. We then have

$$V = \{a_0 - 2a_0x + a_2x^2 + \frac{1}{2}a_2x^3 : a_0, a_2 \in \mathbb{R}\}.$$

We hypothesize that the dimension of V will be 2, because there are only two variables making up the set; however, we will confirm this after we have found the basis. We form, somewhat arbitrarily, the set

$$S = \{1 - 2x + 2x^2 + x^3, 1 - 2x + 4x^2 + 2x^3\}.$$

Consider $S' = \{[1, -2, 2, 1], [1, -2, 4, 2]\}$. To find span (S'), we consider

$$\begin{bmatrix} 1 & -2 & 2 & 1 \\ 1 & -2 & 4 & 2 \end{bmatrix} \xrightarrow[\text{PDEE}]{} \begin{bmatrix} 1 & -2 & 0 & 0 \\ 0 & 0 & 1 & \frac{1}{2} \end{bmatrix}$$

so span $(S') = \{ [c_1, -2c_1, c_2, \frac{1}{2}c_2] : c_1, c_2 \in \mathbb{R} \}$ and

span
$$(S) = \left\{ c_1 - 2c_1x + c_2x^2 + \frac{1}{2}c_2x^3 \right\} = V.$$

Since span (S) = V, we check if S is linearly independent by considering

$$\begin{bmatrix} 1 & 1 \\ -2 & -2 \\ 2 & 4 \\ 1 & 2 \end{bmatrix} \xrightarrow{\mathsf{RREF}} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

All columns have pivots, so S is linearly independent, spans V, and is therefore a basis for V. Since there are 2 elements in S, dim V=2.

4.9 Lecture 28: October 31, 2022

4.9.1 Coordinatization: Part I

Consider the following definition.

Definition 4.9.1: © Ordered Bases

An ordered basis of an n dimensional vector space V is an n-tuple of vectors $(\vec{v}_1, \dots, \vec{v}_n)$ such that the set $\{\vec{v}_1, \dots, \vec{v}_n\}$ is a basis for V.

Consider the following example.

Example 4.9.1: * Ordered Bases of \mathbb{R}^3

The tuples

$$B_1 = ([1, 0, 0], [0, 1, 0], [0, 0, 1])$$

and

$$B_2 = ([0, 1, 0], [1, 0, 0], [0, 0, 1])$$

are two distinct ordered bases of \mathbb{R}^3 .

The concept of ordering allows us to define coordinates with respect to a basis.

Definition 4.9.2: © Coordinates With Respect to a Basis

Let $B = (\vec{v}_1, ..., \vec{v}_n)$ be an ordered basis for a vector space V. Suppose that

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

for $c_1, \ldots, c_n \in \mathbb{F}$. Then,

$$[\overrightarrow{w}]_B = [c_1, \ldots, c_n].$$

The quantity $[\vec{w}]_B$ is the coordinatization of \vec{w} with respect to B, or equivalently, " \vec{w} expressed in B coordinates."

Consider the following examples.

Example 4.9.2: * Coordinatization 1

Let $\vec{w} = [1, 2, 3] \in \mathbb{R}^3$. Then,

$$[\vec{w}]_{B_1=([1,0,0],[0,1,0],[0,0,1])} = [1, 2, 3].$$

Example 4.9.3: * Coordinatization 2

Let $\vec{w} = [1, 2, 3] \in \mathbb{R}^3$. Then,

$$[\vec{w}]_{B_2=([0,1,0],[1,0,0],[0,0,1])} = [2,1,3].$$

Example 4.9.4: * Coordinatization 3

Find the matrix associated to

$$[-2, 6, -1, -11]_B$$

where

$$B = \left(\begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \right).$$

We have the matrix

$$-2\begin{bmatrix}1 & -1\\0 & 0\end{bmatrix}+6\begin{bmatrix}1 & 0\\1 & 0\end{bmatrix}-1\begin{bmatrix}0 & 0\\0 & 1\end{bmatrix}-11\begin{bmatrix}0 & 1\\1 & 0\end{bmatrix}=\begin{bmatrix}4 & -9\\-5 & -1\end{bmatrix}.$$

We now present the general algorithm for find the coordinatization of a vector with respect to an ordered basis.

Theorem 4.9.1: Coordinatization

Let V be a nontrivial subspace of \mathbb{R}^n and let $B = (\vec{v}_1, ..., \vec{v}_k)$ be an ordered basis for V, and let $\vec{v} \in \mathbb{R}^n$. The following steps will find $[\vec{v}]_B$, if it exists:

- 1. Form the system $[A|\vec{v}]$ by using the vectors in B as the columns for A, in order. Use \vec{v} as a column as well.
- 2. Row reduce to obtain the system $[C|\vec{w}]$ in reduced row echelon form.
- 3. If the system has no solutions, $\vec{v} \notin \text{span}(B)$ and coordinatization is not possible. We note that B is not a basis if this happens.
- 4. Otherwise, $\vec{v} \in \text{span}(B) = V$. Eliminate all zeroes consisting solely of zeroes in $[C|\vec{w}]$ to obtain $[I_k|\vec{y}]$. Then, $[\vec{v}]_B = \vec{y}$.

Theorem 4.9.2: Properties of Coordinatization

Let $B = (\vec{v}_1, ..., \vec{v}_n)$ be an ordered basis of V and $\vec{w}_1, ..., \vec{w}_k \in V$ and $c_1, ..., c_k \in \mathbb{F}$. Then,

- 1. $[\vec{w}_1 + \vec{w}_2]_B = [\vec{w}_1]_B + [\vec{w}_2]_B$.
- 2. $[c_1\vec{w}_1]_B = c_1[\vec{w}_1]_B$.
- 3. $[c_1\vec{w}_1 + \dots + c_k\vec{w}_k]_B = c_1[\vec{w}_1]_B + \dots + c_k[\vec{w}_k]_B$.

Example 4.9.5: * * Coordinatization 4

Let B be an ordered basis of a subspace of \mathbb{R}^n where

$$B = ([2, 6, -4, 2], [-2, -4, -8, 1], [0, 1, -6, -2]).$$

Let $\vec{v} = [-2, -4, -8, -2]$. Find $[\vec{v}]_B$.

To find $[\vec{v}]_B$, we form

$$\begin{bmatrix} 2 & -2 & 0 & | & -2 \\ 6 & -4 & 1 & | & -4 \\ -4 & -8 & -6 & | & -8 \\ 2 & 1 & -2 & | & -2 \end{bmatrix}$$

and row reduce to get

$$\begin{bmatrix} 1 & 0 & 0 & | & -\frac{3}{7} \\ 0 & 1 & 0 & | & \frac{4}{7} \\ 0 & 0 & 1 & | & \frac{6}{7} \\ 0 & 0 & 0 & | & 0 \end{bmatrix}.$$

Thus, $[\vec{v}]_B = [-\frac{3}{7}, \frac{4}{7}, \frac{6}{7}].$

Example 4.9.6: * * Coordinatization 5

Let B be an ordered basis of a subspace of \mathbb{R}^n where

$$B = ([1, 0, 16, 0], [0, 1, -6, 0], [0, 0, 0, 1]).$$

Let $\vec{v} = [-2, -4, -8, -2]$. Find $[\vec{v}]_B$.

To find $[\vec{v}]_B$, we form

$$\begin{bmatrix} 1 & 0 & 0 & | & -2 \\ 0 & 1 & 0 & | & -4 \\ 16 & -6 & 0 & | & -8 \\ 0 & 0 & 1 & | & -2 \end{bmatrix}$$

and row reduce to get

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

Thus, $[\vec{v}]_B = [-2, -4, -2].$

4.10 Lecture 29: November 2, 2022

4.10.1 Coordinatization: Part II

We now posit a question: Given two ordered bases B_1 and B_2 , how are $[\vec{v}]_{B_1}$ and $[\vec{v}]_{B_2}$ related? Consider the following definition and theorem.

Definition 4.10.1: Transition Matrices

Suppose that V is a nontrivial vector space with dim V=n. Let B_1 and B_2 be ordered bases of V. Let P be the $n \times n$ matrix whose ith column, for $1 \le i \le n$, equals $[\vec{b}_{1i}]_{B_2}$, where \vec{b}_{1i} is the ith basis vector in B_1 . Then, P is called the transition matrix from B_1 to B_2 .

Theorem 4.10.1: Finding a Transition Matrix

Given $B_1 = (\vec{v}_1, ..., \vec{v}_n)$ and $B_2 = (\vec{w}_1, ..., \vec{w}_n)$. To go from B_1 to B_2 , we form

$$\begin{bmatrix} \vec{w}_1 & \cdots & \vec{w}_n & | & \vec{v}_1 & \cdots & \vec{v}_n \end{bmatrix}$$

and row reduce to get $[I_n|P]$. The matrix P is the transition matrix from the coordinatization with respect to B_1 to the coordinatization with respect to B_2 . Then,

$$P[\overrightarrow{v}]_{B_1} = [\overrightarrow{v}]_{B_2}.$$

Consider the following example.

Example 4.10.1: * Finding a Transition Matrix

Let

 $B_1 = ([1, 1, 0, 1], [1, 0, 0, 1], [0, 1, 0, 1], [1, 1, 1, 0]),$ $B_2 = ([1, 0, 0, 0], [1, 1, 0, 0], [1, 1, 0, 1], [1, 1, 1, 1]).$ Find the associated transition matrix P.

We form

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

and row reduce to get

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

Thus,

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

Example 4.10.2: * Checking a Transition Matrix

Check that the P found in Example 4.10.1 yields the correct result for

$$\vec{v}_1 = [1, 1, 0, 1].$$

We see that $[\vec{v}_1]_{B_1} = [1, 0, 0, 0]$. Then, we have

$$P[\vec{v}_1]_{B_1} = \begin{bmatrix} 0 & 1 & -1 & 0 \\ 0 & -1 & 0 & 1 \\ 1 & 1 & 1 & -1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$
$$= \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$$
$$= [\vec{v}_1]_{B_1}.$$

as desired.

Consider the following theorem.

Theorem 4.10.2: Properties of Transition Matrices

Suppose B_1 , B_2 , and B_3 are ordered bases of \mathbb{R}^n . Then,

1. The transition matrix from B_1 to B_2 is unique.

Proof. Suppose we have $P[\vec{v}]_{B_1} = Q[\vec{v}]_{B_1} = [\vec{v}]_{B_2}$. Then, we have $P[\vec{v}_i]_{B_1} = Q[\vec{v}_i]_{B_1}$ for each $\vec{v}_i \in B_1$. Also, note that for each $\vec{v}_i \in B_1$, $[\vec{v}_i]_{B_1} = \vec{e}_i$, so we have

$$P\vec{e}_i = Q\vec{e}_i$$
.

For any matrix A, we have that $A\vec{e}_i$ is the ith column of A. Therefore, each column of P is equal to the respective corresponding column of Q, so P = Q, as desired.

2. If P is the transition matrix from B_1 to B_2 and Q is the transition matrix from B_2 to B_3 , QP is the transition matrix from B_1 to B_3 .

Proof. We consider

$$QP([\vec{v}_{B_1}]) = Q(P[\vec{v}]_{B_1})$$

$$= Q[\vec{v}]_{B_2}$$

$$= [\vec{v}]_{B_3}.$$

Then, uniqueness implies we have the correct transition matrix QP.

3. If P is the transition matrix from B_1 to B_2 , P^{-1} is the transition matrix from B_2 to B_1 .

Proof. Uniqueness implies that the transition matrix from B_1 to B_1 is I_n . By existence, we have P, the transition matrix from B_1 to B_2 and Q, the transition matrix from B_2 to B_1 . We know that

$$QP = I_n$$

which only occurs if $Q = P^{-1}$.

We now revisit diagonalization, yet again.

Theorem 4.10.3: Diagonalizability, Revisited: Part II

Suppose $A \in \mathcal{M}_{nn}$ and we have found nonsingular P with

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

To find P, we have

$$P = \begin{bmatrix} \vec{v}_1 & \cdots & \vec{v}_n \end{bmatrix}.$$

By Theorem 4.8.2, $B = (\vec{v}_1, ..., \vec{v}_n)$ is an ordered basis of \mathbb{R}^n . Then, P is the transition matrix from B to the standard basis.