

# MATH250 NOTES

# Contents

<b>1</b>	<b>Unit 1</b>	<b>2</b>
1.1	Lecture 1: Systems of Linear Equations [1.1-1.2]	2
1.2	Lecture 2: Vector and Matrix Equations [1.3-1.4]	5
1.3	Lecture 3: Homogenous Systems of Linear Equations [1.5]	8
1.4	Lecture 4: Span of a Set of Vectors [1.3]	11
1.5	Lecture 5: Linear Dependence and Independence [1.7]	15
1.6	Lecture 5: Linear Transformations [1.8]	17
1.7	Lecture 7: The Matrix of a Linear Transformation [1.9]	19
<b>2</b>	<b>Unit 2</b>	<b>24</b>
2.1	Lecture 8: Matrix Operations [2.1]	24
2.2	Lecture 9: Inverse of a Matrix and Characterizations of Invertible Matrices [2.2-2.3]	26
2.3	Lecture 10: Review for Midterm 1	31
<b>3</b>	<b>Unit 3</b>	<b>34</b>
3.1	Lecture 12: Introduction to Determinants [3.1]	34
3.2	Lecture 13: Properties of Determinants [3.1]	35
<b>4</b>	<b>Unit 4</b>	<b>38</b>
4.1	Lecture 14: Vector Spaces and Subspaces [2.8/4.1]	38
4.2	Lecture 15: Column, Row, and Null Spaces [2.9/4.2]	41
4.3	Lecture 16: Linearly Independent Sets, Basis [4.3]	46
4.4	Lecture 17: Coordinates in Vector Spaces [4.4]	49
4.5	Lecture 18: The Dimension of a Vector Space and Change of Basis [4.5-4.6]	50
<b>5</b>	<b>Unit 5</b>	<b>54</b>
5.1	Lecture 19: Eigenvalues and Eigenvectors	54
5.2	Lecture 20: The Characteristic Equation and Complex Eigenvalues [5.2, 5.5]	57
5.3	Lecture 22: Diagonalization [5.3]	58
<b>6</b>	<b>Unit 6</b>	<b>62</b>
6.1	Lecture 23: Orthogonality, Orthogonal Projections [6.1, 6.2, 6.3]	62
6.2	Lecture 24: The Gram-Schmidt Process [6.4]	64
6.3	Lecture 24: The Least-Squares Problem [6.5]	66

# 1 Unit 1

## 1.1 Lecture 1: Systems of Linear Equations [1.1-1.2]

A system of linear equations is when you have more than 1 equation with more than 1 unknown.

### Example 1.1.1

Solve the following system:

$$\begin{cases} x_1 + 2x_2 = 5 \\ 2x_1 + x_2 = 4 \end{cases}$$

**Solution:**

$$\begin{aligned} x_1 + 2x_2 = 5 &\Rightarrow x_1 = 5 - 2x_2 \\ 2x_1 + x_2 = 4 &\Rightarrow 2(5 - 2x_2) + x_2 = 4 \Rightarrow 10 - 3x_2 = 4 \\ &\Rightarrow x_2 = 2 \\ &\Rightarrow x_1 = 1 \end{aligned}$$

Note that the solution to the previous system  $(x_1, x_2) = (1, 2)$  also corresponds to the point of intersection of the lines that each equation represents. This then implies that non-parallel lines have a single solution, parallel lines have no solutions, and scalar multiples of the same line have infinitely many solutions.

### Example 1.1.2: Systems with Different Number of Solutions

1. One Solution:

$$\begin{cases} x_1 + 2x_2 = 5 \\ 2x_1 + x_2 = 4 \end{cases}$$

2. No solution:

$$\begin{cases} x_1 + 2x_2 = 5 \\ 2x_1 + 4x_2 = 6 \end{cases}$$

3. Infinite Solutions

$$\begin{cases} x_1 + 2x_2 = 5 \\ 2x_1 + 4x_2 = 10 \end{cases}$$

### Elementary Row Operations

Consider the system

$$\begin{cases} E_1 \\ E_2 \\ \vdots \\ E_n \end{cases}$$

Where  $E_i$  is the  $i$ -th equation. Then the following operations on the system **do not change the solution**

1. Changing the order of equations:  $E_i \leftrightarrow E_j$  where  $i \neq j$  for the  $i$ -th and  $j$ -th equation
2. Scaling equations:  $E_i \rightarrow \lambda E_i$  where  $\lambda \in \mathbb{R} (\lambda \neq 0)$
3. Combining equations:  $E_i \rightarrow E_i + \lambda E_j$  where  $i \neq j$  and  $\lambda \in \mathbb{R} (\lambda \neq 0)$

**Example 1.1.3**

Solve the following system:

$$\begin{cases} E_1 &= x_1 + 2x_2 = 5 \\ E_2 &= 2x_1 + x_2 = 4 \end{cases}$$

**Solution:**

$$\begin{aligned} -2E_1 &= -2x_1 - 4x_2 = -10 \\ E_2 - 2E_1 &= (2x_1 - x_2) - 2x_1 - 4x_2 = 4 - 10 \\ &\Rightarrow -3x_2 = -6 \\ &\Rightarrow x_2 = 2 \\ &\Rightarrow x_1 = 1 \end{aligned}$$

**Matrices**

A **matrix** is a rectangular array of numbers arranged in rows and columns. A matrix  $A$  has size  $m \times n$  if it has  $m$  **rows** and  $n$  **columns**. A matrix is written as

$$A_{m \times n} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

Where each  $a_{ij}$  is called an element of the matrix.  $i$  denotes the row and  $n$  denotes the column of the element.

With this in mind, note that we can represent the previous system as a matrix:

$$\begin{pmatrix} 1 & 2 & 5 \\ 2 & 1 & 4 \end{pmatrix}$$

where each row represents an equation and each element corresponds to either the coefficient of a variable or the solution. With this system in matrix form we can manipulate the rows as follows:

$$\begin{pmatrix} 1 & 2 & 5 \\ 2 & 1 & 4 \end{pmatrix} \xrightarrow{R_2 \rightarrow R_2 - 2R_1} \begin{pmatrix} 1 & 2 & 5 \\ 0 & -3 & -6 \end{pmatrix}$$

Since each row corresponds to an equation, if we take a look at the bottom row we can see

$$-3x_2 = -6 \Rightarrow x_2 = 2 \Rightarrow x_1 = 1$$

Thus allowing us to reach the same final answer with a lot less hassle.

## General Form of a System of Linear Equations

Consider any system of linear equations in the form

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ \vdots & \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{cases}$$

Now, notice we can rewrite this as a matrix:

$$A = \left( \begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array} \right)$$

Everything to the **left** of the solid line is referred to as the *coefficient matrix*.  $A$  itself is referred to as the **augmented matrix** for the system.

## Row Echelon Form

Given a matrix  $A$ ,  $A_{\text{REF}}$  is an equivalent matrix that satisfies the following properties:

1. All zero rows are below non-zero rows
2. each next leading element is in the column to the right of the previous leading element (called pivots)

Note that the **leading element** of a row is simply the first non-zero element in that row.

A matrix can also be put in RREF (reduced row echelon form) if it is already in REF, each pivot is 1, and the only non-zero element in the pivot column is the pivot. This would be  $A_{\text{RREF}}$ . Now, let's try to combine all we've done by applying basic ERO to a simple matrix to put it in REF.

## Example 1.1.4: Basic REF

Consider the system

$$\begin{cases} x_1 + 2x_2 &= 5 \\ 2x_1 + x_2 &= 4 \end{cases}$$

Find  $(A \mid B)$

**Solution:** It's augmented matrix can be put into REF as follows

$$(A \mid B) = \left( \begin{array}{cc|c} 1 & 2 & 5 \\ 2 & 1 & 4 \end{array} \right) \xrightarrow{R_2 \rightarrow R_2 - 2R_1} \left( \begin{array}{cc|c} 1 & 2 & 5 \\ 0 & -3 & -6 \end{array} \right) = (A \mid B)_{\text{REF}}$$

**Example 1.1.5: Basic RREF**

Consider again the system from ex (1.1.4) and find the RREF of the augmented matrix.

**Solution:** Notice we can put it into RREF as follows:

$$(A \mid B)_{\text{REF}} = \left( \begin{array}{cc|c} 1 & 2 & 5 \\ 0 & -3 & -6 \end{array} \right) \xrightarrow{R_2 \rightarrow -\frac{1}{3}R_2} \left( \begin{array}{cc|c} 1 & 2 & 5 \\ 0 & 1 & 2 \end{array} \right) \xrightarrow{R_1 \rightarrow R_1 - 2R_2} \left( \begin{array}{cc|c} 1 & 0 & 1 \\ 0 & 1 & 2 \end{array} \right) = (A \mid B)_{\text{RREF}}$$

It's important to note that the RREF for any given matrix is *unique*.

**Example 1.1.6: Solve the System**

Solve the system

$$\begin{cases} x_1 + 2x_2 - x_3 = 2 \\ 2x_1 + x_2 + x_3 = 1 \\ x_1 - x_2 + 2x_3 = -1 \end{cases}$$

**Solution:**

$$\begin{cases} x_1 + 2x_2 - x_3 = 2 \\ 2x_1 + x_2 + x_3 = 1 \\ x_1 - x_2 + 2x_3 = -1 \end{cases} = \left( \begin{array}{ccc|c} 1 & 2 & -1 & 2 \\ 2 & 1 & 1 & 1 \\ 1 & -1 & 2 & -1 \end{array} \right) \xrightarrow{\substack{R_2 \rightarrow R_2 - 2R_1 \\ R_3 \rightarrow R_3 - R_1}} \left( \begin{array}{ccc|c} 1 & 2 & -1 & 2 \\ 0 & -3 & 3 & -3 \\ 0 & -3 & 3 & -3 \end{array} \right) \\ \xrightarrow{R_3 \rightarrow R_3 - R_2} \left( \begin{array}{ccc|c} 1 & 2 & -1 & 2 \\ 0 & -3 & 3 & -3 \\ 0 & 0 & 0 & 0 \end{array} \right)$$

Now because there is no pivot in the third column of our matrix,  $x_3$  is what's known as a 'free variable'. All this means is that  $x_3$  can take on any value, we'll notate this by  $x_3 = t$  with  $t \in \mathbb{R}$ .

$$\Rightarrow \begin{cases} x_3 = t \\ -3x_2 + 3x_3 = -3 \\ x_1 + 2x_2 - x_3 = 2 \end{cases} \\ \Rightarrow X_{\text{gen}} = \begin{pmatrix} -t \\ 1+t \\ t \end{pmatrix}$$

**1.2 Lecture 2: Vector and Matrix Equations [1.3-1.4]**

As a reminder from last time, given a system with an augmented matrix  $(A \mid B)$  the system will have

1. no solutions if there is a pivot in the last column of  $(A \mid B)_{\text{REF}}$
2. one solution if  $(A \mid B)_{\text{REF}}$  is  $m \times n$  with  $n$  pivots (and no pivots in the last column)

3. infinite solutions if  $(A \mid B)_{\text{REF}}$  is  $m \times n$  with less than  $n$  pivots (and no pivots in the last column)

### Vectors

For a vector  $\vec{AB}$ , point  $A$  is the tail and point  $B$  is the head. Two vectors,  $\vec{AB}$  and  $\vec{CD}$  are equal if, and only if, their magnitude and directions are equal. Vectors in  $\mathbb{R}^2$  can be notated as  $\begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$  where the point  $(a_1, a_2)$  is the head and the tail is (usually) assumed to the origin. Generally, for  $\mathbb{R}^n$  we have  $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$  which will always have size  $n \times 1$ .

Since we can represent vectors as matrices, we can perform the standard matrix operations on them:

### Matrix Operations

Let  $A$  and  $B$  be matrices. Then

1.  $A + B = C$  where  $A$  and  $B$  are the same size. This addition is defined as  $c_{ij} = a_{ij} + b_{ij}$
2.  $\lambda A = \hat{A}$  where  $\hat{a}_{ij} = \lambda a_{ij}$  for  $\lambda \in \mathbb{R}$
3.  $A + (-A) = \hat{0}$  where  $\hat{0}$  is the zero-matrix which has the same size as  $A$  with each element being 0
4. Consider matrices  $A_{m \times n}$  and  $X_{n \times 1}$ , then  $AX$  is defined as

$$A_{m \times 1} = \begin{pmatrix} a_{11}x_1 & a_{12}x_2 & \dots & a_{1n}x_n \\ a_{21}x_1 & a_{22}x_2 & \dots & a_{2n}x_n \\ \vdots & & \ddots & \vdots \\ a_{m1}x_1 & a_{m2}x_2 & \dots & a_{mn}x_n \end{pmatrix} \text{ if } A_{m \times n} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \text{ and } X_{n \times 1} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

### Example 1.2.1: Basic Matrix Multiplication

Consider the the matrices

$$A_{2 \times 3} = \begin{pmatrix} 1 & -1 & 2 \\ 3 & 4 & 5 \end{pmatrix} \text{ and } B_{3 \times 1} = \begin{pmatrix} 2 \\ 4 \\ 7 \end{pmatrix}$$

Find  $AB$ .

**Solution:**

$$AB_{2 \times 1} = \begin{pmatrix} (1 * 2) + (-1 * 4) + (2 * 7) \\ (3 * 2) + (4 * 4) + (5 * 7) \end{pmatrix} = \begin{pmatrix} 12 \\ 57 \end{pmatrix}$$

It's important to note that an alternative form of defining matrix multiplication is as follows:

$$AX = x_1 \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} + x_2 \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix} + \dots + x_n \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix}$$

This is known as the **linear combination of vectors**.

## Linear Combination of Vectors

Let  $\{\vec{u}_1, \vec{u}_2, \dots, \vec{u}_k\}$  be a set of vectors from  $\mathbb{R}^n$ . Then,

$$c_1\vec{u}_1 + c_2\vec{u}_2 + \dots + c_k\vec{u}_k, \quad (c_1, c_2, \dots, c_k \in \mathbb{R})$$

is called the **linear combination of vectors**  $\vec{u}_1, \vec{u}_2, \dots, \vec{u}_k$ .

We can combine the previous two definitions of matrix multiplication and the linear combination of vectors to get this next fact: if we consider a vector

$$(\vec{u}_k) = \begin{pmatrix} u_{1k} \\ u_{2k} \\ \vdots \\ u_{nk} \end{pmatrix}$$

then,  $c_1\vec{u}_1 + c_2\vec{u}_2 + \dots + c_k\vec{u}_k = AX$  where

$$A = ((\vec{u}_1) \quad (\vec{u}_2) \quad \dots \quad (\vec{u}_k)) \text{ and } X = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_k \end{pmatrix}$$

This is especially important when working with systems of equations as we can represent systems as linear combinations of vectors. Given any general system we can rewrite it as a linear combination of vectors as such

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n &= b_1 \\ \vdots & \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n &= b_1 \end{cases} \Leftrightarrow AX = B$$

$B$  is said to be a *linear combination of columns of  $A$*  if, and only if,  $A$  and  $B$  are **compatible**. For  $A$  and  $B$  to be compatible essentially just means that  $(A \mid B)_{\text{REF}}$  has no pivots in the last column.

## Span

$c_1\vec{v}_1 + c_2\vec{v}_2 + \dots + c_k\vec{v}_k = \text{Span}\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k\}$  where  $c_1, c_2, \dots, c_k$  are all possible numbers

For a single vector  $\vec{v}$ ,  $\text{Span}\{\vec{v}\}$  is simply the set containing all scaled multiples of  $\vec{v}$ .

## Example 1.2.2: Span of Two Vectors

Notice that  $\text{Span}\left\{\begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 3 \\ 4 \end{pmatrix}\right\} = \mathbb{R}^2$ . This implies that any vector from  $\mathbb{R}^2$  can be written as

$$c_1 \begin{pmatrix} 1 \\ 2 \end{pmatrix} + c_2 \begin{pmatrix} 3 \\ 4 \end{pmatrix}$$

We can prove this fact by considering the augmented matrix for  $AX = B$ :

$$\left( \begin{array}{cc|c} 1 & 3 & b_1 \\ 2 & 4 & b_2 \end{array} \right) \xrightarrow{R_2 \rightarrow R_2 - 2R_1} \left( \begin{array}{cc|c} 1 & 3 & b_1 \\ 0 & -2 & b_2 - 2b_1 \end{array} \right)$$

Therefore, since there is no pivote in the last column, this system has a single solution for any vector



$$B = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$$

### Example 1.2.3

Given the following vectors  $\vec{a}_1, \vec{a}_2, \vec{a}_3$  and  $\vec{b}$ , determine if  $\vec{b}$  is a linear combination of  $\vec{a}_1, \vec{a}_2, \vec{a}_3$ .

$$\vec{a}_1 = \begin{pmatrix} 1 \\ -2 \\ 0 \end{pmatrix} \quad \vec{a}_2 = \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix} \quad \vec{a}_3 = \begin{pmatrix} 5 \\ -6 \\ 8 \end{pmatrix} \quad \vec{b} = \begin{pmatrix} 2 \\ -1 \\ 6 \end{pmatrix}$$

We can start by noticing that  $\vec{b} = c_1\vec{a}_1 + c_2\vec{a}_2 + c_3\vec{a}_3 \Leftrightarrow A \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$  where

$A = ((\vec{a}_1) \ (\vec{a}_2) \ (\vec{a}_3))$ . Thus,

$$\left( \begin{array}{ccc|c} 1 & 0 & 5 & 2 \\ -2 & 1 & -6 & -1 \\ 0 & 2 & 8 & 6 \end{array} \right) \xrightarrow{R_2 \rightarrow R_2 + 2R_1} \left( \begin{array}{ccc|c} 1 & 0 & 5 & 2 \\ 0 & 1 & 4 & 3 \\ 0 & 2 & 8 & 6 \end{array} \right) \xrightarrow{R_3 \rightarrow R_3 - 2R_2} \left( \begin{array}{ccc|c} 1 & 0 & 5 & 2 \\ 0 & 1 & 4 & 3 \\ 0 & 0 & 0 & 0 \end{array} \right)$$

Therefore, since there is no pivot in the final column,  $B$  is a linear combination of  $A$ .

## 1.3 Lecture 3: Homogenous Systems of Linear Equations [1.5]

### Homogenous Systems of Linear Equations

A **homogenous system of linear equations** is any system in the form

$$Ax = \hat{0}$$

Thus, any system in the form

$$Ax = B$$

is called a **non-homogenous system**.

For a non-homogenous system we can follow the steps

$$Ax = B \Rightarrow (A \mid B) \xrightarrow{\text{ERO}} (A \mid B)_{\text{REF}} \Rightarrow \text{backwards substitution into the equations}$$

For each non-homogenous system:

1. if the augmented matrix has a pivot in the last column, it has no solutions
2. if  $A$  is  $m \times n$  and the augmented matrix has less than  $n$  pivots, then  $Ax = B$  has an infinite amount of solutions
3. if  $A$  is  $m \times n$  and the augmented matrix has exactly  $n$  pivots then  $Ax = B$  has one solution

For a homogenous system we can follow similar steps:

$$Ax = \hat{0} \Rightarrow (A \mid 0) \xrightarrow{\text{ERO}} A_{\text{REF}} \Rightarrow \text{backwards substitution to the equations}$$

Note that since the solution matrix is simply just the zero matrix, we don't need to consider it in our augmented matrix - that's why you see  $A_{\text{REF}}$ . Unlike non-homogenous systems there are only two options for solutions:

1.  $x = \hat{0}$  is always a solution, this is called the *trivial solution*. If  $A_{\text{REF}}$  is  $m \times n$  large and has  $n$  pivots, then  $x = \hat{0}$  is the **only** solution
2. if  $A_{\text{REF}}$  is  $m \times n$  and has less than  $n$  pivots, then  $Ax = \hat{0}$  has an **infinite** amount of solutions

### Example 1.3.1: Solve the following System

Solve the below system

$$\begin{cases} x_1 + 2x_2 - 3x_3 + x_4 = 0 \\ 2x_2 - x_2 + x_3 - 2x_4 = 0 \\ 3x_1 + x_2 - x_3 + x_4 = 0 \end{cases}$$

We can start by putting this in the form  $Ax = \hat{0}$  where

$$A_{3 \times 4} = \begin{pmatrix} 1 & 2 & -3 & -1 \\ 2 & -1 & 1 & -2 \\ 3 & 1 & -1 & 1 \end{pmatrix}, \quad X_{4 \times 1} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}$$

Next, we need to put  $A$  into REF:

$$A_{3 \times 4} \xrightarrow[\substack{R_2 \rightarrow R_2 - 2R_1 \\ R_3 \rightarrow R_3 - 3R_1}]{\quad} \begin{pmatrix} 1 & 2 & -3 & -1 \\ 0 & -5 & 7 & -4 \\ 0 & -5 & 8 & -2 \end{pmatrix} \xrightarrow{R_3 \rightarrow R_3 - R_2} \begin{pmatrix} 1 & 2 & -3 & -1 \\ 0 & -5 & 7 & -4 \\ 0 & 0 & 1 & 2 \end{pmatrix} = A_{\text{REF}}$$

Since there is no pivot in the fourth column,  $x_4$  must be a free variable. We can rename it as  $x_4 = t$  for  $t \in \mathbb{R}$ . Thus

$$\begin{cases} x_3 + 2x_4 = 0 \Rightarrow x_3 = -2t \\ -5x_2 + 7x_3 - 4x_4 = 0 \Rightarrow x_2 = -\frac{18}{5}t \\ x_1 + 2x_2 - 3x_3 + x_4 = 0 \Rightarrow x_1 = \frac{1}{5}t \end{cases} \Rightarrow X = \begin{pmatrix} \frac{1}{5}t \\ -\frac{18}{5}t \\ -2t \\ t \end{pmatrix} \Leftrightarrow \begin{pmatrix} 1 \\ -18 \\ -10 \\ 5 \end{pmatrix} t$$

Notice that since  $t$  is any arbitrary value we can make  $t = 5t$  in order to eliminate the fractions. As well, the second matrix with  $t$  on the outside is known as the *parametric form* of the solution.

**Theorem 1.1.** *The general solution of a consistent non-homogenous system  $Ax = B$  has the form  $X_{\text{gen}} = X_{\text{genhom}} + X_p$  where  $X_{\text{genhom}}$  is the general solution of the corresponding homogenous system  $Ax = \hat{0}$  and  $X_p$  is any particular solution of  $Ax = B$ .*

*Proof.* Let  $X_1$  and  $X_2$  be any two solutions of  $Ax = B$ . That is,

$$\begin{cases} AX_1 = B \\ AX_2 = B \end{cases} \Rightarrow AX_1 - AX_2 = 0 \Leftrightarrow A(X_1 - X_2) = 0$$

Therefore,  $X_1 - X_2$  is a solution of the homogenous system. With some clever ‘renaming’ we can see:

$$X_1 - X_2 = X_{\text{hom}} \Rightarrow X_1 = X_{\text{hom}} + X_2 \Rightarrow X_{\text{gen}} = X_{\text{genhom}} + X_p$$

□

**Example 1.3.2: Solve the System**

$$\begin{cases} x_1 + 2x_2 - x_3 &= 4 \\ 2x_1 + x_2 - 2x_3 &= 2 \\ x_1 + x_2 - x_3 &= 2 \end{cases}$$

As always, we start by forming the augmented matrix and putting it into REF:

$$\begin{aligned} \left( \begin{array}{ccc|c} 1 & 2 & -1 & -4 \\ 2 & 1 & -2 & 2 \\ 1 & 1 & -1 & 2 \end{array} \right) &\xrightarrow[R_3 \rightarrow R_3 - R_2]{R_2 \rightarrow R_2 - 2R_1} \left( \begin{array}{ccc|c} 1 & 2 & -1 & 4 \\ 0 & -3 & 0 & -6 \\ 0 & -1 & 0 & -2 \end{array} \right) \xrightarrow[R_3 \rightarrow R_3 - \frac{1}{3}R_2]{R_2 \rightarrow -\frac{1}{3}R_2} \left( \begin{array}{ccc|c} 1 & 2 & -1 & 4 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 0 \end{array} \right) \\ &\xrightarrow{R_1 \rightarrow R_1 - 2R_2} \left( \begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 0 \end{array} \right) \end{aligned}$$

Now, since there is no pivot in the third column,  $x_3$  must be a free variable, so we have  $x_3 = t$  for  $t \in \mathbb{R}$ . Using some very basic backwards substitution we can see that  $x_2 = 2$  and  $x_1 = x_3 = t$ . As well, notice that the final form of the augmented matrix is actually in RREF and not REF, because of how easy it was to put the matrix in RREF from REF, there was almost no reason **not** to put it into RREF. We can now write our solution vector,

$$X_{\text{gen}} = \begin{pmatrix} t \\ 2 \\ t \end{pmatrix} = \underbrace{\begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix}}_{X_p} + t \underbrace{\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}}_{X_{\text{genhom}}}$$

**Example 1.3.3: Find the General Homogenous Solution**

Just to prove the point, let's find the general homogenous solution for the above system,  $Ax = \hat{0}$ .

$$A = \begin{pmatrix} 1 & 2 & -1 \\ 2 & 1 & -2 \\ 1 & 1 & -1 \end{pmatrix} \xrightarrow[R_3 \rightarrow R_3 - R_1]{R_2 \rightarrow R_2 - 2R_1} \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} = A_{\text{RREF}}$$

With backwards substitution we can easily see that  $x_3 = t$  for  $t \in \mathbb{R}$ ,  $x_2 = 0$ , and  $x_1 = t$ . So,

$$X_{\text{genhom}} = \begin{pmatrix} t \\ 0 \\ t \end{pmatrix} = t \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$$

Which, as expected, is exactly the same vector we saw in the previous example.

**Example 1.3.4: Solve the System**

$$\begin{cases} x_1 - 2x_2 + x_3 + x_4 &= 0 \\ 2x_1 + x_2 - 2x_3 + x_4 &= 0 \\ 3x_1 - x_2 - x_3 &= 0 \end{cases}$$

Once again, we must put  $A$  into REF:

$$A \xrightarrow[\substack{R_2 \rightarrow R_2 - 2R_1 \\ R_3 \rightarrow R_3 - 3R_1}]{\substack{R_2 \rightarrow R_2 - 2R_1 \\ R_3 \rightarrow R_3 - 3R_1}} \begin{pmatrix} 1 & -2 & 1 & -1 \\ 0 & 5 & -4 & 3 \\ 0 & 5 & -4 & 3 \end{pmatrix} \xrightarrow{R_3 \rightarrow R_3 - R_2} \begin{pmatrix} 1 & -2 & 1 & -1 \\ 0 & 5 & -4 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Now, because there are no pivots in the third and fourth column,  $x_3$  and  $x_4$  are free variables which makes  $x_1$  and  $x_2$  our basic variables  $\therefore x_3 = t, x_4 = s$  for  $s, t \in \mathbb{R}$

$$\begin{cases} 5x_2 - 4x_3 + 3x_4 = 0 \Rightarrow x_2 = \frac{4}{5}x_3 - \frac{3}{5}x_4 = \frac{4}{5}t - \frac{3}{5}s \\ x_1 - 2x_2 + x_3 - x_4 = 0 \Rightarrow x_1 = 2x_2 - x_3 + x_4 = 2(\frac{4}{5}t - \frac{3}{5}s) - t + s \end{cases}$$

Before continuing, it's smart to make  $t = 5t$  and  $s = 5s$  in order to eliminate the fractions. Then we have:

$$X_{\text{gen}} = \begin{pmatrix} 3t - s \\ 4t - 3s \\ 5t \\ 5s \end{pmatrix} = t \begin{pmatrix} 3 \\ 4 \\ 5 \\ 0 \end{pmatrix} + s \begin{pmatrix} -1 \\ -3 \\ 0 \\ 5 \end{pmatrix}$$

## 1.4 Lecture 4: Span of a Set of Vectors [1.3]

### Rank of a Matrix

Let  $A$  be a matrix. Then  $\text{rank}(A)$  is the number of pivots in  $A_{\text{REF}}$ .  $\text{rank}(A)$  is equivalent to the number of non-zero rows in  $A_{\text{REF}}$ .

It follows from the definition of  $\text{rank}(A)$  that the number of free variables in a system is  $n - \text{rank}(A)$ . This quantity is known as  $\text{null}(A)$ .

If  $A$  is a matrix such that  $\text{rank}(A) = n$ , then  $A_{\text{REF}} = I$  where  $I$  is the identity matrix.

### Identity Matrix

The identity matrix,  $I$ , is a square matrix which is automatically in RREF and its only non-zero elements are the pivots. For example

$$I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Another important fact about the rank of matrix is that it can only be as great as the minimum of  $n$  and  $m$  for a matrix of size  $m \times n$ . Formally,

$$\text{rank} \left( \begin{matrix} A \\ m \times n \end{matrix} \right) \leq \min \{m, n\}$$

where  $\min \{x_1, x_2, \dots, x_n\}$  returns the smallest value out of  $x_1, x_2, \dots, x_n$ .

### Span of Vectors from $\mathbb{R}^n$

The span of a set of vectors from  $\mathbb{R}^n$ ,  $\text{Span} \{\vec{u}_1, \vec{u}_2, \dots, \vec{u}_k\}$  is the set of all possible linear combinations of  $\vec{u}_1, \vec{u}_2, \dots, \vec{u}_k$ :

$$AC = c_1\vec{u}_1 + c_2\vec{u}_2 + \dots + c_k\vec{u}_k, \quad (c_1, c_2, \dots, c_k \in \mathbb{R})$$

where

$$A = ((\vec{u}_1) \quad (\vec{u}_2) \quad \dots \quad (\vec{u}_k)) \text{ and } C = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_k \end{pmatrix}$$

The span of the zero vector, is also the zero vector:  $\text{Span}\{\vec{0}\} = \vec{0}$

#### Example 1.4.1: Span of a Single Vector

Consider a vector  $u_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ . Then  $\text{Span}\{\vec{u}_1\} = c\vec{u}_1$  for any  $c \in \mathbb{R}$ . Geometrically, this is the set of vectors in the same, or opposite, direction as  $\vec{u}_1$  with a scaled magnitude. When put together,  $\text{Span}\{\vec{u}_1\}$  gives a line in  $\mathbb{R}^2$ .

#### Example 1.4.2: Span of Two Vectors

Consider  $\vec{u}_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$  and  $\vec{u}_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ . Then,  $\text{Span}\{\vec{u}_1, \vec{u}_2\} = \mathbb{R}^2$  because any vector from  $\mathbb{R}^2$  can be written in the form

$$c_1\vec{u}_1 + c_2\vec{u}_2 = c_1 \begin{pmatrix} 1 \\ 2 \end{pmatrix} + c_2 \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

*Proof.* Consider the system

$$c_1 \begin{pmatrix} 1 \\ 2 \end{pmatrix} + c_2 \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \Leftrightarrow \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \Leftrightarrow A \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}, \quad {}_{2 \times 2} A = \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix}$$

Then, the augmented matrix is

$$\left( \begin{array}{cc|c} 1 & 1 & b_1 \\ 2 & 1 & b_2 \end{array} \right) \xrightarrow{R_2 \rightarrow R_2 - 2R_1} \left( \begin{array}{cc|c} 1 & 1 & b_1 \\ 0 & -1 & b_2 - 2b_1 \end{array} \right) = (A \mid B)_{\text{REF}}$$

$\therefore$  since the augmented matrix in REF has no pivots in the final column, the system is compatible and contains a solution for all vectors  $\begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \in \mathbb{R}$ .  $\square$

Note as well, that we could've written the final justification using the rank of augmented matrix as reasoning: ' $\therefore \text{rank}(A) = 2$ , so the system is compatible...'. As well, note that since  $A$  is a square matrix and  $\text{rank}(A) = 2$ , that also implies that  $A_{\text{RREF}} = I_2$  (it should be pretty easy to put  $A$  into RREF and see this for yourself).

#### Example 1.4.3: Span of Three Vectors

Consider three vectors

$$\vec{u}_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \quad \vec{u}_2 = \begin{pmatrix} 2 \\ 4 \end{pmatrix}, \quad \vec{u}_3 = \begin{pmatrix} 3 \\ 6 \end{pmatrix}$$

and consider the span of  $\vec{u}_1, \vec{u}_2$  and  $\vec{u}_3$ ,  $\text{Span}\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$ . Now notice that  $u_2 = 2u_1$  and  $u_3 = 3u_1$ . This implies that  $\text{Span}\{\vec{u}_1, \vec{u}_2, \vec{u}_3\} = \text{Span}\{\vec{u}_1\}$ .  $u_2$  and  $u_3$  are considered 'redundant' in the span.

*Proof.* Consider the matrix representing the vectors of the generation set:

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \end{pmatrix} \xrightarrow{R_2 \rightarrow R_2 - 2R_1} \begin{pmatrix} 1 & 2 & 3 \\ 0 & 0 & 0 \end{pmatrix}$$

$\therefore$  no pivots in the second and third columns implies that the second and third vectors are redundant.  $\square$

#### Example 1.4.4: Span of Three Vectors (again)

Now consider  $\vec{u}_1, \vec{u}_2, \vec{u}_3$  such that

$$\vec{u}_1 = \begin{pmatrix} 1 \\ 3 \end{pmatrix}, \quad \vec{u}_2 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \quad \vec{u}_3 = \begin{pmatrix} 4 \\ 6 \end{pmatrix}$$

Notice that  $\text{Span}\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$  **must** be reducible since the corresponding matrix has dimensions  $2 \times 3 \Rightarrow \text{rank} A \leq \min 2, 3 = 2$ .  $\therefore$  (at least) one vector is redundant. We can figure out which one as follows

$$A_{2 \times 3} = \begin{pmatrix} 1 & 2 & 4 \\ 3 & 1 & 6 \end{pmatrix} \xrightarrow{R_2 \rightarrow R_2 - 3R_1} \begin{pmatrix} 1 & 2 & 4 \\ 0 & -5 & -6 \end{pmatrix}$$

Thus, no pivot in the third column implies that the third vector is redundant:

$$\therefore \text{Span}\{\vec{u}_1, \vec{u}_2, \vec{u}_3\} = \text{Span}\{\vec{u}_1, \vec{u}_2\}$$

#### Example 1.4.5: Proving a Vector is in a Span

Show that  $\begin{pmatrix} 4 \\ 6 \end{pmatrix} \in \text{Span}\left\{\begin{pmatrix} 1 \\ 3 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \end{pmatrix}\right\}$ . We can start with the system

$$c_1 \begin{pmatrix} 1 \\ 3 \end{pmatrix} + c_2 \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 4 \\ 6 \end{pmatrix} \Rightarrow \begin{pmatrix} 1 & 2 & 4 \\ 3 & 1 & 6 \end{pmatrix} \xrightarrow{R_2 \rightarrow R_2 - 3R_1} \begin{pmatrix} 1 & 2 & 4 \\ 0 & -5 & -6 \end{pmatrix}$$

This gives

$$\begin{cases} -5c_2 = 6 \\ c_1 + 2c_2 = 4 \end{cases} \Rightarrow c_2 = -\frac{6}{5} \text{ and } c_1 = \frac{8}{5}$$

Note that we knew that  $\begin{pmatrix} 4 \\ 6 \end{pmatrix} \in \text{Span}\left\{\begin{pmatrix} 1 \\ 3 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \end{pmatrix}\right\}$  since the system was compatible, we just found the values of  $c_1$  and  $c_2$  such that

$$c_1 \begin{pmatrix} 1 \\ 3 \end{pmatrix} + c_2 \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 4 \\ 6 \end{pmatrix}$$

#### Example 1.4.6: Geometric Interpretation of $\mathbb{R}^3$ Span

Consider vectors  $\vec{u}_1, \vec{u}_2$  and give a geometric interpretation of their span,  $\text{Span}\{\vec{u}_1, \vec{u}_2\}$

$$\vec{u}_1 = \begin{pmatrix} 8 \\ 2 \\ -6 \end{pmatrix}, \quad \vec{u}_2 = \begin{pmatrix} 12 \\ 3 \\ -9 \end{pmatrix}$$

Start by noticing that  $\vec{u}_2 = \frac{3}{2}\vec{u}_1 \therefore u_1 \parallel u_2 \Rightarrow \text{Span}\{\vec{u}_1, \vec{u}_2\}$  must be a line in  $\mathbb{R}^3$  along  $\vec{u}_1$ .

**Example 1.4.7: Geometric Interpretation of  $\mathbb{R}^3$  Span (again)**

Consider vectors  $\vec{u}_1, \vec{u}_2$  and give a geometric interpretation of their span,  $\text{Span}\{\vec{u}_1, \vec{u}_2\}$

$$\vec{u}_1 = \begin{pmatrix} 8 \\ 2 \\ -6 \end{pmatrix}, \quad \vec{u}_2 = \begin{pmatrix} 12 \\ 3 \\ 9 \end{pmatrix}$$

Since  $\vec{u}_1 \nparallel \vec{u}_2$ ,  $\text{Span}\{\vec{u}_1, \vec{u}_2\}$  makes a plane in  $\mathbb{R}^3$  (spanned by  $\vec{u}_1, \vec{u}_2$ )

**Example 1.4.8**

Let  $A$  and  $b$  be matrices such that

$$A_{3 \times 3} = \begin{pmatrix} 1 & 0 & -4 \\ 0 & 3 & -2 \\ -2 & 6 & 3 \end{pmatrix} = ((\vec{u}_1) \quad (\vec{u}_2) \quad (\vec{u}_3)) \text{ and } b = \begin{pmatrix} 4 \\ 1 \\ -4 \end{pmatrix}$$

Now consider  $\mathcal{W} = \text{Span}\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$ .

(a) Is  $b \in \mathcal{W}$ ?

**Solution:** Start by assuming that  $c \in \mathcal{W} \therefore c_1\vec{u}_1 + c_2\vec{u}_2 + c_3\vec{u}_3 = b \Leftrightarrow$

$$\left( \begin{array}{ccc|c} 1 & 0 & -4 & 4 \\ 0 & 3 & -2 & 1 \\ -2 & 6 & 3 & -4 \end{array} \right) \xrightarrow{R_3 \rightarrow R_3 + 2R_1} \left( \begin{array}{ccc|c} 1 & 0 & -4 & 4 \\ 0 & 3 & -2 & 1 \\ 0 & 6 & -5 & 4 \end{array} \right) \xrightarrow{R_3 \rightarrow R_3 - 2R_2} \left( \begin{array}{ccc|c} 1 & 0 & -4 & 4 \\ 0 & 3 & -2 & 1 \\ 0 & 0 & -1 & 2 \end{array} \right)$$

Thus,  $\text{rank}(A_{\text{REF}}) = 3 \therefore$  the system is compatible and  $b \in \mathcal{W}$

(b) Is  $\text{Span}\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$  reducible?

**Solution:**

$$A_{3 \times 3} = \begin{pmatrix} 1 & 0 & -4 \\ 0 & 3 & -2 \\ -2 & 6 & 3 \end{pmatrix} \xrightarrow{R_3 \rightarrow R_3 + 2R_1} \begin{pmatrix} 1 & 0 & -4 \\ 0 & 3 & -2 \\ 0 & 6 & -5 \end{pmatrix} \xrightarrow{R_3 \rightarrow R_3 - 2R_2} \begin{pmatrix} 1 & 0 & -4 \\ 0 & 3 & -2 \\ 0 & 0 & -1 \end{pmatrix}$$

$\text{rank}(A_{\text{REF}}) = 3 \therefore$  the span is irreducible.

(c) Is  $\text{Span}\{\vec{u}_1, \vec{u}_2, \vec{u}_3\} = \text{Span}\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$

**Solution:** Yes, they're both  $\mathbb{R}^3$

## 1.5 Lecture 5: Linear Dependence and Independence [1.7]

### Linear Independence and Dependence

A set of vectors from  $\mathbb{R}^n$ ,  $\{\vec{u}_1, \vec{u}_2, \dots, \vec{u}_k\}$ , is **linearly independent** if

$$c_1\vec{u}_1 + c_2\vec{u}_2 + \dots + c_k\vec{u}_k = 0$$

is satisfied if, and only if,  $c_1 = c_2 = \dots = c_k = 0$ . Otherwise the set is called **linearly dependent**.

Now, notice that the above equation is actually equivalent to the matrix equation  $AC = 0$  where

$$A_{n \times k} = ((\vec{u}_1) \ (\vec{u}_2) \ \dots \ (\vec{u}_k)) \text{ and } C = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_k \end{pmatrix}$$

It should be evident that we can now view this as a homogenous system of linear equations. Thus, if  $\text{rank}(A_{\text{REF}}) = k$  the only solution to the homogenous system is the trivial solution,  $\vec{0}$ . We can also define linear dependence this way: if  $\text{rank}(A_{\text{REF}}) < k$ , the system must have non trivial solutions and the set of vectors must be linearly dependent.

#### Example 1.5.1

Consider the set of vectors below and determine if it is linearly independent or dependent:

$$\left\{ \begin{pmatrix} 1 \\ -2 \\ 4 \end{pmatrix}, \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -2 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix} \right\}$$

**Solution:** This problem is actually fairly easy and requires zero calculations. Consider the matrix formed from these vectors,  $A$ . Notice that  $A$  must have dimensions  $3 \times 4$  and recall that for any  $m \times n$  matrix, the rank of said matrix is always less than or equal to  $\min m, n$ . Applying this to  $A$ , we can see that  $\text{rank}(A_{\text{REF}}) \leq \min 3, 4 = 3$ . Therefore, since there can be only a max of 3 pivots, at least one column will not have a pivot, so the system will have non-trivial solutions. This implies that the set of vectors is linearly **dependent**.

A set of vectors that is linearly dependent contains at least one vector that can be represented as a linear combination of the other vectors in the set. This should also remind you of sets of vectors that have reducible spans.

#### Example 1.5.2

Consider the set of vectors in the above example, and determine which vector(s) are redundant.



**Solution:**

$$A_{3 \times 4} = \begin{pmatrix} 1 & 3 & 1 & 0 \\ -2 & 1 & 2 & 1 \\ 4 & 1 & -1 & 2 \end{pmatrix} \xrightarrow[R_3 \rightarrow R_3 - 4R_1]{R_2 \rightarrow R_2 + 2R_1} \begin{pmatrix} 1 & 3 & 1 & 0 \\ 0 & 7 & 4 & 1 \\ 0 & -11 & -5 & -2 \end{pmatrix} \xrightarrow[R_3 \rightarrow \frac{9}{7}R_3]{R_3 \rightarrow R_3 + \frac{11}{7}R_2} \begin{pmatrix} 1 & 3 & 1 & 0 \\ 0 & 7 & 4 & 1 \\ 0 & 0 & 1 & * \end{pmatrix}$$

$\therefore$  no pivot in the fourth column implies that the fourth vector is redundant.

In the above example, notice that we really didn't even need to calculate the element  $a_{34}$  since it's value has no effect on whether or not the fourth column will have a pivot or not.

### Example 1.5.3

Consider the set of vectors below and find a value of  $h$  such that the set is linearly dependent

$$\left\{ \begin{pmatrix} 1 \\ -3 \\ 2 \end{pmatrix}, \begin{pmatrix} -3 \\ 10 \\ -6 \end{pmatrix}, \begin{pmatrix} 2 \\ -7 \\ h \end{pmatrix} \right\}$$

**Solution:**

$$A_{3 \times 3} = \begin{pmatrix} 1 & 3 & 2 \\ -3 & 10 & -6 \\ 2 & -6 & h \end{pmatrix} \xrightarrow[R_3 \rightarrow R_3 - 2R_1]{R_2 \rightarrow R_2 + 3R_1} \begin{pmatrix} 1 & 3 & 2 \\ 0 & 1 & -1 \\ 0 & 0 & h - 4 \end{pmatrix}$$

Thus, in order for the set to be linearly dependent, one column must *not* have a pivot. This will happen if  $h - 4 = 0 \Rightarrow h = 4$ .

### Example 1.5.4

Determine the column vectors of  $A$  are linearly independent or dependent where  $A =$

$$\begin{pmatrix} -4 & -3 & 0 \\ 0 & -1 & 4 \\ 1 & 0 & 5 \\ 5 & 4 & 6 \end{pmatrix} \xrightarrow{R_3 \leftrightarrow R_1} \begin{pmatrix} 1 & 0 & 5 \\ 0 & -1 & 4 \\ -4 & -3 & 0 \\ 5 & 4 & 6 \end{pmatrix} \xrightarrow[R_4 \rightarrow R_4 - 5R_1]{R_3 \rightarrow R_3 + 4R_1} \begin{pmatrix} 1 & 0 & 5 \\ 0 & -1 & 4 \\ 0 & -3 & 20 \\ 0 & 4 & -9 \end{pmatrix} \xrightarrow[R_3 \rightarrow R_3 - 3R_2]{R_4 \rightarrow R_4 + 4R_2, R_3 \leftrightarrow R - 4} \begin{pmatrix} 1 & 0 & 5 \\ 0 & -1 & 4 \\ 0 & 0 & 7 \\ 0 & 0 & 0 \end{pmatrix}$$

$\therefore$  if  $\vec{u}_1, \vec{u}_2, \vec{u}_3$  are the column vectors, then they are linearly independent because  $\text{rank}(A_{REF}) = 3$

### Example 1.5.5

Determine the value of  $h$  such that the vectors are linearly dependent

$$\left\{ \begin{pmatrix} 1 \\ 5 \\ -3 \end{pmatrix}, \begin{pmatrix} -2 \\ -9 \\ 6 \end{pmatrix}, \begin{pmatrix} 3 \\ h \\ -9 \end{pmatrix} \right\} \Rightarrow A_{3 \times 3} = \begin{pmatrix} 1 & -2 & 3 \\ 5 & -9 & h \\ -3 & 6 & -9 \end{pmatrix} \xrightarrow[R_3 \rightarrow R_3 + 3R_1]{R_2 \rightarrow R_2 - 5R_1} \begin{pmatrix} 1 & -2 & 3 \\ 0 & 1 & h - 15 \\ 0 & 0 & 0 \end{pmatrix}$$

$\therefore$  since there is no pivot in the third column, the vectors will be linearly independent  $\forall h \in \mathbb{R}$ .

## 1.6 Lecture 5: Linear Transformations [1.8]

### Transformations

A linear transformation  $\mathcal{T} : \mathbb{R}^m \rightarrow \mathbb{R}^n$  given by the map  $A \cdot \vec{x} \mapsto \vec{b}$  is a function that can be thought of as finding all vectors  $\vec{x} \in \mathbb{R}^n$  that turn into  $\vec{b} \in \mathbb{R}^m$  when multiplied by  $A$ .  $\mathbb{R}^n$  is called the **domain** and  $\mathbb{R}^m$  is called the **co-domain**.

### Example 1.6.1

Consider a linear map  $\mathcal{T} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$  given by  $A\vec{x} \mapsto \vec{b}$ . For the given matrix  $A$  and vectors  $\vec{u}$  and  $\vec{v}$ , find  $\mathcal{T}(\vec{u})$  and  $\mathcal{T}(\vec{v})$ .

$$A = \begin{pmatrix} 5 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{2} \end{pmatrix}, \quad \vec{u} = \begin{pmatrix} 1 \\ 0 \\ -4 \end{pmatrix}, \quad \vec{v} = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$$

Since  $\mathcal{T}(\vec{x}) = A\vec{x}$ , we have

$$\mathcal{T}(\vec{u}) = A\vec{u} = \begin{pmatrix} 5 \\ 0 \\ -2 \end{pmatrix} \quad \text{and} \quad \mathcal{T}(\vec{v}) = A\vec{v} = \begin{pmatrix} 5a \\ \frac{1}{2}a \\ \frac{1}{2}c \end{pmatrix}$$

### Linear Transformations

A transformation,  $\mathcal{T}(\vec{x})$ , is **linear**, if

$$\mathcal{T}(a\vec{u} + b\vec{v}) = a\mathcal{T}(\vec{u}) + b\mathcal{T}(\vec{v}) \Leftrightarrow A(a\vec{X} + b\vec{Y}) = aA\vec{X} + bA\vec{Y}$$

Note that not all transformations are linear transformations, there is a distinct difference!

### Basic Linear Transformations

- Dilation Transformation:  $\mathcal{T}(\vec{u}) = r\vec{u}$  for  $r \in \mathbb{R}$
- 90° Transformation:  $\mathcal{T}(\vec{u}) = A\vec{u} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \vec{u}$
- Shear Transformation:  $\mathcal{T}(\vec{u}) = A\vec{u} = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \vec{u}$

Note that for the **shear** transformation, the  $y$ -coordinate ( $j$ ) of  $\vec{u}$  stays the same, only the  $x$ -coordinate ( $i$ ) is stretched.

### Example 1.6.2: Mapping to the Zero Vector

Consider the following matrix  $A$  and find all vectors  $\vec{x}$  such that they get mapped to the zero vector when multiplied by  $A$ :

$$A = \begin{pmatrix} 1 & -4 & 7 & -5 \\ 0 & 1 & -4 & 3 \\ 2 & -6 & 6 & -4 \end{pmatrix}_{3 \times 4}$$

We can start by noticing that if we define  $\mathcal{T} : \mathbb{R}^4 \rightarrow \mathbb{R}^3$  such that  $\mathcal{T}(\vec{x}) = A\vec{x}$ , then we just need to

solve  $\mathcal{T}(\vec{x}) = \vec{0} \Rightarrow A\vec{x} = \vec{0}$ :

$$A \xrightarrow{R_3 \rightarrow R_3 - 2R_1} \begin{pmatrix} 1 & -4 & 7 & -5 \\ 0 & 1 & -4 & 3 \\ 0 & 2 & -5 & 6 \end{pmatrix} \xrightarrow{R_3 \rightarrow R_3 - 2R_2} \begin{pmatrix} 1 & -4 & 7 & -5 \\ 0 & 1 & -4 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Thus, since there is no pivot in the third and forth columns,  $x_3 = s$  and  $x_4 = t$  are free variables (for  $s, t, \in \mathbb{R}$ ). Backwards substitution into the equations gives:

$$\begin{cases} x_2 - 4x_3 + 3x_4 = 0 \Rightarrow x_2 = 4s - 3t \\ x_1 - 4x_2 + 7x_3 - 5x_4 = 0 \Rightarrow x_1 = 9s - 7t \end{cases} \Rightarrow X = s \begin{pmatrix} 9 \\ 4 \\ 1 \\ 0 \end{pmatrix} + t \begin{pmatrix} -7 \\ -3 \\ 0 \\ 1 \end{pmatrix}$$

### Example 1.6.3

Consider  $\mathcal{T}$  given by  $A\vec{X} \mapsto \vec{B}$  and determine if  $\vec{b} = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \in \mathcal{T}^{\text{img}}$ , where  $A$  is the same matrix from the above example. Notice that all this questions is really asking is if the system  $A\vec{x} = \vec{b}$  is consistent. We can figure this out fairly easily:

$$(A \mid b) = \left( \begin{array}{cccc|c} 1 & -4 & 7 & -5 & -1 \\ 0 & 1 & -4 & 3 & 1 \\ 2 & -6 & 6 & -4 & 0 \end{array} \right) \xrightarrow{\text{ERO}} \left( \begin{array}{cccc|c} 1 & -4 & 7 & -5 & -1 \\ 0 & 1 & -4 & 3 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{array} \right)$$

$\therefore$  the system is consistent so  $b \in \mathcal{T}^{\text{img}}$

## 1.7 Lecture 7: The Matrix of a Linear Transformation [1.9]

Recall that a function  $\mathcal{T}(X) : \mathbb{R}^n \rightarrow \mathbb{R}^m$  is called a linear transformation if

1.  $\mathcal{T}(X + Y) = \mathcal{T}(X) + \mathcal{T}(Y)$ , and
2.  $\mathcal{T}(kX) = k\mathcal{T}(X)$  for  $k \in \mathbb{R}$

It's important to notice the similarities to the properties of matrix multiplication:

1.  $A_{m \times n} \begin{pmatrix} X \\ Y \end{pmatrix}_{\substack{n \times 1 \\ n \times 1}} = AX + AY$
2.  $A_{m \times n} \begin{pmatrix} kX \\ \end{pmatrix}_{n \times 1} = kAX$

**Theorem 1.2.** Let  $\mathcal{T} : \mathbb{R}^n \rightarrow \mathbb{R}^m$  be a linear transformation. Then, there exists a matrix,  $A_{m \times n}$  such that

$$\mathcal{T}(\vec{x}) = A\vec{x}$$

for all  $\vec{x} \in \mathbb{R}^n$

*Proof.* For any matrix  $X$ ,

$$X = I_n X = x_1 e_1 + x_2 e_2 + \cdots + x_n e_n$$

Then,

$$\begin{aligned} \mathcal{T}(\vec{x}) &= \mathcal{T}(x_1 e_1 + x_2 e_2 + \cdots + x_n e_n) \\ &= \mathcal{T}(x_1 e_1) + \mathcal{T}(x_2 e_2) + \cdots + \mathcal{T}(x_n e_n) \\ &= x_1 \mathcal{T}(e_1) + x_2 \mathcal{T}(e_2) + \cdots + x_n \mathcal{T}(e_n) \\ &= ([\mathcal{T}(e_1)] \quad [\mathcal{T}(e_2)] \cdots [\mathcal{T}(e_n)]) \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \\ &= A\vec{x} \end{aligned}$$

□

Note that  $I_n$  is the identity matrix which is an  $n \times n$  matrix in which every element is a 0 except for elements,  $a_{11}, a_{22}, \dots, a_{nn}$ . Note as well, that  $e_i$  is the  $i$ -th column of the identity matrix. For example, consider  $I_2$  and  $I_3$ :

$$I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \text{ and } I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Then  $e_1$  for  $I_2$  would be  $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ . Similarly,  $e_3$  for  $I_3$  would be  $\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$

### One-to-one and Onto

Consider a linear transformation  $\mathcal{T} : \mathbb{R}^n \rightarrow \mathbb{R}^m$  given by the map  $\vec{x} \mapsto \begin{pmatrix} A \\ m \times n \end{pmatrix} \vec{x}$ . Then,  $\mathcal{T}$  is called **onto** when there exists at least one vector  $\vec{x} \in \mathbb{R}^n$  for *any* vector  $\vec{B} \in \mathbb{R}^m$ .  $\mathcal{T}$  is called **one-to-one** when  $A\vec{x} = \vec{B}$  has only one solution.

One way to think about it is that  $\mathcal{T}$  is **onto** if any vector in  $\mathbb{R}^m$  can be ‘made’ by applying the transformation to a vector  $\vec{x}$ .  $\mathcal{T}$  is **one-to-one** if for any solution to the equation  $A\vec{x} = \vec{B}$ , the matrix  $\vec{x}$  is the sole, unique, solution to that equation. Note as well, that a transformation can be both one-to-one and onto.

#### Determining if a transformation is one-to-one or onto

Consider a linear transformation  $\mathcal{T}(\vec{x}) = A\vec{x}$  for a matrix  $A_{m \times n}$ .

- If  $\text{rank}(A) = m$  (implying that  $m \leq n$ ), then each row in  $A$  contains a pivot which ensures that the augmented matrix  $(A \mid B)_{\text{REF}}$  will have no pivots in the final column. In other words, the system  $A\vec{x} = \vec{B}$  will be consistent for all  $\vec{B} \in \mathbb{R}^m$  and  $\mathcal{T}$  is **onto**.
- If  $\text{rank}(A) = n$  (implying that  $n \leq m$ ), then each column in  $A$  contains a pivot which ensures that there are no free variables in the system  $A\vec{x} = \vec{B}$ . In other words, any solution to  $A\vec{x} = \vec{B}$  will be unique and  $\mathcal{T}$  is **one-to-one**.

#### Example 1.7.1: Determine if a Transformation is One-to-one or Onto

Consider the transformation  $\mathcal{T} : \mathbb{R}^4 \rightarrow \mathbb{R}^3$  given by

$$\mathcal{T}(\vec{x}) = \begin{pmatrix} 2x_1 + x_2 - x_3 + x_4 \\ x_1 + 3x_2 + 5x_3 - 2x_4 \\ 3x_1 - x_2 + 2x_3 + 3x_4 \end{pmatrix}$$

determine if  $\mathcal{T}$  is one-to-one, onto, both, or neither.

**Solution:** We can start by rewriting  $\mathcal{T}$  in the form  $\mathcal{T}(\vec{x}) = A\vec{x}$ :

$$\mathcal{T}(\vec{x}) = \begin{pmatrix} 2 & 1 & -1 & 1 \\ 1 & 3 & 5 & -2 \\ 3 & -1 & 2 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = A\vec{x}$$

Since  $A$  is a  $3 \times 4$  matrix,  $\text{rank}(A) \leq 3 \therefore \mathcal{T}$  cannot be one-to-one. Now, as per usual, in order to determine the rank of  $A$  we must find  $A_{\text{REF}}$ . We can start with  $A \xrightarrow{R_1 \leftrightarrow R_2}$

$$\begin{pmatrix} 1 & 3 & 5 & -2 \\ 2 & 1 & -1 & 1 \\ 3 & -1 & 2 & 3 \end{pmatrix} \xrightarrow[\substack{R_2 \rightarrow R_2 - 2R_1 \\ R_3 \rightarrow R_3 - 3R_1}]{R_1 \leftrightarrow R_2} \begin{pmatrix} 1 & 3 & 5 & -2 \\ 0 & -5 & -11 & 5 \\ 0 & -10 & -13 & 9 \end{pmatrix} \xrightarrow{R_3 \rightarrow R_3 - 2R_2} \begin{pmatrix} 1 & 3 & 5 & -2 \\ 0 & -5 & -11 & 5 \\ 0 & 0 & 9 & -1 \end{pmatrix}$$

Thus,  $\text{rank}(A) = 3 = n$  so  $\mathcal{T}$  is **onto**

#### Rotation Transformations

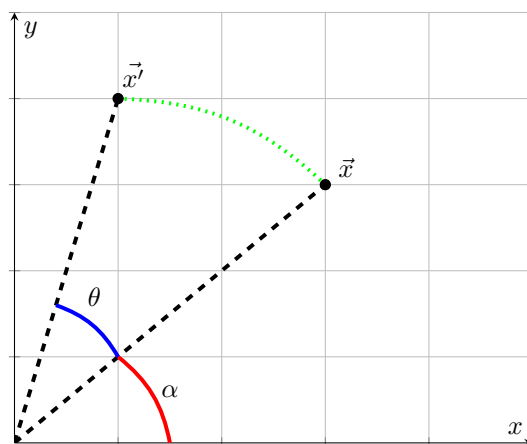
Consider vectors  $\vec{x}$  which has an angle  $\alpha$  from the  $x$ -axis and  $\vec{x}'$  which has an angle  $\alpha + \theta$  from the  $x$ -axis. There exists a transformation  $\mathcal{T}$  such that  $\mathcal{T}(\vec{x}) = \vec{x}'$ . Just like all linear transformations,  $\mathcal{T}$  can be written in the form  $\mathcal{T}(\vec{x}) = A\vec{x}$  where

$$A = \underbrace{\begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}}_{\text{Rotation Matrix}}$$

Thus, if  $\vec{x} = \begin{pmatrix} x \\ y \end{pmatrix}$  and  $\vec{x}' = \begin{pmatrix} x' \\ y' \end{pmatrix}$  then,

$$\mathcal{T}(\vec{x}) = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x' \\ y' \end{pmatrix}$$

If you're curious where the definition of the rotation matrix comes from, consider the two vectors  $\vec{x} = \begin{pmatrix} x \\ y \end{pmatrix}$  and  $\vec{x}' = \begin{pmatrix} x' \\ y' \end{pmatrix}$ :



We can represent  $\vec{x}$  in polar coordinates as

$$x = r \cos \alpha \text{ and } y = r \sin \alpha$$

Then,

$$x' = r \cos(\alpha + \theta) \text{ and } y' = r \sin(\alpha + \theta)$$

This gives,

$$\begin{aligned} x' &= r \cos(\alpha + \theta) = \overbrace{r \cos \alpha}^x \cos \theta - \overbrace{r \sin \alpha}^y \sin \theta = x \cos \theta - y \sin \theta \\ y' &= \overbrace{r \sin \alpha}^y \cos \theta + \overbrace{r \cos \alpha}^x \sin \theta = x \sin \theta + y \cos \theta \\ \Rightarrow \begin{pmatrix} x' \\ y' \end{pmatrix} &= \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \end{aligned}$$

**Example 1.7.2: Prove  $\mathcal{T}$  is a Linear Transformation**

Consider a transformation  $\mathcal{T}$  such that

$$\mathcal{T}(\vec{x}) = \begin{pmatrix} 0 \\ x_1 + x_2 \\ x_2 + x_3 \\ x_3 + x_4 \end{pmatrix}$$

determine if  $\mathcal{T}$  is a linear transformation.

**Solution:** Notice that  $\mathcal{T}(\vec{x}) = A\vec{x}$  for

$$A_{4 \times 4} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix} \text{ and } \vec{x}_{4 \times 1} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}$$

Because  $\mathcal{T}$  can be put in the form  $\mathcal{T}(\vec{x}) = A\vec{x}$ , it must be a linear transformation.

**Example 1.7.3: Determine if  $\mathcal{T}$  is One-to-one or Onto**

Consider the linear transformation  $\mathcal{T}$  from the above example. Determine if  $\mathcal{T}$  is one-to-one, onto, both, or neither.

**Solution:**

$$A_{4 \times 4} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix} \xrightarrow{R_1 \leftrightarrow R_4} \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Thus,  $\text{rank}(A) \neq 4 = m = n$  so  $\mathcal{T}$  is neither one-to-one nor onto.

**Example 1.7.4: Find the Matrix  $A$** 

Given that  $\mathcal{T}$  is a linear transformation such that

$$\mathcal{T}(\vec{x}) = A \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2x_1 - 3x_2 \\ 4x_1 \\ x_1 - x_2 + x_3 \end{pmatrix}$$

determine the value of the matrix  $A$ .

**Solution:** We can find  $A$  as the coefficient matrix of  $\mathcal{T}$ :

$$A = \begin{pmatrix} 2 & -3 & 0 \\ 4 & 0 & 0 \\ 1 & -1 & 1 \end{pmatrix}$$

**Example 1.7.5: Determine if  $\mathcal{T}$  is One-to-one or Onto**

Consider the linear transformation  $\mathcal{T}(\vec{x}) = A\vec{x}$  such that  $A$  is the matrix in the above example. Determine if  $\mathcal{T}$  is one-to-one, onto, both, or neither.

**Solution:** We can easily figure this out by looking at the rank of  $A_{\text{REF}}$ .

$$A \xrightarrow[R_2 \leftrightarrow R_1]{R_2 \rightarrow \frac{1}{4}R_2} \begin{pmatrix} 1 & 0 & 0 \\ 2 & -3 & 0 \\ 1 & -1 & 1 \end{pmatrix} \xrightarrow[R_3 \rightarrow R_3 - R_1]{R_2 \rightarrow R_2 - 2R_1} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & -1 & 1 \end{pmatrix} \xrightarrow[R_2 \rightarrow -\frac{1}{3}R_2]{R_3 \rightarrow R_3 + R_2} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Thus,  $\text{rank}(A_{\text{REF}}) = 3 = m = n$  so  $\mathcal{T}$  is both one-to-one and onto.



## 2 Unit 2

### 2.1 Lecture 8: Matrix Operations [2.1]

Before I start, I must note that I was not able to physically make it to lecture 8, so these notes are simply just a brief overview of concepts covered in lecture.

So far, we've very briefly been introduced to the idea of matrix multiplication through our studies of systems of equation in the form  $A\vec{x} = \vec{b}$  which we know can be written as

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & & \ddots & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$$

in order to represent the system

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n & = b_1 \\ \vdots & \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n & = b_n \end{cases}$$

You'll notice that in order to go from the matrix to the system we simply take each row of the matrix (or each coefficient) and multiply it by the corresponding variable. This is exactly how matrix multiplication is defined.

#### Matrix Multiplication

Consider two matrices  $A_{m \times n}$  and  $B_{n \times k}$  such that

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & & \ddots & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \text{ and } B = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1k} \\ \vdots & & \ddots & \\ b_{n1} & b_{n2} & \cdots & b_{nk} \end{pmatrix}$$

Label the columns in  $B$  as  $\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_k$ . Then their product,  $AB$ , is defined as follows:

$$AB = \begin{pmatrix} A\mathbf{b}_1 & A\mathbf{b}_2 & \cdots & A\mathbf{b}_k \end{pmatrix}_{m \times k}$$

It's extremely important to note that matrix multiplication is only defined two matrices,  $A$  and  $B$ , if, and only if, the number of columns in  $A$  is equal to the number of rows in  $B$ .

One other way to think of matrix multiplication that if  $A$  and  $B$  are defined as above then,

$$AB = \begin{pmatrix} \sum_{i=1}^n a_{1i}b_{i1} & \sum_{i=1}^n a_{1i}b_{i2} & \cdots & \sum_{i=1}^n a_{1i}b_{ik} \\ \sum_{i=1}^n a_{2i}b_{i1} & \sum_{i=1}^n a_{2i}b_{i2} & \cdots & \sum_{i=1}^n a_{2i}b_{ik} \\ \vdots & & \ddots & \\ \sum_{i=1}^n a_{mi}b_{i1} & \sum_{i=1}^n a_{mi}b_{i2} & \cdots & \sum_{i=1}^n a_{mi}b_{ik} \end{pmatrix}$$

Now, granted, this looks very complicated and extremely hard to use. So, why did I bother including this here? Well, if you look closely at each summation, what you'll notice is that each summation is simply the sum of the product of a corresponding **row** from  $A$  with a column from  $B$ . So, for example, the first element in  $AB$  is the sum of each element in the **first row** in  $A$  multiplied by its corresponding element in the **first column** in  $B$ . The second element in  $AB$  is the **first row** in  $A$  multiplied by the **second column** in  $B$ . This is probably best explained with a few examples:

**Example 2.1.1: Basic Matrix Multiplication**

Let  $A$  and  $B$  be defined as below. Determine if  $AB$  is defined, and if it is find it.

$$A_{2 \times 2} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \text{ and } B_{2 \times 2} = \begin{pmatrix} 2 & 4 \\ 1 & 3 \end{pmatrix}$$

**Solution:** First, since  $A$  and  $B$  are both  $2 \times 2$  matrices, we know that  $AB$  must exist since  $A$  and  $B$ , obviously, have the same number of rows and columns as each other. We can then find  $AB$ :

$$AB = \begin{pmatrix} (1 \cdot 2) + (2 \cdot 1) & (1 \cdot 4) + (2 \cdot 3) \\ (3 \cdot 2) + (4 \cdot 1) & (3 \cdot 4) + (4 \cdot 3) \end{pmatrix} = \begin{pmatrix} 4 & 10 \\ 10 & 24 \end{pmatrix}$$

Keep in mind that we also could've found the product  $AB$  using the other definition of matrix multiplication:

$$AB = (A\mathbf{b}_1 \quad A\mathbf{b}_2) = \left( \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \end{pmatrix} \quad \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 4 \\ 3 \end{pmatrix} \right) = \left( \begin{pmatrix} 4 \\ 10 \end{pmatrix} \quad \begin{pmatrix} 10 \\ 24 \end{pmatrix} \right) = \begin{pmatrix} 4 & 10 \\ 10 & 24 \end{pmatrix}$$

**Example 2.1.2: Basic Matrix Multiplication**

Let  $A$  and  $B$  be defined as below. Determine if  $AB$  is defined, and if it is find it.

$$A_{3 \times 2} = \begin{pmatrix} 1 & 3 \\ 5 & 7 \\ 9 & 11 \end{pmatrix} \text{ and } B_{2 \times 2} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

**Solution:** Since  $A$  has 2 columns and  $B$  has 2 rows, we know that  $AB$  exists. Thus,

$$AB = \begin{pmatrix} (1 \cdot 1) + (3 \cdot 0) & (1 \cdot 1) + (3 \cdot 1) \\ (5 \cdot 1) + (7 \cdot 0) & (5 \cdot 1) + (7 \cdot 1) \\ (9 \cdot 1) + (11 \cdot 0) & (9 \cdot 1) + (11 \cdot 1) \end{pmatrix} = \begin{pmatrix} 1 & 4 \\ 5 & 12 \\ 9 & 20 \end{pmatrix}$$

**Theorem 2.1** (Properties of Matrix Multiplication). Consider a matrix  $A_{m \times n}$  and let  $B$  and  $C$  be matrices such that the necessary products are defined.

1. *Associative Law:*  $A(BC) = (AB)C$
2. *Left Distributive Law:*  $A(B + C) = AB + AC$
3. *Right Distributive Law:*  $(B + C)A = BA + CA$
4.  $r(AB) = (rA)B = A(rB)$  for any  $r \in \mathbb{R}$
5. *Identity for Matrix Multiplication:*  $I_m A = A = A I_n$

Recall that  $I_n$  is the  $n \times n$  matrix such that every element is zero except for  $a_{11}, a_{22}, \dots, a_{nn}$ , referred to as the **Identity Matrix**.

It's important to note a few things about the above theorem.

1. There is NO commutative law, thus, you cannot assume  $AB = BA$
2. There is NO cancellation law:  $AB = BC$  does NOT imply that  $A = C$
3. If  $AB$  is the zero matrix, you CANNOT conclude that either  $A = 0$  or  $B = 0$

### The Power of a Matrix

Let  $A$  be any  $n \times n$  matrix. Then  $A^k$  is defined as

$$A^k = \underbrace{A \cdots A}_{k \text{ times}}$$

If  $k = 0$  then  $A^k = I_n$  where  $I_n$  is the identity matrix in  $\mathbb{R}^n$ .

The definition of the power of a matrix should feel extremely similar to the definition of a real number to any power: repeated multiplication.

### Transpose of a Matrix

Let  $A$  be any  $m \times n$  matrix. Then  $A^T$  is the **transpose** of  $A$ , which is a  $n \times m$  matrix whose rows are made from the columns of  $A$  and whose columns are made from the rows.

For example, let  $A$  and  $B$  be defined as

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \text{ and } B = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$$

then,

$$A^T = \begin{pmatrix} a & c \\ b & d \end{pmatrix} \text{ and } B^T = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}$$

**Theorem 2.2** (Properties of the Transpose). *Let  $A$  and  $B$  be matrices such that necessary sums and products are defined. Then,*

1.  $(A^T)^T = A$
2.  $(A + B)^T = A^T + B^T$
3.  $(rA)^T = rA^T$  for any  $r \in \mathbb{R}$
4.  $(AB)^T = B^T A^T$

## 2.2 Lecture 9: Inverse of a Matrix and Characterizations of Invertible Matrices [2.2-2.3]

### Invertible Matrices

An  $n \times n$ ,  $A$ , matrix is called invertible if there exists an  $n \times n$  matrix,  $B$ , such that  $AB = BA = I_n$ . Such  $B$  is called the inverse of  $A$ , denoted  $A^{-1}$ .  $A^{-1}$  is always a unique matrix for any  $A$ .

**Theorem 2.3.** An  $n \times n$  matrix is invertible if, and only if,  $\text{rank}(A) = n$

*Proof.* Consider a matrix  $A_{n \times n}$  such that  $\text{rank}(A_{\text{RREF}}) = n$ . Then  $A_{\text{RREF}} = I_n$ . Assume it takes  $k$  iterations of elementary row operations (ERO) to obtain  $A_{\text{RREF}}$  from  $A$ . Then,  $A_{\text{RREF}} = I_n = E_k \cdots E_2 E_1$ . So,  $I = BA \Leftrightarrow A^{-1} = B$  and  $B^{-1} = A$  and  $B^{-1} = E_1^{-1} E_2^{-1} \cdots E_k^{-1}$ .

If  $A^{-1}$  exists then consider  $A\vec{x} = \vec{b} \Rightarrow \vec{x} = A^{-1}\vec{b}$ . Thus, the system has only one solution and so  $\text{rank}(A) = n$ .  $\square$

If you're confused on why  $\vec{x} = A^{-1}\vec{b}$  implies that the system only has a single unique solution, consider that if  $\vec{x} = A^{-1}\vec{b}$  then we have

$$A\vec{x} = A(A^{-1}\vec{b}) = (AA^{-1})\vec{b} = I\vec{b} = \vec{b}$$

so,  $A^{-1}\vec{b}$  must be a solution. To show that it is unique, consider any solution  $\vec{u}$ . Then,

$$A\vec{u} = \vec{b} \Rightarrow A^{-1}A\vec{u} = A^{-1}\vec{b} \Rightarrow I\vec{u} = A^{-1}\vec{b} \Rightarrow \vec{u} = A^{-1}\vec{b}$$

Then, we know from 1.3, that if a system has a single solution, its augmented matrix (or  $A_{\text{REF}}$ ) has a rank equal to  $n$ . It should follow from this that if  $A^{-1}$  exists the homogenous system,  $A\vec{x} = \vec{0}$  has a single solution, the trivial solution:  $\vec{x} = \vec{0}$ .

#### How to Find the Inverse of a $2 \times 2$ Matrix

Consider  $2 \times 2$  matrices  $A, B$  with rank 2 such that  $AB = I_2$ ,

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \text{ and } B = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$$

Since  $\text{rank}(A) = 2$ ,  $A$  is invertible. Then,

$$AB = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \Rightarrow \begin{pmatrix} a\alpha + b\gamma & a\beta + b\delta \\ c\alpha + d\gamma & c\beta + d\delta \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

This gives,

$$\begin{aligned} \begin{cases} a\alpha + b\gamma = 1 \\ c\alpha + d\gamma = 0 \end{cases} &\Rightarrow \left( \begin{array}{cc|c} a & b & 1 \\ c & d & 0 \end{array} \right) \xrightarrow{\text{ERO}} \left( \begin{array}{cc|c} 1 & 0 & \alpha \\ 0 & 1 & \gamma \end{array} \right)_{\text{RREF}} \\ \begin{cases} a\beta + b\delta = 0 \\ c\beta + d\delta = 1 \end{cases} &\Rightarrow \left( \begin{array}{cc|c} a & b & 0 \\ c & d & 1 \end{array} \right) \xrightarrow{\text{ERO}} \left( \begin{array}{cc|c} 1 & 0 & \beta \\ 0 & 1 & \delta \end{array} \right)_{\text{RREF}} \end{aligned}$$

We can then combine to form one large matrix

$$(A \mid I) \xrightarrow{\text{ERO}} (I \mid A^{-1}) \Rightarrow \left( \begin{array}{cc|cc} 1 & 0 & \alpha & \beta \\ 0 & 1 & \gamma & \delta \end{array} \right)$$

Note that this method can actually be generalized such that for any matrix  $A_{n \times n}$ , augmenting it with the identity matrix and using ERO to find  $(A \mid I)_{\text{RREF}}$  will always result in  $(I \mid A^{-1})$

**Example 2.2.1**

Find the inverse of  $A$  and verify that  $AA^{-1} = I$

$$A = \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix}$$

**Solution:**

$$\begin{aligned} \left( \begin{array}{cc|cc} 1 & 3 & 1 & 0 \\ 2 & 4 & 0 & 1 \end{array} \right) &\xrightarrow{R_2 \rightarrow R_2 - 2R_1} \left( \begin{array}{cc|cc} 1 & 3 & 1 & 0 \\ 0 & -2 & -2 & 1 \end{array} \right) \xrightarrow{R_2 \rightarrow -\frac{1}{2}R_2} \left( \begin{array}{cc|cc} 1 & 3 & 1 & 0 \\ 0 & 1 & 1 & -\frac{1}{2} \end{array} \right) \\ &\xrightarrow{R_1 \rightarrow R_1 - 3R_2} \left( \begin{array}{cc|cc} 1 & 0 & -2 & \frac{3}{2} \\ 0 & 1 & 1 & -\frac{1}{2} \end{array} \right) \\ &\Rightarrow A^{-1} = \begin{pmatrix} -2 & \frac{3}{2} \\ 1 & -\frac{1}{2} \end{pmatrix} = -\frac{1}{2} \begin{pmatrix} 4 & -3 \\ 2 & 1 \end{pmatrix} \end{aligned}$$

Then,

$$AA^{-1} = -\frac{1}{2} \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix} \begin{pmatrix} 4 & -3 \\ -2 & 1 \end{pmatrix} = -\frac{1}{2} \begin{pmatrix} -2 & 0 \\ 0 & -2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I_2$$

Now, you might have noticed that in the previous example,  $A^{-1}$  actually looked extremely similar to  $A$ . In fact,  $A$  and  $A^{-1}$  are made up elements of the exact same magnitude. This phenomenon will occur for the inverses of every  $2 \times 2$  matrix with rank 2 and, in fact, we have a formula that explains the similarity.

**Formula for Finding the Inverse of a  $2 \times 2$  matrix**

Let  $A$  be a  $2 \times 2$  matrix,

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

Then,

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

where the quantity  $ad - bc$  is called the **determinant of  $A$** , known as  $\det A$ .

It should follow from the last line that  $\det A \neq 0 \Leftrightarrow A^{-1}$  exists.

**Example 2.2.2**

For  $A$  below, does  $A^{-1}$  exist? If so, find  $A^{-1}$

$$A = \begin{pmatrix} -4 & 6 \\ 6 & -9 \end{pmatrix}$$

**Solution:** Start by finding  $\det A$ :

$$\det A = (-4 \cdot -9) - (4 \cdot 9) = 36 - 36 = 0$$

Thus, by the fact above,  $A^{-1}$  must not exist. Or, in other words,  $A$  is not invertible.

### Example 2.2.3

For  $A$  below, does  $A^{-1}$  exist? If so, find  $A^{-1}$ .

$$A = \begin{pmatrix} 1 & -5 & -4 \\ 0 & 3 & 4 \\ -3 & 6 & 0 \end{pmatrix}$$

**Solution:** Start by finding the rank of  $A$ :

$$A \xrightarrow{R_3 \rightarrow R_3 + 3R_1} \begin{pmatrix} 1 & -5 & -4 \\ 0 & 3 & 4 \\ 0 & -9 & -12 \end{pmatrix} \xrightarrow{R_3 \rightarrow R_3 + 3R_2} \begin{pmatrix} 1 & -5 & -4 \\ 0 & 3 & 4 \\ 0 & 0 & 0 \end{pmatrix}$$

Therefore,  $\text{rank}(A) < 3$  so  $A$  is not invertible.

### Example 2.2.4

For  $A$  below, does  $A^{-1}$  exist? If so, find  $A^{-1}$ .

$$A = \begin{pmatrix} 1 & -5 & -4 \\ 0 & 3 & 5 \\ -3 & 6 & 0 \end{pmatrix}$$

**Solution:** Start by finding the rank of  $A$ :

$$A \xrightarrow{R_3 \rightarrow R_3 + 3R_1} \begin{pmatrix} 1 & -5 & -4 \\ 0 & 3 & 5 \\ 0 & -9 & -12 \end{pmatrix} \xrightarrow{R_3 \rightarrow R_3 + 3R_2} \begin{pmatrix} 1 & -5 & -4 \\ 0 & 3 & 5 \\ 0 & 0 & 3 \end{pmatrix}$$

Therefore,  $\text{rank}(A) = 3$  so  $A$  is invertible. Now, consider

$$\begin{aligned} (A \mid I_3) &= \left( \begin{array}{ccc|ccc} 1 & -5 & -4 & 1 & 0 & 0 \\ 0 & 3 & 5 & 0 & 1 & 0 \\ -3 & 6 & 0 & 0 & 0 & 1 \end{array} \right) \xrightarrow{R_3 \rightarrow R_3 + 3R_1} \left( \begin{array}{ccc|ccc} 1 & -5 & -4 & 1 & 0 & 0 \\ 0 & 3 & 5 & 0 & 1 & 0 \\ 0 & -9 & -12 & 3 & 0 & 1 \end{array} \right) \\ &\xrightarrow{R_3 \rightarrow R_3 + 3R_2} \left( \begin{array}{ccc|ccc} 1 & -5 & -4 & 1 & 0 & 0 \\ 0 & 3 & 5 & 0 & 1 & 0 \\ 0 & 0 & 3 & 3 & 3 & 1 \end{array} \right) \xrightarrow{\substack{R_2 \rightarrow \frac{1}{3}R_2 \\ R_3 \rightarrow \frac{1}{3}R_3}} \left( \begin{array}{ccc|ccc} 1 & -5 & -4 & 1 & 0 & 0 \\ 0 & 1 & \frac{5}{3} & 0 & \frac{1}{3} & 0 \\ 0 & 0 & 1 & 1 & 1 & \frac{1}{3} \end{array} \right) \\ &\xrightarrow{\substack{R_2 \rightarrow R_2 - \frac{5}{3}R_3 \\ R_1 \rightarrow R_1 + 5R_3}} \left( \begin{array}{ccc|ccc} 1 & -5 & 0 & 5 & 4 & \frac{4}{3} \\ 0 & 1 & 0 & -\frac{5}{3} & -\frac{4}{3} & -\frac{5}{9} \\ 0 & 0 & 1 & 1 & 1 & \frac{1}{3} \end{array} \right) \xrightarrow{R_1 \rightarrow R_1 + 5R_2} \left( \begin{array}{ccc|ccc} 1 & 0 & 0 & -\frac{10}{3} & -\frac{8}{3} & -\frac{13}{9} \\ 0 & 1 & 0 & -\frac{5}{3} & -\frac{4}{3} & -\frac{5}{9} \\ 0 & 0 & 1 & 1 & 1 & \frac{1}{3} \end{array} \right) \\ &\Rightarrow A^{-1} = \begin{pmatrix} -\frac{10}{3} & -\frac{8}{3} & -\frac{13}{9} \\ -\frac{5}{3} & -\frac{4}{3} & -\frac{5}{9} \\ 1 & 1 & \frac{1}{3} \end{pmatrix} \end{aligned}$$

#### Example 2.2.5: Is $A$ invertible?

Let  $A$  be the matrix below. Determine if  $A$  is invertible.

$$A = \begin{pmatrix} -7 & 0 & 4 \\ 3 & 0 & -1 \\ 2 & 0 & 9 \end{pmatrix}$$

**Solution:** Since  $A$  has a 0 column,  $\text{rank}(A) < 3 \therefore A$  is not invertible

#### Example 2.2.6: Is $A$ invertible?

Let  $A$  be the matrix below. Determine if  $A$  is invertible.

$$A = \begin{pmatrix} -1 & -3 & 0 & 1 \\ 3 & 5 & 8 & -3 \\ -2 & -6 & 3 & 2 \\ 0 & -1 & 2 & 1 \end{pmatrix}$$

**Solution:**

$$A \xrightarrow[\substack{R_3 \rightarrow R_3 - 2R_1 \\ R_2 \rightarrow -\frac{1}{4}R_2}]{R_2 \rightarrow R_2 + 3R_1} \begin{pmatrix} -1 & -3 & 0 & 1 \\ 0 & 1 & -2 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & -1 & 2 & 1 \end{pmatrix} \xrightarrow{R_4 \rightarrow R_4 + R_2} \begin{pmatrix} -1 & -3 & 0 & 1 \\ 0 & 1 & -2 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Thus,  $\text{rank}(A) = 4 \therefore A$  is invertible.

## 2.3 Lecture 10: Review for Midterm 1

### Systems of Equations

*Homogenous* systems of equations are systems in the form  $A\vec{x} = \vec{0}$ . These systems always contain at least one solution,  $\vec{x} = \vec{0}$  (known as the trivial solution). *Non-Homogenous* systems of equations are system in the form  $A\vec{x} = \vec{b}$  where  $\vec{b} \neq \vec{0}$ .

#### Example 2.3.1: Solve the System

$$\begin{cases} 2x_1 + x_2 - x_3 = 4 \\ x_1 - 2x_2 + x_3 = 2 \end{cases}$$

**Solution:** We can form an augmented matrix and use elementary row operations to find a solution:

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

Now, because there is no pivot in the third column,  $x_3$  is a free variable: let  $x_3 = 5t$  for  $t \in \mathbb{R}$ . Then, we can back substitute from the matrix to the equations to get:

$$\begin{cases} 5x_2 - 3x_3 = 0 \\ x_1 - 2x_2 + x_3 = 2 \end{cases} \Rightarrow x_2 = 3t \text{ and } x_1 = 2 + t$$

We can formally write our solution as follows:

$$X_{\text{gen}} = \begin{pmatrix} 2+t \\ 3t \\ 5t \end{pmatrix} = t \underbrace{\begin{pmatrix} 1 \\ 3 \\ 5 \end{pmatrix}}_{X_{\text{genhom}}} + \underbrace{\begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix}}_{X_p}$$

Where  $X_{\text{genhom}}$  is the general homogenous solution and  $X_p$  is a particular solution.

A system is **consistent** if there exists at least one solution.  $A\vec{x} = \vec{b}$  is consistent if the augmented matrix,  $(A|\vec{b})_{\text{REF}}$  has no pivots in the final column. Equivalently: if  $\text{rank} \begin{pmatrix} A \\ m \times n \end{pmatrix} = m$  then  $A\vec{x} = \vec{b}$  is always consistent. We can note a few things from this fact. If  $A\vec{x} = \vec{b}$  is consistent, then the span of the columns



of  $A$  is  $\mathbb{R}^m$ . As well, consider a linear transformation  $\mathcal{T} : \mathbb{R}^n \rightarrow \mathbb{R}^m$  given by the map  $\vec{x} \mapsto A\vec{x}$ , then  $\mathcal{T}$  will be onto. If the system  $A\vec{x} = \vec{b}$  has only a single solution, then  $\text{rank} \begin{pmatrix} A \\ m \times n \end{pmatrix} = n$ . Again, consider a linear transformation,  $\mathcal{T} : \mathbb{R}^m \rightarrow \mathbb{R}^n$  given by the mapping  $\vec{x} \mapsto A\vec{x}$ , then  $\mathcal{T}$  will be one-to-one. If a linear transformation is one-to-one **and** onto, then it is called invertible.

Recall that  $\text{rank} \begin{pmatrix} A \\ m \times n \end{pmatrix} \leq \min\{m, n\}$ , so in order for a linear transformation to be both one-to-one and onto,  $m$  must equal  $n$  (in other words:  $A$  must be a square matrix).

### Span/Linear Dependence

Consider a set of vectors,  $\{\vec{u}_1, \vec{u}_2, \dots, \vec{u}_k\}$ . Then, the **span** of that set of vectors, denoted  $\text{Span}\{\vec{u}_1, \vec{u}_2, \dots, \vec{u}_k\}$ , is the set of all possible linear combinations of  $\vec{u}_1, \vec{u}_2, \dots, \vec{u}_k$ :

$$c_1\vec{u}_1 + c_2\vec{u}_2 + \dots + c_k\vec{u}_k = AC, \quad A = ((\vec{u}_1) \quad (\vec{u}_2) \quad \dots \quad (\vec{u}_k)) \quad \text{and} \quad C = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_k \end{pmatrix}$$

The set  $\{\vec{u}_1, \vec{u}_2, \dots, \vec{u}_k\}$  is called the **generating set**.

Consider  $AC = \vec{0}$ . If the only solution to  $AC = \vec{0}$  is the trivial solution, then the vectors  $\vec{u}_1, \vec{u}_2, \dots, \vec{u}_k$  are called **linearly independent**. If there exists a non-trivial solution to  $AC = \vec{0}$ , then the vectors are called **linearly dependent**. If the vectors are linearly dependent, then at least one vector in the generating set is redundant. We can determine which vector is redundant by finding  $A_{\text{REF}}$  and noting that any columns without a pivot correspond to redundant vectors. Notice that we can easily determine if vectors are linearly independent or dependent simply by just finding  $A_{\text{REF}}$ . If  $\text{rank}(A_{\text{REF}}) < n$  then the columns will always be linearly dependent. If  $\text{rank}(A_{\text{REF}}) = n$ , then the columns will always be linearly independent.

#### Example 2.3.2

Determine a value of  $r$  such that the set below is linearly dependent.

$$\left\{ \begin{pmatrix} 1 \\ -1 \\ 3 \end{pmatrix}, \begin{pmatrix} 2 \\ 4 \\ -2 \end{pmatrix}, \begin{pmatrix} r \\ 1 \\ 2 \end{pmatrix} \right\}$$

**Solution:** Consider

$$A = \begin{pmatrix} 1 & 2 & r \\ -1 & 4 & 1 \\ 3 & -2 & 2 \end{pmatrix} \xrightarrow[R_3 \rightarrow R_3 - 3R_1]{R_2 \rightarrow R_2 + R_1} \begin{pmatrix} 1 & 2 & r \\ 0 & 6 & 1+r \\ 0 & -8 & 2-3r \end{pmatrix} \xrightarrow{R_3 \rightarrow R_3 + \frac{8}{6}R_2} \begin{pmatrix} 1 & 2 & r \\ 0 & 6 & 1+r \\ 0 & 0 & \frac{10}{3} - \frac{5}{3}r \end{pmatrix} = A_{\text{REF}}$$

Therefore, in order for  $\text{rank}(A_{\text{REF}})$  to be less than  $n = 3$ ,  $r$  must be equal to 2. So,  $\boxed{r = 2}$

### Inverses

A  $n \times n$  matrix is invertible if, and only if, it has rank  $n$ .

**Example 2.3.3**

Let  $A$  be the matrix below. Find  $A^{-1}$

$$A = \begin{pmatrix} 1 & -2 & 3 \\ 2 & 1 & 1 \\ -1 & 0 & 2 \end{pmatrix}$$

**Solution:** We can find  $A^{-1}$  by augmenting  $A$  with  $I$  and finding  $(A | I)_{\text{RREF}} = (I | A^{-1})$ .

$$\begin{aligned} (A | I) &= \left( \begin{array}{ccc|ccc} 1 & -2 & 3 & 1 & 0 & 0 \\ 2 & 1 & 1 & 0 & 1 & 0 \\ -1 & 0 & 2 & 0 & 0 & 1 \end{array} \right) \xrightarrow[R_3 \rightarrow R_3 + R_1]{R_2 \rightarrow R_2 - 2R_1} \left( \begin{array}{ccc|ccc} 1 & -2 & 3 & 1 & 0 & 0 \\ 0 & 5 & -5 & -2 & 1 & 0 \\ 0 & -2 & 5 & 1 & 0 & 1 \end{array} \right) \\ &\xrightarrow{R_2 \rightarrow \frac{1}{5}R_2} \left( \begin{array}{ccc|ccc} 1 & -2 & 3 & 1 & 0 & 0 \\ 0 & 1 & -1 & -\frac{2}{5} & \frac{1}{5} & 0 \\ 0 & -2 & 5 & 1 & 0 & 1 \end{array} \right) \xrightarrow{R_3 \rightarrow R_3 + 2R_2} \left( \begin{array}{ccc|ccc} 1 & -2 & 3 & 1 & 0 & 0 \\ 0 & 1 & -1 & -\frac{2}{5} & \frac{1}{5} & 0 \\ 0 & 0 & 3 & \frac{1}{5} & \frac{2}{5} & 1 \end{array} \right) \\ &\xrightarrow[R_3 \rightarrow \frac{1}{3}R_3]{R_1 \rightarrow R_1 - R_3} \left( \begin{array}{ccc|ccc} 1 & -2 & 0 & \frac{4}{5} & -\frac{2}{5} & -1 \\ 0 & 1 & -1 & -\frac{2}{5} & \frac{1}{5} & 0 \\ 0 & 0 & 1 & \frac{1}{15} & \frac{2}{15} & \frac{1}{3} \end{array} \right) \xrightarrow{R_2 \rightarrow R_2 + R_3} \left( \begin{array}{ccc|ccc} 1 & -2 & 0 & \frac{4}{5} & -\frac{2}{5} & -1 \\ 0 & 1 & 0 & -\frac{1}{3} & \frac{3}{15} & \frac{1}{3} \\ 0 & 0 & 1 & \frac{1}{15} & \frac{2}{15} & \frac{1}{3} \end{array} \right) \\ &\xrightarrow{R_1 \rightarrow R_1 + 2R_2} \left( \begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{2}{15} & \frac{4}{15} & -\frac{1}{3} \\ 0 & 1 & 0 & -\frac{1}{3} & \frac{3}{15} & \frac{1}{3} \\ 0 & 0 & 1 & \frac{1}{15} & \frac{2}{15} & \frac{1}{3} \end{array} \right) \Rightarrow A^{-1} = \begin{pmatrix} \frac{2}{15} & \frac{4}{15} & -\frac{1}{3} \\ -\frac{1}{3} & \frac{3}{15} & \frac{1}{3} \\ \frac{1}{15} & \frac{2}{15} & \frac{1}{3} \end{pmatrix} \end{aligned}$$

## 3 Unit 3

### 3.1 Lecture 12: Introduction to Determinants [3.1]

Before I start, I must note that I was not able to physically make it to lecture 12, so these notes are simply just a brief overview of concepts covered in lecture.

Back in section (2.2) we mentioned that a  $2 \times 2$  matrix is invertible if, and only if, its **determinant** was non-zero and we defined the determinant of a two by two matrix as

$$\det A = ad - bc, \text{ where } A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

The idea of the determinant is one that we will extend to any  $n \times n$  matrix. Before we introduce a formal definition of and method for finding the determinant, we must first learn some new notation.

#### Submatrix Notation

Let  $A$  be any  $n \times n$  matrix. Then let  $A_{ij}$  denote the submatrix obtained by removing the  $i$ -th row and  $j$ -th column from  $A$ .

For example, consider a  $4 \times 4$  matrix

$$A_{4 \times 4} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 5 \\ 3 & 4 & 5 & 6 \\ 4 & 5 & 6 & 7 \end{pmatrix}$$

Then,

$$A_{11} = \begin{pmatrix} 3 & 4 & 5 \\ 4 & 5 & 6 \\ 5 & 6 & 7 \end{pmatrix}$$

$$A_{22} = \begin{pmatrix} 1 & 3 & 4 \\ 3 & 5 & 6 \\ 4 & 6 & 7 \end{pmatrix}$$

$$A_{34} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 4 & 5 & 6 \end{pmatrix}$$

#### The Determinant of a $n \times n$ Matrix

For  $n \geq 2$  the **determinant** of a  $n \times n$  matrix  $A$ , is

$$\det A = a_{11} \det A_{11} - a_{12} \det A_{12} + \cdots + (-1)^{1+n} a_{1n} \det A_{1n} = \sum_{j=1}^n (-1)^{1+j} a_{1j} \det A_{1j}$$

Note that the terms being added are occasionally written in a slightly different form:

$$\det A = a_{11}C_{11} + a_{12}C_{12} + \cdots + a_{1n}C_{1n}, \text{ where } C_{ij} = (-1)^{i+j} \det A_{ij}$$

The variable  $C_{ij}$  is also known as that  $(i, j)$ -**cofactor** of  $A$ . With this in mind, we introduce the next theorem

**Theorem 3.1.** *The determinant of a  $n \times n$  matrix  $A$  can be computed by a cofactor expansion across any row or down any column. The expansion across any row is*

$$\det A = a_{i1}C_{i1} + a_{i2}C_{i2} + \cdots + a_{in}C_{in}$$

*and the expansion down any column is*

$$\det A = a_{1j}C_{1j} + a_{2j}C_{2j} + \cdots + a_{nj}C_{nj}$$

We also have the following theorem

**Theorem 3.2.** For any triangular matrix  $A$ ,  $\det A$  is the product of all the entries on the main diagonal of  $A$ .

### 3.2 Lecture 13: Properties of Determinants [3.1]

#### Cofactors

Consider a matrix  $A_{n \times n}$ . Then

$$\det A = |A| = \underbrace{\sum_{j=1}^n a_{ij} C_{ij}}_{i\text{-th row expansion}} = \underbrace{\sum_{i=1}^n a_{ij} C_{ij}}_{j\text{-th column expansion}}$$

where  $C_{ij} = (-1)^{i+j} M_{ij}$ .  $M_{ij}$  is the matrix obtained by removing the  $i$ -th row and  $j$ -th column from  $A$ .  $C_{ij}$  is called the **cofactor**.

#### Properties of Determinants

Consider an  $n \times n$  matrix  $A$ , then

1.  $\det A = \det A^T$
2.  $A \xrightarrow{R_i \leftrightarrow R_j} \tilde{A} \Rightarrow \det \tilde{A} = -\det A \quad (i \neq j)$
3.  $A \xrightarrow{R_i \rightarrow \lambda R_i} \tilde{A} \Rightarrow \det \tilde{A} = \lambda \det A \quad (\lambda \neq 0)$
4.  $A \xrightarrow{\text{ERO}} \tilde{A} = \lambda A \Rightarrow \det \tilde{A} = \lambda^n \det A$
5.  $A \xrightarrow{R_i \rightarrow R_i + \lambda R_j} \tilde{A} \Rightarrow \det \tilde{A} = \det A \quad (i \neq j)(\lambda \neq 0)$
6.  $\det(A_{\text{RREF}}) = a_{11} \times a_{22} \times \cdots \times a_{nn}$
7.  $\det A \neq 0 \Leftrightarrow \text{rank}(A) = n$
8.  $\det A \neq 0 \Leftrightarrow A^{-1}$  exists
9.  $\det(AB) = \det(A) \det(B)$
10.  $\det(A^{-1}) = \frac{1}{\det A}$
11. If  $\det A \neq 0$  then,  $A^{-1} = \frac{C^T}{\det A}$  where  $C$  is the matrix of cofactors

One should note that it follows from property (6) that the determinant of any triangular matrix is simply the product of the elements on the main diagonal. You should also note that properties (7) and (8) are equivalent statements since an  $n \times n$  matrix is only invertible if, and only if, it has rank  $n$ .

#### Cramer's Rule

Consider an  $n \times n$  matrix  $A$  and an  $n \times 1$  matrix  $b$  such that  $A\vec{x} = b$  and  $\det A \neq 0$ . Then, the

system has solutions

$$x_1 = \frac{\Delta_1}{\det A}, x_2 = \frac{\Delta_2}{\det A}, \dots, x_n = \frac{\Delta_n}{\det A}$$

where each  $\Delta_k$  is the determinant of the matrix obtained from  $A$  by replacing the  $k$ -th column in  $A$  with  $b$ .

### Example 3.2.1: Solve the system using Cramer's Rule

$$\begin{cases} x_1 + 2x_2 - x_3 &= 4 \\ 2x_1 - x_2 + x_3 &= 3 \\ -x_1 + x_3 &= 6 \end{cases}$$

**Solution:** Start with the following matrices:

$$A = \begin{pmatrix} 1 & 2 & -1 \\ 2 & -1 & 1 \\ -1 & 0 & 1 \end{pmatrix} \quad B = \begin{pmatrix} 4 \\ 3 \\ 6 \end{pmatrix}$$

Then we can easily calculate  $\det A$  as follows:

$$\det A = \begin{vmatrix} 5 & 0 & 1 \\ 2 & -1 & 1 \\ -1 & 0 & 1 \end{vmatrix} = - \begin{vmatrix} 5 & 1 \\ -1 & 1 \end{vmatrix} = -6 \neq 0$$

Not only have we found that  $\det A = -6$  but since it is non-zero we also know we can apply Cramer's rule. Now we can find  $\Delta_1, \Delta_2, \Delta_3$ :

$$\begin{aligned} \Delta_1 &= \begin{vmatrix} 4 & 2 & -1 \\ 3 & -1 & 1 \\ 6 & 0 & 1 \end{vmatrix} = -4 \Rightarrow x_1 = \frac{-4}{-6} = \frac{2}{3} \\ \Delta_2 &= \begin{vmatrix} 1 & 4 & -1 \\ 2 & 3 & 1 \\ -1 & 6 & 1 \end{vmatrix} = -30 \Rightarrow x_2 = \frac{-30}{-6} = 5 \\ \Delta_3 &= \begin{vmatrix} 1 & 2 & 4 \\ 2 & -1 & 3 \\ -1 & 0 & 6 \end{vmatrix} = -40 \Rightarrow x_3 = \frac{-40}{-6} = \frac{20}{3} \end{aligned}$$

Thus,

$$X = \begin{pmatrix} \frac{2}{3} \\ 5 \\ \frac{20}{3} \end{pmatrix}$$

### Example 3.2.2: Find $A^{-1}$ if it exists

$$A = \begin{pmatrix} 5 & 3 \\ 2 & 1 \end{pmatrix}$$

**Solution:** Start with  $\det A = 5 - 6 = -1 \neq 0$  to verify that  $A^{-1}$  does exist. Then we can find the cofactors as follows

$$C_{11} = 1 \quad C_{12} = -2 \quad C_{21} = -3 \quad C_{22} = 5$$

Since  $A$  is  $2 \times 2$  each cofactor will simply be the only element remaining after removing the corresponding row and column, and then changing the sign based on  $(-1)^{i+j}$ .

We can now construct  $A^{-1}$ :

$$C = \begin{pmatrix} 1 & -2 \\ -3 & 5 \end{pmatrix} \Rightarrow C^T = \begin{pmatrix} 1 & -3 \\ -2 & 5 \end{pmatrix} \Rightarrow A^{-1} = - \begin{pmatrix} 1 & -3 \\ -2 & 5 \end{pmatrix} = \begin{pmatrix} -1 & 3 \\ 2 & -5 \end{pmatrix}$$

We can easily verify this is correct using the formula for the inverse of a  $2 \times 2$  matrix given in section (2.2)

We will now note down a few of the important applications of determinants in math. Consider two vectors  $\vec{a} = \langle a_1, a_2, a_3 \rangle$  and  $\vec{b} = \langle b_1, b_2, b_3 \rangle$ . Then their cross product will be

$$\vec{a} \times \vec{b} = \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix}$$

Now, consider a third vector  $\vec{c} = \langle c_1, c_2, c_3 \rangle$  and define the triple product of vectors to be  $\vec{a} \cdot (\vec{b} \times \vec{c})$ . This can also be represented as the determinant of the matrix

$$\begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}$$

### Example 3.2.3: Determine if the Given Set is Linearly Independent

$$\left\{ \begin{pmatrix} 4 \\ 6 \\ 2 \end{pmatrix}, \begin{pmatrix} -6 \\ 0 \\ 6 \end{pmatrix}, \begin{pmatrix} -3 \\ -5 \\ -2 \end{pmatrix} \right\}$$

**Solution:** Recalling property (7), we know that we only have to find the determinant of the matrix with columns made up of vectors from the set and check that the determinant is non-zero.

$$\begin{vmatrix} 4 & -6 & -3 \\ 6 & 0 & -5 \\ 2 & 6 & -2 \end{vmatrix} = 4(30) - (-6)(-2) + (-3)(-36) = 120 - 12 - 108 = 0$$

$\det A = 0 \therefore$  the set of vectors is **not** linearly independent.

## 4 Unit 4

### 4.1 Lecture 14: Vector Spaces and Subspaces [2.8/4.1]

Consider  $\mathbb{R}^n$  and vectors  $\vec{u}, \vec{v} \in \mathbb{R}^n$  and notice that the following properties always hold true:

1.  $\forall \vec{v}, \vec{u} \in \mathbb{R}^n, \vec{u} + \vec{v} \in \mathbb{R}^n$  (Closure under addition)
2.  $\forall \vec{v} \in \mathbb{R}^n, \forall \lambda \in \mathbb{R}, \lambda \vec{v} \in \mathbb{R}^n$  (Closure under scalar multiplication)
3.  $\vec{0} \in \mathbb{R}^n$  (Inclusion of the zero vector)

If you are unaware, the symbol  $\forall$  means ‘for all’ and the symbol  $\in$  means ‘contained in’. So, property (1) would be read ‘for all vectors  $v$  and  $u$  in  $\mathbb{R}^n$ , their sum is also in  $\mathbb{R}^n$ ’.

Now these properties that  $\mathbb{R}^n$  has make it special. In fact because of these three properties we call  $\mathbb{R}^n$  a **vector space**.

#### Subspaces of $\mathbb{R}^n$

Consider a subset of  $\mathbb{R}^n$ ,  $W$ .  $W$  is a subspace of  $\mathbb{R}^n$  if

1.  $\forall \vec{v}, \vec{u} \in W, \vec{u} + \vec{v} \in W$  (Closure under addition)
2.  $\forall \vec{v} \in W, \forall \lambda \in \mathbb{R}, \lambda \vec{v} \in W$  (Closure under scalar multiplication)
3.  $\vec{0} \in W$  (Inclusion of the zero vector)

It’s important to note that the *trivial* subspaces of  $\mathbb{R}^n$  are  $\mathbb{R}^n$  itself and the zero subspace:  $\{\vec{0}\}$

#### Example 4.1.1

If  $V \subseteq \mathbb{R}^2 = \{\vec{v} \in \mathbb{R}^2 \mid \vec{v} \text{ is a unit vector}\}$  (read ‘ $V$  is a subset of  $\mathbb{R}^2$  such that  $V$  is the set of all vectors in  $\mathbb{R}^2$  that are unit vectors’), is  $V$  a subspace of  $\mathbb{R}^2$ ?

**Solution:** If we were to write out  $V$  we would have

$$V = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$$

as these are the only unit vectors in  $\mathbb{R}^2$ . From this it should be very clear that  $\vec{0} \notin V \therefore V$  is **not** a subspace of  $\mathbb{R}^2$ .

#### Example 4.1.2

If  $V \subseteq \mathbb{R}^2 = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2 \mid x_1 + x_2 = 0 \right\}$ , is  $V$  a subspace of  $\mathbb{R}^2$ ?

**Solution:** Consider the parameterization  $x_2 = t$  and  $x_1 = 1 - t$ . Then

$$X = \begin{pmatrix} 1-t \\ t \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} + t \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

Once again, from this parameterization it should be easy to see that  $\vec{0} \notin V \therefore V$  is **not** a subspace of  $\mathbb{R}^2$ .

**Example 4.1.3**

Let  $V = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{R}^3 \mid x_1 + 2x_2 - x_3 = 0 \right\}$ . Is  $V$  a subspace of  $\mathbb{R}^3$ ?

**Solution:** Just like before consider the parameterization  $x_2 = s, x_3 = t, x_1 = -2s + t$ . This gives

$$X = \begin{pmatrix} -2s + t \\ s \\ t \end{pmatrix} = s \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix} + t \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \Rightarrow V = \text{Span} \left\{ \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \right\}$$

From this we can see that all three properties are satisfied thus  $V$  is a subspace of  $\mathbb{R}^3$ .

You may notice that the above example involves an equality to the Span of a set of vectors, despite said span not being mentioned once. Well, it is included because of the following theorem:

**Theorem 4.1.** *The span of a set of vectors from  $\mathbb{R}^n$  is a subspace of  $\mathbb{R}^n$ .*

*Proof.* Consider two vectors  $\vec{v}, \vec{u} \in \mathbb{R}^n$  and the span,  $\text{Span}\{x_1, x_2, \dots, x_k\}$ . Then  $\vec{v}$  and  $\vec{u}$  can be expressed as a linear combination as follows:

$$\begin{aligned} \vec{u} &= c_1 \vec{x}_1 + c_2 \vec{x}_2 + \dots + c_k \vec{x}_k \\ \vec{v} &= \tilde{c}_1 \vec{x}_1 + \tilde{c}_2 \vec{x}_2 + \dots + \tilde{c}_k \vec{x}_k \\ \Rightarrow \vec{u} + \vec{v} &= (c_1 + \tilde{c}_1) \vec{x}_1 + (c_2 + \tilde{c}_2) \vec{x}_2 + \dots + (c_k + \tilde{c}_k) \vec{x}_k \end{aligned}$$

Then,  $\vec{u}, \vec{v}, \vec{u} + \vec{v} \in \text{Span}\{x_1, x_2, \dots, x_k\}$  by the definition of span. □

**Example 4.1.4**

Consider the set of all polynomials of degree 3,  $P_3$ . Is  $P_3$  a vector space?

**Solution:** Let  $P$  be any polynomial of degree 3. Then

$$P(x) = a_0 + a_1x + a_2x^2 + a_3x^3$$

We can have  $P(x) = 0$  if  $\forall x, a_0 = a_1 = a_2 = a_3 = 0$ , therefore  $\vec{0} \in P_3$ .

Now consider another polynomial  $Q$  of degree three:

$$Q(x) = b_0 + b_1x + b_2x^2 + b_3x^3$$

Then,

$$R(x) = P(x) + Q(x) = (a_0 + b_0) + (a_1 + b_1)x + (a_2 + b_2)x^2 + (a_3 + b_3)x^3$$

But,  $R$  is still a degree 3 polynomial, so  $R \in P_3$ , and thus  $P_3$  is closed under addition. We can also consider any scalar  $\lambda \in \mathbb{R}$ . Then

$$Q(x) = \lambda P(x) = \lambda a_0 + \lambda a_1x + \lambda a_2x^2 + \lambda a_3x^3$$

Again,  $Q \in P_3$  because  $Q$  is still a degree 3 polynomial. Thus  $P_3$  is a vector space.



One interesting thing comes when we consider vectors in  $\mathbb{R}^4$  and  $P_3$ . Notice that the general form for a vector in  $\mathbb{R}^4$  is

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

Now, consider the general (vector) form of any polynomial in  $P_3$ :

$$\begin{pmatrix} a_0 \\ a_1x \\ a_2x^2 \\ a_3x^3 \end{pmatrix} = a_0 \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + a_1 \begin{pmatrix} 0 \\ x \\ 0 \\ 0 \end{pmatrix} + a_2 \begin{pmatrix} 0 \\ 0 \\ x^2 \\ 0 \end{pmatrix} + a_3 \begin{pmatrix} 0 \\ 0 \\ 0 \\ x^3 \end{pmatrix}$$

Now, hopefully, you notice the extremely similarities in the structures of these two vector spaces. In fact, the structure of each of these vector spaces is considered identical in mathematics and  $\mathbb{R}^4$  is call **isomorphic** to  $P_3$ . We can show this by defining a bijective map  $\phi : \mathbb{R}^4 \rightarrow P_3$ . Obviously, in this case,  $\phi$  is simply the direct mapping of each corresponding vector.

Now, if we consider these vectors *without* their coefficients, what we get is actually called the **standard basis vectors** since

$$\mathbb{R}^4 = \text{Span} \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\}$$

$$P_3 = \text{Span} \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ x \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ x^2 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ x^3 \end{pmatrix} \right\}$$

#### Example 4.1.5

Consider  $W$  such that  $W = \text{Span} \{ \vec{u}, \vec{v} \}$ . Find  $\vec{u}$  and  $\vec{v}$ .

$$W = \begin{pmatrix} 5b + 2c \\ b \\ c \end{pmatrix} = \text{Span} \{ \vec{u}, \vec{v} \}$$

**Solution:**

$$W = b \begin{pmatrix} 5 \\ 1 \\ 0 \end{pmatrix} + c \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix} = \text{Span} \left\{ \begin{pmatrix} 5 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix} \right\} \Rightarrow \vec{u} = \begin{pmatrix} 5 \\ 1 \\ 0 \end{pmatrix} \text{ and } \vec{v} = \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}$$

We can as well notice that the above span was a subspace of  $\mathbb{R}^3$  since it's span was of a set of vectors contained in  $\mathbb{R}^3$ .

**Example 4.1.6**

Consider  $v_1, v_2, v_3$  and  $W$  as below.

$$v_1 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \quad v_2 = \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix} \quad v_3 = \begin{pmatrix} 4 \\ 2 \\ 6 \end{pmatrix} \quad W = \begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix}$$

(a) Is  $W \in \text{Span}\{v_1, v_2, v_3\}$ ?

**Solution:** Consider  $(A \mid W)$  where  $A = ((\vec{v}_1) \ (\vec{v}_2) \ (\vec{v}_3))$

$$\left( \begin{array}{ccc|c} 1 & 2 & 4 & 3 \\ 0 & 1 & 2 & 1 \\ -1 & 3 & 6 & 2 \end{array} \right) \xrightarrow{R_3 \rightarrow R_3 + R_1} \left( \begin{array}{ccc|c} 1 & 2 & 4 & 3 \\ 0 & 1 & 2 & 1 \\ 0 & 5 & 10 & 5 \end{array} \right) \xrightarrow{R_3 \rightarrow R_3 + 5R_2} \left( \begin{array}{ccc|c} 1 & 2 & 4 & 3 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 0 & 0 \end{array} \right)$$

Thus, the system is consistent so  $W \in \text{Span}\{v_1, v_2, v_3\}$

(b) Is  $\{v_1, v_2, v_3\}$  linearly independent?

**Solution:** From the above part we know that  $v_3 = 2v_2 \therefore$  the set is linearly **dependent**.

**Basis of a Subspace**

The basis of a subspace  $W$  of  $\mathbb{R}^n$  is any set of linearly independent vectors such that its span is  $W$ .

**4.2 Lecture 15: Column, Row, and Null Spaces [2.9/4.2]**

**Theorem 4.2.** The span of any non-empty set of vectors from  $\mathbb{R}^n$  is a subspace of  $\mathbb{R}^n$

**The Basis of a Subspace**

Let  $\mathcal{W}$  be a subspace of any given vector space. The **basis** of  $\mathcal{W}$  is the smallest generating set and the biggest linear independent set whose span is  $\mathcal{W}$ . The **dimension** of  $\mathcal{W}$  is the number of vectors in a basis.

**Rank**

Let  $A$  be a matrix. Then  $\text{rank}(A) = n$  if, and only if,  $A_{\text{REF}}$  has  $n$  pivots and there are  $n$  nonzero rows in  $A_{\text{REF}}$

**Nullity**

The **nullity** of  $A$  is  $\text{nul}(A) = n - \text{rank}(A)$

With these definitions in mind, we introduce the following difference subspaces that are associated with a matrix.

Consider an  $m \times n$  matrix  $A$ . Then we have the following subspaces

- **Column Space:** The column space of  $A$  is defined as

$$\text{Col}(A) = \text{Span}\{\text{columns of } A\}$$

and is a subspace of  $\mathbb{R}^m$

- **Row Space:** The row space of  $A$  is defined as

$$\text{Row}(A) = \text{Span}\{\text{rows of } A\}$$

and is a subspace of  $\mathbb{R}^n$

- **Null Space:** The null space of  $A$  is defined as

$$\text{Nul}(A) = \text{Span}\{\text{set of all solution of } A\vec{x} = 0\}$$

and is a subspace of  $\mathbb{R}^n$

- **Null Space of  $A^T$ :** The null space of  $A^T$  is defined as

$$\text{Nul}(A^T) = \text{Span}\{\text{set of all solutions of } A^T\vec{y} = 0\}$$

and is a subspace of  $\mathbb{R}^m$

It's important to notice the following equalities

$$\text{Col}(A) = \text{Row}(A^T) \text{ and } \text{Row}(A) = \text{Col}(A^T)$$

as these will lead directly into the fundamental equations of linear algebra we will define at the end of this section.

In order to find the basis of a column space of  $A$ , we'll start by finding  $A_{\text{REF}}$ . Once we have  $A_{\text{REF}}$ , any column without a pivot corresponds to a redundant vector in the basis.

#### Example 4.2.1

Consider the matrix  $A$  below. Find  $\mathcal{W} = \text{Col}(A)$  and a basis of  $\mathcal{W}$

$$A = \begin{pmatrix} 1 & -1 & 2 \\ -2 & 2 & -4 \\ 1 & 3 & 2 \\ -1 & 5 & -2 \end{pmatrix}$$

**Solution:** We can easily find

$$\text{Col}(A) = \text{Span} \left\{ \begin{pmatrix} 1 \\ -2 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} -1 \\ 2 \\ 3 \\ 5 \end{pmatrix}, \begin{pmatrix} 2 \\ -4 \\ 2 \\ -2 \end{pmatrix} \right\}$$

In order to find the basis of  $\mathcal{W}$  we must find  $A_{\text{REF}}$ .

$$A \xrightarrow[\substack{R_3 \rightarrow R_3 - R_1 \\ R_4 \rightarrow R_4 + R_1}]{R_2 \rightarrow R_2 + 2R_1} \begin{pmatrix} -1 & -1 & 2 \\ 0 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 4 & 0 \end{pmatrix} \xrightarrow[\substack{R_3 \rightarrow R_3 - R_4}]{R_2 \leftrightarrow R_4} \begin{pmatrix} 1 & -1 & 2 \\ 0 & 4 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = A_{\text{REF}}$$

$A_{\text{REF}}$  has pivots in the first and second column, but no pivot in the third column. Thus, the third vector is redundant and we can choose a basis to be

$$\left\{ \begin{pmatrix} 1 \\ -2 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} -1 \\ 2 \\ 3 \\ 5 \end{pmatrix} \right\}$$

In order to find a basis of a row space of  $A$ , we'll start by finding  $A_{\text{REF}}$ . Each nonzero column in  $(A_{\text{REF}})^T$  is a vector in the basis.

#### Example 4.2.2

Consider the matrix  $A$  below. Find  $\mathcal{W} = \text{Row}(A)$  and a basis of  $\mathcal{W}$ .

$$A = \begin{pmatrix} 1 & -1 & 2 \\ -2 & 2 & -4 \\ 1 & 3 & 2 \\ -1 & 5 & -2 \end{pmatrix}$$

**Solution:** We can easily find

$$\text{Row}(A) = \text{Span} \left\{ \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}, \begin{pmatrix} -2 \\ 2 \\ -4 \end{pmatrix}, \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix}, \begin{pmatrix} -1 \\ 5 \\ -2 \end{pmatrix} \right\} = \mathcal{W}$$

Now, in order to determine a basis of  $\mathcal{W}$  we have to find  $A_{\text{REF}}$ . But recall that  $\text{Row}(A) = \text{Col}(A^T)$ , so we can simply take  $A_{\text{REF}}$  from (4.2) and transpose it:

$$(A_{\text{REF}})^T = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -1 & 4 & 0 & 0 \\ 2 & 0 & 0 & 0 \end{pmatrix}$$

we can easily see the nonzero columns and find a basis to be

$$\left\{ \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}, \begin{pmatrix} 0 \\ 4 \\ 0 \end{pmatrix} \right\}$$

### Example 4.2.3

Consider the matrix  $A$  below. Find  $\mathcal{W} = \text{Nul}(A)$  and a basis of  $\mathcal{W}$ .

$$A = \begin{pmatrix} 1 & -1 & 2 \\ -2 & 2 & -4 \\ 1 & 3 & 2 \\ -1 & 5 & -2 \end{pmatrix}$$

**Solution:** We know that we can find solutions to  $A\vec{x} = 0$  by analyzing  $A_{\text{REF}}$ . Recall that

$$A_{\text{REF}} = \begin{pmatrix} 1 & -1 & 2 \\ 0 & 4 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

No pivot in the third column tells us that  $x_3 = t$  is a free variable. We can also see that  $x_2 = 0$  and that  $x_1 = -2x_3 = -2t$ . Thus,

$$X = \begin{pmatrix} -2t \\ 0 \\ t \end{pmatrix} = t \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix}$$

Thus,

$$\text{Nul}(A) = \text{Span} \left\{ \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix} \right\}$$

Because the generating set is only a single vector, it is also the basis for  $\text{Nul}(A)$ .

**Example 4.2.4**

Consider the matrix  $A$  below. Find  $\mathcal{W} = \text{Nul}(A^T)$  and a basis of  $\mathcal{W}$ .

$$A = \begin{pmatrix} 1 & -1 & 2 \\ -2 & 2 & -4 \\ 1 & 3 & 2 \\ -1 & 5 & -2 \end{pmatrix}$$

**Solution:** We must find all the solutions to  $A^T \vec{y} = 0$  which means we need to find  $(A^T)_{\text{REF}}$

$$A^T = \begin{pmatrix} 1 & -2 & 1 & -1 \\ -1 & 2 & 3 & 5 \\ 2 & -4 & 2 & -2 \end{pmatrix} \xrightarrow[R_3 \rightarrow R_3 - 2R_1]{R_2 \rightarrow R_2 + R_1} \begin{pmatrix} 1 & -2 & 1 & -1 \\ 0 & 0 & 4 & 4 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

No pivot in the second and fourth columns tells us that  $y_2 = t$  and  $y_4 = s$  are free variables. Solving the system we get

$$Y = \begin{pmatrix} 2s + 2t \\ t \\ -s \\ s \end{pmatrix} = t \begin{pmatrix} 2 \\ 1 \\ 0 \\ 0 \end{pmatrix} + s \begin{pmatrix} 2 \\ 0 \\ -1 \\ 1 \end{pmatrix}$$

Thus,

$$\text{Nul}(A^T) = \text{Span} \left\{ \begin{pmatrix} 2 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ -1 \\ 1 \end{pmatrix} \right\}$$

One thing you should notice that is the number of vectors in a basis of a subspace also corresponds to the number of nonzero rows in  $A_{\text{REF}}$ . Connecting this with the definitions of dimension and rank we can see that

$$\dim[\text{Row}(A)] = \text{rank}(A) = \dim[\text{Col}(A^T)] \quad (1)$$

Now if we recall from earlier that

$$\text{Col}(A) = \text{Row}(A^T)$$

and

$$\dim[\text{Col}(A)] = \text{rank}(A) = \dim[\text{Row}(A)]$$

we get

$$\text{rank}(A) = \text{rank}(A^T) \quad (2)$$

which is the *first* fundamental equation of linear algebra.

We now have the following two facts

$$\dim[\text{Nul}(A)] = n - \text{rank}(A) \text{ and } \dim[\text{nul}(A^T)] = m - \text{rank}(A)$$

Combining these with equation (10) we get

$$\dim[\text{Row}(A)] + \dim[\text{Nul}(A)] = n \quad (3)$$

$$\dim[\text{Col}(A)] + \dim[\text{Nul}(A^T)] = m \quad (4)$$

Combining all of these we get

## Fundamental Equations of Linear Algebra

Below are the three fundamental equations of linear algebra

$$\text{rank}(A) = \text{rank}(A^T) \quad (5)$$

$$\dim[\text{Row}(A)] + \dim[\text{Nul}(A)] = n \quad (6)$$

$$\dim[\text{Col}(A)] + \dim[\text{Nul}(A^T)] = m \quad (7)$$

## Example 4.2.5

Consider the matrix  $A$  below. Find  $\text{Nul}(A)$ .

$$A_{3 \times 5} = \begin{pmatrix} 1 & 5 & -4 & -3 & 1 \\ 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

**Solution:** Notice that  $A = A_{\text{REF}}$ . With this observation we can get that  $\text{rank}(A) = 2$ ,  $\text{nul}(A) = m - \text{rank}(A) = 5 - 2 = 3$ ,  $\text{Nul}(A) \subset \mathbb{R}^5$ , and  $\dim[\text{Nul}(A)] = 3$ . Now from  $A_{\text{REF}}$  we can see that  $x_3 = r$ ,  $x_4 = s$ ,  $x_5 = t$  are free variables. Thus,

$$X = r \begin{pmatrix} -6 \\ 2 \\ 1 \\ 0 \\ 0 \end{pmatrix} + s \begin{pmatrix} 8 \\ -1 \\ 0 \\ 1 \\ 0 \end{pmatrix} + t \begin{pmatrix} -1 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \Rightarrow \text{Nul}(A) = \text{Span} \left\{ \begin{pmatrix} -6 \\ 2 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 8 \\ -1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\}$$

And, once again, because the generating set is composed of the solutions to  $A\vec{x} = 0$ , we know that it is linearly independent and thus forms a basis.

## 4.3 Lecture 16: Linearly Independent Sets, Basis [4.3]

## Linear Independence

A set of vectors  $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k\}$  is **linearly independent** if the equation

$$c_1\vec{v}_1 + c_2\vec{v}_2 + \dots + c_k\vec{v}_k, \quad c_i \in \mathbb{R}$$

is satisfied if, and only if,  $c_1 = c_2 = \dots = c_k = 0$ . Otherwise, the set is **linearly dependent**.

We can determine if a set is linearly independent by considering the equation  $AC = 0$  where

$$A_{n \times k} = \begin{pmatrix} | & | & \dots & | \\ \vec{v}_1 & \vec{v}_2 & \dots & \vec{v}_k \\ | & | & \dots & | \end{pmatrix} \text{ and } C_{k \times 1} = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_k \end{pmatrix}.$$

Then,

1. if  $\text{rank}(A) = k$ , then  $AC = 0$  has only the trivial solution  $C = 0$ , so the set is linearly *independent*.
2. if  $\text{rank}(A) < k$ , then  $AC = 0$  has an infinite amount of solutions so the set is linearly *dependent*.

$\text{Span}\{\vec{u}_1, \vec{u}_2, \dots, \vec{u}_k\}$  is a subspace of  $\mathbb{R}^n$  if  $\vec{u}_1, \vec{u}_2, \dots, \vec{u}_k \in \mathbb{R}^n$

### Basis of a Subspace

The basis of a subspace is the smallest linearly independent generating set whose span is the subspace.

#### Example 4.3.1

Find a basis of  $\mathcal{W}$

$$\mathcal{W} = \text{Span} \left\{ \begin{pmatrix} 1 \\ -2 \\ 3 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} \right\}$$

**Solution:** Consider the matrix whose columns are the vectors in the generating set of  $\mathcal{W}$ :

$$A = \begin{pmatrix} 1 & 1 & 2 & 1 \\ -2 & 1 & 1 & 0 \\ 3 & -1 & 3 & 2 \end{pmatrix} \xrightarrow[R_3 \rightarrow R_3 - 3R_1]{R_2 \rightarrow R_2 + 2R_1} \begin{pmatrix} 1 & 1 & 2 & 1 \\ 0 & 3 & 5 & 2 \\ 0 & -4 & -3 & -1 \end{pmatrix} \xrightarrow{R_3 \rightarrow 3R_3 + 4R_2} \begin{pmatrix} 1 & 1 & 2 & 1 \\ 0 & 3 & 5 & 2 \\ 0 & 0 & 11 & 5 \end{pmatrix}$$

Now, since there isn't a pivot in the fourth column, the fourth vector in the generating set is redundant. Thus, one possible basis for  $\mathcal{W}$  is

$$\left\{ \begin{pmatrix} 1 \\ -2 \\ 3 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix} \right\}$$

You may also note that  $\mathcal{W}$  from the above example is a subspace of  $\mathbb{R}^3$  and so  $\text{Span}\{\mathcal{W}\} = \mathbb{R}^3$ .

### Standard Basis of $\mathbb{R}^n$

The **standard basis** of  $\mathbb{R}^n$  is the set

$$\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$$

where  $\mathbf{e}_k$  is the  $k$ -th column of  $I_n$  (the identity matrix in  $\mathbb{R}^n$ ).

#### Example 4.3.2

Consider

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

(a) Determine if  $\mathcal{S}$  is linearly independent.



**Solution:** By definition of linear independence we must have

$$\begin{aligned} c_1(1 - x^2 + 2x^3) + c_2(1 + x) + c_3(x^2) + c_4(x^4) &= 0 \\ \Rightarrow (c_1 + c_2) + (c_2)x + (-c_1 + c_3)x^2 + (c_1)x^3 + (c_4)x^4 &= 0 \\ \Rightarrow c_1 = c_2 = c_3 = c_4 = 0 \end{aligned}$$

Thus,  $\mathcal{S}$  is linearly independent.

(b) Determine if  $\text{Span}\{\mathcal{S}\} = \mathbb{P}_4$

**Solution:** Since there are only four vectors,  $\text{Span}\{\mathcal{S}\} \neq \mathbb{P}_4$ .

Our answer in part b of the above example follows directly from the fact that  $\mathbb{P}_n \cong \mathbb{R}^{n+1}$  ( $\mathbb{P}_n$  is *isomorphic* to  $\mathbb{R}^{n+1}$ ), and in order for a set of vectors in  $\mathbb{R}^n$  to span  $\mathbb{R}^n$ , there must be exactly  $n$  vectors in the generating set. Now, because of this fact, we actually have an easier way to solve part a, rather than splitting up each term by degrees.

We can notice that the standard basis for  $\mathbb{P}_4$  is

$$\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ x \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ x^2 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ x^3 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ x^4 \end{pmatrix} \right\}$$

since any polynomial in  $\mathbb{P}_4$  can be made up as a linear combination of these vectors. If we modify this basis slightly to get

$$\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, x \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, x^2 \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, x^3 \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, x^4 \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\}$$

then we can see how we can represent  $\mathcal{S}$  in the following manner:

$$\mathcal{S} = \left\{ \begin{pmatrix} 1 \\ 0 \\ -1 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} \right\}.$$

Now, consider the matrix whose columns are the vectors from  $\mathcal{S}$ :

$$A = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \xrightarrow[R_4 \rightarrow R_4 - 2R_1]{R_3 \rightarrow R_3 + R_1} \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & -2 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \xrightarrow[R_4 \rightarrow R_4 + 2R_2]{R_3 \rightarrow R_3 - R_2} \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \xrightarrow[R_4 \leftrightarrow R_5]{R_4 \rightarrow R_4 - 2R_3} \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

So,  $\text{rank}(A) = 4 = k$  so  $\mathcal{S}$  is linearly independent.

#### 4.4 Lecture 17: Coordinates in Vector Spaces [4.4]

**Theorem 4.3** (Uniqueness Theorem). *Consider a basis  $\{\vec{b}_1, \vec{b}_2, \dots, \vec{b}_n\}$  for  $\mathbb{R}^n$  and let  $\vec{v} \in \mathbb{R}^n$ . Then if*

$$\vec{v} = c_1 \vec{b}_1 + c_2 \vec{b}_2 + \dots + c_n \vec{b}_n, \quad (8)$$

*$c_1, c_2, \dots, c_n$  is a unique set of numbers for  $\vec{v}$ .*

*Proof.* Suppose  $\vec{v} \in \mathbb{R}^n$  and  $\mathbb{R}^n$  has a basis  $\{\vec{b}_1, \vec{b}_2, \dots, \vec{b}_n\}$ . Now, assume that there are two different representations of  $\vec{v}$ :

$$\begin{aligned} \vec{v} &= c_1 \vec{b}_1 + c_2 \vec{b}_2 + \dots + c_n \vec{b}_n \\ \vec{v} &= \tilde{c}_1 \vec{b}_1 + \tilde{c}_2 \vec{b}_2 + \dots + \tilde{c}_n \vec{b}_n. \end{aligned}$$

Then,

$$0 = (c_1 - \tilde{c}_1) \vec{b}_1 + (c_2 - \tilde{c}_2) \vec{b}_2 + \dots + (c_n - \tilde{c}_n) \vec{b}_n.$$

However, this is only possible if, and only if,  $c_i = \tilde{c}_i$  for  $0 \leq i \leq n$ . Thus,  $c_1, c_2, \dots, c_n$  is unique.  $\square$

Consider equation (8) in the form

$$\vec{v} = \hat{B}C \text{ where } \hat{B} = \begin{pmatrix} | & | & & | \\ \vec{b}_1 & \vec{b}_2 & \dots & \vec{b}_n \\ | & | & & | \end{pmatrix} \text{ and } C = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix}.$$

We now re-label  $C$  to be  $[v]_B$  and call it the **vector coordinate of  $\vec{v}$  in  $B$** . Then, since  $\hat{B}$  is invertible,

$$\vec{v} = \hat{B}[v]_B \Rightarrow [v]_B = \hat{B}^{-1}\vec{v}$$

##### Example 4.4.1

Find the vector coordinate of  $\vec{v}$  in  $\hat{B}$  if

$$\hat{B} = \left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix} \right\} \text{ and } \vec{v} = \begin{pmatrix} 1 \\ -4 \\ 4 \end{pmatrix}$$

**Solution:** We can find  $(\hat{B} | \vec{v})$  and use it to solve  $\hat{B}[\vec{v}]_B = \vec{v}$ .

$$(\hat{B} | \vec{v}) = \left( \begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 1 & -1 & 2 & -4 \\ 1 & 1 & 2 & 4 \end{array} \right) \xrightarrow[R_3 \rightarrow R_3 - R_1]{R_2 \rightarrow R_2 - R_1} \left( \begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & -2 & 1 & -5 \\ 0 & 0 & 1 & 3 \end{array} \right) = (\hat{B} | \vec{v})_{\text{REF}}.$$

Using this we can solve the system

$$\begin{bmatrix} c_3 & = & 3 \\ -2c_2 + c_3 & = & -5 \Rightarrow c_2 = 4 \\ c_1 + c_2 + c_3 & = & 1 \Rightarrow c_1 = -6 \end{bmatrix} \Rightarrow [\vec{v}]_B = \begin{pmatrix} -6 \\ 4 \\ 3 \end{pmatrix}$$

**Example 4.4.2**

Consider the set  $\hat{B}$  below as a basis for  $\mathbb{P}_2$ . Find the coordinates of  $\mathbf{p}(t) = 3 + t - 6t^2$  relative to  $\hat{B}$ .

$$\hat{B} = \left\{ \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 2 \\ -2 \\ 1 \end{pmatrix} \right\}$$

**Solution:** This time around, we will find  $\hat{B}^{-1}$  and use it to solve for  $[\vec{v}]_{\hat{B}}$ .

$$\begin{aligned} \left( \begin{array}{ccc|ccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 0 & 1 & -2 & 0 & 1 & 0 \\ -1 & -1 & 1 & 0 & 0 & 1 \end{array} \right) &\xrightarrow{R_3 \rightarrow R_3 + R_1} \left( \begin{array}{ccc|ccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 0 & 1 & -2 & 0 & 1 & 0 \\ 0 & -1 & 3 & 1 & 0 & 1 \end{array} \right) \\ &\xrightarrow{R_3 \rightarrow R_3 + R_2} \left( \begin{array}{ccc|ccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 0 & 1 & -2 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 \end{array} \right) \\ &\xrightarrow{R_1 \rightarrow R_1 + R_2} \left( \begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & -2 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 \end{array} \right) \\ &\xrightarrow{R_3 \rightarrow R_2 + 2R_3} \left( \begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 2 & 3 & 2 \\ 0 & 0 & 1 & 1 & 1 & 1 \end{array} \right) \\ &\Rightarrow \hat{B}^{-1} = \begin{pmatrix} 1 & 1 & 0 \\ 2 & 3 & 2 \\ 1 & 1 & 1 \end{pmatrix}. \end{aligned}$$

Thus,

$$[\vec{v}]_{\hat{B}} = \hat{B}^{-1}\vec{v} = \begin{pmatrix} 4 \\ -3 \\ -2 \end{pmatrix}$$

**4.5 Lecture 18: The Dimension of a Vector Space and Change of Basis [4.5-4.6]**

**Theorem 4.4.** If a vector space  $V$  has a basis  $\mathcal{B} = \{\vec{b}_1, \dots, \vec{b}_n\}$ , then any set in  $V$  containing more than  $n$  vectors must be linearly dependent.

**Theorem 4.5.** If a vector  $V$  has a basis of  $n$  vectors, then every basis of  $V$  must consist of exactly  $n$  vectors.

If a vector space  $V$  is spanned by a finite set, then  $V$  is said to be have finite dimension, and the dimension of  $V$ , written as  $\dim V$  is the number of vectors in a basis for  $V$ . If  $V$  is not spanned by a finite set, then  $V$  is said to be infinite-dimensional.

The dimension of the zero vector space,  $\{\vec{0}\}$  is defined to be 0. The standard basis for  $\mathbb{R}^n$  has  $n$  vectors, so  $\dim[\mathbb{R}^n] = n$ . The standard basis for  $\mathbb{P}_n$  has  $n + 1$  vectors, so  $\dim[\mathbb{P}_2] = n + 1$ .

**Theorem 4.6.** *Let  $H$  be a subspace of a finite dimensional vector space  $V$ . Any linearly independent set in  $H$  can be expanded to a basis for  $H$ . Also,  $H$  is finite dimensional and  $\dim H \leq \dim V$ .*

**Theorem 4.7.** *Let  $V$  be a  $p$ -dimensional vector space with  $p \geq 1$ . Any linearly independent set of exactly  $p$  elements in  $V$  is automatically a basis for  $V$ . Any set of exactly  $p$  elements that spans  $V$  is automatically a basis for  $V$ .*

The rank of an  $m \times n$  matrix  $A$ ,  $\text{rank}(A)$ , is the dimension of the column space. The nullity of  $A$ ,  $\text{nul}(A)$ , is the dimension of the null space.

The nullity of  $A$  is also equal to the number of free variables in the homogenous system  $A\vec{x} = \vec{0}$ .

$\text{rank}(A) + \text{nul}(A) = \text{the number of columns in } A$ .

From these above theorems and facts we can extend our understanding of invertible matrices.

**Theorem 4.8.** *Let  $A$  be an  $n \times n$  matrix. If  $A$  is invertible then,*

1.  $\text{Col}(A)$  is a basis of  $\mathbb{R}^n$
2.  $\text{Col}(A) = \mathbb{R}^n$
3.  $\text{rank}(A) = n$
4.  $\text{nul}(A) = 0$
5.  $\text{Nul}(A) = \{0\}$

### Change of Coordinates

Let  $\mathcal{B} = \{b_1, \dots, b_n\}$  and  $\mathcal{C} = \{c_1, \dots, c_n\}$  be basis of a vector space  $V$ . Then, there exists a unique  $n \times n$  matrix  ${}_{\mathcal{C} \leftarrow \mathcal{B}}P$  such that

$$[\vec{x}]_{\mathcal{C}} = {}_{\mathcal{C} \leftarrow \mathcal{B}}P [\vec{x}]_{\mathcal{B}}.$$

The columns of  ${}_{\mathcal{C} \leftarrow \mathcal{B}}P$  are the  $\mathcal{C}$ -coordinate vectors of the vectors in the basis  $\mathcal{B}$ . That is,

$${}_{\mathcal{C} \leftarrow \mathcal{B}}P = ([b_1]_{\mathcal{C}} \quad [b_2]_{\mathcal{C}} \quad \cdots \quad [b_n]_{\mathcal{C}}).$$

The matrix  ${}_{\mathcal{C} \leftarrow \mathcal{B}}P$  is called the change-of-coordinates matrix from  $\mathcal{B}$  to  $\mathcal{C}$ .

### Example 4.5.1

Consider the bases  $\mathcal{B}$  and  $\mathcal{C}$  below. Find  ${}_{\mathcal{C} \leftarrow \mathcal{B}}P$ .

$$\mathcal{B} = \left\{ \begin{pmatrix} -9 \\ 1 \end{pmatrix}, \begin{pmatrix} -5 \\ -1 \end{pmatrix} \right\} \text{ and } \mathcal{C} = \left\{ \begin{pmatrix} 1 \\ -4 \end{pmatrix}, \begin{pmatrix} 3 \\ -5 \end{pmatrix} \right\}$$

**Solution:** Consider the vectors in  $\mathcal{B}$  to be  $b_1$  and  $b_2$  and similarly for  $\mathcal{C}$  to be  $c_1$  and  $c_2$ . We must find  $[b_1]_{\mathcal{C}}$  and  $[b_2]_{\mathcal{C}}$ . If we let  $[b_1]_{\mathcal{C}} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$  and  $[b_2]_{\mathcal{C}} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$  we should have the following relationship:

$$\begin{aligned} b_1 &= c_1 x_1 + c_2 x_2 = ((c_1) \ (c_2)) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \\ b_2 &= c_1 y_1 + c_2 y_2 = ((c_1) \ (c_2)) \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}. \end{aligned}$$

Now, we can solve both these systems simultaneously by row reducing the matrix

$$((c_1) \ (c_2) \mid (b_1) \ (b_2)).$$

We do this as follows

$$\begin{aligned} ((c_1) \ (c_2) \mid (b_1) \ (b_2)) &= \left( \begin{array}{cc|cc} 1 & 3 & -9 & -5 \\ -4 & -5 & 1 & -1 \end{array} \right) \\ &\xrightarrow{R_2 \rightarrow R_2 + 4R_1} \left( \begin{array}{cc|cc} 1 & 3 & -9 & -5 \\ 0 & 7 & -35 & -21 \end{array} \right) \\ &\xrightarrow{R_2 \rightarrow \frac{1}{7}R_2} \left( \begin{array}{cc|cc} 1 & 3 & -9 & -5 \\ 0 & 1 & -5 & -3 \end{array} \right) \\ &\xrightarrow{R_1 \rightarrow R_1 - 3R_2} \left( \begin{array}{cc|cc} 1 & 0 & 6 & 4 \\ 0 & 1 & -5 & -3 \end{array} \right). \end{aligned}$$

Thus,

$$[b_1]_{\mathcal{C}} = \begin{pmatrix} 6 \\ -5 \end{pmatrix} \text{ and } [b_2]_{\mathcal{C}} = \begin{pmatrix} 4 \\ -3 \end{pmatrix}.$$

So,

$$\boxed{{}_{\mathcal{C} \leftarrow \mathcal{B}} P = \begin{pmatrix} 6 & 4 \\ -5 & -3 \end{pmatrix}}$$

The method used here holds in most cases. In general we have

$$((c_1) \ (c_2) \mid (b_1) \ (b_2)) \sim \left( I \mid {}_{\mathcal{C} \leftarrow \mathcal{B}} P \right)$$

#### Example 4.5.2

Consider the bases for  $\mathbb{R}^2$  given by  $\mathcal{B} = \{b_1, b_2\}$  and  $\mathcal{C} = \{c_1, c_2\}$  below.

$$\mathcal{B} = \left\{ \begin{pmatrix} 1 \\ -3 \end{pmatrix}, \begin{pmatrix} -2 \\ 4 \end{pmatrix} \right\} \text{ and } \mathcal{C} = \left\{ \begin{pmatrix} -7 \\ 9 \end{pmatrix}, \begin{pmatrix} -5 \\ 7 \end{pmatrix} \right\}$$

(a) Find  ${}_{\mathcal{B} \leftarrow \mathcal{C}} P$

**Solution:** Consider

$$\left( (b_1) \quad (b_2) \mid (c_1) \quad (c_2) \right).$$

Then,

$$\begin{aligned} & \left( \begin{array}{cc|cc} 1 & -2 & -7 & -5 \\ -3 & 4 & 9 & 7 \end{array} \right) \\ & \xrightarrow{R_2 \rightarrow R_2 + 3R_1} \left( \begin{array}{cc|cc} 1 & -2 & -7 & -5 \\ 0 & -2 & -12 & -8 \end{array} \right) \\ & \xrightarrow[R_1 \rightarrow R_1 + R_2]{R_2 \rightarrow -\frac{1}{2}R_2} \left( \begin{array}{cc|cc} 1 & 0 & 5 & 3 \\ 0 & 0 & 6 & 4 \end{array} \right). \end{aligned}$$

Thus,

$$\boxed{{}_B P_{B \leftarrow C} = \begin{pmatrix} 5 & 3 \\ 6 & 4 \end{pmatrix}}$$

(b) Find  ${}_C P_{C \leftarrow B}$

**Solution:** Before we start consider the relationship

$${}_C P_{C \leftarrow B} [b_1]_B = [b_1]_C$$

and multiply by  $\left({}_C P_{C \leftarrow B}\right)^{-1}$  and notice that we now have

$$[b_1]_B = \left({}_C P_{C \leftarrow B}\right)^{-1} [b_1]_C.$$

Therefore,  ${}_B P_{B \leftarrow C} = \left({}_C P_{C \leftarrow B}\right)^{-1}$ ! We can apply this property to our problem to see

$${}_B P_{B \leftarrow C} = \left({}_C P_{C \leftarrow B}\right)^{-1} = \frac{1}{2} \begin{pmatrix} 4 & -3 \\ -6 & 5 \end{pmatrix}$$

## 5 Unit 5

### 5.1 Lecture 19: Eigenvalues and Eigenvectors

Let  $A$  be any  $n \times n$  matrix. A non-zero vector  $\vec{S}$  is called an **eigenvector** of  $A$  is  $A\vec{S} = \lambda\vec{S}$  for  $\lambda \in \mathbb{C}$ . Such  $\lambda$  is called the corresponding **eigenvalue** of  $A$ .

#### Example 5.1.1

Given  $A$  and  $\vec{S}$  below, verify  $\vec{S}$  is an eigenvector of  $A$  and find the corresponding eigenvalue of  $A$ .

$$A = \begin{pmatrix} -5 & -4 \\ 8 & 7 \end{pmatrix} \text{ and } \vec{S} = \begin{pmatrix} 1 \\ -2 \end{pmatrix}$$

**Solution:**

$$\begin{aligned} A\vec{S} &= \begin{pmatrix} -5 & -4 \\ 8 & 7 \end{pmatrix} \begin{pmatrix} 1 \\ -2 \end{pmatrix} \\ &= \begin{pmatrix} 3 \\ -6 \end{pmatrix} \\ &= 3\vec{S} \\ &\Rightarrow \boxed{\lambda = 3} \end{aligned}$$

#### Example 5.1.2

Given  $A$  below and  $\lambda = 2$ , find the eigenvector

$$A = \begin{pmatrix} 5 & -1 \\ 3 & 1 \end{pmatrix}$$

**Solution:** Since  $\lambda$  is an eigenvalue it must satisfy the equation

$$A\vec{S} = 2\vec{S} \Rightarrow A\vec{S} - 2\vec{S} = 0$$

However, since 2 is a scalar we can't factor out  $\vec{S}$ . We can fix this by multiplying by the identity matrix  $I$ . This gives

$$I(A\vec{S} - 2\vec{S}) = AI\vec{S} - 2I\vec{S} = (A - 2I)\vec{S} = 0.$$

Since  $\vec{S}$  is non-zero by definition, we know that  $A - 2I = 0$ . So,

$$\begin{pmatrix} 5 & -1 \\ 3 & 1 \end{pmatrix} - 2 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 3 & -1 \\ 3 & -1 \end{pmatrix} \xrightarrow{R_2 \rightarrow R_2 - R_1} \begin{pmatrix} 3 & -1 \\ 0 & 0 \end{pmatrix} \\ \Rightarrow X = t \begin{pmatrix} 1 \\ 3 \end{pmatrix}.$$

Thus, our eigenvector is

$$\begin{pmatrix} 1 \\ 3 \end{pmatrix}$$

You may notice that in the above example the solution to  $A - 2I$  was

$$t \begin{pmatrix} 1 \\ 3 \end{pmatrix}.$$

Since this generates vectors for all non-zero values of  $t$  we call this the **eigenspace**.

Its important to realize that for any  $n \times n$  matrix  $A$ , if  $\det A = 0$  then  $\det A - \lambda I = 0$ . If  $\det A = 0$  then  $\text{rank}(A) < n$  implies that the equation  $A\mathbf{x} = 0$  has *non-trivial* solutions which is exactly what we are in search of. From this discussion we can now learn about the characteristic equation.

### The Characteristic Equation

The **characteristic equation** of an  $n \times n$  matrix  $A$  is

$$\det A - \lambda I = 0$$

and its solutions are the eigenvalues of  $A$ .

It follows from the above definition that any characteristic equation will be a polynomial of degree  $n$  and thus  $A$  will have  $n$  eigenvectors (including multiplicity).

To see this in example, lets determine the characteristic equation for any  $2 \times 2$  matrix. Let  $A$  be the  $2 \times 2$  matrix below

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

We can find the characteristic equation using the formula provided in the definition to give,

$$\det A - \lambda I_2 = \begin{vmatrix} a - \lambda & b \\ c & d - \lambda \end{vmatrix} = (a - \lambda)(d - \lambda) - bc = \lambda^2 - \lambda(a + d) + (ad - bc). \quad (9)$$

This process can be repeated for any  $n \times n$  matrix to find its general characteristic equation.



## How to find Eigenvectors and Eigenvalues

Let  $A$  be an  $n \times n$  matrix. To find the eigenvalues and eigenvectors of  $A$ :

1. Find the eigenvalues by solving  $\det(A - \lambda I) = 0$  (the characteristic equation)
2. For each eigenvalue, find the corresponding eigenvector(s)  $y$  solving  $A - \lambda I = 0$ .

Its also important to note that for a triangular matrix (upper or lower), its eigenvalues are simply the entires on the main diagonal.

## Example 5.1.3

Find the eigenvalues and eigenvectors of  $A$  below

$$A = \begin{pmatrix} 5 & -1 \\ 3 & 1 \end{pmatrix}$$

**Solution:** We can find the characteristic equation by solving  $\det A - \lambda I = 0$ . This gives

$$\begin{vmatrix} 5 - \lambda & -1 \\ 3 & 1 - \lambda \end{vmatrix} = \lambda^2 - 6\lambda + 8 = 0 \Rightarrow (\lambda - 4)(\lambda - 2) = 0 \Rightarrow \lambda_1 = 4, \lambda_2 = 2.$$

Thus we can find our eigenvectors:

$$\begin{aligned} \Rightarrow A - 4I &= \begin{pmatrix} 1 & -1 \\ 3 & -3 \end{pmatrix} \xrightarrow{R_2 \rightarrow R_2 - 3R_1} \begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix} \Rightarrow X = t \begin{pmatrix} 1 \\ 1 \end{pmatrix} \\ \Rightarrow A - 2I &= \begin{pmatrix} 3 & -1 \\ 3 & -1 \end{pmatrix} \xrightarrow{R_2 \rightarrow R_2 - R_1} \begin{pmatrix} 3 & -1 \\ 0 & 0 \end{pmatrix} \Rightarrow X = t \begin{pmatrix} 1 \\ 3 \end{pmatrix} \end{aligned}$$

Therefore for all  $t \neq 0$  our eigenvectors are

$$t \begin{pmatrix} 1 \\ 1 \end{pmatrix} \text{ and } t \begin{pmatrix} 1 \\ 3 \end{pmatrix}$$

## Eigenspace

The **eigenspace** is the null space of  $A - \lambda I$  without the zero vectors

## Example 5.1.4

If  $S_1 = \begin{pmatrix} 1 \\ 3 \end{pmatrix}$  for  $\lambda_1 = 2$  and  $S_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$  for  $\lambda_2 = 4$  what are the eigenspaces?

**Solution:** The eigenspace associated with  $\lambda_1$  is

$$\text{Span} \left\{ \begin{pmatrix} 1 \\ 3 \end{pmatrix} \right\} \setminus \{ \vec{0} \}.$$

The eigenspace associated with  $\lambda_2$  is

$$\text{Span} \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\} \setminus \{ \vec{0} \}.$$

Note that the symbol  $\setminus$  is the set minus symbol and all it means is exclusion from the set. So  $\text{Span} \{x\} \setminus \{ \vec{0} \}$  would be read as ‘the span of the vector  $\vec{x}$  without the zero vector’.

**Theorem 5.1.** *If  $\vec{v}_1, \dots, \vec{v}_r$  are the eigenvectors that correspond to the eigenvalues  $\lambda_1, \dots, \lambda_r$  for an  $n \times n$  matrix then the set  $\{\vec{v}_1, \dots, \vec{v}_r\}$  is linearly independent.*

## 5.2 Lecture 20: The Characteristic Equation and Complex Eigenvalues [5.2, 5.5]

Let  $A$  be an  $n \times n$  matrix. Then  $A$  is invertible if, and only if, 0 is **not** an eigenvalue of  $A$ .

### The Characteristic Equation

The **characteristic equation** is  $\det(A - \lambda I) = 0$ .

With this definition, we can now think of an eigenvalue of  $A$  as a number  $\lambda \in \mathbb{C}$  if, and only if,  $\lambda$  satisfies the characteristic equation. If you recall how to find eigenvalues, from (5.1), this should make logical sense.

If  $A$  is  $n \times n$  then  $\det(A - \lambda I)$  will be a polynomial of degree  $n$ , called the **characteristic polynomial**.

### Similarity

If  $A$  and  $B$  are  $n \times n$  matrices, then  $A$  is **similar to**  $B$  if there is an invertible matrix  $P$  such that  $P^{-1}AP = B$ , or equivalently,  $A = PBP^{-1}$ . Changing  $A$  into  $P^{-1}AP$  is called a **similarity transform**.

It's worth mentioning that  $A$  being similar to  $B$  implies that  $B$  is similar to  $A$ , but we just say that  $A$  is similar to  $B$ .

**Theorem 5.2.** *If  $n \times n$  matrices  $A$  and  $B$  are similar, then they have the same characteristic polynomial and hence the same eigenvalues with the same multiplicity.*

### Example 5.2.1

Let  $A$  be the matrix below. Find the eigenvalues of  $A$ , and find a basis for each eigenspace.

$$A = \begin{pmatrix} 0.5 & -0.6 \\ 0.75 & 1.1 \end{pmatrix}$$

**Solution:** Using the formula from equation (9) we can find the characteristic polynomial to be  $\lambda^2 - 1.6\lambda + 1 = 0$ . A quick glance should tell us that this is not factorable so using the quadratic equation we have

$$\lambda = \frac{1.6 \pm \sqrt{(-1.6)^2 - 4}}{2} = .8 \pm .6i.$$

Now, we follow the steps to find eigenvectors as normal

$$A - (.8 - .6i)I = \begin{pmatrix} -.3 + .6i & -.6 \\ .75 & .3 + .6i \end{pmatrix} \Rightarrow \begin{cases} (-.3 + .6i)x_1 - .6x_2 = 0 \\ .75x_1 + (.3 + .6i)x_2 = 0 \end{cases}$$

Since  $\lambda = .8 - .6i$  is an eigenvalue we know that the solutions to the homogenous system are non-trivial (i.e.  $X \neq 0$ ). This allows us to write  $x_1$  and  $x_2$  in terms of each other. We can use this to write

$$x_1 = (-.4 - .8i)x_2$$

from the second equation in the system. Thus our eigenvector (and basis) will be

$$S_1 = \begin{pmatrix} -2 - 4i \\ 5 \end{pmatrix}$$

Because complex numbers always come in conjugate pairs we can easily see that

$$S_2 = \begin{pmatrix} -2 + 4i \\ 5 \end{pmatrix}$$

If you aren't convinced, you can apply the same process as above using  $\lambda = .8 + .6i$ .

### 5.3 Lecture 22: Diagonalization [5.3]

Recall that a diagonal matrix is any matrix in the form

$$\begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & & \ddots & \\ 0 & \cdots & 0 & a_{nn} \end{pmatrix}.$$

In other words, a diagonal matrix is a matrix whos only non-zero entries lie on the main diagonal. For any diagonal matrix,  $D$ , we have the following fact:

$$D = \begin{pmatrix} a_{11} & \cdots & 0 \\ 0 & \ddots & \vdots \\ 0 & \cdots & a_{nn} \end{pmatrix} \Rightarrow D^k = \begin{pmatrix} (a_{11})^k & \cdots & 0 \\ 0 & \ddots & \vdots \\ 0 & \cdots & (a_{nn})^k \end{pmatrix}. \quad (10)$$

With this in mind, consider the following example:

**Example 5.3.1**

Let  $A$ ,  $P$ , and  $D$  be given below. Find a formula for  $A^k$  if  $A = PDP^{-1}$ .

$$A = \begin{pmatrix} 7 & 2 \\ -4 & 1 \end{pmatrix}, \quad D = \begin{pmatrix} 5 & 0 \\ 0 & 3 \end{pmatrix}, \quad \text{and } P = \begin{pmatrix} 1 & 1 \\ -1 & -2 \end{pmatrix}$$

**Solution:** Since  $A = PDP^{-1}$  it would be a good first step to find  $P^{-1}$ . We can do this quickly using the formula for the inverse of a  $2 \times 2$  matrix:

$$P^{-1} = \frac{1}{-1} \begin{pmatrix} -2 & -1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ -1 & -1 \end{pmatrix}.$$

Now before we jump straight to the formula for  $A^k$  let's consider a few concrete examples.

$$\begin{aligned} A^2 &= (PDP^{-1})(PDP^{-1}) = PDP^{-1}PDP^{-1} = PDDP^{-1} = PD^2P^{-1} \\ A^3 &= (PDP^{-1})(PD^2P^{-1}) = PDP^{-1}PD^2P^{-1} = PDD^2P^{-1} = PD^3P^{-1} \end{aligned}$$

Using these two examples, we can pretty clearly generalize for the pattern

$$A^k = PD^kP. \tag{11}$$

Now we can apply equation (10) to equation (11) in matrix form to see that we have

$$A^k = \begin{pmatrix} 1 & 1 \\ -1 & -2 \end{pmatrix} \begin{pmatrix} 5^k & 0 \\ 0 & 3^k \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -1 & -2 \end{pmatrix}.$$

Working our way from left to right multiplying the matrices will give us that

$$A^k = \begin{pmatrix} 2 \cdot 5^k - 3^k & 5^k - 3^k \\ 2 \cdot 3^k - 2 \cdot 5^k & 2 \cdot 3^k - 5^k \end{pmatrix}$$

**Diagonalizable Matrices**

A square matrix  $A$  is said to be **diagonalizable** if, and only if,  $A$  is similar to some diagonal matrix, that is  $A = PDP^{-1}$  for some invertible matrix  $P$  and some diagonal matrix  $D$ .

**Theorem 5.3.** *An  $n \times n$  matrix  $A$  is diagonalizable if, and only if,  $A$  has  $n$  linearly independent eigenvectors. In fact  $A = PDP^{-1}$  with  $D$  a diagonal matrix, if and only if, the columns of  $P$  are  $n$  linearly independent eigenvectors of  $A$ . In this case the diagonal entries of  $D$  are eigenvalues of  $A$  that correspond, respectively, to the eigenvectors in  $P$ .*

**Example 5.3.2**

Diagonalize the matrix below with characteristic equation  $-\lambda^3 - 3\lambda^2 + 4$

$$A = \begin{pmatrix} 1 & 3 & 3 \\ -3 & -5 & -3 \\ 3 & 3 & 1 \end{pmatrix}$$

**Solution:** We can start by finding the eigenvalues. Notice that  $-\lambda^3 - 3\lambda^2 + 4 = -(\lambda - 1)(\lambda + 2)^2$  so the eigenvalues are  $\lambda = 1$  and  $\lambda = -2$  (multiplicity 2). Next we can find our eigenvectors by solving  $A - \lambda I$  for each value of  $\lambda$ .

$$\begin{aligned} \lambda = 1 : \vec{v}_1 &= \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} \\ \lambda = -2 : \vec{v}_2 &= \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \text{ and } \vec{v}_3 = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \end{aligned}$$

We can now apply Theorem (5.3) to find  $P$ :

$$P = (\vec{v}_1 \quad \vec{v}_2 \quad \vec{v}_3) = \begin{pmatrix} 1 & -1 & -1 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$$

Once again, following Theorem (5.3) we can find  $D$  according to the order of eigenvectors used to construct  $P$ :

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

You'll notice that we repeated  $\lambda = -2$  twice because of its multiplicity. And thus,  $D$  is our diagonalized matrix!

**Theorem 5.4.** *An  $n \times n$  matrix with  $n$  distinct eigenvalues is diagonalizable.*

**Example 5.3.3**

Determine if  $A$  below is diagonalizable.

$$A = \begin{pmatrix} 5 & -8 & 1 \\ 0 & 0 & 7 \\ 0 & 0 & 2 \end{pmatrix}$$

**Solution:** Since  $A$  is an (upper) triangular matrix and square, we know that its eigenvalues are the diagonal entries with  $\lambda_1 = 5, \lambda_2 = 0$  and  $\lambda_3 = 2$ . Since  $A$  is a  $3 \times 3$  matrix with 3 distinct eigenvalues ( $\lambda_1 \neq \lambda_2 \neq \lambda_3$ ),  $A$  is diagonalizable.

So far we've been working with  $n \times n$  matrices that have  $n$  distinct eigenvalues and eigenvectors  $\vec{v}_1$  through  $\vec{v}_n$ . This makes  $P$  automatically invertible since its columns are linearly independent (by Theorem (5.1)). However, if a matrix has less than  $n$  distinct eigenvalues, we can still construct  $P$  in such a way to make it diagonalizable.

**Theorem 5.5.** *Let  $A$  be an  $n \times n$  matrix whose distinct eigenvalues are  $\lambda_1, \dots, \lambda_p$ .*

- 1. For  $1 \leq k \leq p$ , the dimension of the eigenspace for  $\lambda_k$  is less than or equal to the multiplicity of the eigenvalue  $\lambda_k$*
- 2. The matrix  $A$  is diagonalizable if, and only if, the sum of the dimensions of the eigenspaces equals  $n$ , and this happens if, and only if,*
  - (a) the characteristic polynomial factors completely into linear factors and*
  - (b) the dimension of the eigenspace for  $\lambda_k$  equal the multiplicity of  $\lambda_k$*
- 3. If  $A$  is diagonalizable and  $\mathcal{B}_k$  is a basis for the eigenspace corresponding to  $\lambda_k$  for each  $k$ , then the total collection of vectors in the sets  $\mathcal{B}_1, \dots, \mathcal{B}_p$  forms an eigenvector basis for  $\mathbb{R}^n$ .*

## 6 Unit 6

### 6.1 Lecture 23: Orthogonality, Orthogonal Projections [6.1, 6.2, 6.3]

**Theorem 6.1.** Let  $\vec{u}, \vec{v}, \vec{w} \in \mathbb{R}^n$  and  $c \in \mathbb{R}$ . Then,

1.  $\vec{u} \cdot \vec{v} = \vec{v} \cdot \vec{u}$
2.  $(\vec{u} + \vec{v}) \cdot \vec{w} = \vec{u} \cdot \vec{w} + \vec{v} \cdot \vec{w}$
3.  $(c\vec{u}) \cdot \vec{v} = c(\vec{u} \cdot \vec{v}) = \vec{u} \cdot (c\vec{v})$
4.  $\vec{u} \cdot \vec{u} \geq 0$ , and  $\vec{u} \cdot \vec{u} = 0$  if, and only if,  $\vec{u} = 0$

#### Length of a Vector

The **length** (or **norm**) of a vector  $\vec{v}$ , is the nonnegative scalar  $\|\vec{v}\|$  defined by

$$\|\vec{v}\| = \sqrt{\vec{v} \cdot \vec{v}} = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2} \text{ and } \|\vec{v}\|^2 = \vec{v} \cdot \vec{v}$$

A vector who has a length of 1 is called a *unit vector*. For any given vector, you can turn it into a unit vector by scaling it by the reciprocal of its magnitude (or length). If  $\vec{v}$  is any vector, then  $\hat{v} = \frac{1}{\|\vec{v}\|}\vec{v}$  is a unit vector in the same direction as  $\vec{v}$

#### Distance Between Two Vectors

For  $\vec{u}, \vec{v} \in \mathbb{R}^n$ , the **distance between  $\vec{u}$  and  $\vec{v}$** , written as  $\text{dist}(\vec{u}, \vec{v})$  is the length of the vector  $\vec{u} - \vec{v}$ . That is

$$\text{dist}(\vec{u}, \vec{v}) = \|\vec{u} - \vec{v}\|$$

#### Orthogonal Vectors

Two vector  $\vec{u}, \vec{v} \in \mathbb{R}^n$  are orthogonal to each other if  $\vec{u} \cdot \vec{v} = 0$

**Theorem 6.2** (The Pythagoren Theorem). Two vectors  $\vec{u}$  and  $\vec{v}$  are orthogonal if, and only if,  $\|\vec{u} + \vec{v}\|^2 = \|\vec{u}\|^2 + \|\vec{v}\|^2$

If a vector  $\vec{z}$  is orthogonal to every vector in a subspace  $W$  of  $\mathbb{R}^n$ , then  $\vec{z}$  is said to be **orthogonal to  $W$** . The set of all vectors  $\vec{z}$  that are orthogonal to  $W$  is called the **orthogonal complement** of  $W$  and is denoted by  $W^\perp$ . A vector  $\vec{x}$  is in  $W^\perp$  if, and only if,  $\vec{x}$  is orthogonal to every vector in a set that spans  $W$ .  $W^\perp$  is a subspace of  $\mathbb{R}^n$ .

**Theorem 6.3.** Let  $A$  be an  $m \times n$  matrix. The orthogonal complement of the row space of  $A$  is the null space of  $A$ , and the orthogonal complement of the column space of  $A$  is the null space of  $A^T$ :

$$(\text{Row}(A))^\perp = \text{Nul}(A) \text{ and } (\text{Col}(A))^\perp = \text{Nul}(A^T)$$

## Angle Between Two Vectors

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

$$\vec{u} \cdot \vec{v} = \|\vec{u}\| \|\vec{v}\| \cos \theta$$

where  $\theta$  is the angle between  $\vec{u}$  and  $\vec{v}$ .

## Orthogonal Sets

A set of vectors  $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_p\}$  in  $\mathbb{R}^n$  is said to be an **orthogonal set** if each pair of distinct vectors from the set is orthogonal, that is, if  $\vec{v}_i \cdot \vec{v}_j = 0$  for  $i \neq j$ .

A set  $\{\vec{u}_1, \dots, \vec{u}_p\}$  is an **orthonormal set** if it is an orthogonal set of unit vectors.

**Theorem 6.4.** *If  $S = \{\vec{u}_1, \dots, \vec{u}_p\}$  is an orthogonal set of nonzero vectors in  $\mathbb{R}^n$ , then  $S$  is linearly independent and hence is a basis for the subspace spanned by  $S$ .*

## Orthogonal Basis

An **orthogonal basis** for a subspace  $W$  of  $\mathbb{R}^n$  is a basis for  $W$  that is also an orthogonal set.

If  $W$  is such a subspace spanned by an orthonormal set  $\{\vec{u}_1, \dots, \vec{u}_p\}$ , then the set is an **orthonormal basis** for  $W$ . The standard basis for  $\mathbb{R}^n$  is an orthonormal basis.

**Theorem 6.5.** *Let  $\{\vec{u}_1, \dots, \vec{u}_p\}$  be an orthogonal basis for a subspace  $W$  of  $\mathbb{R}^n$ . For each  $\vec{y}$  in  $W$  the weights in the linear combination*

$$\vec{y} = c_1 \vec{u}_1 + \dots + c_p \vec{u}_p$$

are given by

$$c_j = \frac{\vec{y} \cdot \vec{u}_j}{\vec{u}_j \cdot \vec{u}_j} \text{ for } j = 1, \dots, p$$

The orthogonal projection of  $\vec{y}$  onto  $\vec{u}$  is  $\text{proj}_{\vec{u}} \vec{y} = \frac{\vec{y} \cdot \vec{u}}{\vec{u} \cdot \vec{u}} \vec{u} = \hat{y}$

**Theorem 6.6.** *An  $m \times n$  matrix  $U$  has orthonormal columns if, and only if,  $U^T U = I$*

**Theorem 6.7.** *Let  $U$  be an  $m \times n$  matrix with orthonormal columns, and let  $\vec{x}, \vec{y} \in \mathbb{R}^n$ . Then,*

1.  $\|U\vec{x}\| = \|\vec{x}\|$
2.  $(U\vec{x}) \cdot (U\vec{y}) = \vec{x} \cdot \vec{y}$
3.  $(U\vec{x}) \cdot (U\vec{y}) = 0$  if, and only if,  $\vec{x} \cdot \vec{y} = 0$ .

An **orthogonal matrix** is a square matrix  $U$  such that  $U^{-1} = U^T$ .

**Theorem 6.8** (The Orthogonal Decomposition Theorem). *Let  $W$  be a subspace of  $\mathbb{R}^n$ . Then each*



$\vec{y} \in \mathbb{R}^n$  can be written uniquely in the form

$$\vec{y} = \hat{y} + \vec{z}$$

where  $\hat{y} \in W$  and  $\vec{z} \in W^\perp$ . In fact, if  $\{\vec{u}_1, \dots, \vec{u}_p\}$  is any orthogonal basis of  $W$ , then

$$\hat{y} = \frac{\vec{y} \cdot \vec{u}_1}{\vec{u}_1 \cdot \vec{u}_1} \vec{u}_1 + \dots + \frac{\vec{y} \cdot \vec{u}_p}{\vec{u}_p \cdot \vec{u}_p} \vec{u}_p$$

and  $\vec{z} = \vec{y} - \hat{y}$ .

If  $\vec{y} \in W = \text{Span}\{\vec{u}_1, \dots, \vec{u}_p\}$  then  $\text{proj}_W \vec{y} = \vec{y}$

**Theorem 6.9** (Best Approximation Theorem). *Let  $W$  be a subspace of  $\mathbb{R}^n$ , let  $\vec{y} \in \mathbb{R}^n$ , and let  $\hat{y}$  be the orthogonal projection of  $\vec{y}$  onto  $W$ . Then  $\hat{y}$  is the closest point in  $W$  to  $\vec{y}$  in the sense that*

$$\|\vec{y} - \hat{y}\| < \|\vec{y} - \vec{v}\|$$

for all distinct  $\vec{v} \in W$

**Theorem 6.10.** *If  $\{\vec{u}_1, \dots, \vec{u}_p\}$  is an orthonormal basis for a subspace  $W$  of  $\mathbb{R}^n$ , then*

$$\text{proj}_W \vec{y} = (\vec{y} \cdot \vec{u}_1) \vec{u}_1 + (\vec{y} \cdot \vec{u}_2) \vec{u}_2 + \dots + (\vec{y} \cdot \vec{u}_p) \vec{u}_p$$

If  $U = \begin{pmatrix} \vec{u}_1 & \vec{u}_2 & \dots & \vec{u}_p \end{pmatrix}$ , then

$$\text{proj}_W \vec{y} = UU^T \vec{y}$$

for all  $\vec{y} \in \mathbb{R}^n$ .

## 6.2 Lecture 24: The Gram-Schmidt Process [6.4]

**Theorem 6.11** (The Gram-Schmidt Process). *Given a basis  $\{\vec{x}_1, \vec{x}_2, \dots, \vec{x}_p\}$  for a nonzero subspace  $W$  of  $\mathbb{R}^n$ , define*

$$\begin{aligned} \vec{v}_1 &= \vec{x}_1 \\ \vec{v}_2 &= \vec{x}_2 - \frac{\vec{x}_2 \cdot \vec{v}_1}{\vec{v}_1 \cdot \vec{v}_1} \vec{v}_1 \\ \vec{v}_3 &= \vec{x}_3 - \frac{\vec{x}_3 \cdot \vec{v}_1}{\vec{v}_1 \cdot \vec{v}_1} \vec{v}_1 - \frac{\vec{x}_3 \cdot \vec{v}_2}{\vec{v}_2 \cdot \vec{v}_2} \vec{v}_2 \\ &\vdots \\ \vec{v}_p &= \vec{x}_p - \frac{\vec{x}_p \cdot \vec{v}_1}{\vec{v}_1 \cdot \vec{v}_1} \vec{v}_1 - \frac{\vec{x}_p \cdot \vec{v}_2}{\vec{v}_2 \cdot \vec{v}_2} \vec{v}_2 - \dots - \frac{\vec{x}_p \cdot \vec{v}_{p-1}}{\vec{v}_{p-1} \cdot \vec{v}_{p-1}} \vec{v}_{p-1} \end{aligned}$$

Then  $\{\vec{v}_1, \dots, \vec{v}_p\}$  is an orthogonal basis for  $W$ . In addition,

$$\text{Span}\{\vec{v}_1, \dots, \vec{v}_k\} = \text{Span}\{\vec{x}_1, \dots, \vec{x}_k\} \quad \text{for } 1 \leq k \leq p$$

**Example 6.2.1**

Let  $\vec{x}_1, \vec{x}_2$ , and  $\vec{x}_3$  be given below such that  $\{\vec{x}_1, \vec{x}_2, \vec{x}_3\}$  is a basis for a subspace  $W$  of  $\mathbb{R}^4$ . Construct an orthogonal basis for  $W$ .

$$\vec{x}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \vec{x}_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \vec{x}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}$$

**Solution:** We will use the Gram-Schmidt process with  $\vec{x}_1, \vec{x}_2$ , and  $\vec{x}_3$ .

$$\vec{v}_1 = \vec{x}_1$$

$$\vec{v}_2 = \vec{x}_2 - \frac{\vec{x}_2 \cdot \vec{v}_1}{\vec{v}_1 \cdot \vec{v}_1} \vec{v}_1 = \vec{x}_2 - \frac{3}{4} \vec{v}_1 = \begin{pmatrix} -3/4 \\ 1/4 \\ 1/4 \\ 1/4 \end{pmatrix}$$

$$\vec{v}_2' = 4\vec{v}_2 = \begin{pmatrix} -3 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

$$\vec{v}_3 = \vec{x}_3 - \frac{\vec{x}_3 \cdot \vec{v}_1}{\vec{v}_1 \cdot \vec{v}_1} \vec{v}_1 - \frac{\vec{x}_3 \cdot \vec{v}_2'}{\vec{v}_2' \cdot \vec{v}_2'} \vec{v}_2' = \vec{x}_3 - \frac{2}{4} \vec{v}_1 - \frac{2}{12} \vec{v}_2' = \begin{pmatrix} 0 \\ -2/3 \\ 1/3 \\ 1/3 \end{pmatrix}$$

Thus our orthogonal basis is

$$\left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} -3 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ -2/3 \\ 1/3 \\ 1/3 \end{pmatrix} \right\}$$

You form an orthonormal basis by taking an orthogonal basis in its entirety and normalizing each  $\vec{u}_k$  to make it a unit vector, thus resulting in an orthonormal basis.

**Theorem 6.12** (*QR Factorization*). *If  $A$  is an  $m \times n$  matrix with linearly independent columns, then  $A$  can be factored as  $A = QR$  where  $Q$  is an  $m \times n$  matrix whose columns form an orthonormal basis for  $\text{Col}(A)$  and  $R$  is an  $n \times n$  upper triangular invertible matrix with positive entries on its diagonal.*

**Example 6.2.2**

Find a  $QR$  factorization of  $A$  below

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

**Solution:** Notice that the columns of  $A$  are actually  $\vec{x}_1, \vec{x}_2$ , and  $\vec{x}_3$  from example (6.2). We will scale  $\vec{v}_3$  to  $\vec{v}_3' = 3\vec{v}_3$  to simplify calculations. We can then normalize each of the three vectors to obtain the columns of  $Q$ :

$$Q = \begin{pmatrix} 1/2 & -3/\sqrt{12} & 0 \\ 1/2 & 1/\sqrt{12} & -2/\sqrt{6} \\ 1/2 & 1/\sqrt{12} & 1/\sqrt{6} \\ 1/2 & 1/\sqrt{12} & 1/\sqrt{6} \end{pmatrix}$$

By Theorem (6.6) we know that  $Q^T Q = I$ . Thus, if  $A = QR$  then left multiplying by  $Q^T$  gives:  $Q^T A = Q^T Q R = R$ . So,

$$R = \begin{pmatrix} 1/2 & 1/2 & 1/2 & 1/2 \\ -3/\sqrt{12} & 1/\sqrt{12} & 1/\sqrt{12} & 1/\sqrt{12} \\ 0 & -2/\sqrt{6} & 1/\sqrt{6} & 1/\sqrt{6} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 3/2 & 1 \\ 0 & 3/\sqrt{12} & 2/\sqrt{12} \\ 0 & 0 & 2/\sqrt{6} \end{pmatrix}$$

### 6.3 Lecture 24: The Least-Squares Problem [6.5]

#### Least-Square Solution

If  $A$  is  $m \times n$  and  $\vec{b}$  is in  $\mathbb{R}^n$ , a **least-squares solution** of  $A\vec{x} = \vec{b}$  is an  $\hat{x}$  in  $\mathbb{R}^n$  such that

$$\|\vec{b} - A\hat{x}\| \leq \|\vec{b} - A\vec{x}\|$$

for all  $\vec{x}$  in  $\mathbb{R}^n$

**Theorem 6.13.** *The set of least-square solutions of  $A\vec{x} = \vec{b}$  coincides with the nonempty set of solutions of the normal equation  $A^T A\vec{x} = A^T \vec{b}$*

#### Example 6.3.1

Find the least-squares solution of the inconsistent system  $A\vec{x} = \vec{b}$  with

$$A = \begin{pmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{pmatrix}, \quad \vec{b} = \begin{pmatrix} 2 \\ 0 \\ 11 \end{pmatrix}$$

**Solution:** We can start by finding  $A^T A$  and  $A^T \vec{b}$ :

$$\begin{aligned} A^T A &= \begin{pmatrix} 4 & 0 & 1 \\ 0 & 2 & 1 \end{pmatrix} \begin{pmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 17 & 1 \\ 1 & 5 \end{pmatrix} \\ A^T \vec{b} &= \begin{pmatrix} 4 & 0 & 1 \\ 0 & 2 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ 0 \\ 11 \end{pmatrix} = \begin{pmatrix} 19 \\ 11 \end{pmatrix} \end{aligned} \quad (12)$$

Now, it should be clear that columns of  $A^T A$  are linearly independent and thus  $A^T A$  is invertible. We can leverage this fact to solve for  $\hat{x}$  in the following manner:

$$A^T A \hat{x} = A^T \vec{b} \Rightarrow \hat{x} = (A^T A)^{-1} A^T \vec{b} \quad (13)$$

since left multiplying by  $(A^T A)^{-1}$  on the left side of the equation simply just gives the identity matrix. In order to use equation (13) we must first find  $(A^T A)^{-1}$ . Since we are dealing with a  $2 \times 2$  matrix we can simply just use the formula for the inverse of a  $2 \times 2$  (2.2). This gives

$$(A^T A)^{-1} = \frac{1}{84} \begin{pmatrix} 5 & -1 \\ -1 & 17 \end{pmatrix} \quad (14)$$

Thus, we can apply (13) with (12) and (14)

$$\hat{x} = \frac{1}{84} \begin{pmatrix} 5 & -1 \\ -1 & 17 \end{pmatrix} \begin{pmatrix} 19 \\ 11 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

**Theorem 6.14.** *Let  $A$  be any  $m \times n$  matrix. The following statements are logically equivalent:*

1. *The equation  $A\vec{x} = \vec{b}$  has a unique least-squares solution for each  $\vec{b}$  in  $\mathbb{R}^m$*
2. *The columns of  $A$  are linearly independent*
3. *The matrix  $A^T A$  is invertible*

*When these statements are true, the least-squares solution  $\hat{x}$  is given by*

$$\hat{x} = (A^T A)^{-1} A^T \vec{b}$$

### Least-Square Error

If  $\hat{x}$  is the least-square solution to the system  $A\vec{x} = \vec{b}$ , then the distance from  $\vec{b}$  to  $A\hat{x}$  is called the **least-square error** of this approximation, given by

$$\|\vec{b} - A\hat{x}\|$$

For example, in order to find the the least-square error of example (6.3) we first find  $A\hat{x}$  to be

$$A\hat{x} = \begin{pmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 4 \\ 4 \\ 3 \end{pmatrix}$$

Thus,

$$\vec{b} - A\hat{x} = \begin{pmatrix} 2 \\ 0 \\ 11 \end{pmatrix} - \begin{pmatrix} 4 \\ 4 \\ 3 \end{pmatrix} = \begin{pmatrix} -2 \\ -4 \\ 8 \end{pmatrix} \Rightarrow \|\vec{b} - A\hat{x}\| = \sqrt{84}$$

This tells us that for every vector  $\vec{x}$  in  $\mathbb{R}^2$  its distance between  $\vec{b}$  and  $A\vec{x}$  is **at least**  $\sqrt{84}$ .

**Theorem 6.15.** *Given an  $m \times n$  matrix  $A$  with linearly independent columns, let  $A = QR$  be a QR factorization of  $A$  (6.12). The, for each  $\vec{b}$  in  $\mathbb{R}^m$ , the equation  $A\vec{x} = \vec{b}$  has a unique least-squares solution, given by*

$$\hat{x} = R^{-1}Q^T\vec{b}$$