201-NYC-05-E (Enriched Linear Algebra I) Lecture Notes

Tristan Martin

Fall 2017

Enriched 201-NYC-05 (Enriched Linear Algebra I), instructed by Matthew Egan, Yariv Barsheshat, Christopher Turner, and Dominic Lemelin. These notes probably contain many typos.

Contents

1 Introduction to Linear Systems and How to Solve Them (Egan) 4

1.1 Linear equations and linear systems . . . 4

1.2 Introduction to augmented matrices . . . 5

2 Applications of Manipulations that Leave the Solution Set Unchanged (Egan) 6

3 More on Row Reduction (Egan) 9

3.1 REF and RREF . . . 9

3.2 Row vectors, column vectors and linear dependance . . . 10

3.3 Matrix rank and linear systems . . . 10

4 Applications of Row Reduction and Introduction to Homogenous Systems (Egan) 13

4.1 Some applications . . . 13

4.1.1 Network flow . . . 13

4.1.2 Balancing chemical reactions . . . 13

4.2 Homogenous systems and solutions to linear systems . . . 14

5 Introduction to Geometric Vectors (Lemelin) 16

5.1 Vectors: the geometric case . . . 16

6 Notation (Turner) 17

7 Dot Product, Projections, Orthogonals and the Cross Product (Barsheshat) 18

7.1 Dot product . . . 18

7.2 Projection of a vector onto another vector . . . 20

7.3 Orthogonal projection . . . 20

7.4 Cross product . . . 21

8 Lines and Planes (Barsheshat) 23

8.1 Lines in $\mathbb{R}^2$ . . . 23

8.2 Lines in $\mathbb{R}^3$ . . . 23

8.3 Parallel lines . . . 24

8.4 Planes in $\mathbb{R}^3$ . . . 25
8.4.1 Parallel planes . . . 27

9 Span and Linear Combination, Dependence and Independence (Barsheshat) 29

10 Span and Bases (Barsheshat) 31
10.1 More on span and linear combination, dependence and independence . . . 31
10.2 Basis for $\mathbb{R}^m$ . . . 32

11 Subspaces and Span (Barsheshat) 34

12 Basis and Dimension for a Subspace (Barsheshat) 36
12.1 Basis for a subspace . . . 36
12.2 Dimension of a subspace . . . 36

13 Exercises on Span (Barsheshat) 38

14 Subspaces of Matrices (Barsheshat) 40

15 Matrix Operations (Barsheshat) 42
15.1 Matrix addition and scalar multiplication . . . 42
15.2 Matrix multiplication . . . 42

16 Column and Row Representations of Matrix Multiplication and the Properties of the Operation (Barsheshat) 44
16.1 Partitioning matrices . . . 44
16.2 Column representation of matrix multiplication . . . 44
16.3 Row representation of matrix multiplication . . . 44
16.4 Column-row representation . . . 44
16.5 Algebraic proterties of matrix addition and scalar multiplication . . . 45
16.6 Matrix exponentiation . . . 46
16.7 Transpose of matrix . . . 46
16.8 Properties of matrix multiplication . . . 46

17 Linear Systems With Matrices (Barsheshat) 47

18 Matrix Inverses (Barsheshat) 49
18.1 Using row-reduction to find inverses . . . 50

19 Gauss-Jordian Algorithm for Finding Inverses (Barsheshat) 51

20 More on Matrix Inverses 54

21 Determinants (Barsheshat) 55
21.1 Determinants for 3×3 matrices . . . 56

22 More on Determinants (Barsheshat) 57
22.1 Generalized cofactor expansions . . . 57

23 Even More On Determinants (Barsheshat) 59
23.1 Triangular Matrices . . . 59
23.2 Determinants of elements matrices . . . 59
23.3 Cramer's Rule . . . 60

24 Complex Numbers (Barsheshat) 62
24.1 Operations on complex numbers 62
24.2 Polar form of copmlex numbers 63

25 More on Complex Numbers (Barsheshat) 65
25.1 De Moivre's formula 65
25.2 Basic polynomial of complex numbers 66

26 Integrative Activity: Electrical Circuits (Barsheshat) 67

27 Enriched Material: Eigenvalues and Eigenvectors (Barsheshat) 68
27.1 Finding eigenvalues of A 68

§1 Introduction to Linear Systems and How to Solve Them (Egan)

§1.1 Linear equations and linear systems

The main field of study for this course will be **linear equation** which have the form

$$a_1x_1 + a_2x_2 + a_3x_3 + \cdots + a_nx_n = b, \tag{1}$$

$a_1, a_2, a_3, \cdots, a_n, b \in \mathbb{R}$ and $x_1, , x_n$ are unknowns. A **solution** of a linear sysrem is a set of values $\{s_1, \cdots, s_n\}$. that make the equation true when we replace $\{x_1, \cdots, x_n\}$ with $\{s_1, \cdots, s_n\}$.

Definition 1.1. A linear system of linear equations is a collection of m linear equations in n unknowns.

It thereby has the form

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_2 + a_{22}x_2 + a_{23}x_3 + \cdots + a_{2n}x_n = b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + \cdots + a_{3n}x_n = b_3 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + a_{m3}x_3 + \cdots + a_{mn}x_n = b_m \end{cases} \tag{2}$$

Definition 1.2. A solution to a linear system is a set of values $\{s_1, \cdots, s_n\}$ that is a solution for all equations in the system.

Example 1.3

Consider the 2×2 system

$$\begin{cases} x_1 - 2x_2 = -1 \\ -x_1 + 3x_2 = 3 \end{cases} \tag{3}$$

It has unique solution $(3, 2)$ and the system is therefore consistent.

Example 1.4

Consider the 2×2 system

$$\begin{cases} x_1 - 2x_2 = -1 \\ -x_1 + 2x_2 = 3 \end{cases} \tag{4}$$

It has no solution (i.e., the system is inconsistent).

Example 1.5

Consider the 2×2 system

$$\begin{cases} x_1 - 2x_2 = -1 \\ -x_1 + 2x_2 = 1 \end{cases} \tag{5}$$

It has infinitely many solutions and is hence a consistent system.

Theorem 1.6

Every linear system of equations has either 0,1 or ∞ solutions.

Furthermore, to describe the complete solution set for example 5, we use parameters to write a out **general solution**. If we define $x_2 = t \in \mathbb{R}$, we can write x_1 as

$$\begin{aligned} x_1 &= 2x_1 - 1 && (6) \\ &= 2t - 1. && (7) \end{aligned}$$

The former is called the parametric eqautions of a line. The general solution is written as

$$\begin{cases} x_1 = 2t - 1 \\ \quad x_2 = t \end{cases} \tag{8}$$

Choose a value for t to obtain a particular solution. For instance $t = 0$ gives solution $(-1, 0)$ and $t = 1$ gives $(1, 1)$.

§1.2 Introduction to augmented matrices

Now, consider larger linear systems like the following 3×3 system:

$$\begin{cases} x_1 - 2x_2 + x_3 = 0 \\ 2x_2 - 8x_3 = 8 \\ -4x_1 + 5x_2 + 9x_3 = -9 \end{cases} \tag{9}$$

To solve such a system we first need to strip away all but the coefficients and the constant to obtain the **augmented matrix of the system**:

$$\left(\begin{array}{ccc|c} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ -4 & 5 & 9 & -9 \end{array}\right). \tag{10}$$

Remark 1.7. *What manipulations of the system will leave the solution set unchanged?* We can:

- interchange any two rows (equations)
- multiply a row (equation) by a non-zero constant
- replace a row (equation) by itself plus a multiple of another row (equation)

§2 Applications of Manipulations that Leave the Solution Set Unchanged (Egan)

Continuing our example from lecture 1, we will use these properties as a means to solve the system. We will soon learn there is an algorithm (i.e., Gauss-Jordan elimination) that allows us to apply these transformations to *any* augmented matrix to obtain a solution set to the original system of equations. For now, we will gloss over the definition of **row echelon form (REF)** and **reduced row echelon form (RREF)**, but from the upcoming examples, it will be pretty clear.

Now, back to our augmented matrix from lecture 1. For now, just focus on understanding how the discussed manipulations are applied (noting that R_i indicates the ith row), the desired result will be later studied.

$$\left(\begin{array}{ccc|c} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ -4 & 5 & 9 & -9 \end{array}\right) \xrightarrow{R_3+4R_1} \left(\begin{array}{ccc|c} \mathbf{1} & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ 0 & -3 & 13 & -9 \end{array}\right) \tag{11}$$

$$\left(\begin{array}{ccc|c} \mathbf{1} & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ 0 & -3 & 13 & -9 \end{array}\right) \xrightarrow{\frac{1}{2}R_2} \left(\begin{array}{ccc|c} \mathbf{1} & -2 & 1 & 0 \\ 0 & \mathbf{1} & -4 & 4 \\ 0 & -3 & 13 & -9 \end{array}\right) \tag{12}$$

$$\left(\begin{array}{ccc|c} \mathbf{1} & -2 & 1 & 0 \\ 0 & \mathbf{1} & -4 & 4 \\ 0 & -3 & 13 & -9 \end{array}\right) \xrightarrow{R_3-3R_2} \left(\begin{array}{ccc|c} \mathbf{1} & -2 & 1 & 0 \\ 0 & \mathbf{1} & -4 & 4 \\ 0 & 0 & \mathbf{1} & 3 \end{array}\right) \tag{13}$$

$$\left(\begin{array}{ccc|c} \mathbf{1} & -2 & 1 & 0 \\ 0 & \mathbf{1} & -4 & 4 \\ 0 & 0 & \mathbf{1} & 3 \end{array}\right) \xrightarrow[R_1-R_3]{R_2+4R_3} \left(\begin{array}{ccc|c} \mathbf{1} & -2 & 0 & -3 \\ 0 & \mathbf{1} & 0 & 16 \\ 0 & 0 & \mathbf{1} & 3 \end{array}\right) \tag{14}$$

$$\left(\begin{array}{ccc|c} \mathbf{1} & -2 & 0 & -3 \\ 0 & \mathbf{1} & 0 & 16 \\ 0 & 0 & \mathbf{1} & 3 \end{array}\right) \xrightarrow{R_1+2R_2} \left(\begin{array}{ccc|c} \mathbf{1} & 0 & 0 & 29 \\ 0 & \mathbf{1} & 0 & 16 \\ 0 & 0 & \mathbf{1} & 3 \end{array}\right) \tag{15}$$

This matrix is in RREF. Note that these manipulations cannot be done simultaneously, but one after the other. Hence the notation $\xrightarrow[R_1-R_3]{R_2+4R_3}$ indicates that the operation R_2+4R_3 was first conduction, then followed by $R_1 - R_3$. From this matrix, the systems solution is obvious, whereas with our augmented matrix, the solution was not. We therefore find:

$$\begin{cases} 1 \cdot x_1 + 0 \cdot x_2 + 0 \cdot x_3 = x_1 = 29 \\ 0 \cdot x_1 + 1 \cdot x_2 + 0 \cdot x_3 = x_2 = 16 \\ 0 \cdot x_1 + 0 \cdot x_2 + 1 \cdot x_3 = x_3 = 3 \end{cases} \tag{16}$$

Note that the **1** are called pivot entries or leading coefficients. Technically, a pivot entry does not necessarily need to be 1, but depending on the textbook you use, some authors are quite pernickety about solely using 1s as pivot entries. Hence, to accommodate the most textbooks possible, I will only be using 1s as pivot entries in these notes, but be aware that it is not required.

We can write our solution in another form using **column vectors** (objects we will study later on in the course):

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 29 \\ 16 \\ 3 \end{pmatrix}. \tag{17}$$

Note that we can verify our answer by plugging in our solution into the original. This is quite trivial, however, so I will leave it for you to do as an exercise.

Exercise 2.1. Find the augmented matrix of the following system and try to row reducing it:

$$\begin{cases} x_2 - 4x_3 = 8 \\ 2x_1 - 3x_2 + 2x_3 = 1 \\ 5x_1 - 8x_2 + 7x_3 = 1 \end{cases} \tag{18}$$

We therefore have the following augmented matrix which can be reduced in the following way:

$$\left(\begin{array}{ccc|c} 0 & 1 & -4 & 8 \\ 2 & -3 & 2 & 1 \\ 5 & -8 & 7 & 1 \end{array}\right) \xrightarrow{R_1 \leftrightarrow R_2} \left(\begin{array}{ccc|c} 2 & -3 & 2 & 1 \\ 0 & \mathbf{1} & -4 & 8 \\ 5 & -8 & 7 & 1 \end{array}\right) \tag{19}$$

$$\left(\begin{array}{ccc|c} 2 & -3 & 2 & 1 \\ 0 & \mathbf{1} & -4 & 8 \\ 5 & -8 & 7 & 1 \end{array}\right) \xrightarrow{R_3 - \frac{5}{2}R_1} \left(\begin{array}{ccc|c} 2 & -3 & 2 & 1 \\ 0 & \mathbf{1} & -4 & 8 \\ 0 & -\frac{1}{2} & 2 & -\frac{3}{2} \end{array}\right) \tag{20}$$

$$\left(\begin{array}{ccc|c} 2 & -3 & 2 & 1 \\ 0 & \mathbf{1} & -4 & 8 \\ 0 & -\frac{1}{2} & 2 & -\frac{3}{2} \end{array}\right) \xrightarrow{R_3 + \frac{1}{2}R_2} \left(\begin{array}{ccc|c} 2 & -3 & 2 & 1 \\ 0 & \mathbf{1} & -4 & 8 \\ 0 & 0 & 0 & -\frac{5}{2} \end{array}\right) \tag{21}$$

We stop here. Notice the last row is equivalent to $0 \cdot x_1 + 0 \cdot x_2 + 0 \cdot x_3 = 5/2$. No x_1, x_2, x_3 satisfies this equations. The system is therefore inconsistent.

Exercise 2.2. Find the augmented matrix of the following system and try to row reducing it:

$$\begin{cases} 3x_2 - 6x_3 + 6x_4 + 4x_5 = -5 \\ 3x_1 - 7x_2 + 8x_3 - 5x_4 + 8x_5 = 9 \\ 3x_1 - 9x_2 + 12x_3 - 9x_4 + 6x_5 = 15 \end{cases} \tag{22}$$

We therefore have the following augmented matrix which can be reduced in the following way:

$$\left(\begin{array}{ccccc|c} 0 & 3 & 6 & 6 & 4 & -5 \\ 3 & -7 & 8 & -5 & 8 & 9 \\ 3 & -9 & 12 & -9 & 6 & 15 \end{array}\right) \xrightarrow{R_1 \leftrightarrow R_3} \left(\begin{array}{ccccc|c} 3 & -9 & 12 & -9 & 6 & 15 \\ 3 & -7 & 8 & -5 & 8 & 9 \\ 0 & 3 & -6 & 6 & 4 & -5 \end{array}\right) \tag{23}$$

$$\left(\begin{array}{ccccc|c} 3 & -9 & 12 & -9 & 6 & 15 \\ 3 & -7 & 8 & -5 & 8 & 9 \\ 0 & 3 & -6 & 6 & 4 & -5 \end{array}\right) \xrightarrow{R_2 - R_1} \left(\begin{array}{ccccc|c} 3 & -9 & 12 & -9 & 6 & 15 \\ 0 & 2 & -4 & 4 & 2 & -6 \\ 0 & 3 & -6 & 6 & 4 & -5 \end{array}\right) \tag{24}$$

$$\left(\begin{array}{ccccc|c} 3 & -9 & 12 & -9 & 6 & 15 \\ 0 & 2 & -4 & 4 & 2 & -6 \\ 0 & 3 & -6 & 6 & 4 & -5 \end{array}\right) \xrightarrow[\frac{1}{2}R_2]{\frac{1}{3}R_1} \left(\begin{array}{ccccc|c} \mathbf{1} & -3 & 4 & -3 & 2 & 5 \\ 0 & \mathbf{1} & -2 & 2 & 1 & -3 \\ 0 & 3 & -6 & 6 & 4 & -5 \end{array}\right) \tag{25}$$

$$\left(\begin{array}{ccccc|c} \mathbf{1} & -3 & 4 & -3 & 2 & 5 \\ 0 & \mathbf{1} & -2 & 2 & 1 & -3 \\ 0 & 3 & -6 & 6 & 4 & -5 \end{array}\right) \xrightarrow{R_3 - 2R_2} \left(\begin{array}{ccccc|c} \mathbf{1} & -3 & 4 & -3 & 2 & 5 \\ 0 & \mathbf{1} & -2 & 2 & 1 & -3 \\ 0 & 0 & 0 & 0 & \mathbf{1} & 4 \end{array}\right) \tag{26}$$

Note that the last pivot entry is further away, that's ok. This matrix is still REF. If we keep going, we can row reduce it until it's in RREF.

$$\left(\begin{array}{ccccc|c} \mathbf{1} & -3 & 4 & -3 & 2 & 5 \\ 0 & \mathbf{1} & -2 & 2 & 1 & -3 \\ 0 & 0 & 0 & 0 & \mathbf{1} & 4 \end{array}\right) \xrightarrow[R_2-R_3]{R_1-2R_3} \left(\begin{array}{ccccc|c} \mathbf{1} & -3 & 4 & -3 & 0 & 3 \\ 0 & \mathbf{1} & -2 & 2 & 1 & -7 \\ 0 & 0 & 0 & 0 & \mathbf{1} & 4 \end{array}\right) \tag{27}$$

$$\left(\begin{array}{ccccc|c} \mathbf{1} & -3 & 4 & -3 & 0 & 3 \\ 0 & \mathbf{1} & -2 & 2 & 1 & -7 \\ 0 & 0 & 0 & 0 & \mathbf{1} & 4 \end{array}\right) \xrightarrow{R_1+3R_2} \left(\begin{array}{ccccc|c} \mathbf{1} & 0 & -2 & 3 & 0 & -24 \\ 0 & \mathbf{1} & -2 & 2 & 0 & -7 \\ 0 & 0 & 0 & 0 & \mathbf{1} & 4 \end{array}\right) \tag{28}$$

This RREF augmented matrix represents the linear system

$$\begin{cases} x_1 - 2x_3 + 3x_4 = -24 \\ x_2 - 2x_3 + 2x_4 = -7 \\ x_5 = 4 \end{cases} \tag{29}$$

We now assign parameters to free variables: let $x_3 = s$ and $x_4 = t$. We can thereby express the general solution of the system

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} -24 + 2s - 3t \\ -7 + 2s - 2t \\ s \\ t \\ 4 \end{pmatrix} \tag{30}$$

$$= \begin{pmatrix} -24 \\ -7 \\ 0 \\ 0 \\ 4 \end{pmatrix} + t \begin{pmatrix} 2 \\ 2 \\ 1 \\ 0 \\ 0 \end{pmatrix} + s \begin{pmatrix} -3 \\ -2 \\ 0 \\ 1 \\ 0 \end{pmatrix} \tag{31}$$

§3 More on Row Reduction (Egan)

§3.1 REF and RREF

Our goal in applying row reduction is to obtain RREF of the augmented matrix

Definition 3.1. A matrix A is in row echelon form if

1. zero rows are at the bottom
2. every non-zero row begins with a 1
3. leading 1s (echelon) from top left to bottom right
4. entries below leading 1s are 0.

Add one more condition, and we get reduced row echelon form.

Definition 3.2. A matrix A is in reduced row echelon form if

1. zero rows are at the bottom
2. every non-zero row begins with a 1
3. leading 1s (echelon) from top left to bottom right
4. entries below leading 1s are 0
5. entries above leading 1s are 0

Exercise 3.3. Which of the following matrices are in RREF and REF?

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{pmatrix} \qquad B = \begin{pmatrix} 1 & 0 & 2 & 5 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 7 \end{pmatrix} \qquad C = \begin{pmatrix} 0 & 1 & 7 \\ 0 & 0 & 6 \end{pmatrix} \tag{32}$$

$$D = \begin{pmatrix} 1 & 5 & 7 & 3 \\ 0 & 1 & 6 & 4 \end{pmatrix} \qquad E = \begin{pmatrix} 1 & 0 & 7 \\ 0 & 1 & 5 \\ 0 & 0 & 0 \end{pmatrix} \tag{33}$$

D is in REF and E, in RREF.

Theorem 3.4

Every $m \times n$ matrix A has a unique RREF R that is row equivalent to A: $A \sim R$.

Given a system with augmented matrix $A \sim R$, how do we write a general solution?

Example 3.5

$$A \sim R = \left(\begin{array}{ccc|c} 1 & 0 & 3 & 7 \\ 0 & 1 & 2 & 1 \end{array}\right) \tag{34}$$

We thereby have

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -3s+7 \\ -2s+1 \\ s \end{pmatrix} = s\begin{pmatrix} -3 \\ -2 \\ 1 \end{pmatrix} + \begin{pmatrix} 7 \\ 1 \\ 0 \end{pmatrix} \tag{35}$$

§3.2 Row vectors, column vectors and linear dependance

Given the following augmented matrix A, what can we conclude about the relationship between the equations of the linear system?

$$A = \left(\begin{array}{ccc|c} 1 & 2 & -1 & 2 \\ 2 & 5 & 2 & -1 \\ 7 & 17 & 5 & -1 \end{array}\right) \xrightarrow[R_3-7R_1]{R_2-2R_1} \left(\begin{array}{ccc|c} \mathbf{1} & 2 & -1 & 2 \\ 0 & 1 & 4 & -5 \\ 0 & 3 & 12 & -15 \end{array}\right) \xrightarrow{R_3-3R_2} \left(\begin{array}{ccc|c} \mathbf{1} & 2 & -1 & 2 \\ 0 & \mathbf{1} & 4 & -5 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

We see that row 3 in REF is a row of 0s. It must therefore be a linear combination of the other row. Let's recap what happened. Looking at the operations conducted on the second row, we have: R_3 was replaced by $R_3 - 7R_1$, hence $R_3 - 7R_1 = 3(R_2 - 2R_1)$. Hence, $R_3 = R_1 + 3R_2$.

Now consider columns of RREF(A):

$$\left(\begin{array}{ccc|c} \mathbf{1} & 2 & -1 & 2 \\ 0 & \mathbf{1} & 4 & -5 \\ 0 & 0 & 0 & 0 \end{array}\right) \xrightarrow{R_1-2R_2} \left(\begin{array}{ccc|c} \mathbf{1} & 0 & -9 & 12 \\ 0 & \mathbf{1} & 4 & -5 \\ 0 & 0 & 0 & 0 \end{array}\right) \tag{36}$$

In RREF, we have (noting that C_i indicates the ith column) $C_3 = -9C_1 + 4C_2$ and $C_4 = 12C_1 - 5C_2$; these as invariants.

Definition 3.6. A set of vectors $(\vec{u}_1, \vec{u}_2, \cdots, \vec{u}_n)$ is **linear dependant (LD)** if one of the vectors $\vec{u}_k$ is a linear combination of the others.

In the example above, the row vectors $\{\vec{r}_1, \vec{r}_2, \vec{r}_3\}$ are LD, since $\vec{r}_3 = \vec{r}_1 + 3\vec{r}_2$. The column vectors are also LD since $\vec{c}_3 = -9\vec{c}_1 + 4\vec{c}_2$. Note that when one vectors is a linear combination of the others, so are the others.

Example 3.7

We can express the linear combination of $\vec{r}_1$ in terms of $\vec{r}_2$ and $\vec{r}_3$, but we can also $\vec{r}_2$ in terms of $\vec{r}_1$ and $\vec{r}_3$:

$$\vec{r}_1 = -3\vec{r}_2 + \vec{r}_3 \tag{37}$$

$$\vec{r}_2 = -\frac{1}{3}\vec{r}_1 + \frac{1}{3}\vec{r}_3 \tag{38}$$

§3.3 Matrix rank and linear systems

Definition 3.8. The rank of a matrix A is rank(A) and is defined as the number of leading ones in REF/RREF of A. Note that rank$(A) \leq m$ and rank$(A) \leq n$

For instance,

$$A = \left(\begin{array}{ccc|c} 1 & 2 & -1 & 2 \\ 2 & 5 & 2 & -4 \\ 7 & 17 & 5 & -1 \end{array}\right) \sim \left(\begin{array}{ccc|c} \mathbf{1} & 2 & -1 & 2 \\ 0 & \mathbf{1} & 4 & -5 \\ 0 & 0 & 0 & 0 \end{array}\right) \tag{39}$$

Hence, rank$(A) = 2$. Note that

$$C = \begin{pmatrix} 1 & 2 & -1 \\ 2 & 5 & 2 \\ 7 & 17 & 5 \end{pmatrix} \tag{40}$$

is called the coefficient matrix. If $\text{rank}(A) = \text{rank}(C)$, then the system is consistent since there is no leading 1s in the last column, whereas if $\text{rank}(A) = \text{rank}(C) + 1$, the system is inconsistent. In the above example,

$$C = \begin{pmatrix} 1 & 2 & -1 \\ 2 & 5 & 2 \\ 7 & 17 & 5 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix} \tag{41}$$

Hence, for this system, $\text{rank}(A) = \text{rank}(C) = 2$. Hence, the system is constitent. However, how to we differentiate between consistent systems with 1 solution and ones with infinitely many? If consistent, then

- $\text{rank}(A) = n$ of C[1] indicates that the system has a unique solution
- $\text{rank}(A) < n$ of C indicates that the system has infinitely many solutions

In summary, consider three systems with following ranks, where b is a column vector wherein the entries are the constants of the linear equations of the systems,

System	$\text{rank}(C)$	$\text{rank}(C\|b)$	n	**Number of solutions**
First	2	2	2	1 *(consistent)*
Second	1	2	2	0 *(inconsistent)*
Third	1	1	2	∞ *(consistent)*

Furthemore, a system with more variables than equations (i.e., $n > m$) is called an **underdetermined system**.

Theorem 3.9
An underdetermined system has either no solutions or infinitely many solutions.

Proof. Since $\text{rank}(C) \leq m$ and $n > m$, then $\text{rank}(C) < n$, and thus, as showed in the table above, the system can only have 0 solutions or infinitely many solutions.

I'm also putting a more involved proof to this theorem here as it seems appropriate and will give students a sneak-peek at material which will be seen later in the course. Simply look over it, and come back once the rank-nullity theorem has been covered.

As seen above $\text{rank}(A) \leq m$. Furthermore, $n > m$ for an underdetermined system. By the rank-nullity theorem, $\text{Null}(A) = n - \text{rank}(A) \geq n - m > 0$. Hence, $\dim \text{Null}(A) > 0$. Thus, if the system has solution $\overrightarrow{X}_\mu$, then any vector $\overrightarrow{X}_\mu + \overrightarrow{X}_0$ is a solution if $\overrightarrow{X}_0 \in \text{Null}(A)$, where there are infinitely many vectors $\overrightarrow{X}_0$. The system can therefore have no solutions or infinitely many. $\square$

[1] Note that n of C corresponds to the number of variables of the linear system.

Example 3.10

For what values of $k \in \mathbb{R}$ does the augmented matrix A represent **(1)** a consistent a linear system and **(2)** an inconsistent system?

$$A = \left(\begin{array}{ccc|c} 1 & 2 & -2 & 4 \\ 3 & -1 & 1 & 2 \\ 2 & -3 & 3 & k \end{array}\right) \sim \left(\begin{array}{ccc|c} 1 & 2 & -2 & 4 \\ 0 & -7 & 7 & -10 \\ 0 & 0 & 0 & k+2 \end{array}\right) \tag{42}$$

1. For the system to be consistent $\text{rank}(A) = \text{rank}(C)$. Hence, $k = -2$, yielding infinitely many solution. The general solution would require one free variable.

2. For the system to be inconsistent $\text{rank}(A) = \text{rank}(C) + 1$. Hence $k \neq -2$.

Exercise 3.11. For what values of $k \in \mathbb{R}$ does the augmented matrix A represent **(1)** a linear system with infinitely many solution, **(2)** a linear system with 1 solution and **(3)** a linear system with no solutions?

$$A = \left(\begin{array}{ccc|c} 1 & 2 & -3 & 4 \\ 3 & -1 & 5 & 2 \\ 4 & 1 & k^2-4 & k+2 \end{array}\right) \sim \left(\begin{array}{ccc|c} 1 & 2 & -3 & 4 \\ 0 & -7 & 14 & -10 \\ 0 & 0 & k^2-16 & k-4 \end{array}\right) \tag{43}$$

1. The only way to let the system to have infinitely many solutions by setting k if by creating a row of zeros. Hence, $k^2 - 16 = 0$ and $k - 4 = 0$. Hence $k = 4$ yields a system with infinitely many solutions.

2. For the system to have a single solution $k^2 - 16 \neq 0$. Hence, $k \neq \pm 4$

3. For the system to be inconsistent, $k^2 - 16 = 0$ and $k - 4 \neq 0$. Hence, $k = -4$

Notice that we could have done this another way: we could have easily determined that $k = -4$ yields no solutions and $k = 4$ yields infinitely many solutions. Hence, by theorem 1.6 we could have therefore said that $k \neq \pm 4$ (since a system can only have either 0,1, or ∞ solutions and since $k^2 - 16$ is defined for all $\mathbb{R}$) gives a matrix representing a linear system with a unique.

§4 Applications of Row Reduction and Introduction to Homogenous Systems (Egan)

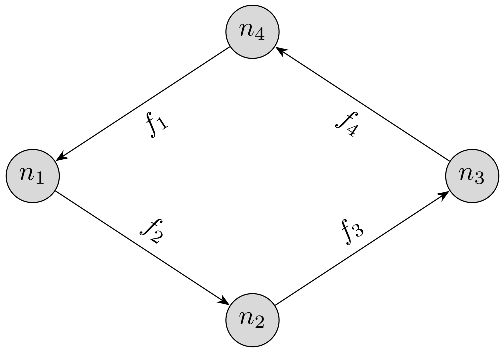
§4.1 Some applications

The following are some applications of what we have learned so far.

§4.1.1 Network flow

Definition 4.1. A network is a collection of nodes withs inbetween them.

For instance, consider the following traffic flow diagram, where $f_1, \cdots, f_4$ indicate flow:



Exercise 4.2. *If node n_1 has an input of 100 cars per unit time and n_4, an input of 200 cars per unit time, find f_i (where $i = 1, 2, 3$) in terms of f_4.* This system gives the following augmented matrix

$$A = \left(\begin{array}{cccc|c} 1 & -1 & 0 & 0 & -100 \\ 0 & 1 & -1 & 0 & 200 \\ 0 & 0 & 1 & -1 & 100 \\ 1 & 0 & 0 & 1 & -200 \end{array}\right) \sim \left(\begin{array}{cccc|c} 1 & 0 & 0 & -1 & 200 \\ 0 & 1 & 0 & -1 & 300 \\ 0 & 0 & 1 & -1 & 100 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right) \tag{44}$$

We therefore have (for $f_4 \geq 0$)

$$f_1 = f_4 + 200 \tag{45}$$
$$f_2 = f_4 + 300 \tag{46}$$
$$f_3 = f_4 + 100 \tag{47}$$

Note that highest flow is f_2.

§4.1.2 Balancing chemical reactions

We can use the method previously shown to balance chemical equations. For instance, what coefficients x, y, z, w will balance the following reaction?

$$xC_8H_8 + yO_2 \to zCO_2 + wH_2O \tag{48}$$

Assuming the conservation of matter, we have

$$\begin{cases} 8x = z \\ 8x = 2w \\ 2y = 2z + w \end{cases} \tag{49}$$

The augmented matrix is thus

$$A = \left(\begin{array}{cccc|c} 8 & 0 & -1 & 0 & 0 \\ 8 & 0 & 0 & -2 & 0 \\ 0 & 2 & -2 & -1 & 0 \end{array}\right) \xrightarrow{\frac{1}{8}R_1} \left(\begin{array}{cccc|c} 1 & 0 & -1/8 & 0 & 0 \\ 8 & 0 & 0 & -2 & 0 \\ 0 & 2 & -2 & -1 & 0 \end{array}\right) \tag{50}$$

$$\left(\begin{array}{cccc|c} 1 & 0 & -1/8 & 0 & 0 \\ 8 & 0 & 0 & -2 & 0 \\ 0 & 2 & -2 & -1 & 0 \end{array}\right) \xrightarrow{R_2-8R_1} \left(\begin{array}{cccc|c} 1 & 0 & -1/8 & 0 & 0 \\ 0 & 0 & 1 & -2 & 0 \\ 0 & 2 & -2 & -1 & 0 \end{array}\right) \tag{51}$$

$$\left(\begin{array}{cccc|c} 1 & 0 & -1/8 & 0 & 0 \\ 0 & 0 & 1 & -2 & 0 \\ 0 & 2 & -2 & -1 & 0 \end{array}\right) \xrightarrow[\frac{1}{2}R_2]{R_2\leftrightarrow R_3} \left(\begin{array}{cccc|c} 1 & 0 & -1/8 & 0 & 0 \\ 0 & 1 & -1 & -1/2 & 0 \\ 0 & 0 & 1 & -2 & 0 \end{array}\right) \tag{52}$$

$$\left(\begin{array}{cccc|c} 1 & 0 & -1/8 & 0 & 0 \\ 0 & 1 & -1 & -1/2 & 0 \\ 1 & 0 & 1 & -2 & 0 \end{array}\right) \xrightarrow[R_1+\frac{1}{8}R_3]{R_2+R_3} \left(\begin{array}{cccc|c} 1 & 0 & 0 & -1/4 & 0 \\ 0 & 1 & 0 & -5/2 & 0 \\ 0 & 0 & 1 & -2 & 0 \end{array}\right) \tag{53}$$

$$\tag{54}$$

Hence,

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = w \begin{pmatrix} 1/4 \\ 5/2 \\ 2 \end{pmatrix} \tag{55}$$

We need the smallest integer solution. Thus, using $w = 4$, we have $x = 1, y = 10, z = 8$:

$$C_8H_8 + 10O_2 \to 8CO_2 + 4H_2O \tag{56}$$

§4.2 Homogenous systems and solutions to linear systems

Definition 4.3. A linear system is homogenous if constants are 0, or in other words $\overrightarrow{b} = \overrightarrow{0}$, where $\overrightarrow{0}$ is the 0 vector.

Remark 4.4. Homogenous systems are always consistent; they always have the following trivial solution:

$$\overrightarrow{X} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \overrightarrow{0} \tag{57}$$

Hence, homogenous systems either only have $\overrightarrow{X} = \overrightarrow{0}$ as a solution or have infinitely many solutions.

Every linear system with augmented matrix $(C|\overrightarrow{b})$ has an associated homogenous system with augmented matrix $(C|\overrightarrow{0})$.

Problem 4.5. Consider the following system:

$$\begin{cases} 2x_1 + 4x_2 + x_3 - 2x_4 = 1 \\ -2x_1 - 4x_2 + x_3 - 5x_4 = 3 \\ 4x_1 + 8x_2 + 4x_3 - x_4 = 6 \\ 2x_1 + 4x_2 + 3x_3 + x_4 = 5 \end{cases} \tag{58}$$

At this point, you should be able to solve the system. You can try this yourself. The solution is

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = s \begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \end{pmatrix} + t \begin{pmatrix} 7/4 \\ 0 \\ -3/2 \\ 1 \end{pmatrix} + \begin{pmatrix} -1/2 \\ 0 \\ 2 \\ 0 \end{pmatrix} \tag{59}$$

where $x_2 = s$ and $x_4 = t$. What do you suppose the solution to the associated homogenous system would look like? Try it out, solve the following system:

$$\begin{cases} 2x_1 + 4x_2 + x_3 - 2x_4 = 0 \\ -2x_1 - 4x_2 + x_3 - 5x_4 = 0 \\ 4x_1 + 8x_2 + 4x_3 - x_4 = 0 \\ 2x_1 + 4x_2 + 3x_3 + x_4 = 0 \end{cases} \tag{60}$$

Notice that RREF(A) is not the same, but RREF(C) is; this is trivial. The solution to the homogenous system is thus

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = s \begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \end{pmatrix} + t \begin{pmatrix} 7/4 \\ 0 \\ -3/2 \\ 1 \end{pmatrix} \tag{61}$$

where $x_2 = s$ and $x_4 = t$. Notice how the solution subspace passes through the origin; this is a consequence of remark 4.4. Let's compare the solution sets. $s = t = 0$ gives $\overrightarrow{X} \neq \overrightarrow{0}$ in the non-homogenous system, whereas it yields the trivial solution in the homogeneous system. However, the parameter portion of the solutions are equivalent.

Theorem 4.6

Suppose an augmented matrix A has a particular solution $\overrightarrow{X}$, then any solution $\overrightarrow{x}_\mu$ to the system must look like

$$\overrightarrow{x_\mu} = \overrightarrow{x}_0 + \overrightarrow{x}_\nu, \tag{62}$$

where $\overrightarrow{x}_\nu$ is a column vector with entries $\in R$ and $\overrightarrow{x}_0$ is a solution of the associated homogenous system.

§5 Introduction to Geometric Vectors (Lemelin)

Notice: *At this point, linear progression of the course material kind of goes out the window. Professor Egan was injured and was subsequently replaced by mutliple professors. We now switch gears and begin material on vectors.*

§5.1 Vectors: the geometric case

Later in this course, and especially in Linear Algebra II and other courses, vectors will be seen to be many, many different things. For now, we content ourselfs by knowing vectors are an element of a vector space, a concept which will be studied later on. For now, we consider the special case where the space is $\mathbb{R}^2$ or $\mathbb{R}^3$. We will look at the particular case where vectors are geometric objects.

Geometric vectors have magnitude and direction, whereas a scalar $\in \mathbb{R}$ (an later $\in \mathbb{C}$), however, only has magnitude. All vectors have two types of operations

1. Scalar multiplication: stretches or shrinks the vectors by keeping the same direction or exactly the opposite direction. Two vectors are parallel when they are a scalar multiple of each other: $\overrightarrow{u} || \overrightarrow{v}$ if $\overrightarrow{u} = k\overrightarrow{v}$, where $k \in \mathbb{R}$.

   Note that $\overrightarrow{0}$ has magnitude 0, but by convention its direction is perpendicular to any other vector, including being perpendicular to itself.

2. Vector addition (subtraction is only a particular case of addition)

Vector addition has the following properties:

- To add vectors $\overrightarrow{w}$ and $\overrightarrow{v}$ ($\overrightarrow{w} + \overrightarrow{v}$), we move $\overrightarrow{v}$ so that its starting point is at the endpoint of $\overrightarrow{w}$.
- Vector addition is commutative $\overrightarrow{w} + \overrightarrow{v} = \overrightarrow{v} + \overrightarrow{w}$. As an exercise, show that this gives a parallelogram.

Given a point P in $\mathbb{R}^2$ or in $\mathbb{R}^3$, the position vector $\overrightarrow{P}$ of the point P is the vector starting at the origin and extending to P. This vector is in **standard position**. We want a useful formula to describe a vector between two points P and Q: $\overrightarrow{PQ} = \overrightarrow{Q} - \overrightarrow{P}$.

§6 Notation (Turner)

This lecture was pretty short and focused on vecture nomenclature.

In $\mathbb{R}^2$, a vector is a list of two compotent:

$$\begin{aligned}\overrightarrow{v} &= \langle v_1, v_2 \rangle = [v_1, v_2] = (v_1, v_2) &(63)\\ &= \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = [v_1, v_2]^{\mathrm{T}}, &(64)\end{aligned}$$

where T indicates the transpose of the row vector. These are all appropriate ways of expressing a vector in $\mathbb{R}^2$. Furthermore, a vector $\overrightarrow{v}$ may be also labeled $\overline{v}$, $\vec{v}$, $\underline{v}$, $\mathbf{v}$.

Definition 6.1. For a matrix A, the transpose respectively maps the ith row and jth column element of A^{T} to the jth row and ith column element of A: $[A^{\mathrm{T}}]_{ij} = A_{ji}$.

The magnitude of a vector $\overrightarrow{\omega} \in \mathbb{R}^n$ is $||\overrightarrow{\omega}|| = \sqrt{\omega_1^2 + \cdots + \omega_n^2}$. Furthermore, the addition of two vectors $\overrightarrow{v}$ and $\overrightarrow{w}$ is expressed as

$$\overrightarrow{v} + \overrightarrow{w} = \langle v_1 + w_1, \cdots, v_n + w_n \rangle. \tag{65}$$

Proposition 6.2

If $k \in \mathbb{R}$ and $\overrightarrow{v} \in \mathbb{R}^n$, then the length of $k\overrightarrow{v}$ is $|k|$ times the length of $\overrightarrow{v}$, $||\overrightarrow{v}||$. This can be shown by assuming the statement is true:

$$\begin{aligned}||k\overrightarrow{v}|| &= |k|||\overrightarrow{v}|| &(66)\\ &= |k|\sqrt{v_1^2 + v_2^2 + v_3^2 + \cdots + v_n^2} &(67)\\ &= \sqrt{k^2}\sqrt{v_1^2 + v_2^2 + v_3^2 + \cdots + v_n^2} &(68)\\ &= \sqrt{k^2v_1^2 + k^2v_2^2 + k^2v_3^2 + \cdots + k^2v_n^2} &(69)\\ &= \langle kv_1, kv_2, kv_3, \cdots, kv_n \rangle &(70)\\ &= ||k\overrightarrow{v}|| &(71)\end{aligned}$$

Finally, here's a nifty that will come in handy when doing vector geometry:

Theorem 6.3

Given three points, they form a triangle if and only if the points are not colinear.

§7 Dot Product, Projections, Orthogonals and the Cross Product (Barsheshat)

§7.1 Dot product

Definition 7.1. $\vec{u} \cdot \vec{v} = \sum_{i=1}^{n} u_i v_i$. This operation is called the **dot product** (or scalar product) of vectors $\vec{u}$ and $\vec{v}$.

The dot product has the following properties:

- commutativity: $\vec{u} \cdot \vec{v} = \vec{v} \cdot \vec{u}$
- distributivity over vector addition: $\vec{u} \cdot (\vec{w} + \vec{v}) = \vec{u} \cdot \vec{w} + \vec{u} \cdot \vec{v}$
- Bilinearity: $\vec{u} \cdot (k\vec{w} + \vec{v}) = k(\vec{u} \cdot \vec{w}) + (\vec{u} \cdot \vec{v})$
- Scalar multiplication: $(k_1\vec{u}) \cdot (k_2\vec{v}) = k_1 k_2(\vec{u} \cdot \vec{v})$
- Not associative because the operation $\vec{u} \cdot (\vec{v} \cdot \vec{w})$ makes no sense: what is in the parentheses is a scalar and you cannot compute the dot product between a scalar and a vector.
- Orthogonality: Two non-zero vectors $\vec{u}$ and $\vec{v}$ are orthogonal if and only if their scalar prouct is 0 (we will shortly see this concretely).
- $\vec{u} \cdot \vec{u} = ||\vec{u}||^2$

Theorem 7.2

We can interpret the dot product geometrically (at least in a familiar way in $\mathbb{R}^2$ and $\mathbb{R}^3$): $\vec{u} \cdot \vec{v} = ||\vec{u}||||\vec{v}|| \cos\theta$, where θ is the angle between the two vectors.

Proof. Given two vectors $\vec{u}$ and $\vec{v}$ separated by angle θ, they form a triangle with a third side $\vec{u} - \vec{v}$. Applying the law of cosines

$$||\vec{u} - \vec{v}||^2 = ||\vec{u}||^2 + ||\vec{v}||^2 - 2||\vec{u}||||\vec{v}|| \cos\theta \tag{72}$$
$$(\vec{u} - \vec{v}) \cdot (\vec{u} - \vec{v}) = ||\vec{u}||^2 + ||\vec{v}||^2 - 2||\vec{u}||||\vec{v}|| \cos\theta \tag{73}$$
$$\vec{u} \cdot \vec{u} - 2\vec{u} \cdot \vec{v} + \vec{v} \cdot \vec{v} = ||\vec{u}||^2 + ||\vec{v}||^2 - 2||\vec{u}||||\vec{v}|| \cos\theta \tag{74}$$
$$||\vec{u}||^2 - 2\vec{u} \cdot \vec{v} + ||\vec{v}||^2 = ||\vec{v}||^2 + ||\vec{v}||^2 - 2||\vec{u}||||\vec{v}|| \cos\theta \tag{75}$$
$$\vec{u} \cdot \vec{v} = ||\vec{u}||||\vec{v}|| \cos\theta \tag{76}$$

□

Now, we can calculate the angle θ between two vectors $\vec{u}$ and $\vec{v}$ by

$$\theta = \arccos\left(\frac{\vec{u} \cdot \vec{v}}{||\vec{u}||||\vec{v}||}\right) \tag{77}$$

Note that the $\arccos\theta$ function maps $[-1, 1] \to [0, \pi/2]$.

Theorem 7.3

For two vectors $\vec{u}$ and $\vec{v}$ in $\mathbb{R}^n$, we have $||\vec{u} \pm \vec{v}||^2 = ||\vec{u}||^2 + ||\vec{v}||^2 \pm 2(\vec{u} \cdot \vec{v})$

Proof.

$$||\overrightarrow{u} \pm \overrightarrow{v}||^2 = \sum_{i=1}^{n}(u \pm v)^2 \tag{78}$$

$$= \sum_{i=0}^{n} u_i^2 \pm 2\sum_{i=0}^{n} u_i v_i + \sum_{i=0}^{n} v_i^2 \tag{79}$$

$$= ||u||^2 \pm 2(\overrightarrow{u} \cdot \overrightarrow{v}) + ||\overrightarrow{v}||^2 \tag{80}$$

□

Theorem 7.4

If $\overrightarrow{v} \in \mathbb{R}^m$ is orthogonal to every other vector, then $\overrightarrow{v} = \overrightarrow{0}$

Proof. Suppose $\overrightarrow{v}$ and $\overrightarrow{w}$ are non-zero, orthogonal vectors, which are not linearly independent: $\overrightarrow{v} = k\overrightarrow{w}$ for some scalar $k \neq 0$. Since $\overrightarrow{v} \cdot \overrightarrow{w} = 0$, then

$$\overrightarrow{v} \cdot \overrightarrow{w} = k(\overrightarrow{w} \cdot \overrightarrow{w}) = k||\overrightarrow{w}|| = 0 \tag{81}$$

. Since $k \neq 0$, then $||\overrightarrow{w}|| = 0$, which contradicts the fact that $\overrightarrow{w}$ is non-zero. □

Theorem 7.5

For two vectors $\overrightarrow{u}$ and $\overrightarrow{v}$ in $\mathbb{R}^2$ and $\mathbb{R}^3$ (and technically in $\mathbb{R}^n$), we have the Cauchy-Schwarz inequality $|\overrightarrow{u} \cdot \overrightarrow{v}| \leq ||\overrightarrow{u}||||\overrightarrow{v}||$

Proof. For $\mathbb{R}^2$ and $\mathbb{R}^3$, we use our geometric intuition and begin with

$$|\cos\theta| \leq 1 \tag{82}$$

Multiplying both sides by $||\overrightarrow{u}||||\overrightarrow{v}||$, we have

$$||\overrightarrow{u}||||\overrightarrow{v}|||\cos\theta| \leq ||\overrightarrow{u}||||\overrightarrow{v}|| \tag{83}$$

Hence,

$$|\overrightarrow{u} \cdot \overrightarrow{v}| \leq ||\overrightarrow{u}||||\overrightarrow{v}|| \tag{84}$$

For $\mathbb{R}^n$, the inequality may be thusly proved: if inequality is true, then

$$|\overrightarrow{u} \cdot \overrightarrow{v}| \leq ||\overrightarrow{u}||||\overrightarrow{v}|| \tag{85}$$

$$|u_1v_1 + \cdots + u_nv_n| \leq \sqrt{u_1^2 + \cdots + u_n^2}\sqrt{v_1^2 + \cdots + v_n^2} \tag{86}$$

$$\leq \sqrt{\left(\sum_{i=1}^{n} u_i^2\right)\left(\sum_{i=1}^{n} v_i^2\right)} \tag{87}$$

$$\left|\sum_{i=1}^{n} u_i v_i\right|^2 \leq \left(\sum_{i=1}^{n} u_i^2\right)\left(\sum_{i=1}^{n} v_i^2\right) \tag{88}$$

$$\left(\sum_{i=1}^{n} u_i v_i\right)^2 \leq \left(\sum_{i=1}^{n} u_i^2\right)\left(\sum_{i=1}^{n} v_i^2\right) \tag{89}$$

$$\tag{90}$$

Now, we consider the function:

$$f(x) = \sum_{i=1}^{n} (u_i x + v_i)^2. \tag{91}$$

We have $(u_i x + v_i) \geq 0 \Rightarrow f(x) \geq 0$, $\forall x \in \mathbb{R}$. We develop $f(x)$,

$$\begin{aligned} f(x) &= \sum_{i=1}^{n} (u_i^2 x^2 + 2u_i v_i x + v_i^2) &(92)\\ &= \sum_{i=1}^{n} (u_i^2) x^2 + 2\sum_{i=1}^{n} (u_i v_i) x + \sum_{i=1}^{n} v_i^2 &(93) \end{aligned}$$

Solving the quadratic equation, noting that $\Delta \leq 0$

$$\begin{aligned} \left(2\sum_{i=1}^{n} (u_i v_i)\right)^2 &\leq 4\left(\sum_{i=1}^{n} u_i^2\right)\left(\sum_{i=1}^{n} v_i^2\right) &(94)\\ \left(\sum_{i=1}^{n} u_i v_i\right)^2 &\leq \left(\sum_{i=1}^{n} u_i^2\right)\left(\sum_{i=1}^{n} v_i^2\right) &(95) \end{aligned}$$

$\square$

§7.2 Projection of a vector onto another vector

Imagine two vectors seperated by angle, identical to the situation seen above. If we were to shine a light straight above these vectors, perpendicular to $\vec{v}$ for instance, the shadown cast by $\vec{u}$ onto $\vec{v}$ is called the projection of $\vec{u}$ onto $\vec{v}$, which is written as $\text{proj}_{\vec{v}}\vec{u}$ and has magnitude

$$\begin{aligned} ||\text{proj}_{\vec{v}}\vec{u}|| &= ||\vec{u}|||\cos\theta| &(96)\\ &= ||\vec{u}||\frac{|\vec{u}\cdot\vec{v}|}{||\vec{u}||||\vec{v}||} &(97)\\ &= \frac{|\vec{u}\cdot\vec{v}|}{||\vec{v}||} &(98) \end{aligned}$$

Furthermore, as $\text{proj}_{\vec{v}}\vec{u}$ is in the same direction (or exactly opposite to) $\vec{v}$, we write:

$$\begin{aligned} \text{proj}_{\vec{v}}\vec{u} &= \frac{\vec{u}\cdot\vec{v}}{||\vec{v}||}\left(\frac{\vec{v}}{||\vec{v}||}\right) &(99)\\ &= \frac{(\vec{u}\cdot\vec{v})}{||\vec{v}||^2}\vec{v} &(100) \end{aligned}$$

If $0 < \theta < \pi/2$, then $\text{proj}_{\vec{v}}\vec{u} > 0$ and if $\pi/2 < \theta < \pi$, then $\text{proj}_{\vec{v}}\vec{u} < 0$. However, if $\theta = \pi/2$, then $\text{proj}_{\vec{v}}\vec{u} = 0$ because of the dot product.

§7.3 Orthogonal projection

Definition 7.6. The orthogonal projection of $\vec{u}$ on $\vec{v}$, written as $\text{orth}_{\vec{v}}\vec{u}$, is $\text{orth}_{\vec{v}}\vec{u} = \vec{u} - \text{proj}_{\vec{v}}\vec{u}$, or

$$\text{orth}_{\vec{v}}\vec{u} = \vec{u} - \frac{(\vec{u}\cdot\vec{v})}{||\vec{v}||^2}\vec{v} \tag{101}$$

Note that $||\text{orth}_{\vec{v}}\vec{u}||^2 + ||\text{proj}_{\vec{v}}\vec{u}||^2 = ||\vec{u}||^2$.

Exercise 7.7. Show that $\text{orth}_{\vec{v}}\vec{u}$ and $\text{proj}_{\vec{v}}\vec{u}$ are perpendicular. We calculate the dot product

$$\begin{aligned}
(\text{proj}_{\vec{v}}\vec{u}) \cdot (\text{orth}_{\vec{v}}\vec{u}) &= \left(\frac{(\vec{u}\cdot\vec{v})}{||\vec{v}||^2}\vec{v}\right) \cdot \left(\vec{u} - \frac{(\vec{u}\cdot\vec{v})}{||\vec{v}||^2}\vec{v}\right) && (102)\\
&= \frac{(\vec{u}\cdot\vec{v})}{||\vec{v}||^2}(\vec{u}\cdot\vec{v}) - \frac{(\vec{u}\cdot\vec{v})}{||\vec{v}||^2}\frac{(\vec{u}\cdot\vec{v})}{||\vec{v}||^2}(\vec{v}\cdot\vec{v}) && (103)\\
&= \frac{(\vec{u}\cdot\vec{v})}{||\vec{v}||^2}(\vec{u}\cdot\vec{v}) - \frac{(\vec{u}\cdot\vec{v})}{||\vec{v}||^2}\frac{(\vec{u}\cdot\vec{v})}{||\vec{v}||^2}||\vec{v}||^2 && (104)\\
&= \frac{(\vec{u}\cdot\vec{v})}{||\vec{v}||^2}(\vec{u}\cdot\vec{v}) - \frac{(\vec{u}\cdot\vec{v})}{||\vec{v}||^2}(\vec{u}\cdot\vec{v}) && (105)\\
&= 0 && (106)
\end{aligned}$$

Hence, $\text{orth}_{\vec{v}}\vec{u}$ and $\text{proj}_{\vec{v}}\vec{u}$ are perpendicular.

§7.4 Cross product

Definition 7.8. $\vec{u} \times \vec{v} = \langle u_2v_3 - u_3v_2, u_3v_1 - u_1v_3, u_1v_2 - u_2v_1\rangle$, which corresponds to

$$\vec{u} \times \vec{v} = \begin{vmatrix} \hat{\imath} & \hat{\jmath} & \hat{k} \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix}, \tag{107}$$

where the expression on the right is called the determinant (an operation we will see later on in the course). Note that in higher mathematics classes the notation is $\vec{u} \wedge \vec{v}$.

The vector $\vec{u} \times \vec{v}$ is perpendicular to both $\vec{u}$ and $\vec{v}$ and satisfies

$$||\vec{u} \times \vec{v}|| = ||\vec{u}||||\vec{v}|| \sin\theta, \tag{108}$$

which corresponds to the area of the parallelogram with sides $\vec{u}$ and $\vec{v}$. The cross product is anticommutative (i.e., $\vec{u} \times \vec{v} = -(\vec{v} \times \vec{u})$), distributive over multiplication, and bilinear. Furthermore $\vec{u} \times \vec{u} = 0$ and

$$\begin{aligned}
&\hat{\imath} \times \hat{\imath} = 0 && \hat{\jmath} \times \hat{\jmath} = 0 && \hat{k} \times \hat{k} = 0 && (109)\\
&\hat{\jmath} \times \hat{k} = \hat{\imath} && \hat{k} \times \hat{\imath} = \hat{\jmath} && \hat{\imath} \times \hat{\jmath} = \hat{k} && (110)
\end{aligned}$$

Exercise 7.9. Show that if three vectors $\vec{u}, \vec{v}, \vec{w}$ are coplanar, then their triple scalar product is 0. We can show this easily: $|\vec{u} \cdot (\vec{v} \times \vec{w})| = ||\vec{u}||||\vec{v}||||\vec{w}||| \sin\theta|| \cos\phi|$, where $\phi = \pi/2$, and thus the scalar triple product is 0.

Exercise 7.10. Prove that $\vec{u} \times \vec{v}$ is perpendicular to $\vec{u}$ and $\vec{v}$. We compute the dot product of these two vectors:

$$\begin{aligned}
\vec{u} \cdot (\vec{u} \times \vec{v}) &= \langle u_1, u_2, u_3\rangle \cdot \langle u_2v_3 - u_3v_2, u_3v_1 - u_1v_3, u_1v_2 - u_2v_1\rangle && (111)\\
&= u_1u_2v_3 - u_1u_3v_2 + u_2u_3v_1 - u_2u_1v_3 + u_3u_1v_2 - u_3u_2v_1 && (112)\\
&= 0 && (113)
\end{aligned}$$

Similarly for $\vec{v}$.

What we have just computed is scalar triple product which corresponds to

$$(\overrightarrow{u} \times \overrightarrow{v}) \cdot \overrightarrow{w} = \begin{vmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix} = \overrightarrow{w} \cdot (\overrightarrow{u} \times \overrightarrow{v}) = \begin{vmatrix} w_1 & w_2 & w_3 \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix} \tag{114}$$

and has the following property: $(\overrightarrow{u} \times \overrightarrow{v}) \cdot \overrightarrow{w} = -(\overrightarrow{u} \times \overrightarrow{w}) \cdot \overrightarrow{v} = -(\overrightarrow{v} \times \overrightarrow{u}) \cdot \overrightarrow{w} = (\overrightarrow{v} \times \overrightarrow{w}) \cdot \overrightarrow{u}$. Note that $|(\overrightarrow{u} \times \overrightarrow{v}) \cdot \overrightarrow{w}|$ corresponds to the volume of the parallelepipeds formed by these three vectors. Note that any permutation of these vectors yields the same result.

Theorem 7.11
$\overrightarrow{a} \times (\overrightarrow{b} \times \overrightarrow{c}) = \overrightarrow{b}(\overrightarrow{a} \cdot \overrightarrow{c}) - \overrightarrow{c}(\overrightarrow{a} \cdot \overrightarrow{b})$. This is known as Lagrange's formula.

Proof.

$$\begin{aligned}
\overrightarrow{a} \times (\overrightarrow{b} \times \overrightarrow{c}) &= \langle a_1, a_2, a_3 \rangle \times \langle b_2c_3 - b_3c_2, b_3c_1 - b_1c_3, b_1c_2 - b_2c_1 \rangle && (115) \\
&= a_2(b_1c_2 - b_2c_1)\hat{\imath} - a_3(b_3c_1 - b_1c_3)\hat{\imath} + a_3(b_2c_3 - b_3c_2)\hat{\jmath} && (116) \\
&\quad - a_1(b_1c_2 - b_2c_1)\hat{\jmath} + a_1(b_3c_1 - b_1c_3)\hat{k} - a_2(b_2c_3 - b_3c_2)\hat{k} && (117) \\
&= a_2b_1c_2\hat{\imath} - a_2b_2c_1\hat{\imath} - a_3b_3c_1\hat{\imath} + a_3b_1c_3\hat{\imath} + a_3b_2c_3\hat{\jmath} - a_3b_3c_2\hat{\jmath} && (118) \\
&\quad - a_1b_1c_2\hat{\jmath} + a_1b_2c_1\hat{\jmath} + a_1b_3c_1\hat{k} - a_1b_1c_3\hat{k} - a_2b_2c_3\hat{k} + a_2b_3c_2\hat{k} && (119) \\
&= \overrightarrow{b}(\overrightarrow{a} \cdot \overrightarrow{c}) - \langle a_1b_1c_1, a_2b_2c_2, a_3b_3c_3 \rangle && (120) \\
&\quad - \overrightarrow{c}(\overrightarrow{a} \cdot \overrightarrow{b}) + \langle a_1b_1c_1, a_2b_2c_2, a_3b_3c_3 \rangle && (121) \\
&= \overrightarrow{b}(\overrightarrow{a} \cdot \overrightarrow{c}) - \overrightarrow{c}(\overrightarrow{a} \cdot \overrightarrow{b}) && (122)
\end{aligned}$$

□

§8 Lines and Planes (Barsheshat)

§8.1 Lines in $\mathbb{R}^2$

The standard equation of a line taught in high schools is $y = ax + b$, where a is the slope and b is the y-intercept. However, we could rewrite this equation in vector notation such that it generalizes nicely in higher dimensions. Consider a point P and Q on a line L. We have $\overrightarrow{OP}$ and $\overrightarrow{OQ}$. Let $\overrightarrow{d} = \overrightarrow{PQ}$ and $\overrightarrow{OP} = \overrightarrow{p}$ for notational simplicity. As a result, the equation of the line L can be written as

$$\overrightarrow{r}(t) = \overrightarrow{p} + t\overrightarrow{d} = \overrightarrow{OX}, \tag{123}$$

where point X is a point on the line. This can be rewritten in vector function form

$$\overrightarrow{r}(t) = \langle x_0 + td_1, y_0 + td_2 \rangle, \tag{124}$$

where $\overrightarrow{p} = \langle x_0, y_0 \rangle$ (which is a point on the line) and $d = \langle d_1, d_2 \rangle$ (which the direction vector of the line).

Exercise 8.1. *Say we have a line described by*

$$\begin{bmatrix} r_1(t) \\ r_2(t) \end{bmatrix} = \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} + t \begin{bmatrix} d_1 \\ d_2 \end{bmatrix}, \tag{125}$$

convert from vector function form to the form $y = ax + b$, where a and b will be in terms of x_0, y_0 and d_1, d_2. We have

$$a = \frac{\Delta y}{\Delta x} = \frac{d_2}{d_1} \tag{126}$$

and

$$y_0 = \frac{d_2}{d_1} x_0 + b \tag{127}$$

$$\Rightarrow b = y_0 - \frac{d_2}{d_1} x_0 \tag{128}$$

§8.2 Lines in $\mathbb{R}^3$

This notation generalizes well in higher dimensions. In n dimensions, the equation is

$$\overrightarrow{r}(t) = \langle x_1, x_2, \cdots, x_n \rangle + t\langle d_1, d_2, \cdots, d_n \rangle. \tag{129}$$

For now, however, we are generally concerned with lines in $\mathbb{R}^3$, written in:

- Vector function form: $\overrightarrow{r}(t) = \langle x_0 + td_1, y_0 + td_2, z_0 + td_3 \rangle$
- Parametric equation form
$$\begin{cases} x = x_0 + td_1 \\ y = y_0 + td_2 \\ z = z_0 + td_3 \end{cases}$$
- Symmetric equation form
$$t = \frac{x - x_0}{d_1} = \frac{y - y_0}{d_2} = \frac{z - z_0}{d_3}$$

Exercise 8.2. *Find another parametric equation describing the line $\vec{r}(t)$.*

$$\vec{r}(t) = \begin{bmatrix} r_1(t) \\ r_2(t) \\ r_3(t) \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} + t \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix} \tag{130}$$

Suppose we introduce a new parameter s such that $t = 3 - 2s$, thereby yielding

$$\vec{r}(t) = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} + (3 - 2s) \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix} \tag{131}$$

$$= \begin{bmatrix} 7 \\ -3 \\ -1 \end{bmatrix} + s \begin{bmatrix} -4 \\ 2 \\ 0 \end{bmatrix} \tag{132}$$

Theorem 8.3

The distance between two skew lines is

$$D = ||\mathrm{proj}_{\vec{n}} \overrightarrow{P_1P_2}||, \tag{133}$$

where P_1 is on the first line, P_2 is on the second line, and $\vec{n} = \vec{d}_1 \times \vec{d}_2$.

Theorem 8.4

The distance between a line L and a point $M_1(x, y, z)$ is

$$D = \frac{||\overrightarrow{M_0M_1} \times \vec{d}\,||}{||\vec{d}\,||}, \tag{134}$$

where M_0 is on the line and $\vec{d}$ is the direction vector of the line.

Proof. Let $s\vec{d}$ (where $s \in \mathbb{R}$) and $\overrightarrow{M_0M_1}$ have $\vec{M}_0$ as an origin. These vectors form the sides of a parallelogram. The endpoints of $s\vec{d}$ and $\overrightarrow{M_0M_1}$ are joined by vector $\vec{D}$ which is orthogonal to $s\vec{d}$. Hence, $||\vec{D}||$ is the distance between the point and the line. Note that

$$A = ||\overrightarrow{M_0M_1} \times s\vec{d}\,||, \tag{135}$$

is the area of the parallelogram. However, A is also equivalent to $A = ||s\vec{d}\,||||\vec{D}||$. As a result,

$$D = ||\vec{D}|| = \frac{|||\overrightarrow{M_0M_1} \times s\vec{d}\,||}{||s\vec{d}\,||} = \frac{|||\overrightarrow{M_0M_1} \times \vec{d}\,||}{||\vec{d}\,||} \tag{136}$$

□

§8.3 Parallel lines

Theorem 8.5

Parallel lines $L_1 // L_2$ have $\vec{d}_1 = k\vec{d}_2$, where $k \in \mathbb{R}$.

Moreover,

1. (in $\mathbb{R}^2$) $L_1 // L_2 \Leftrightarrow L_1$ and L_2 never intersect (L_1 is distinct from L_2)
2. (in $\mathbb{R}^3$) $L_1 // L_2$ (and $L_1 \neq L_2$) $\Leftrightarrow L_1$ and L_2 do not intersect

Proof. We will prove that parallel lines do not intersect. Assume $L_1 // L_2$, where

$$L_1 : = \vec{r}(t) = \vec{p}_1 + t\,\vec{d}_1 \tag{137}$$
$$L_2 : = \vec{r}(s) = \vec{p}_2 + s\,\vec{d}_2. \tag{138}$$

Without loss of generality, we may assume the same direction vector. Assuming distict, we may say that $\vec{p}_2 \notin L_1$:

$$\Rightarrow \vec{p}_2 \neq r(t) = \vec{p}_1 + td_1 \tag{139}$$
$$\Rightarrow \vec{p}_2 - \vec{p}_1 \neq td_1,\ t \in \mathbb{R}. \tag{140}$$

We want to show there are no values for s and t such that $\vec{r}(t) = \vec{r}(s)$. Assume $\exists s, t$ such that $\vec{r}(t) = \vec{r}(s)$:

$$\vec{p}_1 + t\,\vec{d} = \vec{p}_2 + s\,\vec{d} \tag{141}$$
$$(\vec{p}_2 - \vec{p}_1) = (t-s)\,\vec{d},\ (t-s) \in \mathbb{R}. \tag{142}$$

This contradicts the fact that L_1 and L_2 are distinct. □

Theorem 8.6

The distance between two parallel lines is

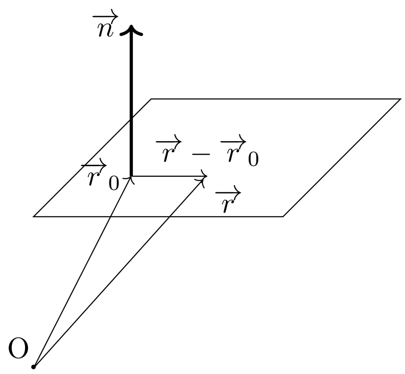
$$D = \frac{||\overrightarrow{P_1P_2} \times \vec{d}\,||}{||\,\vec{d}\,||}, \tag{143}$$

where P_1 is on the first line, P_2 is on the second line, and $\vec{d}$ is the direction vector of one of the lines. Furthermore, D can be expressed as

$$D = ||\text{orth}_{\vec{d}}\overrightarrow{P_1P_2}|| \tag{144}$$

§8.4 Planes in $\mathbb{R}^3$

Given an initial point $P(x_0, y_0, z_0)$ on a plane (with $\vec{r}_0 = \overrightarrow{OP}$) and a normal vector $\vec{n} = \langle a, b, c\rangle$, then any other vector $\vec{r} = \langle x, y, z\rangle$ which starts at the origin and ends at a point on the plane satisfies the equation $\vec{n} \cdot (\vec{r} - \vec{r}_0) = 0$.



Note that $\vec{n}$ is not unique; we can describe a plane in $\mathbb{R}^3$ with ∞ many $\vec{n}$'s. In $\mathbb{R}^n$, $\vec{n} \cdot (\vec{r} - \vec{r}_0) = 0$ defines an $n-1$ dimensional hyperplane.
Subbing $\vec{n} = \langle a, b, c\rangle \neq \vec{0}$, $r = \langle x, y, z\rangle$ and $\vec{r}_0 = \langle x_0, y_0, z_0\rangle$ into the normal vector equations gives

$$\langle a, b, c\rangle \cdot \langle x - x_0, y - y_0, z - z_0\rangle = 0 \tag{145}$$
$$a(x - x_0) + b(y - y_0) + c(z - z_0) = 0 \tag{146}$$
$$ax + by + cz = ax_0 + by_0 + cz_0 = -d, \tag{147}$$

which is in scalar form.
On the other hand, a different formalism can be used to describe planes: given an initial vetcor $\vec{r}_0$ and 2 vectors on the plane, say $\vec{u}, \vec{v}$ (where $\vec{u}, \vec{v} \neq \vec{0}$; $\vec{u} \neq k\vec{v}$), then any vector ending on the plane (say $\vec{r}$ from standard position) can be written as

$$\vec{r} = \vec{r}_0 + s\vec{u} + t\vec{v}, \ \ s, t \in \mathbb{R}, \tag{148}$$

where $\vec{n} = \vec{u} \times \vec{v}$. Note that the parametric form (with 2 parameters) always defines a 2-dimensional plane in $\mathbb{R}^n$. We must ensure that $\vec{u}, \vec{v} \neq \vec{0}$ and $\vec{u} \neq k\vec{v}$, however.

Example 8.7

Conisder the initial vector $\vec{r}_0 = \langle 2, 1, -1\rangle$ and $\vec{n} = \langle 1, 1, 2\rangle$. The equation of the plane in scalar form is

$$\vec{n} \cdot (\vec{r} - \vec{r}_0) = 0 \tag{149}$$
$$\langle 1, 1, 2\rangle \cdot \langle x - 2, y - 1, z + 1\rangle = 0 \tag{150}$$
$$x + y + 2z = 1 \tag{151}$$

Moreover, we can write out solutions to scalar form equation of the plane in parametric form:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 - s - 2t \\ s \\ t \end{bmatrix} \tag{152}$$
$$= \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + s\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + t\begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix}, \ \ s, t \in \mathbb{R}, \tag{153}$$

where $y = s, z = t$.

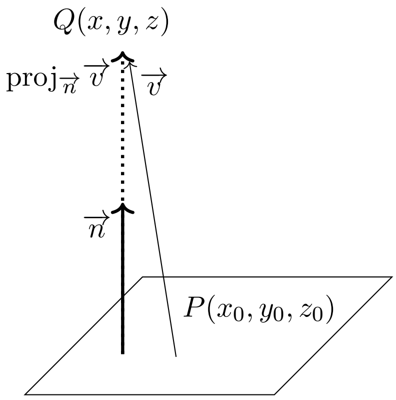
Theorem 8.8

The shortest distance D between some point $Q(x, y, z)$ from some plane described by $ax + by + cz + d = 0$ is

$$D = \frac{|ax + by + cz + d|}{\sqrt{a^2 + b^2 + c^2}}, \tag{154}$$

assuming P is not on the plane.

Proof. Consider the following diagram:



We have

$$\begin{aligned}
D &= ||\text{proj}_{\overrightarrow{n}} \overrightarrow{v}|| && (155)\\
&= \left\| \frac{\overrightarrow{n} \cdot \overrightarrow{v}}{||\overrightarrow{n}||^2} \overrightarrow{n} \right\| && (156)\\
&= \frac{|\overrightarrow{n} \cdot \overrightarrow{v}|}{||\overrightarrow{n}||} && (157)\\
&= \frac{|a(x - x_0) + b(y - y_0) + c(z - z_0)|}{\sqrt{a^2 + b^2 + c^2}} && (158)\\
&= \frac{|ax + by + cz - ax_0 - by_0 - cz_0|}{\sqrt{a^2 + b^2 + c^2}} && (159)\\
&= \frac{|ax + by + cz + d|}{\sqrt{a^2 + b^2 + c^2}} && (160)
\end{aligned}$$

□

§8.4.1 Parallel planes

In $\mathbb{R}^3$, if π_1 and π_2 are parallel and distinct, then they never intersect: $\pi_1 // \pi_2$ $(\pi_1 \neq \pi_2) \Leftrightarrow \pi_1, \pi_2$ never intersect.

Theorem 8.9

In $\mathbb{R}^3$, if two planes π_1, π_2 are not parallel, then they intersect to form a line.

Exercise 8.10. *If two planes are not parallel, they interset to form a line. Assuming*

$$\begin{aligned}
\pi_1 &: a_1x + b_1y + c_1z + d_1 = 0 && (161)\\
\pi_2 &: a_2x + b_2y + c_2z + d_2 = 0 && (162)
\end{aligned}$$

Find a formula for the line formed by the intersection of these planes. The line must have a direction common to both planes: $\overrightarrow{d} = \overrightarrow{n}_1 \times \overrightarrow{n}_2$. The rest is trivial.

Example 8.11

Find the intersection of π_1 and π_2

$$\pi_1 : 2x - y + z = 3 \tag{163}$$
$$\pi_2 : x + 2y + 3z = 0 \tag{164}$$

We have the augmented matrix A which, after row reduction, yields a matrix R in RREF:

$$A = \begin{pmatrix} 1 & 2 & 3 & 0 \\ 2 & -1 & 1 & 3 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 1 & 6/5 \\ 0 & 1 & 1 & -3/5 \end{pmatrix} = R \tag{165}$$

Hence,

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 6/5 - t \\ -3/5 - t \\ t \end{bmatrix} = \begin{bmatrix} 6/5 \\ -3/5 \\ 0 \end{bmatrix} + t \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}, \tag{166}$$

where $z = t \in \mathbb{R}$.
However, a faster solution would have utilized the cross product of the normal vectors:

$$\overrightarrow{n}_1 \times \overrightarrow{n}_1 = \langle 2, -1, 1 \rangle \times \langle 1, 2, 3 \rangle \tag{167}$$
$$= -5\hat{\imath} - 5\hat{\jmath} + 5\hat{k} \tag{168}$$
$$= -5\langle 1, 1, -1 \rangle \tag{169}$$

Hence, $\overrightarrow{d} = \langle 1, , 1, -1 \rangle$. We would simply need a point on the line to be able to describe it mathematically. Out of simplicity, we will look for a point with $z = 0$:

$$2x - y = 3 \tag{170}$$
$$x + 2y = 0 \tag{171}$$
$$\Rightarrow 2(-2y) - y = 3 \tag{172}$$
$$-4y - y = 3 \tag{173}$$
$$y = -3/5 \tag{174}$$
$$\Rightarrow x = 6/5 \tag{175}$$

Hence, we find an equivalent equation. Note that the equations given by the two methods do not always match, because we could use any point on the line as $\overrightarrow{r}_0$ and any multiple of the direction vector. Both, however, will be correct and describe the same line *(if your work is correct)*.

§9 Span and Linear Combination, Dependence and Independence (Barsheshat)

Definition 9.1. A **vector set** is defined as $\mathcal{V} = \{\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_n\} \subseteq \mathbb{R}^m$.

Definition 9.2. A **linear combination** of vectors $\vec{v}_1, \cdots, \vec{v}_n \in \mathbb{R}^m$ is simply a vector sum of arbitrary scalar multiples of these vectors. Note that n is not the dimension of the vector space.

Vector $\vec{v}$ is a linear combination of $\vec{v}_1, \cdots, \vec{v}_n$ if

$$\vec{v} = a_1\vec{v}_1 + \cdots + a_n\vec{v}_n = \sum_{i=1}^{n} a_i\vec{v}_i \tag{176}$$

Definition 9.3. The **linear span** of a set of vectors $\{\vec{v}_1, \cdots, \vec{v}_n\} \subseteq \mathbb{R}^m$ denoted by

$$\text{span}\left(\{\vec{v}_1, \cdots, \vec{v}_n\}\right) = \left\{\sum_{i=1}^{n} a_i\vec{v}_i : a_i \in \mathbb{R}\right\} \tag{177}$$

The properties of the linear span of a set of vectors are:

- $\vec{0} \in \text{span}\left(\{\vec{v}_1, \cdots, \vec{v}_n\}\right)$ or $\text{span}\left(\{\vec{v}_1, \cdots, \vec{v}_n\}\right) = \left\{\vec{0}\right\}$
- $\vec{v}_1, \cdots, \vec{v}_n \in \text{span}\left(\{\vec{v}_1, \cdots, \vec{v}_n\}\right)$
- closure under scalar multiplication: if $\vec{v} \in \text{span}\left(\{\vec{v}_1, \cdots, \vec{v}_n\}\right)$, then $a\vec{v} \in \text{span}\left(\{\vec{v}_1, \cdots, \vec{v}_n\}\right)$ (for any scalar a)
- closure under addition: if $\vec{u}, \vec{v} \in \text{span}\left(\{\vec{v}_1, \cdots, \vec{v}_n\}\right)$, then $\vec{u}+\vec{v} \in \text{span}\left(\{\vec{v}_1, \cdots, \vec{v}_n\}\right)$

Example 9.4
If $\vec{e}_1 = \hat{\imath}, \vec{e}_2 = \hat{\jmath}, \vec{e}_3 = \hat{k}$, then $\text{span}\left(\vec{e}_1, \vec{e}_2, \vec{e}_3\right)$ corresponds to what space?
$\text{span}\left(\vec{e}_1, \vec{e}_2, \vec{e}_3\right) = \mathbb{R}^3$ because any vector $\vec{v} = \langle a, b, c\rangle \in \mathbb{R}^3$ can be written as $\vec{v} = a\vec{e}_1 + b\vec{e}_2 + c\vec{e}_3$

Exercise 9.5. *Let $\vec{v}_1 = \langle 1, 2\rangle, \vec{v}_2 = \langle -1, 3\rangle, \vec{v}_3 = \langle 1, 1\rangle$, find:*

1. $\text{span}\left(\{\vec{v}_1, \vec{v}_2\}\right)$
2. $\text{span}\left(\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}\right)$

For the first exercise, let $\vec{v} = \langle a, b\rangle \in \mathbb{R}^2$. We want to show that $\vec{v} = x\vec{v}_1 + y\vec{v}_2$, for some $x, y \in \mathbb{R}$. Thus

$$\begin{pmatrix} a \\ b \end{pmatrix} = x\begin{pmatrix} 1 \\ 2 \end{pmatrix} = y\begin{pmatrix} -1 \\ 3 \end{pmatrix} \Rightarrow A = \left(\begin{array}{cc|c} 1 & -1 & a \\ 2 & 5 & b \end{array}\right) \tag{178}$$

If this system has a solution, then $\text{span}\left(\{\vec{v}_1, \vec{v}_2\}\right) = \mathbb{R}^2$. We have

$$A = \left(\begin{array}{cc|c} 1 & -1 & a \\ 2 & 5 & b \end{array}\right) \sim \left(\begin{array}{cc|c} 1 & 0 & \frac{3a+b}{5} \\ 0 & 1 & \frac{-2a+b}{5} \end{array}\right) = R \tag{179}$$

Hence, by setting $x = \frac{3a+b}{5}$ and $y = \frac{-2a+b}{5}$, we obtain

$$\overrightarrow{v} = \langle a, b\rangle = x\overrightarrow{v}_1 + y\overrightarrow{v}_2 \tag{180}$$

As a result, $\operatorname{span}(\{\overrightarrow{v}_1, \overrightarrow{v}_2\}) = \mathbb{R}^2$.

For the second exercise, we have $\operatorname{span}(\{\overrightarrow{v}_1, \overrightarrow{v}_2, \overrightarrow{v}_3\}) = \mathbb{R}^2$, because it has to include $\operatorname{span}(\{\overrightarrow{v}_1, \overrightarrow{v}_2\}) = \mathbb{R}^2$): $\operatorname{span}(\{\overrightarrow{v}_1, \overrightarrow{v}_2\}) \subseteq \operatorname{span}(\{\overrightarrow{v}_1, \overrightarrow{v}_2 \overrightarrow{v}_3\})$. In other words, $\overrightarrow{v}_3$ did not add any extra information.

Definition 9.6. A set of vectors $\{\overrightarrow{v}_1, \cdots, \overrightarrow{v}_n\}$ is **linearaly independent** if the homogeneous system defined by $\sum_{i=1}^n a_i\overrightarrow{v}_i = \overrightarrow{0}$ (where the scalars are the variables of the system) has only the trivial solution (i.e., $a_i = 0$). If there is any non-trivial solution, the set of vectors is linearly dependent.

Exercise 9.7. *Show that* $\overrightarrow{v}_1 = \langle 1, 2\rangle, \overrightarrow{v}_2 = \langle -1, 3\rangle, \overrightarrow{v}_3 = \langle 1, 1\rangle$ *are linearly dependent.*
We have

$$a_1\overrightarrow{v}_1 + a_2\overrightarrow{v}_2 + a_3\overrightarrow{v}_3 = \overrightarrow{0} \tag{181}$$

$$a_1\langle 1, 2\rangle + a_2\langle -1, 3\rangle + a_3\langle 1, 1\rangle = \overrightarrow{0}. \tag{182}$$

Hence,

$$\begin{cases} a_1 - a_2 + a_3 = 0 \\ 2a_1 + 3a_2 + a_3 = 0 \end{cases} \tag{183}$$

This system is an underdetermined homogeneous system. Hence, it must have infinite solutions. As a result, $\overrightarrow{v}_1 = \langle 1, 2\rangle, \overrightarrow{v}_2 = \langle -1, 3\rangle, \overrightarrow{v}_3 = \langle 1, 1\rangle$ are linearly dependent.

We could have also used the result from the previous example such that

$$\overrightarrow{v} = \langle a, b\rangle = \frac{3a+b}{5}\overrightarrow{v}_1 + \frac{-2a+b}{5}\overrightarrow{v}_2. \tag{184}$$

We can thereby have (setting $a, b = 1$)

$$\overrightarrow{v}_3 = \langle 1, 1\rangle = \frac{4}{5}\overrightarrow{v}_1 - \frac{1}{5}\overrightarrow{v}_2 \to \frac{4}{5}\overrightarrow{v}_1 - \frac{1}{5}\overrightarrow{v}_2 - \overrightarrow{v}_3 = \overrightarrow{0}. \tag{185}$$

This is a non-trivial solution.

§10 Span and Bases (Barsheshat)

§10.1 More on span and linear combination, dependence and independence

Proposition 10.1

Let $\mathcal{V} = \{\vec{v}_1, \cdots, \vec{v}_n\}$. The following two statements are then equivalent:

- $\mathcal{V}$ is linearly independent
- no vector in $\mathcal{V}$ can be expressed as a linear combination of the other vectors in the set.

Proof. We will prove $\neg a \Rightarrow \neg b$ and $\neg b \Rightarrow \neg a$.

We begin by showing $\neg a \Rightarrow \neg b$. If $\mathcal{V}$ is linearly dependent, then there is a non-trivial solution to:

$$a_1\vec{v}_1 + \cdots a_n\vec{v}_n = \vec{0}. \tag{186}$$

Without loss of generality, assume $a_1 \neq 0$, then we have:

$$a_1\vec{v}_1 = -a_2\vec{v}_2 - \cdots - a_n\vec{v}_n \tag{187}$$

$$\vec{v}_1 = \left(\frac{-a_2}{a_1}\right)\vec{v}_2 + \left(\frac{-a_3}{a_1}\right)\vec{v}_3 + \cdots + \left(\frac{-a_n}{a_1}\right)\vec{v}_n. \tag{188}$$

b is therefore false (i.e., $\neg b$ is true).

Now, we show $\neg b \Rightarrow \neg a$. If, say $\vec{v}_1$, could be expressed as a linear combination of the other vectors, then

$$\vec{v} = b_2\vec{v}_2 + \cdots + b_n\vec{v}_n, \tag{189}$$

for some $b_2, \cdots, b_n \in \mathbb{R}$. We can rearrange the previous equation as follows

$$\vec{v}_1 - b_2\vec{v}_2 - \cdots - b_n\vec{v}_n = \vec{0}. \tag{190}$$

Setting $a_1 = 1$ and $a_i = -b_i$ for $i \geq 2$, we find a non-trivial solution to the independence equation. Thus, $\mathcal{V}$ is not linearly independent (i.e., $\neg a$ is true). $\square$

Before continuing, here are some additional propertites of span:

- $\vec{0} \in \text{span}(\mathcal{V})$
- $\vec{v}_i \in \text{span}(\mathcal{V})$, for $i = 1, \cdots, n$
- **closure under adition/scalar multiplication**: if $\vec{u}, \vec{v} \in \text{span}(\mathcal{V})$ and $a, b \in \mathbb{R}$, then $(a\vec{u} + b\vec{v}) \in \text{span}(\mathcal{V})$.
- if $\mathcal{V} \subseteq \mathcal{U}$, then $\text{span}(\mathcal{V}) \subseteq \text{span}(\mathcal{U})$.

The following are properties of linear independence,

- if $\vec{0} \in \mathcal{V}$, then $\mathcal{V}$ is not independent
- any singleton set is independent if its constituent vector is non-zero
- for any set of two vectors, saying it's independent is equivalent to saying that the vectors are not scalar multiples of eachother
- if $\mathcal{V} \subseteq \mathcal{U}$, and $\mathcal{U}$ is linearly independent, then so is $\mathcal{V}$

Proposition 10.2

In $\mathbb{R}^m$, the maximum size of a set of independent vectors is m.

Proof. Consider the system $\sum_{i=1}^{n} a_i \overrightarrow{v}_i = \overrightarrow{0}$. This corresponds to a coefficient matrix

$$C = [\overrightarrow{v}_1 \cdots \overrightarrow{v}_n] \tag{191}$$

C has m rows and n columns. Hence, from our knowledge of row reduction, if $n > m$, then the homogeneous system has infinitely many solutions which makes the vector set dependent. As a result, for a vector set to be independent, $n \leq m$. □

Example 10.3

If $\{\overrightarrow{u}, \overrightarrow{v}, \overrightarrow{w}\}$ is independent, show that $\{\overrightarrow{u}+\overrightarrow{v}, \overrightarrow{u}+\overrightarrow{w}, \overrightarrow{v}+\overrightarrow{w}\}$ is also independent. We begin by setting up the system with the vectors in the second set. Assume

$$a(\overrightarrow{u}+\overrightarrow{v}) + b(\overrightarrow{u}+\overrightarrow{w}) + c(\overrightarrow{v}+\overrightarrow{w}) = \overrightarrow{0} \tag{192}$$

$$(a+b)\overrightarrow{u} + (a+c)\overrightarrow{v} + (b+c)\overrightarrow{w} = \overrightarrow{0} \tag{193}$$

Because of the independence of $\overrightarrow{u}, \overrightarrow{v}, \overrightarrow{w}$, we can conclude that $a+b=0, a+c=0, b+c=0$. We have a homogeneous system. You can check that the only solution to the system is the trivial solution. This thereby shows that $\{\overrightarrow{u}+\overrightarrow{v}, \overrightarrow{u}+\overrightarrow{w}, \overrightarrow{v}+\overrightarrow{w}\}$ is independent.

§10.2 Basis for $\mathbb{R}^m$

Definition 10.4. A set of vectors $\mathcal{V} = \{\overrightarrow{v}_1, \cdots, \overrightarrow{v}_n\}$ is called a basis for $\mathbb{R}^m$ if any vector in $\mathbb{R}^m$ can be expressed as a linear combinations of vectors in $\mathcal{V}$: $\operatorname{span}(\mathcal{V}) = \mathbb{R}^m$ ($\mathcal{V}$ is a spanning set for $\mathbb{R}^m$).

This is equivalent to:

$$\underbrace{\operatorname{span}(\mathcal{V}) = \mathbb{R}^m}_{(n\geq m)} \text{ and } \underbrace{\mathcal{V} \text{ is linearly independent}}_{(n\leq m)} \tag{194}$$

Notice the following fact: if $\mathcal{V}$ is a basis for $\mathbb{R}^m$, then $|\mathcal{V}| = m$ (i.e., $\mathcal{V}$ has m elements).

Example 10.5

$\left\{\hat{\imath}, \hat{\jmath}, \hat{k}\right\}$ is called the standard basis for $\mathbb{R}^3$. More generally, if we define $\overrightarrow{e}_i \in \mathbb{R}^m$ to be the vector with all zeros except for a 1 in the i'th coordinate, then $\{\overrightarrow{e}_1, \cdots, \overrightarrow{e}_m\}$ is called the stadard (orthonormal) basis for $\mathbb{R}^m$.

Exercise 10.6. *Consider vectors $\overrightarrow{v}_1 = \langle 2,0,1\rangle$, $\overrightarrow{v}_2 = \langle -1,-1,0\rangle$, $\overrightarrow{v}_3 = \langle 1,-1,0\rangle$:*

1. *express $\hat{\imath}, \hat{\jmath}, \hat{k}$ in terms of $\overrightarrow{v}_1, \overrightarrow{v}_2, \overrightarrow{v}_3$.*
2. *use above and the work shown to argue that $\overrightarrow{v}_1, \overrightarrow{v}_2, \overrightarrow{v}_3$ form a basis.*

We can easily see that:

$$\hat{\imath} = \frac{1}{2}(\overrightarrow{v}_3 - \overrightarrow{v}_2) \tag{195}$$

$$\hat{\jmath} = -\frac{1}{2}(\overrightarrow{v}_2 + \overrightarrow{v}_3) \tag{196}$$

$$\hat{k} = \overrightarrow{v}_1 + \overrightarrow{v}_2 - \overrightarrow{v}_3 \tag{197}$$

However, if it were not as obvious, we could have used row reduction. Consider the system

$$x\overrightarrow{v}_1 + y\overrightarrow{v}_2 + z\overrightarrow{v}_3 = \langle a, b, c\rangle. \tag{198}$$

This system translates to the following augmented matrix

$$A = \left(\begin{array}{ccc|c} 2 & -1 & 1 & a \\ 0 & -1 & -1 & b \\ 1 & 0 & 0 & c \end{array}\right) \sim \left(\begin{array}{ccc|c} 1 & 0 & 0 & c \\ 0 & 1 & 0 & \frac{-a-b+2c}{2} \\ 0 & 0 & 1 & \frac{a-b-2c}{2} \end{array}\right) = R \tag{199}$$

The system thereby has a <u>unique</u> solution. To determine the coefficients we need for $\hat{\imath}$, we set $a = 1$ and $b, c = 0$. This yields:

$$\langle a, b, c\rangle = c\overrightarrow{v}_1 + \left(\frac{-a-b+2c}{2}\right)\overrightarrow{v}_2 + \left(\frac{a-b-2c}{2}\right)\overrightarrow{v}_3 \tag{200}$$

$$\hat{\imath} = \langle 1, 0, 0\rangle = \left(\frac{-1-0+0}{2}\right)\overrightarrow{v}_2 + \left(\frac{1-0-0}{2}\right)\overrightarrow{v}_3 \tag{201}$$

$$\hat{\imath} = \frac{1}{2}(\overrightarrow{v}_3 - \overrightarrow{v}_2) \tag{202}$$

Similarly for $\hat{\jmath}$ and $\hat{k}$. Furthermore, we have shown that there is exactly one way of expressing any vector $\overrightarrow{v} = \langle a, b, c\rangle \in \mathbb{R}^3$ as a linear combination of $\overrightarrow{v}_1, \overrightarrow{v}_2, \overrightarrow{v}_3$. Thus $\{\overrightarrow{v}_1, \overrightarrow{v}_2, \overrightarrow{v}_3\}$ is a basis for $\mathbb{R}^3$.

§11 Subspaces and Span (Barsheshat)

$|\mathcal{V}| = m$ if $\mathcal{V}$ is a basis of $\mathbb{R}^m$ is a consequence of the following two propositions:

Proposition 11.1

If $\text{span}(\mathcal{V}) = \mathbb{R}^m$, then $|\mathcal{V}| \geq m$

Proof. INSERT PROOF □

Proposition 11.2

If $\mathcal{V} \subseteq \mathbb{R}^m$, then $|\mathcal{V}| \leq m$

Proof. INSERT PROOF □

Based on the two previous propositions one can say that a basis is both a minimal spanning set or a maximal independent set. The standard basis for $\mathbb{R}^m : \{\vec{e}_i\}$ where $\vec{e}_i$ has 0's as every component except for the i'th component, which is 1.

Exercise 11.3. *Describe the span of the following three vectors:* $\vec{v}_1 = \langle 1, 0, 2\rangle$, $\vec{v}_2 = \langle 2, 1, -1\rangle$, $\vec{v}_3 = \langle 0, 1, -5\rangle$*:*

1. *geometrically (what object is this?)*
2. *as a span of two vectors*

Geometrically, it's a plane with normal $\vec{v}_1 \times \vec{v}_2$ and goes through the origin. The span of three vectors is any vector $\vec{v} = \langle x, y, z\rangle = a\vec{v}_1 + b\vec{v}_2 + c\vec{v}_3$. Treating a, b, c as the variables of the system, we obtain the following augmented matrix

$$A = \left(\begin{array}{ccc|c} 1 & 2 & 0 & x \\ 0 & 1 & 1 & y \\ 2 & -1 & -5 & z \end{array}\right) \sim \left(\begin{array}{ccc|c} 1 & 2 & 0 & x \\ 0 & 1 & 1 & y \\ 0 & 0 & 0 & -2x + 5y + z \end{array}\right) \tag{203}$$

Hence, for this system to be consistent, we must have $-2x + 5y + z = 0$. This yields the equation of a plane (which goes through the origin) with normal $\vec{n} = \langle -2, 5, 1\rangle$ (this answers part 1). For part 2 of the exercise, we find $\vec{v}_2 - 2\vec{v}_1 = \vec{v}_3$. As a result, $\vec{v}_3 \in \text{span}(\{\vec{v}_1, \vec{v}_2\})$ and thus $\text{span}(\{\vec{v}_1, \vec{v}_2\}) = \text{span}(\{\vec{v}_1, \vec{v}_2, \vec{v}_3\})$.

Definition 11.4. A set $\mathcal{V}$ is called a **subspace** of $\mathbb{R}^m$ if $\mathcal{V}$ is a non-empty set closed under addition and scalar multiplication [2].

Properties of subspaces:

- $\vec{0} \in \mathcal{V}$ for any subspace $\mathcal{V}$ since $\mathcal{V}$ is not empty $\Rightarrow \vec{v} \in \mathcal{V} \Rightarrow 0\vec{v} = \vec{0} \in \mathcal{V}$. In other words, any subspace $\mathcal{V} \subseteq \mathbb{R}^m$ can be written as a span of n **linearly independent** vectors with $n \leq m$.

[2] Note the following:

- **closure under addition**: if $\vec{u}, \vec{v} \in \mathcal{V}$, then $\vec{u} + \vec{v} \in \mathcal{V}$.
- **closure under multiplication**: if $\vec{v} \in \mathcal{V}$ and $t \in \mathbb{R}^m$, then $t\vec{v} \in \mathcal{V}$.
- $\mathcal{V}$ is not empty

- Any span of any collection of vectors is a subspace of $\mathbb{R}^m$.
- Any subspace $\mathcal{V} \subseteq \mathbb{R}^m$ can be described as a span of n vectors with $n \leq m$ (tricky to prove, yet important to understand).
- $\left\{\overrightarrow{0}\right\}$ is called the trivial subspace.
- $\mathbb{R}^m$ is a subspace of itself.

Theorem 11.5

For any set of vectors $\mathcal{V} = \{v_1, \cdots, v_n\}$ in a vector space V (which will be $\mathbb{R}^n$ for now), span$(\mathcal{V})$ is a subspace of $\mathbb{R}^n$.

Proof. Let $\overrightarrow{u}, \overrightarrow{w} \in \text{span}(\mathcal{V})$ and $k \in \mathbb{R}$. Then, there exists $c_1, \cdots, c_n, a_1, \cdots, a_n \in \mathbb{R}$ such that

$$\overrightarrow{u} = c_1\overrightarrow{v}_1 + \cdots + c_n\overrightarrow{v}_n \tag{204}$$
$$\overrightarrow{w} = a_1\overrightarrow{v}_1 + \cdots + a_n\overrightarrow{v}_n \tag{205}$$

Notice that

$$\overrightarrow{u} + \overrightarrow{w} = (c_1 + a_1)\overrightarrow{v}_1 + \cdots + (c_n + a_n)\overrightarrow{v}_n. \tag{206}$$

Hence, span is closed under addition. Moreover,

$$k\overrightarrow{u} = (kc_1)\overrightarrow{v}_1 + \cdots (kc_n)\overrightarrow{v}_n. \tag{207}$$

As a result, span is closed under scalar multiplication. Finally, span is not empty as it contains the 0 vector, which is obtained by setting the coefficients in the linear combination to 0. Since span is a non-empty set closed under addition and scalar multiplication, it is a subspace of $\mathbb{R}^n$. $\square$

Exercise 11.6. *Are the following subsets subspaces? Justify carefully.*

1. $\{\overrightarrow{v} \in \mathbb{R}^m : \overrightarrow{u} \cdot \overrightarrow{v} = 0\}$ *for some vector* $\overrightarrow{u} \in \mathbb{R}^m$
2. $\left\{\langle a, b, c\rangle \in \mathbb{R}^3 : b = a + c + 1\right\}$

For part 1, we show that $\mathcal{V}$ is closed under addition and multiplication. If $\overrightarrow{v} \in \mathcal{V}$ and $\overrightarrow{w} \in \mathcal{V}$, then $\overrightarrow{v} \cdot \overrightarrow{u} = 0$ and $\overrightarrow{w} \cdot \overrightarrow{u} = 0$. Hence, $(\overrightarrow{v} + \overrightarrow{w}) \cdot u = 0$. $\mathcal{V}$ is closed under addition. Now, for multiplication, this can be shown as follows: if $\overrightarrow{v} \in \mathcal{V} \Rightarrow \overrightarrow{u} \cdot \overrightarrow{v} = 0 \Rightarrow \overrightarrow{u} \cdot (t\overrightarrow{v}) \Rightarrow t(\overrightarrow{u} \cdot \overrightarrow{v}) = 0$. Hence, $t\overrightarrow{v} \in \mathcal{V}$.

For part 2, we see that the zero vector is not in $\mathcal{V}$, hence it is not a subspace. We could have also shown that $\left\{\langle a, b, c\rangle \in \mathbb{R}^3 : b = a + c + 1\right\}$ is not closed under addition. Choose $\overrightarrow{v} = \langle 0, 1, 0\rangle, \overrightarrow{u} = \langle 1, 2, 0\rangle$. Then, $\overrightarrow{u} + \overrightarrow{v} = \langle 1, 3, 0\rangle \notin \left\{\langle a, b, c\rangle \in \mathbb{R}^3 : b = a + c + 1\right\}$. We could have also shown that it's also not closed under multiplication. You can try this on your own.

§12 Basis and Dimension for a Subspace (Barsheshat)

§12.1 Basis for a subspace

Given a subspace $\mathcal{V} \subseteq \mathbb{R}^m$, we say that $B \subseteq \mathcal{V}$ is a **basis** for $\mathcal{V}$ if

1. $\text{span}(B) = \mathcal{V}$
2. B is linearly independent

The follwing are properties of bases

- If B is a basis for $\mathcal{V} \subseteq \mathbb{R}^m$, then $|B| \leq n$, where $|B|$ is the cadinality of B.
- **If B and B' are both bases for $\mathcal{V} \subseteq \mathbb{R}^m$, then** $|B| = |B'|$.
- Every subspace has a basis (*very hard to prove*)

Exercise 12.1. *What is the basis of $\mathcal{V} = \left\{\overrightarrow{0}\right\} \subseteq \mathbb{R}^m$?* We have $B = \{\}$, which is the empty set.

§12.2 Dimension of a subspace

Definition 12.2. Given a vector subspace $\mathcal{V} \subseteq \mathbb{R}^m$, the **dimension** of $\mathcal{V}$, $\dim(\mathcal{V})$, is defined to be the number of elements in any basis of $\mathcal{V}$.

- $\left\{\overrightarrow{0}\right\}$ has dimension 0
- $\text{span}\{\overrightarrow{v}\}$ for any dimension 1 ($\overrightarrow{v} \neq \overrightarrow{0}$)
- $\text{span}\{\overrightarrow{v}_1, \cdots, \overrightarrow{v}_n\}$ has dimension $\leq n$
- $\mathbb{R}^m$ has dimension m

Exercise 12.3. *Determine (using definition) whether each subset is a subspace. If possible, state the dimension and give the basis.*

1. $\mathcal{V} = \left\{\langle a, 0, 0\rangle \in \mathbb{R}^3 : a \in \mathbb{R}\right\}$
2. $\mathcal{V} = \left\{\langle a, 1, 1\rangle \in \mathbb{R}^3 : a \in \mathbb{R}\right\}$
3. $\mathcal{V} = \left\{\langle a, b, c\rangle \in \mathbb{R}^3 : b = a - c\right\}$

1. Yes. $\mathcal{V}$ is closed under addition and scalar multiplication. The basis for $\mathcal{V} : B = \{\langle 1, 0, 0\rangle\} \Rightarrow$ dimension 1
2. No, since when adding $\overrightarrow{u} = \langle 1, 1, 1\rangle$ and $\overrightarrow{v} = \langle 2, 1, 1\rangle$ (both in $\mathcal{V}$) we get $\overrightarrow{u} + \overrightarrow{v} = \langle 3, 2, 2\rangle \notin \mathcal{V}$ ($\mathcal{V}$ is not closed under addition)
3. Yes. Note the closure under addition: $\overrightarrow{u} = \langle a_1, b_1, c_1\rangle, \overrightarrow{v} = \langle a_2, b_2, c_2\rangle \Rightarrow \overrightarrow{u} + \overrightarrow{v} = \langle a_1 + a_2, b_1 + b_1, c_2 + c_2\rangle \Rightarrow b_1 + b_2 = (a_1 - c_1) + (a_2 - c_2) = (a_1 + a_2) - (c_1 - c_2) \Rightarrow \overrightarrow{u} + \overrightarrow{v} \in \mathbb{R}$. Also, note closure under scalar multiplication: if $\overrightarrow{v} \in \mathcal{V} \Rightarrow \overrightarrow{v} = \langle a, b, c\rangle$, where $b = a - c$. Then $t\overrightarrow{v} = \langle ta, tb, tc\rangle \Rightarrow tb = t(a - c) = ta - tc$. Thus, $t\overrightarrow{v} \in \mathcal{V}$. Furthermore, we have

$$B = \{\langle 1, 1, 0\rangle, \langle 0, -1, 1\rangle\} \Rightarrow \dim(\mathcal{V}) = |B| = 2. \tag{208}$$

Let's prove that B is a basis for $\mathcal{V}$. We need to show that

a) B is independent

b) $\text{span}(B) = \mathcal{V}$

We want to show that $a\langle 1,1,0\rangle + b\langle 0,-1,1\rangle = \overrightarrow{0}$ implies $a, b = 0$. The previous equation is equivalent to $a\overrightarrow{v}_1 = -b\overrightarrow{v}_1 \Rightarrow \overrightarrow{v}_1 = \left(\frac{-b}{a}\right)\overrightarrow{v}_2, a \neq 0$ which is an impossible situation. Clearly only possible if $b = 0$, since $\overrightarrow{v}_2$ is not a scalar multiple of $\overrightarrow{v}_1 \Rightarrow a, b = 0$. Hence B is independent. Next, we show $\text{span}(B) = \mathcal{V}$. We have span $= \{s\overrightarrow{v}_1 + t\overrightarrow{v}_2 : s, t \in \mathbb{R}\}$ and we need to show it is the same as $B = \{\langle 1,1,0\rangle, \langle 0,-1,1\rangle\}$.

$$s\overrightarrow{v}_1 + t\overrightarrow{v}_2 = \langle x, y, z\rangle \tag{209}$$

$$\langle s, s, 0\rangle + \langle 0, -t, t\rangle = \langle x, y, z\rangle \tag{210}$$

We thereby have the system

$$\begin{cases} x = s \\ y = s - t \Rightarrow y = x - z \\ z = t \end{cases} \tag{211}$$

Hence, $\langle x, y, z\rangle = \langle x, x - z, z\rangle$.

Exercise 12.4. Express $\langle 5, 7, -2\rangle$ as a linear combinations of the vectors in B. From our previous work

$$\langle 5, 7, -2\rangle = s\overrightarrow{v}_1 + t\overrightarrow{v}_2 = 5\overrightarrow{v}_1 - 2\overrightarrow{v}_2 \tag{212}$$

§13 Exercises on Span (Barsheshat)

Exercise 13.1. *Consider* $\mathcal{V} = \operatorname{span}\left(\left\{\overrightarrow{a} = \langle 1,0,1,0\rangle, \overrightarrow{b} = \langle 2,-1,1,1\rangle\right\}\right)$. $\mathcal{V}$ *is a 2-dimensional space (i.e., a plane that passes through the origin). Find the orthonormal basis* B *for* $\mathcal{V}$, $B = \{\overrightarrow{v}_1, \overrightarrow{v}_2\}$ *such that*

- $||\overrightarrow{v}_1|| = ||\overrightarrow{v}_2|| = 1$ *(normal)*
- $\overrightarrow{v}_1 \cdot \overrightarrow{v}_2 = 0$ *(orthogonal)*
- $\overrightarrow{v}_1$ *is in the direction of* $\langle 1,0,1,0\rangle$

We begin with $\overrightarrow{v}_1 = s\overrightarrow{a} = \langle s,0,s,0\rangle$ which gives $s^2 + s^2 = 1$ because $||\overrightarrow{v}_1|| = 1$. Hence $s = \pm\sqrt{2}/2$. Take

$$\overrightarrow{v}_1 = \langle \frac{\pm\sqrt{2}}{2}, 0, \frac{\pm\sqrt{2}}{2}, 0\rangle \tag{213}$$

For $\overrightarrow{v}_2$: $\overrightarrow{v}_1 \cdot \overrightarrow{v}_2 = 0$, $||\overrightarrow{v}_2||^2 = 1$ and $\overrightarrow{v}_2 \in \mathcal{V}$. Note that $||\overrightarrow{v}_2||^2 = 1$ is a non-linear equation. Let $\overrightarrow{v}_2 = \langle x,y,z,w\rangle$. We have:

$$\overrightarrow{v}_1 \cdot \overrightarrow{v}_2 = 0 \tag{214}$$

$$\frac{\sqrt{2}}{2}x + \frac{\sqrt{2}}{2}z = 0 \tag{215}$$

$$x + z = 0 \tag{216}$$

Furthermore, we have:

$$\overrightarrow{v}_2 \in \mathcal{V} \tag{217}$$

$$\overrightarrow{v}_2 = s\overrightarrow{a} + t\overrightarrow{b} \tag{218}$$

$$\langle x,y,z,w\rangle = \langle s,0,s,0\rangle + \langle 2t,-t,t,t\rangle \tag{219}$$

The rest of the solution is pretty trivial, you can try this yourself.

There's a second technique for finding $\overrightarrow{v}_2$ called Gram-Schmidt orthogonalization which works as follows for $\mathbb{R}^4$. We have $||\overrightarrow{v}_1|| = 1$. Consider $\operatorname{orth}_{\overrightarrow{v}_1}\overrightarrow{b}$ gives

$$\operatorname{orth}_{\overrightarrow{v}_1}\overrightarrow{b} = \overrightarrow{b} - \frac{\overrightarrow{b} \cdot \overrightarrow{v}_1}{\overrightarrow{v}_1 \cdot \overrightarrow{v}_1}\overrightarrow{v}_1 \tag{220}$$

$\operatorname{orth}_{\overrightarrow{v}_1}\overrightarrow{b}$ satisfies all conditions for $\overrightarrow{v}_2$ except for the norm, so take

$$\overrightarrow{v}_2 = \frac{\operatorname{orth}_{\overrightarrow{v}_1}\overrightarrow{b}}{\left\|\operatorname{orth}_{\overrightarrow{v}_1}\overrightarrow{b}\right\|} \tag{221}$$

Now, we have

$$B = \{\overrightarrow{v}_1, \overrightarrow{v}_2\} \tag{222}$$

with $||\overrightarrow{v}_1|| = ||\overrightarrow{v}_2|| = 1$, $\overrightarrow{v}_1 \cdot \overrightarrow{v}_2 = 0 \Rightarrow B$ is independent, but does $\operatorname{span}(B) = \mathcal{V}$? We know $\overrightarrow{v}_1, \overrightarrow{v}_2 \in \mathcal{V} \Rightarrow \operatorname{span}(\{\overrightarrow{v}_1, \overrightarrow{v}_2\}) \subseteq \mathcal{V} = \operatorname{span}\left(\left\{\overrightarrow{a}, \overrightarrow{b}\right\}\right)$. Finally, to show that $\operatorname{span}(B) = \mathcal{V}$, you need to show that $\overrightarrow{a} \in \operatorname{span}(\{\overrightarrow{v}_1, \overrightarrow{v}_2\})$ and $\overrightarrow{b} \in \operatorname{span}(\{\overrightarrow{v}_1, \overrightarrow{v}_2\})$ and $\mathcal{V} \subseteq \operatorname{span}(\{\overrightarrow{v}_1, \overrightarrow{v}_2\})$. The rest you can do.

Exercise 13.2. *Consider* $\mathcal{V} = \{\langle 1,2,0,0\rangle, \langle 2,0,1,1\rangle, \langle -1,-6,1,1\rangle, \langle 6,8,1,1\rangle\}$. Let's write $\overrightarrow{v}_1, \cdots, \overrightarrow{v}_4$ as columns in a 4×4 matrix.

$$A = \begin{pmatrix} 1 & 2 & -1 & 6 \\ 2 & 0 & -6 & 8 \\ 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & -3 & 4 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = R \tag{223}$$

We clearly see that $\overrightarrow{c}_3 = -3\overrightarrow{c}_1 + \overrightarrow{c}_2$ and $\overrightarrow{c}_4 = 4\overrightarrow{c}_1 + \overrightarrow{c}_2$. Hence, by a theorem which will be covered in the following classes,

$$\overrightarrow{v}_3 = -3\overrightarrow{v}_1 + \overrightarrow{v}_2, \overrightarrow{v}_4 = 4\overrightarrow{v}_1 + \overrightarrow{v}_1 \tag{224}$$

With these relations, we see that relations $\overrightarrow{v}_3$ and $\overrightarrow{v}_4$ can be seen as redundant, this $\mathcal{V} = \operatorname{span}(\{\overrightarrow{v}_1, \overrightarrow{v}_2\})$.

§14 Subspaces of Matrices (Barsheshat)

If A is an $m \times n$ matrix then A can be viewed as a collection of m row vectors $\in \mathbb{R}^n$ or n column vectors $\in \mathbb{R}^m$.

Definition 14.1. $\text{row}(A)$ = linear span of the row vectors $\subseteq \mathbb{R}^n$

Definition 14.2. $\text{col}(A)$ = linear span of the column vectors $\subseteq \mathbb{R}^m$

Moreover, note that

$$\text{rank}(A) = \dim(\text{row}(A)) = \dim(\text{col}(A)) \tag{225}$$

Theorem 14.3

If A and R are two matrices such that A can be transformed into R through elementary row operations, then $\text{row}(A) = \text{col}(A)$

Proof. INSERT PROOF □

Theorem 14.4

If there exists a dependency relation between the column vectors of a certain matrix A, then carrying out row operation does not change this dependency relation:

$$A = [\overrightarrow{c}_1 | \cdots | \overrightarrow{c}_2] \sim R = \left[\overrightarrow{d}_1 | \cdots | \overrightarrow{d}_2\right] \tag{226}$$

and

$$\sum_{i=1}^{n} a_i \overrightarrow{c}_i = \overrightarrow{0} \Leftrightarrow \sum_{i=1}^{n} a_i \overrightarrow{d}_i = \overrightarrow{0} \tag{227}$$

Proof. INSERT PROOF □

Theorem 14.5

Main theorem about bases *(basis theorem)*: If $\mathcal{S}$ is a subspace of $\mathbb{R}^n$, then all bases of $\mathcal{S}$ have the same size (i.e., the same number of vectors).

Proof. Suppose $B = \{\overrightarrow{v}_1, \cdots, \overrightarrow{v}_r\}$ and $C = \{\overrightarrow{u}_1, \cdots, \overrightarrow{u}_s\}$ are bases for $\mathcal{S}$:

$$\text{span}(B) = \mathcal{S} = \text{span}(C), \tag{228}$$

B and C are independent. Without loss of generality, assume $r < s$. We will show that the assumption $r < s$ forces the set C to be dependent. Consider

$$\sum_{i=1}^{s} c_i \overrightarrow{u}_i = \overrightarrow{0} \tag{229}$$

Since B and C span the same subspace, we can represent all vectors in C as the following linear combinations of vectors in B:

$$\overrightarrow{u}_1 = a_{11}\overrightarrow{v}_1 + a_{12}\overrightarrow{v}_1 + \cdots + a_{1r}\overrightarrow{v}_r \tag{230}$$
$$\overrightarrow{u}_2 = a_{21}\overrightarrow{v}_1 + a_{22}\overrightarrow{v}_1 + \cdots + a_{2r}\overrightarrow{v}_r \tag{231}$$
$$\overrightarrow{u}_3 = a_{31}\overrightarrow{v}_1 + a_{32}\overrightarrow{v}_1 + \cdots + a_{3r}\overrightarrow{v}_r \tag{232}$$
$$\vdots \tag{233}$$
$$\overrightarrow{u}_s = a_{s1}\overrightarrow{v}_1 + a_{s2}\overrightarrow{v}_1 + \cdots + a_{sr}\overrightarrow{v}_r \tag{234}$$

We can thereby plug these equations into the previous equation, giving:

$$c_1(a_{11}\overrightarrow{v}_1 + a_{12}\overrightarrow{v}_1 + \cdots + a_{1r}\overrightarrow{v}_r) + \cdots + c_s(a_{s1}\overrightarrow{v}_1 + a_{s2}\overrightarrow{v}_1 + \cdots + a_{sr}\overrightarrow{v}_r) = \overrightarrow{0} \tag{235}$$
$$(c_1a_{11} + \cdots + c_sa_{s1})\overrightarrow{v}_1 + \cdots + (c_1a_{1r} + \cdots + c_sa_{sr})\overrightarrow{v}_r = \overrightarrow{0} \tag{236}$$

Since B is independent (it is a basis), this means the above coefficients are all 0:

$$c_1a_{11} + \cdots + c_sa_{s1} = 0 \tag{237}$$
$$\vdots \tag{238}$$
$$c_1a_{1r} + \cdots + c_sa_{sr} = 0 \tag{239}$$

The previous set of equations is a homogenous system of equations with s variables and r equations. Since $s > r$, there must be at least 1 non-trivial solution to the system for $c_1, \cdots, c_s$. This means that C is dependent which is a contradiction. $\square$

§15 Matrix Operations (Barsheshat)

§15.1 Matrix addition and scalar multiplication

For an $m \times n$ matrix $A = [a_{ij}]$, an $m \times n$ $B = [b_{ij}]$ and $t \in \mathbb{R}$, then

$$A + B = [a_{ij} + b_{ij}] \tag{240}$$
$$tA = [ta_{ij}] \tag{241}$$

Subtraction is a combination of addition and scalar multiplication by -1.

§15.2 Matrix multiplication

Suppose A is a row matrix (i.e., $1 \times n$), $A = [a_1, a_2, \cdots, a_n]$ and B is column matrix (i.e., $n \times 1$)

$$B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}, \tag{242}$$

then AB is a 1×1 matrix equal to the dot product of A and B.

Example 15.1

Consider

$$A = [1, 2, 3, 4] \tag{243}$$
$$B = \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix}. \tag{244}$$

Then $AB = [1(1) + 2(-1) + 3(1) + 4(-1)] = [-2]$

More generally, if A is $m \times n$ and B is $n \times r$, then AB is of size $m \times r$ and described by

$$AB = [c_{ij}], \tag{245}$$

where c_{ij} is the dot product of the i'th row of A with the j'th column of B. In fact, if A is an $n \times m$ matrix and B is an $m \times p$ matrix, then their product is

$$\boxed{AB = [c_{ij}],\ \textbf{where } c_{ij} = \sum_{k=1}^{m} a_{ik} b_{kj}} \tag{246}$$

For instance, consider

$$A = [1, 2, 3, 4] \tag{247}$$
$$B = \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix}. \tag{248}$$

Then

$$BA = \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix} [1, 2, 3, 4] = \begin{pmatrix} 1(1) & 1(2) & 1(3) & 1(4) \\ -1(1) & -1(2) & -1(3) & -1(4) \\ 1(1) & 1(2) & 1(3) & 1(4) \\ -1(1) & -1(2) & -1(3) & -1(4) \end{pmatrix} \tag{249}$$

Not only is matrix multiplication not commutative, but often, flipping the order makes the operation undefined. Furthermore, if $AB = AC$, then B does not necessarily equal C. There is no "cancelation property " in matrix multiplication.
The identity is a matrix of size $n \times n$. Simply put, it is a square matrix with 1's along its main diagonal and 0's elsewhere.

Proposition 15.2

For any $n \times n$ matrix A

$$I_n A = A I_n = A \tag{250}$$

Proof. The rows (in order) of I_n are the same as the columns of I_n. When regarded as vector in $\mathbb{R}^n$, they are simply $\{\overrightarrow{e}_1, \overrightarrow{e}_2, \cdots, \overrightarrow{e}_n\}$. Furthermore, for any vectors $\overrightarrow{v} = \langle v_1, v_2, \cdots, v_n \rangle \in \mathbb{R}^n$, $\overrightarrow{v} \cdot \overrightarrow{e}_i = v_i$. If $A = [a_{ij}]$ and $I_n A = [c_{ij}]$, then

$$c_{ij} = \sum_{k=1}^{n} (\overrightarrow{e}_i)_k \cdot a_{kj} = a_{ij} \tag{251}$$

Similarly for AI_n. $\square$

§16 Column and Row Representations of Matrix Multiplication and the Properties of the Operation (Barsheshat)

§16.1 Partitioning matrices

It's often useful to regard matrices as being partitioned into sub-matrices. For instance,

$$M = \left(\begin{array}{ccc|cc} 1 & 0 & 2 & 1 & 1 \\ 2 & 0 & 1 & -1 & 2 \\ \hline 3 & 2 & 1 & 0 & 0 \\ 1 & 2 & 3 & 2 & 1 \end{array}\right) = \left(\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right), \tag{252}$$

where

$$A = \begin{pmatrix} 1 & 0 & 2 \\ 2 & 0 & 1 \end{pmatrix} \tag{253}$$

$$B = \begin{pmatrix} 1 & 1 \\ -1 & 2 \end{pmatrix} \tag{254}$$

$$C = \begin{pmatrix} 3 & 2 & 1 \\ 1 & 2 & 3 \end{pmatrix} \tag{255}$$

$$D = \begin{pmatrix} 0 & 0 \\ 2 & 1 \end{pmatrix} \tag{256}$$

§16.2 Column representation of matrix multiplication

Given A, an $m \times n$ matrix, and B, an $n \times r$ matrix, then consider partitioning B into its columns:

$$B = \left[\vec{b}_1, \vec{b}_2, \cdots, \vec{b}_r\right]. \tag{257}$$

Then

$$AB = A\left[\vec{b}_1, \vec{b}_2, \cdots, \vec{b}_r\right] = \left[A\vec{b}_1, A\vec{b}_2, \cdots, A\vec{b}_r\right] \tag{258}$$

§16.3 Row representation of matrix multiplication

Suppose A is $m \times n$, B is $n \times r$, and we partition A into rows

$$A = \begin{pmatrix} \vec{a}_1 \\ \vec{a}_2 \\ \vdots \\ \vec{a}_m \end{pmatrix}. \tag{259}$$

As a result, we have

$$AB = \begin{pmatrix} \vec{a}_1 B \\ \vec{a}_2 B \\ \vdots \\ \vec{a}_m B \end{pmatrix} \tag{260}$$

§16.4 Column-row representation

A is $m \times n$

$$A = [\vec{a}_1, \cdots, \vec{a}_n] \tag{261}$$

and B is $n \times r$

$$B = \begin{pmatrix} \overrightarrow{b}_1 \\ \overrightarrow{b}_2 \\ \vdots \\ \overrightarrow{b}_n \end{pmatrix} \tag{262}$$

We thereby have

$$AB = [\overrightarrow{a}_1, \cdots, \overrightarrow{a}_n] \begin{bmatrix} \overrightarrow{b}_1 \\ \overrightarrow{b}_2 \\ \vdots \\ \overrightarrow{b}_n \end{bmatrix} = \overrightarrow{a}_1 \overrightarrow{b}_1 + \overrightarrow{a}_2 \overrightarrow{b}_2 + \cdots + \overrightarrow{a}_n \overrightarrow{b}_n \tag{263}$$

Exercise 16.1. *If*

$$A = \begin{pmatrix} 1 & 2 & 2 \\ 0 & 1 & 1 \end{pmatrix}, B = \begin{pmatrix} 3 & 2 \\ 0 & 1 \\ 2 & 2 \end{pmatrix} \tag{264}$$

find AB using column-row representation. We have

$$\overrightarrow{a}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \tag{265}$$

$$\overrightarrow{a}_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \overrightarrow{a}_3 \tag{266}$$

$$\overrightarrow{b}_1 = [3, 2] \tag{267}$$

$$\overrightarrow{b}_2 = [0, 1] \tag{268}$$

$$\overrightarrow{b}_3 = [2, 2] \tag{269}$$

$$\tag{270}$$

Then

$$AB = \begin{pmatrix} 7 & 6 \\ 2 & 3 \end{pmatrix} \tag{271}$$

§16.5 Algebraic proterties of matrix addition and scalar multiplication

Assume A, B, C are matrices of same size, assume α and β are scalars $\in \mathbb{R}$, then

1. Commutativity: $A + B = B + A$
2. Associativity: $A + (B + C) = (A + B) + C$
3. The zero matrix is the "identity" of matrix addition: $A + 0 = A$
4. $A + (-A) = 0$
5. Distributivity: $\alpha(A + B) = \alpha A + \alpha B$ or $A(\alpha + \beta) = \alpha A + \beta A$.
6. Identity: $1A = A$
7. Scalar multiplication associativity: $\alpha(\beta A) = (\alpha\beta)A$

§16.6 Matrix exponentiation

When A and B are both $n \times n$ (square matrices), then AB is also $n \times n$. However, we have the special case $A = B$ and we then define $A^2 = AA$ (or more generally $A^n = \underbrace{AAA \cdots A}_{n}$).

Furthermore,

$$A^r A^n = A^{r+n} \tag{272}$$

$$(A^r)^s = A^{rs} \tag{273}$$

$$A^0 = I_n \tag{274}$$

§16.7 Transpose of matrix

If A is $m \times n$, then A^T is defined as the $n \times m$ matrix whose rows are the columns of A (in order) in coordinate notation:

$$(A^T)_{ij} = (A)_{ji} : \text{elements along the main diagonal are left unchanged} \tag{275}$$

The following are the properties of matrix transpose:

1. $(A^T)^T = A$
2. $(A + B)^T = A^T + B^T$
3. $(kA)^T = kA^T$, where $k \in \mathbb{R}$
4. $(AB)^T = B^T A^T$
5. For square matrices, $(A^r)^T = (A^T)^r$ where $r > 0$

Furthermore, note that a symmetric matrix has $A^T = A$ and a skew-symmetric (or anti-symmetric) matrix has $A^T = -A$

§16.8 Properties of matrix multiplication

Let $k \in \mathbb{R}$ and A, B, C be matrices whose sizes are such that the following operations are well defined. Then:

1. Associativity: $A(BC) = (AB)C$
2. Left-distributivity: $A(B + C) = AB + AC$
3. Right-distributivity: $(A + B)C = AC + BC$
4. $k(AB) = A(kB)$
5. $I_m A = A = AI_n$ where A is of size $m \times n$

Proof of 2. Let A_i be the i'th row of A and b_j and c_j are the j'th column of B and C, respectively. Then,

$$[A(B + C)]_{ij} = A_i \cdot (b_j + c_j) \tag{276}$$

$$= A_i \cdot b_j + A_i \cdot b_j \tag{277}$$

$$= [AB]_{ij} + [AC]_{ij} \tag{278}$$

$$= [AB + AC]_{ij} \tag{279}$$

□

§17 Linear Systems With Matrices (Barsheshat)

Suppose A is an $m \times n$ matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & & a_{2n} \\ \vdots & & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} = [a_{ij}]. \tag{280}$$

Let $\overrightarrow{x}$ be an arbitrary vector

$$\overrightarrow{v} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}. \tag{281}$$

Let $\overrightarrow{b}$ be

$$\overrightarrow{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}. \tag{282}$$

. Then the system of equations

$$\begin{cases} a_{11}x_1 + a_{12}x_1 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_1 + \cdots + a_{2n}x_n = b_2 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_1 + \cdots + a_{mn}x_n = b_m \end{cases} \tag{283}$$

can be represented by the equation $A\overrightarrow{x} = \overrightarrow{b}$, where A is $m \times n$, $\overrightarrow{x}$ is $n \times 1$ and therefore $\overrightarrow{b}$ is $m \times 1$. For a homogenous system, we can clearly write $A\overrightarrow{x} = \overrightarrow{0}$ (if A is $m \times n$, then $\overrightarrow{x} \in \mathbb{R}^n$ and $\overrightarrow{0} \in \mathbb{R}^m$).
Given an $m \times n$ matrix A, the nullspace of A, denoted null(A), is the set of all vectors $\overrightarrow{x} \in \mathbb{R}^n$ such that $A\overrightarrow{x} = 0$ (the set of all solutions of the homogeneous system):

$$\text{null}(A) = \left\{ \overrightarrow{x} \in \mathbb{R}^n : A\overrightarrow{x} = \overrightarrow{0} \right\}. \tag{284}$$

Note that if null(A) is a vector subspace, then it is

1. closed under addition
2. closed under scalar multiplication
3. not empty: $A\overrightarrow{0} = \overrightarrow{0}$

Proof. We begin by showing that null(A) is closed under addition. Suppose $\overrightarrow{u}, \overrightarrow{v} \in$ null(A). Thus

$$A(\overrightarrow{u} + \overrightarrow{v}) = A\overrightarrow{u} + A\overrightarrow{v} = \overrightarrow{0} + \overrightarrow{0} = \overrightarrow{0}. \tag{285}$$

Hence, $\overrightarrow{u} + \overrightarrow{v} \in$ null(A). Now, we show that null(A) is closed under scalar mutliplication. If $\overrightarrow{v} \in$ null(A) and $t \in \mathbb{R}$, then $A(t\overrightarrow{v}) = t(A\overrightarrow{v}) = t\overrightarrow{0} = \overrightarrow{0}$. Hence, $t\overrightarrow{v} \in$ null(A). □

Proposition 17.1

A basis for the column space of A is composed of the columns in A that are associated with pivot columns in RREF(A). On the other hand, a basis for the row space of A is comprised of the pivot rows of RREF(A).

Example 17.2

Let

$$A = \begin{pmatrix} 1 & 2 & 0 \\ 2 & -1 & 1 \\ 3 & 11 & -1 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 2/5 \\ 0 & 1 & -1/5 \\ 0 & 0 & 0 \end{pmatrix} \tag{286}$$

Hence, the basis for row(A) is $B = \{\langle 1, 0, 2/5\rangle, \langle 0, 1, -1/5\rangle\}$. Moreover, the basis for col(A) is $C = \{\langle 1, 2, 3\rangle, \langle 2, 1, 11\rangle\}$. Moreover the basis for the nullspace is

$$B = \left\{ \begin{pmatrix} -2/5 \\ 1/5 \\ 1 \end{pmatrix} \right\}. \tag{287}$$

As a result, null(A) is

$$\text{null}(A) = \text{span}\left(\left\{ \begin{pmatrix} -2/5 \\ 1/5 \\ 1 \end{pmatrix} \right\} \right) \tag{288}$$

Theorem 17.3

Rank-Nullity Theorem: The rank of a matrix A and it's nullity, which is the dimension of the nullspace of A, are thusly related:

$$\text{rank}(A) + \text{nullity}(A) = n \tag{289}$$

Proposition 17.4

If $\text{null}(A) = \left\{ \vec{0} \right\}$, then the column vectors of A $\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_n$, which span the column space of A

$$\text{col}(A) = \text{span}\left(\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_n\right), \tag{290}$$

are linearly independent and are the basis for the span.

Proposition 17.5

The reduced row echelon form R has the same null space as the original matrix A: $\text{null}(A) = \text{null}(R)$.

§18 Matrix Inverses (Barsheshat)

Definition 18.1. An $n \times n$ (square) matrix A is said to be invertible if there exists another matrix B such that $AB = BA = I$. B is usually labelled as $B = A^{-1}$ and called the inverse of A.

If $A = [a]$ is a 1×1 matrix, then A is invertible if $a \neq 0$. The inverse will be $A^{-1} = [a^{-1}]$.

Example 18.2

If

$$A = \begin{pmatrix} 1 & 2 \\ -1 & 1 \end{pmatrix}, \tag{291}$$

find A^{-1}. Let

$$A^{-1} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}. \tag{292}$$

Then,

$$\begin{pmatrix} 1 & 2 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{293}$$

As a result,

$$\begin{pmatrix} a+2c & b+2d \\ c-a & d-b \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{294}$$

The rest of the problem involves solving a system of equations, which you can do on your own. The answer, however, is

$$A^{-1} = \begin{pmatrix} 1/3 & -2/3 \\ 1/3 & 1/3 \end{pmatrix} \tag{295}$$

Theorem 18.3

If A is invertible and B is invertible, show that AB is invertible.

Proof. Multiply AB by $B^{-1}A^{-1}$:

$$\begin{aligned} (B^{-1}A^{-1})(AB) &= ((B^{-1}A^{-1})A)B && (296) \\ &= (B^{-1}(A^{-1}A))B && (297) \\ &= (B^{-1}I)B && (298) \\ &= B^{-1}B && (299) \\ &= I && (300) \end{aligned}$$

Hence AB is invertible and $(AB^{-1}) = B^{-1}A^{-1}$. □

Theorem 18.4

If A is invertible, then A^{-1} is unique.

Proof. Suppose B and B' are inverses of A. Then

$$AB = BA = A = AB' = B'A. \tag{301}$$

We begin with $B'A = I$,

$$\begin{aligned} B'A &= I && (302)\\ (B'A)B &= IB && (303)\\ B'(AB) &= IB && (304)\\ B'I &= IB && (305)\\ B' &= B, && (306) \end{aligned}$$

which thereby implies that A^{-1} is unique. $\square$

The following are properties of matrix inverses:

- $(AB)^{-1} = B^{-1}A^{-1}$
- $(kA)^{-1} = \frac{1}{k}A^{-1}$ for $k \neq 0$
- $(A^n)^{-1} = (A^{-1})^n$ for $n \in \mathbb{R}$
- $(A^T)^{-1} = (A^{-1})^T$
- The system $A\overrightarrow{x} = \overrightarrow{b}$ has exactly one solution ($\overrightarrow{x} = A^{-1}\overrightarrow{b}$).

§18.1 Using row-reduction to find inverses

Given a certain square matrix A, we can both check if A is invertible and find its inverse with the same procedure. We begin by writing out A and I side by side. For this example, let

$$A = \begin{pmatrix} 3 & 1 & 2 \\ 1 & 2 & -1 \\ 2 & -1 & 1 \end{pmatrix}. \tag{307}$$

We consequently have

$$[A|I] = \left(\begin{array}{ccc|ccc} 3 & 1 & 0 & 1 & 0 & 0 \\ 0 & 2 & -1 & 0 & 1 & 0 \\ 2 & -1 & 1 & 0 & 0 & 1 \end{array}\right) \tag{308}$$

Now, we row reduce until A has been transformed into I. If a row of zeros is obtained, then A is not invertible. In our example,

$$[A|I] = \left(\begin{array}{ccc|ccc} 3 & 1 & 0 & 1 & 0 & 0 \\ 0 & 2 & -1 & 0 & 1 & 0 \\ 2 & -1 & 1 & 0 & 0 & 1 \end{array}\right) \sim \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & -1 & -1 \\ 0 & 1 & 0 & -2 & 3 & 3 \\ 0 & 0 & 1 & -4 & 5 & 6 \end{array}\right) = [I|B]. \tag{309}$$

The matrix B we obtain is actually A^{-1}. You can check that $AB = BA = I$.

§19 Gauss-Jordian Algorithm for Finding Inverses (Barsheshat)

We must begin by discussing elementary matrices.

Definition 19.1. An elementary matrix is a matrix which corresponds to a row operation:

1. Row swap matrix: T_{ij} is by definition the matrix that swaps row i and row j.

$$T_{ij} = \begin{pmatrix} 1 & \cdots & \cdots & 0 \\ \vdots & \ddots & 1 & \vdots \\ \vdots & 1 & \ddots & \vdots \\ 0 & \cdots & \cdots & 1 \end{pmatrix} \tag{310}$$

Consider the 3×3 example

$$T_{12} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{311}$$

and the 5×5 example

$$T_{24} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \tag{312}$$

Note that if A has size $n \times n$, then $T_{ij}A$ is simply the matrix A with rows i and j swapped.

2. The second operation is multiplying a row by a scalar: $R_i \to mR_i$ where $m \neq 0$. $S_i(m)$ is the diagonal matrix with 1's along the diagonal expect for the i'th row which has an m.

$$S_i(m) \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & \ddots & & & 0 \\ \vdots & & m & & \vdots \\ 0 & & & \ddots & 0 \\ 0 & 0 & \cdots & 0 & 1 \end{pmatrix} \tag{313}$$

3. The final row operation is adding a scalar multiple of one row to another: $R_1 \to R_i + mR_j$, where $m \neq 0$.

$$E_{ij}(m) = \begin{pmatrix} 1 & & & & \\ & 1 & \cdots & m & \\ & & 1 & & \\ & & & \ddots & \\ & & & & 1 \end{pmatrix}, \tag{314}$$

where m is in the j'th position in the i'th row.

The following are properties of elementary matrices:

- If A is an $n \times n$ matrix, and E is any elementary matrix with same dimension as A, then EA is simply the matrix A with the corresponding row operation carried out.

- All elementary matrix are invertable:

$$T_{ij}^{-1} = T_{ij} \tag{315}$$
$$(S_i(m))^{-1} = S_i(1/m) \tag{316}$$
$$(E_{ij}(m))^{-1} = E_{ij}(-m) \tag{317}$$

Suppose I apply the row operations $E_1, E_2, \cdots E_k$ in that order to an $n \times n$ matrix A to obtain the identity matrix, then,

$$I_n = E_k E_{k-1} \cdots E_1 A. \tag{318}$$

Therefore, $E_k E_{k-1} \cdots E_1 = A^{-1}$. Also,

$$A = (A^{-1})^{-1} \tag{319}$$
$$= (E_k E_{k-1} \cdots E_1)^{-1} \tag{320}$$
$$= E_1^{-1} E_2^{-1} \cdots E_{k-1}^{-1} E_k^{-1}. \tag{321}$$

Theorem 19.2

Invertability theorem: Let A be an $n \times n$ matrix. Then the following statements are all equivalent:

1. A is invertable (A^{-1} exists)
2. A is a product of elementary matrices
3. $\text{rank}(A) = n$ (A has full rank)
4. A^T is invertable: $(A^T)^{-1} = (A^{-1})^T$
5. Columns (or) rows of A form a basis of $\mathbb{R}^n$.
6. $A\vec{x} = \vec{0}$ has only the trivial solution ($\vec{x} = \vec{0}$)
7. $A\vec{x} = \vec{b}$ has exactly one solution for each $\vec{b}$.

Proof. (1) $\Leftrightarrow$ (2): from what we saw with Guass-Jordan algorithm. A is invertable iff A is reduced to identity with row operations, i.e., $A = E_1 E_2 \cdots E_k$ (a product of elementary matrices).
(2) $\Leftrightarrow$ (3) if A is a product of elementary matrices, then $RREF(A) = I \Rightarrow$ all rows were independent $\Rightarrow \text{rank}(A) = n$.
(3) $\Leftrightarrow$ (4) straightforward since $\text{rank}(A) = \text{rank}(A^T)$.
(5) is clearly equivalent to (3)
(1) $\Leftrightarrow$ (6) if $A\vec{x} = \vec{0}$, and A is invertable, then

$$A^{-1}(A\vec{x}) = A^{-1}\vec{0} = \vec{0} \tag{322}$$
$$\Rightarrow (A^{-1}A)\vec{x} = \vec{x} = \vec{0} \tag{323}$$

(5) $\Leftrightarrow$ (7) First off, if $A\vec{x} = \vec{0}$ only has trivial solution, but suppose $A\vec{x} = \vec{b}$ (for some

$\overrightarrow{b}$) has more than one soultion, i.e., $\overrightarrow{x}_1 \neq \overrightarrow{x}_2$ but $A\overrightarrow{x}_1 = A\overrightarrow{x}_2 = \overrightarrow{b}$.

$$A\overrightarrow{x}_1 - A\overrightarrow{x}_2 = \overrightarrow{0} \tag{324}$$

$$A(\overrightarrow{x}_1 - \overrightarrow{x}_2) = \overrightarrow{0} \tag{325}$$

$\overrightarrow{x}_1 - \overrightarrow{x}_2$ is a solution to the homogenous system, thereby condraticting our original statement. □

§20 More on Matrix Inverses

Theorem 20.1

If A is an $m \times n$ matrix then

1. $\text{rank}(A^T A) = \text{rank}(A)$
2. $A^T A$ is invertible $\Leftrightarrow \text{rank}(A) = n$

Proof. By the rank-nullity theorem,

$$\text{rank}(A) + \text{mull}(A) = n = \text{rank}(A^T A) + \text{nullity}(A^T A). \tag{326}$$

Therefore we must show that $\text{nullity}(A) = \text{nullity}(A^T A)$. In fact, we will show an even stranger statement: $\text{null}(A) = \text{null}(A^T A)$. The first step is proving that

$$\text{nullity}(A) \subseteq \text{nullity}(A^T A) \tag{327}$$

Suppose $\overrightarrow{x} \in \text{null}(A)$, then $A\overrightarrow{x} = \overrightarrow{0}$. As a result,

$$A\overrightarrow{x} = \overrightarrow{0} \tag{328}$$

$$A^T(A\overrightarrow{x}) = A^T\overrightarrow{0} \tag{329}$$

$$(A^T A)\overrightarrow{x} = \overrightarrow{0} \tag{330}$$

Hence $\overrightarrow{x} \in \text{null}(A^T A)$. The second step involves showing that $\text{null}(A) \leq \text{null}(A^T A)$. Suppose $\overrightarrow{x} \in \text{null}(A^T A)$. Then,

$$(A^T A)\overrightarrow{x} = \overrightarrow{0} \tag{331}$$

$$\overrightarrow{x}^T(\overrightarrow{A^T A \overrightarrow{x}}) = 0 \tag{332}$$

$$(\overrightarrow{x}^T A^T) \cdot (A\overrightarrow{x}) = 0 \tag{333}$$

$$(A\overrightarrow{x})^T(A\overrightarrow{x}) = 0 \tag{334}$$

$$\Rightarrow ||A\overrightarrow{x}||^2 = 0 \tag{335}$$

$$||A\overrightarrow{x}|| = 0 \tag{336}$$

$$A\overrightarrow{x} = \overrightarrow{0} \tag{337}$$

Moreover, we now use the rank-nullity theorem in conjunction with the inveritbility theorem to show the second statement of the theorem: $A^T A$ is invertible $\Leftrightarrow \text{rank}(A^T) = n \Leftrightarrow \text{rank}(A) = n$.
We thereby obtain:

1. $\text{rank}(A^T A) = \text{rank}(A)$
2. $A^T A$ is invertible $\Leftrightarrow \text{rank}(A) = n$

. □

§21 Determinants (Barsheshat)

The determinant can be though of as a function whose domain (set of input) is the set of all square matrices, and whose range (set of output) is simply $\mathbb{R}$: $f : A_{m\times n} \to \mathbb{R}$.

Definition 21.1. If A is 2×2,

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc \tag{338}$$

Theorem 21.2

A (which is $n \times n$) is invertable iff $\det(A) \neq 0$

Proof. Let

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \tag{339}$$

and

$$\overrightarrow{u} = \begin{pmatrix} a \\ c \end{pmatrix}, \overrightarrow{v} = \begin{pmatrix} b \\ d \end{pmatrix} \tag{340}$$

Then,

$$\text{proj}_{\overrightarrow{u}} \overrightarrow{v} = \left(\frac{\overrightarrow{u} \cdot \overrightarrow{v}}{||\overrightarrow{u}||^2} \right) \overrightarrow{u} \tag{341}$$

$$= \left(\frac{ab + cd}{a^2 + c^2} \right) \langle a, c \rangle \tag{342}$$

Using this, we calculate the orthogonal

$$\text{orth}_{\overrightarrow{u}} = \langle a, c \rangle - \left(\frac{ab + cd}{a^2 + c^2} \right) \langle a, c \rangle \tag{343}$$

Now, you can easily verify that

$$||\overrightarrow{u}||||\text{ortho}_{\overrightarrow{u}} \overrightarrow{v}|| = |ad - bc| \tag{344}$$

□

We will derive a formula for the inverse of a square 2×2 matrix. Let

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \tag{345}$$

We apply the Gauss-Jordian algorithm (note that a or $c \neq 0$). Without loss of generality, assume for simplicity $a \neq 0$. Hence,

$$A = \left(\begin{array}{cc|cc} a & b & 1 & 0 \\ c & d & 0 & 1 \end{array}\right) \sim \left(\begin{array}{cc|cc} 1 & 0 & \frac{d}{ad-bc} & -\frac{b}{ad-bc} \\ 0 & 1 & -\frac{c}{ad-bc} & \frac{a}{ad-bc} \end{array}\right) \tag{346}$$

As a result,

$$A^{-1} = \frac{1}{\det(A)} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \tag{347}$$

§21.1 Determinants for 3×3 matrices

Suppose

$$A = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} \tag{348}$$

We give a recursive formula:

$$\det(A) = a\begin{vmatrix} e & f \\ h & i \end{vmatrix} - b\begin{vmatrix} d & f \\ g & i \end{vmatrix} + c\begin{vmatrix} d & e \\ g & h \end{vmatrix} \tag{349}$$

$$= a(ei - fh) - b(di - fg) + c(db - eg) \tag{350}$$

$$= aei - afh - bdi + bfg + cbd - ceg \tag{351}$$

$$= aei + bfg + cdh - afh - bdi - ceg \tag{352}$$

Notice the pattern is the same as that of the cross product:

$$\begin{matrix} a & b & c & a & b & c \\ d & e & f & d & e & f \\ g & h & i & g & h & i \end{matrix} \tag{353}$$

and

$$\begin{matrix} a & b & c & a & b & c \\ d & e & f & d & e & f \\ g & h & i & g & h & i \end{matrix} \tag{354}$$

As a result, $\det(A) = aei + bfg + cdh - afh - bdi - ceg$

Exercise 21.3. *Let $\overrightarrow{u} = \langle u_1, u_2, u_3 \rangle$, $\overrightarrow{v} = \langle v_1, v_2, v_3 \rangle$, $\overrightarrow{w} = \langle w_1, w_2, w_3 \rangle$. Now that the scalar tripple product $\overrightarrow{u} \cdot (\overrightarrow{v} \times \overrightarrow{w})$ is the determinant*

$$\begin{vmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix} \tag{355}$$

and is also equivalent to the volume of a parallelpiped. Show that this proves the equivalence between invertability and non-zero determinants for 3×3. This exercise is left to the reader.

Before continuing, here are some of the properties of determinants. Assuming A, B are $n \times n$ matrices, we have

1. $\det(A) = \det(A^T)$
2. $\det(A^{-1}) = \frac{1}{\det(A)}$
3. If A is equivalent to B through adding multiples of rows to other rows, then $\det(A) = \det(B)$
4. $\det(AB) = \det(A)\det(B)$

§22 More on Determinants (Barsheshat)

Definition 22.1. Minors: Given a matrix $A = [a_{ij}]$, A_{ij} is calld the (i, j)-minor of A, and it is simply the matrix obtained by removing the i-th row and the j-th column of A.

Back to 3×3 determinants:

$$\begin{aligned} \det(A) &= a_{11}\det(A_{11}) - a_{12}\det(A_{12}) + a_{13}\det(A_{13}) \\ &= \sum_{j=1}^{3}(-1)^{j+1}a_{1j}\det(A_{1j}) \quad (*) \end{aligned}$$

Definition 22.2. Cofactor: Given a square $(n \times n)$ matrix A, we can define the (i, j) - cofcctor of A, denoted C_{ij} as follows:

$$C_{ij} = (-1)^{i+j}\det(A_{ij}) \tag{356}$$

And $(*)$ becomes

$$\det(A) = \sum_{j=1}^{3} a_{1j}C_{1j} \tag{357}$$

Now on to $n \times n$ matrices: If A is $n \times n$, C_{ij} is the (i, j) cofactor of A, then $\boxed{\det(A) = \sum_{j=1}^{n} a_{1j} \cdot C_{1j}}$ (recursive definition of $n \times n$ determinant).

Example 22.3

$$A = \tag{358}$$

Find $\det(A)$ using above formula.

Solution. Since $a_{12} = a_{14} = 0$, we only need to find C_{11} and C_{13}:

$$C_{11} = (-1)^{1+1}\det(A_{11}) = 1\cdot \tag{359}$$

$\rightarrow \det(A) = a_{11} \cdot C_{11} + a_{12} \cdot C_{12} + a_{13} \cdot C_{13} + a_{14} \cdot C_{14} = 1 \cdot 1 = \boxed{1}$

§22.1 Generalized cofactor expansions

It turns out we can "expand" out determinant calcalation along *any row* or *any column* we want:

- If A is $n \times n$:

$$\begin{aligned} \det(A) &= \sum_{j=1}^{n} a_{1j}C_{1j} \quad \text{(cofactor expansion along first row)} \\ &= \sum_{j=1}^{n} a_{ij}Cij \quad \text{(cofactor expansion along i-th row)} \\ &= \sum_{i=1}^{n} a_{ij}C_{ij} \quad \text{(cofactor expansion along j-th column)} \end{aligned}$$

where C_{ij} is (i, j)-cofactor and a_{ij} is element in i-th row and j-th column.

Example 22.4

$$A = \tag{360}$$

Find determinant using any ow or column:

$\det(A) = a_{14} \cdot C_{14} + a_{24} \cdot C_{24} + a_{34} \cdot C_{34} + a_{44} \cdot C_{44}$

$$\begin{aligned}
\text{4-th column expansion} &= 1 \cdot \\
&= (-1) \\
&= (-1)
\end{aligned}$$

The following are propertities of determinants. Assuming all matrices are square, and all oeprations are well-defined, we have:

1. $\det(A) \neq 0$ iff A is invertible
2. $\det(A) = \det(A^T)$
3. $\det(I_n) = 1$
4. If A is $n \times n$, and $k \in \mathbb{R}$, $\det(kA) = k^n \det(A)$
5. $\det(\text{AB}) = \det(A) \cdot \det(B)$
6. $\det(A^{-1}) = \frac{1}{\det(A)}$

§23 Even More On Determinants (Barsheshat)

§23.1 Triangular Matrices

Definition 23.1. An upper triangular (square) matrix is a matrix which has only 0's below the main diagonal, e.g.,

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 2 & 7 \\ 0 & 0 & 1 \end{pmatrix} \tag{361}$$

Definition 23.2. A lower triangular (square) matrix is a matrix which has only 0's above the main diagonal, e.g.,

$$B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 2 & 3 \end{pmatrix} \tag{362}$$

Note that any diagonal matrix and the $n \times n$ 0 matrix are upper and lower triangluar.

Example 23.3

Compute determinants of the triangler matrices A and B defined above. We have:

$$\det(A) = \begin{vmatrix} 1 & 2 & 3 \\ 0 & 2 & 7 \\ 0 & 0 & 1 \end{vmatrix} = 1 \cdot \begin{vmatrix} 2 & 7 \\ 0 & 1 \end{vmatrix} = 2 \tag{363}$$

Moreover,

$$\det(B) = \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 2 & 3 \end{vmatrix} = 1 \cdot \begin{vmatrix} 1 & 0 \\ 2 & 3 \end{vmatrix} = 3 \tag{364}$$

Theorem 23.4

If A is upper or lower triangluar, then $\det(A)$ is simply the product of the elements along a main diagonal.

§23.2 Determinants of elements matrices

There are three types of elementary matrices:

1. *(Row Wwapping)* If E is an elementary matrix obtained by swapping two rows of I, then $\det(E) = -1$. For instance,

$$\begin{vmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{vmatrix} = -1 \tag{365}$$

2. *(Multiplication of a Row by k)* If E is obtained by multiplying a row of I by a scalar k, then $\det(E) = k$. For instance,

$$\begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & k \end{vmatrix} = 1 \cdot 1 \cdot k = k \tag{366}$$

3. *(Adding a Scalar Multiple of One Row to Another)* If E is obtained by adding a scalar multiple of one row to another, then $\det(E) = \det(I) = 1$. For instance,

$$\begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & m \\ 0 & 0 & 1 \end{vmatrix} = 1 \cdot 1 \cdot 1 = 1 \tag{367}$$

Exercise 23.5. *A quicker way to fin determinants:*

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 1 \\ -1 & 2 & 1 \end{pmatrix} \tag{368}$$

We have,

$$\det(A) = \begin{vmatrix} 1 & 2 & 3 \\ 2 & 1 & 1 \\ -1 & 2 & 1 \end{vmatrix} \tag{369}$$

$$= \begin{vmatrix} 1 & 2 & 3 \\ 0 & -3 & -2 \\ 0 & 4 & 4 \end{vmatrix} \tag{370}$$

$$= -\begin{vmatrix} 1 & 2 & 3 \\ 0 & 4 & 4 \\ 0 & -3 & -2 \end{vmatrix} \tag{371}$$

$$= -4\begin{vmatrix} 1 & 2 & 3 \\ 0 & 1 & 1 \\ 0 & -3 & -2 \end{vmatrix} \tag{372}$$

$$= -4\begin{vmatrix} 1 & 2 & 3 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{vmatrix} \tag{373}$$

$$= (-4) \cdot 1 \cdot 1 \cdot 1 \tag{374}$$

$$= -4 \tag{375}$$

Recap: To compute determinants, on ecan apply row operations, making sure to keep track of how determinants changes, until we obtain an upper or lower triangular matrix, then simply multiply along the diagonal (as well as any changes to the determinant).

§23.3 Cramer's Rule

Suppose A is $n \times n$ and invertible. Consider the system $A\overrightarrow{x} = \overrightarrow{b}$, which has the unique solution $\overrightarrow{x} = A^{-1}\overrightarrow{b}$.

Definition 23.6. Given A $(n \times n)$ and a column vector $\overrightarrow{b}$ $(n \times 1)$, then $A_i(\overrightarrow{b})$ is the $n \times n$ matrix where column i of A is replaced by $\overrightarrow{b}$.

Example 23.7

If

$$A = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 0 \\ 0 & 2 & 1 \end{pmatrix} \tag{376}$$

and

$$\overrightarrow{b} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \tag{377}$$

then

$$A_2(\overrightarrow{b}) = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 0 \\ 0 & 3 & 1 \end{pmatrix} \tag{378}$$

Cramer's Rule: If A is an invertible $n \times n$ matrix, then the solution to $A\overrightarrow{x} = \overrightarrow{b}$ is given by

$$\boxed{x_i = \frac{\det(A_i(\overrightarrow{b}))}{\det(A)}} \tag{379}$$

for $i = 1, \cdots, n$.

Exercise 23.8. *Solve* $A\overrightarrow{x} = \overrightarrow{b}$*, using Cramer's rule, given*

$$A = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 0 \\ 0 & 2 & 1 \end{pmatrix} \tag{380}$$

and

$$\overrightarrow{b} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \tag{381}$$

We have $\det(A) = 1$, $\det(A_1(\overrightarrow{b})) = -2$, $\det(A_2(\overrightarrow{b})) = 2$, $\det(A_3(\overrightarrow{b})) = -1$. Hence,

$$\overrightarrow{x} = \begin{pmatrix} -2 \\ 2 \\ -1 \end{pmatrix} \tag{382}$$

§24 Complex Numbers (Barsheshat)

Definition 24.1. A complex number, z, is an number of the form $\boxed{z = a + bi}$ (rectangular form) where $a, b \in \mathbb{R}$, and i is called the imaginary unit satisfying $i^2 = -1$.

The set of complex numbers is denoted $\mathbb{C}$:

$$\mathbb{C} = \left\{a + bi : a, b \in \mathbb{R}, i^2 = -1\right\}. \tag{383}$$

Given $z = a + bi$, a is called the real part and b is called the imaginary part:

$$a = \mathrm{Re}(z) \tag{384}$$
$$b = \mathrm{Im}(z) \tag{385}$$

Given any real number, say $x \in R$ we conventionally vew x as complex number as wll by associating: $x = x + 0i$. In this sense, we can view $\mathbb{R} \subseteq \mathbb{C}$. Hence,

$$\mathbb{R} = \{z \in \mathbb{C} : \mathrm{Im}(z) = 0\} \tag{386}$$

§24.1 Operations on complex numbers

Addition/Subtraction: Given $z_1 = a_1 + b_1 i$ and $z_2 = a_2 + b_2 i$, then we have:

$$z_1 + z_2 = (a_1 + a_2) + (b_1 + b_2)i \tag{387}$$
$$z_1 - z_2 = (a_1 - a_2) + (b_1 - b_2)i \tag{388}$$

With these definitions, the properties of addition and subtraction of real numbers (e.g., commutativity, associativity) carry over to complex numbers. Moreover, viewing z_1 and z_2 as *vectors* in the complex plane, then $z_1 + z_2$ corresponds to vector addition.
Multiplication: To define multiplication, we will assume that the “natural” properties (i.e., commutativity, associativity, distributivity) along with $i^2 = -1$. Consider $z_1 = a_1 + b_1 i$ and $z_2 = a_2 + b_2 i$ once more. Then

$$\begin{aligned} z_1 \cdot z_2 &= (a_1 + b_1 i) \cdot (a_2 + b_2 i) && (389)\\ &= a_1 \cdot a_2 + a_1(b_2 i) + (b_1 i)a_2 + (b_1 i)(b_2)i && (390)\\ &= a_1 \cdot a_2 + b_1 \cdot b_2(i^2) + (a_1 b_2 + a_2 b_1)i && (391)\\ &= (a_1 \cdot a_2 - b_1 \cdot b_2) + (a_1 b_2 + a_2 b_1)i && (392) \end{aligned}$$

Exercise 24.2. *Find the unique complex number such that*

$$z \cdot (3 + 2i) = 13 = 13 + 0i. \tag{393}$$

Let $z = a + bi$. Then,

$$(a + bi) \cdot (3 + 2i) = 13 + 0i \tag{394}$$
$$(3a - 2b) + (2a + 3b) = 13 + 0i. \tag{395}$$

Thus, we have the system

$$\begin{cases} 3a - 2b = 13 \\ 2a + 3b = 0 \end{cases} \tag{396}$$

Solving such a system is trivial at this point in the course. The answer is $a = 3$, $b = -2$: $z = 3a - 2b$.

Definition 24.3. Given a complex number $z = a + bi$, we define the conjugate of z, denoted z^* or $\overline{z}$, is defined as $\overline{z} = a - bi$, $\text{Re}(z) = \text{Re}(\overline{z})$, $\text{Im}(z) = -\text{Im}(\overline{z})$. The geometric interpretation of conjugation corresponds to a reflection with respect to the real axis.

Definition 24.4. The modulus, or magnitude, of a complez number $z = a + bi$, is denoted $|z|$ and defined as

$$|z| = \sqrt{a^2 + b^2}. \tag{397}$$

The following are properties of conjugation, given complex numbers z and w:

1. $\overline{(z + w)} = \overline{z} + \overline{w}$
2. $\overline{(z \cdot w)} = \overline{z} \cdot \overline{w}$
3. $\overline{(\overline{z})} = z$
4. $z \cdot \overline{z} = |z|^2$.
5. $|\overline{z}| = |z|$

Exercise 24.5. *Given a complex number $z = a_1 + b_1 i \neq 0 + 0i$, find the* <u>*unique*</u> *complex number w such that $z \cdot w = w \cdot z = 1 + 0i$.* Recall $z \cdot \overline{z} = |z|^2 > 0$.

$$\frac{z \cdot \overline{z}}{|z|^2} = 1, \;\; w = z^{-1} = \frac{\overline{z}}{|z|^2}. \tag{398}$$

<u>Division:</u> Given $w \in \mathbb{C}, z \in \mathbb{C}, z \neq 0 + 0i$, we define

$$\boxed{\frac{w}{z} = w \cdot z^{-1} = \frac{w \cdot \overline{z}}{|z|^2}} \tag{399}$$

Exercise 24.6. Given $z = 1 + i$,

1. Find z^{-1}
2. Find $\frac{w}{z}$, where $w = 3 + 2i$

We have,

$$z^{-1} = \frac{\overline{z}}{|z|^2} = \frac{1 - i}{1^2 + 1^2} = \frac{1}{2} - \frac{1}{2}i. \tag{400}$$

Moreover, we have

$$\frac{w}{z} = w \cdot z^{-1} = (3 + 2i)\left(\frac{1}{2} - \frac{1}{2}i\right) = (1/2)(5 - i) \tag{401}$$

§24.2 Polar form of copmlex numbers

Given $z = a + bi$, we can represent this as an arrow on the complex plane. We can take the angle θ (known as the argument of z: $\theta = \arg(z)$) in standard position. Hence,

$$a = |z| \cos\theta \tag{402}$$
$$b = |z| \sin\theta \tag{403}$$

We can consequently express z in terms of θ:

$$\begin{aligned} z &= a + bi && (404)\\ &= |z|\cos\theta + |z|i\sin\theta && (405)\\ &= |z|(\cos\theta + i\sin\theta) && (406)\\ &= |z|e^{i\theta} = |z|\operatorname{cis}\theta && (407) \end{aligned}$$

Converting from polar to rectangular form is straightforward given the relationships expressed about. To convert from rectangular to polar is trickier. Given $z = a + bi$, we have $|z| = \sqrt{a^2 + b^2}$. Moreover,

$$\tan\theta = \frac{b}{a}, \tag{408}$$

assuming $a, r \neq 0$. Then, use quadrants/signs of a and b to determine θ properly. If $a = 0$, then $\theta = \pi/2$ (if $b > 0$) or $\theta = 3\pi/2$ (if $b < 0$). If $r = 0$, then θ can be anything, but for simplicity we take $\theta = 0$ by convention.

§25 More on Complex Numbers (Barsheshat)

The product of two complex numbers, $z_1 = r_1(\cos\theta_1 + i\sin\theta_1)$ and $z_2 = r_1(\cos\theta_2 + i\sin\theta_2)$, expressed in polar form is

$$\begin{aligned} z_1 \cdot z_2 &= r_1 r_2(\cos\theta_1 + i\sin\theta_1)(\cos\theta_2 + i\sin\theta_2) && (409)\\ &= r_1 r_2(\cos\theta_1\cos\theta_2 + i\cos_1\sin\theta_2 + i\sin\theta_1\cos\theta_2 - \sin\theta_1\sin\theta_2) && (410)\\ &= r_1 r_2(\cos(\theta_1+\theta_2) + i\sin(\theta_1+\theta_2)) && (411)\end{aligned}$$

Moreover, the product $z_1 \cdot z_2^{-1}$ is

$$\begin{aligned} z_1 \cdot z_2^1 &= z_1 \cdot \frac{\overline{z_2}}{|z|^2} && (412)\\ &= r_1 r_2(\cos\theta_1 + i\sin\theta_2)\frac{(\cos\theta_2 - i\sin\theta_2)}{r_2^2(\cos^2\theta_2 + \sin^2\theta_2)} && (413)\\ &= \frac{r_1}{r_2}(\cos\theta_1\cos\theta_2 - i\cos\theta_1\sin\theta_2 + i\sin\theta_1\cos\theta_2 + \sin\theta_1\sin\theta_2) && (414)\\ &= \frac{r_1}{r_2}(\cos(\theta_1-\theta_2) + i\sin(\theta_1-\theta_2)) && (415)\end{aligned}$$

Exercise 25.1. Consider $f : \mathbb{C} \to \mathbb{C}$ where $f(z) = iz$. Describe the action of f. We know $i = \cos(\pi/2) + i\sin(\pi/2)$ and $z = \cos\theta + i\sin\theta$. Hence,

$$f(z) = (\cos(\pi/2) + i\sin(\pi/2))(\cos\theta + i\sin\theta) = r_1(\cos(\theta+\pi/2) + i\sin(\theta+\pi/2)) \quad (416)$$

f rotates z by $\pi/2$ counter-clockwise.

§25.1 De Moivre's formula

Consider z_1 and z_2. By setting $z_1 = z_2$, we find the powers of a single complex number:

$$\begin{aligned} z_1 \cdot z_2 &= r_1 r_2 \operatorname{cis}(\theta_1+\theta_2) && (417)\\ z^2 &= r^2 \operatorname{cis}(2\theta) && (418)\\ z^3 &= r^3 \operatorname{cis}(3\theta) && (419)\\ z^4 &= r^4 \operatorname{cis}(4\theta) && (420)\\ &\vdots && (421)\\ z^n &= r^n \operatorname{cis}(n\theta) && (422)\end{aligned}$$

We can notably use De Moivres formula to find the **roots of complex numbers**. In other words, if z^n is known, we could use this formula when to find the possible values of z. Suppose $z^n = r^n \operatorname{cis}\theta = r^n(\operatorname{cis}(\theta + k2n)), k \in NN$. Recall that $z^n = r^n \operatorname{cis}(n\theta)$, where $\theta \equiv$ argument. Going backwards, we take the n^{th} root of r (positive) and the angle divided by n, then

$$z^{1/n} = r^{1/n}\operatorname{cis}\left(\frac{\theta + 2k\pi}{n}\right), k = 0, 1, 2, \cdots, n-1 \quad (423)$$

and gives exactly n distinct roots.

§25.2 Basic polynomial of complex numbers

Recall that if $ax^2 + bx + c = 0$ for $a, b, c \in \mathbb{R}$, then

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}. \tag{424}$$

If we allow x to be complex, then we can adjust the formula:

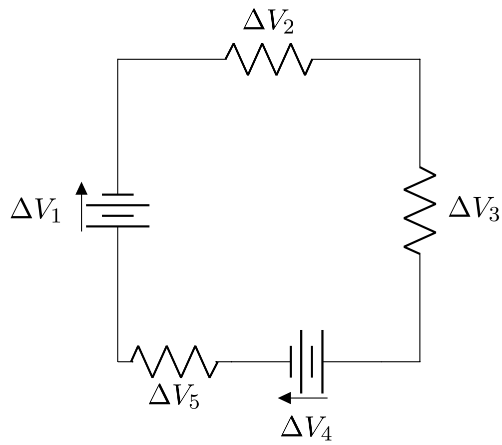
$$x = \frac{-b \pm \sqrt{\Delta}}{2a}, \Delta \geq 0 \tag{425}$$

$$x = \frac{-b \pm \sqrt{|\Delta| i}}{2a}, \Delta < 0 \tag{426}$$

§26 Integrative Activity: Electrical Circuits (Barsheshat)

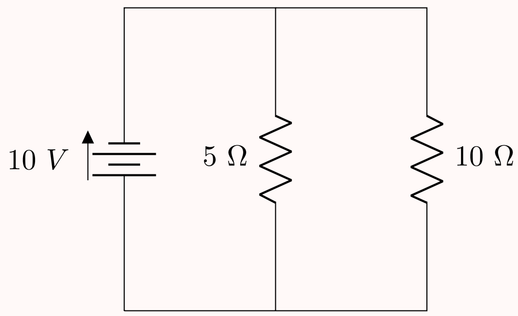
The following are the laws that govern electrical circuits. ΔV is the potential difference, I is the current and R is the resistance.

1. Ohm's law: $\Delta V = IR$
2. Kirchhoff's laws
   - Kirchhoff's current law (junction or node rule): The total current flowing into a node is euqal to the current flowing out. Assuming a + sign for current flowing in, - sign for current flowing out: $\sum I = 0$.
   - Kirchhoff's voltage law (loop rule): The sum of all voltage differences around a loop is 0, or $\sum \Delta V = 0$



Example 26.1

Solve for the currents in the following circuit.



Two loops constitue the circuit:

1. Starting at negative terminal of the battery: $\Delta V_1 = 10V, \Delta V_2 = -I_2 R_1 = -(5\ \Omega)I_2 \Rightarrow = 10V - (5\Omega)I_2 = 0 \Rightarrow I_2 = 2\ A$
2. We have $\Delta V_1 = -I_3 R_2 = -(10\ \Omega)I_3 \Rightarrow \Delta V_2 = (5\ \Omega)I_2$. As a result, $(-10\ \Omega)I_3 + (5\Omega)I_2 = 0 \Rightarrow (-10\ \Omega)I_3 + 10\ V = 0 \Rightarrow I_3 = 1\ A$

§27 Enriched Material: Eigenvalues and Eigenvectors (Barsheshat)

Definition 27.1. Given a square matrix $n \times n$ matrix A, an eigenvector/eigenvalue equation is: $\boxed{A\overrightarrow{v} = \lambda\overrightarrow{v}}$, $\overrightarrow{v} \in \mathbb{R}^n, \lambda \in \mathbb{R}$. In this case, $\overrightarrow{v}$ is an eigenvector of A associated with eigenvalue λ.

Remark 27.2. $\overrightarrow{0}$ is trivially an eigenvector of any square matrix A, associated with any eigenvalue λ ($A\overrightarrow{0} = \lambda\overrightarrow{0}$ is always true). Generally, when looking at eigenvectors and eigenvalues, we ignore $\overrightarrow{0}$ but not necessarily the 0 eigenvalue.

Remark 27.3. Any $\overrightarrow{v} \in \mathbb{R}^n$ is an eigenvector of I with eigenvalue 1 ($I\overrightarrow{v} = 1\overrightarrow{v} = \overrightarrow{v}$).

Definition 27.4. Given a square matrix A and $\lambda \in \mathbb{R}$, then $S_\lambda(A)$ is the **λ-eigenspace of** A is the set of all vectors which are eigenvectors of A associated with eigenvalues λ.

$$S_\lambda(A) = \{\overrightarrow{v} \in \mathbb{R}^n : A\overrightarrow{v} = \lambda\overrightarrow{v}\}. \tag{427}$$

Theorem 27.5
$S_\lambda(A)$ is a subspace of $\mathbb{R}^n$

Proof. If $\overrightarrow{v} \in S_\lambda(A)$:

$$\begin{aligned}
&\Leftrightarrow A\overrightarrow{v} = \lambda\overrightarrow{v} && (428)\\
&\Leftrightarrow A\overrightarrow{v} = (\lambda I)\overrightarrow{v} && (429)\\
&\Leftrightarrow A\overrightarrow{v} - \lambda I\overrightarrow{v} = \overrightarrow{0} && (430)\\
&\Leftrightarrow (A - \lambda I)\overrightarrow{v} = \overrightarrow{0} && (431)\\
&\Leftrightarrow \overrightarrow{v} \in \text{null}(A - \lambda I) && (432)
\end{aligned}$$

Thus, $S_\lambda(A) = \text{null}(A - \lambda I)$ and is therefore a subspace. □

Terminology: We say that λ is an eigenvalue of A is $\exists\overrightarrow{v} \neq \overrightarrow{0}$, such that $A\overrightarrow{v} = \lambda\overrightarrow{v}$.

§27.1 Finding eigenvalues of A

If $\overrightarrow{v} \neq \overrightarrow{0}$ and $A\overrightarrow{v} = \lambda\overrightarrow{v} \Leftrightarrow (A - \lambda I)\overrightarrow{v} = \overrightarrow{0} \Leftrightarrow \det(A - \lambda I) = 0$. The last expression is the characteristic polynomial in λ.

Exercise 27.6. Consider

$$P = \frac{1}{3}\begin{pmatrix} 2 & -1 & 1 \\ -1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix} \tag{433}$$

We have $P^T = P$ (P is symmetric), $P^2 = P$ (P is a projection matrix). *Verify that for any*

$$\overrightarrow{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{R}^3 \tag{434}$$

that $(P\overrightarrow{x}) = \overrightarrow{y}$ *is contained in the plane* $x + y - z = 0$*. By calculating* $\det(P - \lambda I)$ *for arbitrary* $\lambda \in \mathbb{R}$*, solve* $\det(P - \lambda I) = 0$ *to determine the eigenvalues of* P*.* We have

$$P - \lambda I = \begin{pmatrix} 2/3 & -1/3 & 1/3 \\ -1/3 & 2/3 & 1/3 \\ 1/3 & 1/3 & 2/3 \end{pmatrix} - \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix}$$

$$= \begin{pmatrix} (2/3 - \lambda) & -1/3 & 1/3 \\ -1/3 & (2/3 - \lambda) & 1/3 \\ 1/3 & 1/3 & (2/3 - \lambda) \end{pmatrix}$$

$$\det(P - \lambda I) = (2/3 - \lambda) \begin{vmatrix} (2/3 - \lambda) & 1/3 \\ 1/3 & (2/3 - \lambda) \end{vmatrix} + \frac{1}{3} \begin{vmatrix} -1/3 & 1/3 \\ 1/3 & (2/3 - \lambda) \end{vmatrix} + \frac{1}{3} \begin{vmatrix} -1/3 & (2/3 - \lambda) \\ 1/3 & 1/3 \end{vmatrix}$$

$$\begin{aligned} 0 &= (2/3 - \lambda)[(2/3 - \lambda)^2 - 9] + 1/3[-1/3(2/3 - \lambda) - 1/9] + 1/3[-1/9 - 1/3(2/3 - \lambda)] \\ &= (2/3 - \lambda)(1/3 - \lambda)(1 - \lambda) + 2/3[-1/3(2/3 - \lambda) - 1/9] \\ &= (2/3 - \lambda)(1/3 - \lambda)(1 - \lambda) - 2/9(1 - \lambda) \\ &= (1 - \lambda)\left[(2/3 - \lambda)(1/3 - \lambda) - 2/9\right] \\ &= (1 - \lambda)(1 - \lambda)\lambda \end{aligned}$$

Hence, $\lambda = \{0, 1\}$: projection matrices always have 0 and 1 as eigenvalues. Hence, $S_{\lambda_1}(P) = S_1(P) = \{\langle x, y, z\rangle : x + y - z = 0\}$. In other words, if $\overrightarrow{v}$ is on the plane, then $P\overrightarrow{v} = \overrightarrow{v}$. Moreover, we have $S_0(P) = \operatorname{span}\{\langle 1, 1, -1\rangle\}$, where $\langle 1, 1, -1\rangle$ is the normal vector of the plane.