MA1508E Notes

Jansen Ken Pegrasio

February 2025

1 Week 1

1.1 What is an Augmented Matrix?

A linear system:

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

can be expressed uniquely as an augmented matrix:

$$\left[\begin{array}{cccc|c} a_{11} & a_{12} & ... & a_{1n} & b_1 \\ a_{21} & a_{22} & ... & a_{2n} & b_2 \\ a_{31} & a_{32} & ... & a_{3n} & b_3 \\ \vdots & & & & \\ a_{m1} & a_{m2} & ... & a_{mn} & b_m \end{array}\right]$$

1.2 Row-Echelon Form (REF)

An augmented matrix is in row-echelon form (REF) if

1. If zero rows (rows with all entries 0) exists, they are at the bottom of the matrix
2. The leading entries (the first nonzero entry of the row counting from left) are strictly further to the right as we move down the rows

1.3 Reduced Row-Echelon Form (RREF)

An augmented matrix is in reduced row-echelon form (RREF) if

1. The matrix is in REF
2. The leading entries are 1

3. In each pivot column (a column containing a leading entry), all entries except the leading entry is 0

1.4 Number of Solutions from REF

1. No solution (Inconsistent): if a row of zero before the bar and a non zero number after the bar
2. Unique solution: if all columns of coefficient matrix are pivot columns
3. Infinitely many solutions: if there is a non-pivot column in the augmented matrix before the bar

Note: It is not possible to have a number of solutions equal a finite number greater than 1

Tips: Assign parameters to the variables corresponding to the non-pivot columns in the LHS of the augmented matrix. As a consequence, the number of parameters needed is equal to the number of non-pivot columns in the LHS of the augmented matrix.

1.5 Elementary Row Operations

There are 3 types of elementary row operations:

1. Exchanging 2 rows, $R_i \leftrightarrow R_j$
2. Adding a multiple of a row to another, $R_i + cR_j, c \in \mathbb{R}$ and $i \neq j$
3. Multiplying a row by a nonzero constant, $aR_j, a \neq 0$

With these elementary row operations in mind, let's define what it means for two augmented matrices to be row equivalent.

Definition 1.5.1: Two augmented matrices are row equivalent if one can be obtained from the other by elementary row operations.

Corollary 1.5.2: Two linear systems have the same solutions if their augmented matrices are row equivalent.

1.6 Gaussian and Gauss-Jordan Elimination

1.6.1 Gaussian Elimination

Step 1: Locate the leftmost column that does not consist entirely of zeros

Step 2: Interchange the top row with another row, if necessary, to bring a nonzero entry to the top of the column found in step 1

Step 3: For each row below the top row, add a suitable multiple of the top row to it so that the entry below the leading entry of the top row becomes 0

Step 4: Now cover the top row in the augmented matrix and begin again with Step 1 applied to the submatrix that remains

By following along from Step 1 to 4, the entire matrix will end up to be in its row-echelon form.

1.6.2 Gauss-Jordan Elimination

Step 5: Multiply a suitable constant to each row so that all leading entries become 1

Step 6: Beginning with the last nonzero row and working upward, add suitable multiples of each row to the rows above to introduce zeros above the leading entries.

By performing Step 5 and 6, the entire matrix will then be in its reduced row-echelon form.

2 Week 2

2.1 Special types of Matrices

2.1.1 Vectors

There are two different types of vectors:

1. A $n \times 1$ matrix is called a (column) vector
2. A $1 \times n$ matrix is called a (row) vector.

2.1.2 Zero matrices

Zero matrices are matrices where all entries equal to 0, denoted as $\mathbf{0}_{m \times n}$

2.1.3 Square matrices

Square matrices are matrices which have the same number of rows and columns. Some definitions related to square matrices:

1. A size of $n \times n$ matrix is a square matrix of **order** n
2. The entries $a_{ii},\ i = 1, 2, 3, ..., n$, are called **diagonal entries** of the square matrix.

2.1.4 Diagonal matrices

Diagonal matrices are defined to be (a_{ij}), where $a_{ij} = 0$ for $i \neq j$. They are sometimes denoted as:

$$\operatorname{diag}(d_1, d_2, \ldots, d_n) = \begin{pmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n \end{pmatrix}$$

2.1.5 Scalar matrices

Scalar matrices are basically diagonal matrices that have the same diagonal value. They are sometimes denoted as:

$$\operatorname{diag}(c, c, \ldots, c) = \begin{pmatrix} c & 0 & \cdots & 0 \\ 0 & c & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & c \end{pmatrix}$$

2.1.6 Identity matrices

Identity matrices are basically scalar matrices that have the diagonal value equal to 1. They are usually denoted as

$$I_n = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

2.1.7 Triangular matrices

1. **Upper triangular** $A = (a_{ij})$, $a_{ij} = 0$ for all $i > j$:

$$A = \begin{pmatrix} * & * & \cdots & * \\ 0 & * & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & * \end{pmatrix}$$

2. **Strictly upper triangular** $A = (a_{ij})$, $a_{ij} = 0$ for all $i \geq j$:

$$A = \begin{pmatrix} 0 & * & \cdots & * \\ 0 & 0 & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}$$

3. **Lower triangular** $A = (a_{ij})$, $a_{ij} = 0$ for all $i < j$:

$$A = \begin{pmatrix} * & 0 & \cdots & 0 \\ * & * & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ * & * & \cdots & * \end{pmatrix}$$

4. **Strictly lower triangular** $A = (a_{ij})$, $a_{ij} = 0$ for all $i \leq j$:

$$A = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ * & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ * & * & \cdots & 0 \end{pmatrix}$$

2.2 Other Interesting Properties of Matrices

2.2.1 Multiplication and Transpose

$$(AB)^T = B^T A^T$$

2.2.2 Symmetric

A square matrix A is symmetric if and only if $A^T = A$.

3 Tutorial 1

3.1 Question 1(a)(i)

First, let's express the linear system in augmented matrix form, and reduce it to its row-echelon form.

$$\left[\begin{array}{ccc|c} 3 & 2 & -4 & 3 \\ 2 & 3 & 3 & 15 \\ 5 & -3 & 1 & 14 \end{array}\right] \xrightarrow{R_2 - \frac{2}{3}R_1} \left[\begin{array}{ccc|c} 3 & 2 & -4 & 3 \\ 0 & \frac{5}{3} & \frac{17}{3} & 13 \\ 5 & -3 & 1 & 14 \end{array}\right]$$
$$\xrightarrow{R_3 - \frac{5}{3}R_1} \left[\begin{array}{ccc|c} 3 & 2 & -4 & 3 \\ 0 & \frac{5}{3} & \frac{17}{3} & 13 \\ 0 & -\frac{19}{3} & \frac{23}{3} & 9 \end{array}\right] \xrightarrow{3R_2}\xrightarrow{3R_3} \left[\begin{array}{ccc|c} 3 & 2 & -4 & 3 \\ 0 & 5 & 17 & 39 \\ 0 & -19 & 23 & 27 \end{array}\right]$$
$$\xrightarrow{R_3 + \frac{19}{5}R_2} \left[\begin{array}{ccc|c} 3 & 2 & -4 & 3 \\ 0 & 5 & 17 & 39 \\ 0 & 0 & \frac{438}{5} & \frac{876}{5} \end{array}\right] \xrightarrow{5R_3} \left[\begin{array}{ccc|c} 3 & 2 & -4 & 3 \\ 0 & 5 & 17 & 39 \\ 0 & 0 & 438 & 876 \end{array}\right]$$

Now, let's reduce it further to its reduced row-echelon form.

$$\left[\begin{array}{ccc|c} 3 & 2 & -4 & 3 \\ 0 & 5 & 17 & 39 \\ 0 & 0 & 438 & 876 \end{array}\right] \xrightarrow{\frac{1}{438}R_3} \left[\begin{array}{ccc|c} 3 & 2 & -4 & 3 \\ 0 & 5 & 17 & 39 \\ 0 & 0 & 1 & 2 \end{array}\right]$$

$$\xrightarrow{R_2-17R_3} \left[\begin{array}{ccc|c} 3 & 2 & -4 & 3 \\ 0 & 5 & 0 & 5 \\ 0 & 0 & 1 & 2 \end{array}\right] \xrightarrow{\frac{1}{5}R_2} \left[\begin{array}{ccc|c} 3 & 2 & -4 & 3 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 2 \end{array}\right]$$

$$\xrightarrow{R_1-2R_2}\xrightarrow{R_1+4R_3} \left[\begin{array}{ccc|c} 3 & 0 & 0 & 9 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 2 \end{array}\right] \xrightarrow{\frac{1}{3}R_1} \left[\begin{array}{ccc|c} 1 & 0 & 0 & 3 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 2 \end{array}\right]$$

Therefore, we now have the solution of the linear equations which are $x_1 = 3, x_2 = 1, x_3 = 2$.

3.2 Question 3

Let's represent the linear system to the augmented matrix form. Then, reduce it to its row-echelon form.

$$\left[\begin{array}{ccc|c} a & 0 & b & 2 \\ a & a & 4 & 4 \\ 0 & a & 2 & b \end{array}\right] \xrightarrow{R_2-R_1} \left[\begin{array}{ccc|c} a & 0 & b & 2 \\ 0 & a & 4-b & 2 \\ 0 & a & 2 & b \end{array}\right] \xrightarrow{R_3-R_2} \left[\begin{array}{ccc|c} a & 0 & b & 2 \\ 0 & a & 4-b & 2 \\ 0 & 0 & b-2 & b-2 \end{array}\right]$$

The key indicator to determine the number of solutions and parameters is in the leading entries. There are 4 cases that are possible:

1. $a = 0$ and $b = 2$
2. $a = 0$ and $b \neq 2$
3. $a \neq 0$ and $b = 2$
4. $a \neq 0$ and $b \neq 2$

For case (1), here is how the matrix would look like:

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

As we can see here, there are two non-pivot columns, so there will be infinitely many solutions with two arbitrary parameters. (For **part d**).

For case (2), here is how the matrix would look like:

$$\left[\begin{array}{ccc|c} 0 & 0 & b & 2 \\ 0 & 0 & 4-b & 2 \\ 0 & 0 & b-2 & b-2 \end{array}\right] \xrightarrow{\frac{1}{b-2}R_3} \left[\begin{array}{ccc|c} 0 & 0 & b & 2 \\ 0 & 0 & 4-b & 2 \\ 0 & 0 & 1 & 1 \end{array}\right]$$

$$\xrightarrow{R_1-bR_3} \left[\begin{array}{ccc|c} 0 & 0 & 0 & 2-b \\ 0 & 0 & 4-b & 2 \\ 0 & 0 & 1 & 1 \end{array}\right] \xrightarrow{R_2-(4-b)R_3} \left[\begin{array}{ccc|c} 0 & 0 & 0 & 2-b \\ 0 & 0 & 0 & b-2 \\ 0 & 0 & 1 & 1 \end{array}\right]$$

Because $b \neq 2$, looking at the first and second row, we could conclude that the system is consistent. (For **part a**)

For case (3), here is how the matrix would look like:

$$\left[\begin{array}{ccc|c} a & 0 & 2 & 2 \\ 0 & a & 2 & 2 \\ 0 & 0 & 0 & 0 \end{array}\right] \xrightarrow{\frac{1}{a}R_1} \xrightarrow{\frac{1}{a}R_2} \left[\begin{array}{ccc|c} 1 & 0 & \frac{2}{a} & \frac{2}{a} \\ 0 & 1 & \frac{2}{a} & \frac{2}{a} \\ 0 & 0 & 0 & 0 \end{array}\right]$$

The third column is a non-pivot column, so let's assume $z = p$. Using this assumption, $x = \frac{2}{a} - \frac{2}{a}p$ and $y = \frac{2}{a} - \frac{2}{a}p$. We could therefore conclude that the system has infinitely many solutions with one arbitrary parameter. (For **part c**)

For case (4), let's reduce the matrix to its reduced row-echelon form.

$$\left[\begin{array}{ccc|c} a & 0 & b & 2 \\ 0 & a & 4-b & 2 \\ 0 & 0 & b-2 & b-2 \end{array}\right] \xrightarrow{\frac{1}{b-2}R_3} \left[\begin{array}{ccc|c} a & 0 & b & 2 \\ 0 & a & 4-b & 2 \\ 0 & 0 & 1 & 1 \end{array}\right]$$

$$\xrightarrow{R_1-bR_3} \left[\begin{array}{ccc|c} a & 0 & 0 & 2-b \\ 0 & a & 4-b & 2 \\ 0 & 0 & 1 & 1 \end{array}\right] \xrightarrow{R_2-(4-b)R_3} \left[\begin{array}{ccc|c} a & 0 & 0 & 2-b \\ 0 & a & 0 & b-2 \\ 0 & 0 & 1 & 1 \end{array}\right]$$

$$\xrightarrow{\frac{1}{a}R_1} \left[\begin{array}{ccc|c} 1 & 0 & 0 & \frac{2-b}{a} \\ 0 & a & 0 & b-2 \\ 0 & 0 & 1 & 1 \end{array}\right] \xrightarrow{\frac{1}{a}R_2} \left[\begin{array}{ccc|c} 1 & 0 & 0 & \frac{2-b}{a} \\ 0 & 1 & 0 & \frac{b-2}{a} \\ 0 & 0 & 1 & 1 \end{array}\right]$$

In this case, the linear system will have unique solution $x_1 = \frac{2-b}{a}$, $x_2 = \frac{b-2}{a}$, and $x_3 = 1$ where $a \neq 0$ and $b \neq 2$. (For **part b**)

3.3 Question 4

3.3.1 Part a

Let's assume the coefficient of each chemical substance!

$$x_1CO_2 + x_2H_2O \to x_3C_6H_{12}O_6 + x_4O_2$$

Now, let's formulate the linear system!

$$\begin{cases} x_1 - 6x_3 = 0 & \text{For C} \\ 2x_1 + x_2 - 6x_3 - 2x_4 = 0 & \text{For O} \\ 2x_2 - 12x_3 = 0 & \text{For H} \end{cases}$$

Now, let's represent this linear system as an augmented matrix!

$$\left[\begin{array}{cccc|c} 1 & 0 & -6 & 0 & 0 \\ 2 & 1 & -6 & -2 & 0 \\ 0 & 2 & -12 & 0 & 0 \end{array}\right] \to \left[\begin{array}{cccc|c} 1 & 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 6 & -1 & 0 \end{array}\right]$$

The fourth column is a non-pivot column, so let's assume $x_4 = p$. Therefore, using this information, we can find that $x_1 = p$, $x_2 = p$, and $x_3 = \frac{1}{6}p$.

Because $\frac{1}{6}p$ must be a positive integer. The smallest integer we can put in to p is 6. With this, $x_1 = 6$, $x_2 = 6$, $x_3 = 1$, $x_4 = 6$. Therefore, the chemical equation will be:

$$6CO_2 + 6H_2O \to C_6H_{12}O_6 + 6O_2$$

3.3.2 Part b

Let's assume the coefficient of each chemical substance!

$$x_1Fe_2O_3 + x_2Al \to x_3Al_2O_3 + x_4Fe$$

Now, let's formulate the linear system!

$$\begin{cases} 2x_1 - x_4 = 0 & \text{For Fe} \\ 3x_1 - 3x_3 = 0 & \text{For O} \\ x_2 - 2x_3 = 0 & \text{For Al} \end{cases}$$

Now, let's represent this linear system as an augmented matrix!

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

The fourth column is a non-pivot column, so let's assume $x_4 = p$. Therefore, using this information, we can find that $x_1 = \frac{1}{2}p$, $x_2 = p$, and $x_3 = \frac{1}{2}p$.

Because $\frac{1}{2}p$ must be a positive integer. The smallest integer we can put in to p is 2. With this, $x_1 = 1$, $x_2 = 2$, $x_3 = 1$, $x_4 = 2$. Therefore, the chemical equation will be:

$$Fe_2O_3 + 2Al \to Al_2O_3 + 2Fe$$

3.4 Question 5

$$\begin{cases} x_1 + x_3 = 800 & \text{Analyzing Point A} \\ x_1 - x_2 + x_4 = 200 & \text{Analyzing Point B} \\ x_2 - x_5 = 500 & \text{Analyzing Point C} \\ x_3 + x_6 = 750 & \text{Analyzing Point D} \\ x_4 + x_6 - x_7 = 600 & \text{Analyzing Point E} \\ x_5 - x_7 = -50 & \text{Analyzing Point F} \end{cases}$$

Now, let's represent this linear system as an augmented matrix!

$$\left[\begin{array}{ccccccc|c} 1 & 0 & 1 & 0 & 0 & 0 & 0 & 800 \\ 1 & -1 & 0 & 1 & 0 & 0 & 0 & 200 \\ 0 & 1 & 0 & 0 & -1 & 0 & 0 & 500 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 750 \\ 0 & 0 & 0 & 1 & 0 & 1 & -1 & 600 \\ 0 & 0 & 0 & 0 & 1 & 0 & -1 & -50 \end{array}\right] \rightarrow \left[\begin{array}{ccccccc|c} 1 & 0 & 0 & 0 & 0 & -1 & 0 & 50 \\ 0 & 1 & 0 & 0 & 0 & 0 & -1 & 450 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 750 \\ 0 & 0 & 0 & 1 & 0 & 1 & -1 & 600 \\ 0 & 0 & 0 & 0 & 1 & 0 & -1 & -50 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

3.4.1 Part a

There is no enough information to find unique solution for $x_1, x_2, x_3, x_4, x_5, x_6, x_7$ because there are two non-pivot columns, which are the sixth and seventh, indicating that we need 2 arbitrary parameters to define the solutions.

3.4.2 Part b

Assume $x_6 = p$ and $x_7 = q$! We will define x_1, x_2, x_3, x_4, x_5 using this information!

$$\begin{aligned} x_1 &= p + 50 \\ x_2 &= q + 450 \\ x_3 &= 750 - p \\ x_4 &= 600 - p + q \\ x_5 &= q - 50 \end{aligned}$$

It is known that $x_6 = 50$ and $x_7 = 100$. Thus, $x_1 = 100$, $x_2 = 550$, $x_3 = 700$, $x_4 = 650$, $x_5 = 50$.

3.4.3 Part c

The road between junctions A and B is x_1. So, we just need to remove all the existence of x_1 from the matrix.

$$\left[\begin{array}{ccccccc|c} 0 & 0 & 1 & 0 & 0 & 0 & 0 & 800 \\ 0 & -1 & 0 & 1 & 0 & 0 & 0 & 200 \\ 0 & 1 & 0 & 0 & -1 & 0 & 0 & 500 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 750 \\ 0 & 0 & 0 & 1 & 0 & 1 & -1 & 600 \\ 0 & 0 & 0 & 0 & 1 & 0 & -1 & -50 \end{array}\right] \rightarrow \left[\begin{array}{ccccccc|c} 0 & 1 & 0 & 0 & 0 & 0 & -1 & 450 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 800 \\ 0 & 0 & 0 & 1 & 0 & 0 & -1 & 650 \\ 0 & 0 & 0 & 0 & 1 & 0 & -1 & -50 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & -50 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

From R_5, we could conclude that $x_6 = -50$ where it would not possible to have a negative number of traffic flow. Thus, it is not possible!

4 Question 6

$$\begin{cases} T_1 = \dfrac{60+100+T_2+T_3}{4} \rightarrow 4T_1 = 60+100+T_2+T_3 \\ T_2 = \dfrac{100+T_1+40+T_4}{4} \rightarrow 4T_2 = 100+T_1+40+T_4 \\ T_3 = \dfrac{T_1+60+T_4+0}{4} \rightarrow 4T_3 = T_1+60+T_4 \\ T_4 = \dfrac{T_2+T_3+40+0}{4} \rightarrow 4T_4 = T_2+T_3+40 \end{cases}$$

Put constant in the RHS, T_1, T_2, T_3, T_4 in LHS.

$$\begin{cases} \quad 4T_1 - T_2 - T_3 = 160 \\ -T_1 + 4T_2 - T_4 = 140 \\ -T_1 + 4T_3 - T_4 = 60 \\ -T_2 - T_3 + 4T_4 = 40 \end{cases}$$

Let's represent this linear system as an augmented matrix, then reduce it to its reduced row-echelon form!

$$\left[\begin{array}{cccc|c} 4 & -1 & -1 & 0 & 160 \\ -1 & 4 & 0 & -1 & 140 \\ -1 & 0 & 4 & -1 & 60 \\ 0 & -1 & -1 & 4 & 40 \end{array}\right] \rightarrow \left[\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 65 \\ 0 & 1 & 0 & 0 & 60 \\ 0 & 0 & 1 & 0 & 40 \\ 0 & 0 & 0 & 1 & 35 \end{array}\right]$$

Using the information from the matrix, we can conclude that $T_1 = 65$, $T_2 = 60$, $T_3 = 40$, $T_4 = 35$.

5 Week 3

5.1 Matrix Equation

A linear system in standard form

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + ... + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + ... + a_{2n}x_n = b_2 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + a_{m3}x_3 + ... + a_{mn}x_n = b_m \end{cases}$$

can be expressed as a **matrix equation**

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

Technical terms, we call $(a_{ij})_{m \times n}$ the coefficient matrix, $(x_i)_{n \times 1}$ the variable vector, and $(b_i)_{m \times 1}$ the constant vector.

Moreover, we can also express the linear system as a **vector equation**:

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

We can also express the vector equation above as:

$$x_1\mathbf{a_1} + x_2\mathbf{a_2} + ... + x_n\mathbf{a_n} = \mathbf{b}$$

where a_i is called the coefficient vector for variable x_i for $1 \leq i \leq n$

5.2 Homogeneous Linear Systems

A homogeneous linear system $Ax = 0$ is always consistent, since the zero vector is a solution.

5.2.1 Trivial solution

The zero vector is called the **trivial solution**

5.2.2 Nontrivial solution

A nonzero solution to a homogeneous system is called a **nontrivial solution**.

5.2.3 A Theorem

A homogeneous linear system $\mathbf{Ax} = \mathbf{0}$ has infinitely many solutions if and only if it has a nontrivial solution.

Proof: Let $\mathbf{u} \neq 0$ be a solution to $\mathbf{Ax} = \mathbf{0}$, so $\mathbf{Au} = \mathbf{0}$. Then, for any $\mathbf{s} \in \mathbb{R}$:

$$\mathbf{A}(\mathbf{su}) = \mathbf{s}(\mathbf{Au}) = \mathbf{s} \cdot 0 = 0$$

By this, we can conclude that $\mathbf{su}$ is also a solution to $\mathbf{Ax} = \mathbf{0}$ – the linear system has infinitely many solutions.

5.2.4 A Lemma

Let $\mathbf{v}$ be a particular solution $\mathbf{Ax} = \mathbf{b}$, and $\mathbf{u}$ be a particular solution to the homogeneous system $\mathbf{Ax} = \mathbf{0}$ with the same coefficient matrix $\mathbf{A}$. Then, $\mathbf{v} + \mathbf{u}$ is also a solution to $\mathbf{Ax} = \mathbf{b}$.

Proof: Because $\mathbf{v}$ is a particular solution of $\mathbf{Ax} = \mathbf{b}$, this means $\mathbf{Av} = \mathbf{b}$.

With the same argument, this means $\mathbf{Au} = \mathbf{0}$. If we add this two equation together, we will have:

$$\begin{aligned} \mathbf{Av} + \mathbf{Au} &= \mathbf{b} + \mathbf{0} \\ \mathbf{A}(\mathbf{v} + \mathbf{u}) &= \mathbf{b} \end{aligned}$$

Therefore, by this, we can deduce that $\mathbf{v} + \mathbf{u}$ is also a solution to $\mathbf{Ax} = \mathbf{b}$.

5.2.5 A Lemma

Suppose $\mathbf{v_1}$ and $\mathbf{v_2}$ are solutions to the linear system $\mathbf{Ax} = \mathbf{b}$. Then, $\mathbf{v_1} - \mathbf{v_2}$ is a solution to the homogeneous linear system $\mathbf{Ax} = \mathbf{0}$ with the same coefficient matrix.

Proof: Because $\mathbf{v_1}$ and $\mathbf{v_2}$ are solutions to the linear system $\mathbf{Ax} = \mathbf{b}$, we have $\mathbf{Av_1} = \mathbf{b}$ and $\mathbf{Av_2} = \mathbf{b}$.

$$\mathbf{A}(\mathbf{v_1} - \mathbf{v_2}) = \mathbf{Av_1} - \mathbf{Av_2} = \mathbf{b} - \mathbf{b} = \mathbf{0}$$

By this chain of calculation, we can prove that $\mathbf{v_1} - \mathbf{v_2}$ is indeed a solution to $\mathbf{Ax} = \mathbf{0}$.

5.3 Submatrices

Let $\mathbf{A}$ be an $m \times n$ matrix,

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

Thus, the rows of $\mathbf{A}$ are:

$$\begin{aligned} r_1 &= \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \end{pmatrix} \\ r_2 &= \begin{pmatrix} a_{21} & a_{22} & \cdots & a_{2n} \end{pmatrix} \\ &\vdots \\ r_m &= \begin{pmatrix} a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \end{aligned}$$

The columns of $\mathbf{A}$ are:

$$\mathbf{c_1} = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} \quad \mathbf{c_2} = \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix} \quad \cdots \quad \mathbf{c_n} = \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix}$$

We will use these rows and columns notation for defining block matrix multiplication.

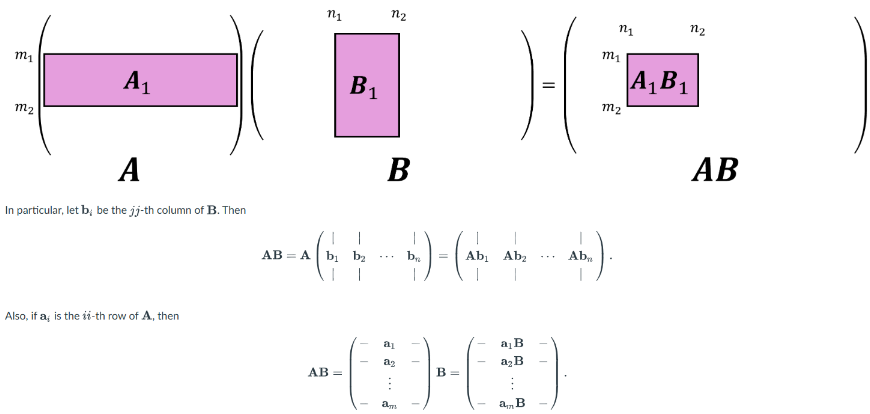

Figure 1: Matrix Block Multiplication

Note that we use **columns notation** if we **post-multiplying** something, whereas we use **rows notation** if we **pre-multiplying** something. This follows the property of matrix multiplication.

5.4 Combined Linear Systems

Note: Both matrix and vector equations are just another way to represent augmented matrix. Therefore, instead of choosing to represent it as matrix and vector equation, we can also represent it as an augmented matrix.

p linear systems with the same coefficient matrix $\mathbf{A} = (a_{ij})_{m\times n}$, for $k = 1, 2, ..., p$.

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + ... + a_{1n}x_n = b_{1k} \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + ... + a_{2n}x_n = b_{2k} \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + a_{m3}x_3 + ... + a_{mn}x_n = b_{mk} \end{cases}$$

Can be expressed as a combined augmented matrix:

$$\left(\begin{array}{cccc|c|c|c|c} a_{11} & a_{12} & ... & a_{1n} & b_{11} & b_{12} & & b_{1p} \\ a_{21} & a_{22} & ... & a_{2n} & b_{21} & b_{22} & & b_{2p} \\ a_{31} & a_{32} & ... & a_{3n} & b_{31} & b_{32} & & b_{3p} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & ... & a_{mn} & b_{m1} & b_{m2} & & b_{mp} \end{array}\right)$$

5.5 Inverse of Matrices

5.5.1 Pitfall 1

For matrices $\mathbf{A}, \mathbf{B}, \mathbf{C}$ with appropriate sizes, if $\mathbf{AB} = \mathbf{AC}$ and $\mathbf{A} \neq \mathbf{0}$ is not the zero matrix, we could not conclude that $\mathbf{B} = \mathbf{C}$.

Counter Example:

$$\begin{pmatrix} 0 & 1 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 3 & 4 \\ 6 & 8 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 2 & 5 \\ 3 & 4 \end{pmatrix}$$

But,

$$\begin{pmatrix} 1 & 1 \\ 3 & 4 \end{pmatrix} \neq \begin{pmatrix} 2 & 5 \\ 3 & 4 \end{pmatrix}$$

5.5.2 Problem of Inverse

We know that $\mathbf{AB} \neq \mathbf{BA}$. Suppose that $\frac{\mathbf{1}}{\mathbf{A}}$ exists such that $\frac{\mathbf{1}}{\mathbf{A}}\mathbf{A} = \mathbf{I}$. Is it true that $\mathbf{A}\frac{\mathbf{1}}{\mathbf{A}} = \mathbf{I}$?
Answer: We are not sure about that.

5.5.3 Definition of Inverse Matrix

It turns out that this problem is irrelevant when we only consider square matrices; we will elaborate more about this in the corollary. Therefore, we define a $n \times n$ square matrix $\mathbf{A}$ is invertible if there exists a square matrix $\mathbf{B}$ of the same size such that $\mathbf{AB} = \mathbf{I_n} = \mathbf{BA}$. Otherwise, we call the matrix non-invertible.

5.5.4 An Observation

It turns out that the problem of inverse only occurs when the matrix is not a square matrix.

Let $\mathbf{A}$ be a $n \times m$ matrix, and $\mathbf{B}$ be a $m \times n$ matrix. As a result, according to the multiplication rule of matrices, $\mathbf{A} \times \mathbf{B}$ will be a matrix $n \times n$, while $\mathbf{B} \times \mathbf{A}$ will be a matrix $m \times m$.

If $n \neq m$, two matrices with different sizes will never be the same, so $\mathbf{A}$ will never be equal to $\mathbf{B}$. Otherwise, if $n = m$, we are considering special case where the matrices is a square matrix, which we talk in Chapter 5.5.3.

5.5.5 Uniqueness of Inverse

If $\mathbf{B}$ and $\mathbf{C}$ are both inverses of a square matrix $\mathbf{A}$, then $\mathbf{B} = \mathbf{C}$.

Proof: According to the associative property of matrix multiplication, we know that $(\mathbf{BA})\mathbf{C} = \mathbf{B}(\mathbf{AC})$ regardless of what the dimension of $\mathbf{A}, \mathbf{B}$, and $\mathbf{C}$. Because both $\mathbf{B}$ and $\mathbf{C}$ are the inverse of $\mathbf{A}$, pre-multiplying or post-multiplying $\mathbf{B}$ or $\mathbf{C}$ to A will result in it to be the identity matrix, I.

$$\begin{aligned}(\mathbf{BA})\mathbf{C} &= \mathbf{B}(\mathbf{AC}) \\ \mathbf{IC} &= \mathbf{BI} \\ \mathbf{C} &= \mathbf{B}\end{aligned}$$

5.5.6 Invertibility and Linear System

Suppose $\mathbf{A}$ is an $n \times n$ invertible square matrix. Then, for any $n \times 1$ vector $\mathbf{b}$, $\mathbf{Ax} = \mathbf{b}$ has a unique solution.

Proof: We will first prove the **existence** of the solution. Let $\mathbf{u} = \mathbf{A^{-1}b}$, so $\mathbf{Au} = \mathbf{A}(\mathbf{A^{-1}b}) = \mathbf{Ib} = \mathbf{b}$; the solution exists. Then, we will prove the **uniqueness** of the solution. Let $\mathbf{v}$ also be the solution of $\mathbf{Ax} = \mathbf{b}$. Thus,

$$\begin{aligned}\mathbf{Av} &= \mathbf{b} = \mathbf{Au} \\ \mathbf{A^{-1}}(\mathbf{Av}) &= \mathbf{A^{-1}}(\mathbf{Au}) \\ \mathbf{Iv} &= \mathbf{Iu} \\ \mathbf{v} &= \mathbf{u}\end{aligned}$$

5.5.7 A Colorrary for Invertibility and Linear System

Suppose $\mathbf{A}$ is invertible. Then, the trivial solution is the only solution to the homogeneous system $\mathbf{Ax} = \mathbf{0}$.

Proof: Combine the ideas in Chapter 5.2.3 and Chapter 5.5.6

5.5.8 Algorithms to Compute Inverse

Suppose $\mathbf{A}$ is an invertible $n \times n$ matrix. By uniqueness of the inverse, there must be a unique solution to $\mathbf{AX} = \mathbf{I}$. Let's expand it to its matrix form.

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nn} \end{pmatrix} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

Break it further!

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} x_{11} \\ x_{21} \\ \vdots \\ x_{n1} \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} x_{12} \\ x_{22} \\ \vdots \\ x_{n2} \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}$$

$$\vdots$$

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} x_{1n} \\ x_{2n} \\ \vdots \\ x_{nn} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

From Chapter 5.4, we know that we can represent this as an augmented matrix and solve the equation! So, we will have the augmented matrix:

$$\left(\begin{array}{cccc|cccc} a_{11} & a_{12} & ... & a_{1n} & 1 & 0 & ... & 0 \\ a_{21} & a_{22} & ... & a_{2n} & 0 & 1 & ... & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & 0 & 0 & \cdots & 1 \end{array}\right)$$

If we then reduce this matrix to its reduced row-echelon form, we will get the RHS of the matrix equal to the $n \times n$ matrix $\mathbf{X}$, which is indeed the inverse matrix of $\mathbf{A}$.

5.5.9 A Property of Inverse Matrix

Let $\mathbf{A}$ and $\mathbf{B}$ be an invertible matrix of order n, then $(\mathbf{AB})$ is invertible with inverse $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$.

Proof: We will prove this by showing that $(\mathbf{AB})\mathbf{B}^{-1}\mathbf{A}^{-1}$ is the identity matrix.

$$(\mathbf{AB})\mathbf{B}^{-1}\mathbf{A}^{-1} = \mathbf{A}(\mathbf{BB}^{-1})\mathbf{A}^{-1} = \mathbf{AIA}^{-1} = \mathbf{AA}^{-1} = \mathbf{I}$$

5.6 Elementary Matrices

5.6.1 Definition of Elementary Matrices

A square matrix $\mathbf{E}$ of order n is called an **elementary matrix** if it can be obtained from the identity matrix $\mathbf{I}_n$ by performing a **single** elementary row operation

$$\mathbf{I}_n \xrightarrow{r} \mathbf{E}$$

where r is an elementary row operation. The elementary row operation is said to be the row operation corresponding to the elementary matrix.

5.6.2 Three types of Elementary Matrices

1. When r is in the form of $R_x + cR_y$, the elementary matrix is the identity matrix with the entries at x, y being changed to c.
2. When r is in the form of $R_x \leftrightarrow R_y$, the elementary matrix is the identity matrix with row x and y are swapped.
3. When r is in the form of cR_x, the elementary matrix is the identity matrix with row x multiplied by c.

6 Tutorial 2

6.1 Question 1(a)

To show if $\mathbf{I} - \mathbf{A}$ is invertible and have an inverse of $\mathbf{I} + \mathbf{A}$, we will prove that $(\mathbf{I} - \mathbf{A})(\mathbf{I} + \mathbf{A})$ is equal to the identity matrix, $\mathbf{I}$.

$$\begin{aligned}(\mathbf{I} - \mathbf{A})(\mathbf{I} + \mathbf{A}) &= \mathbf{I^2} - \mathbf{AI} + \mathbf{IA} - \mathbf{A^2} \\ &= \mathbf{I} - \mathbf{A} + \mathbf{A} - \mathbf{A^2} \\ &= \mathbf{I} - \mathbf{A^2} \\ &= \mathbf{I}\end{aligned}$$

Because the multiplication of these two matrices result to be an identity matrix, we could conclude that $\mathbf{I} - \mathbf{A}$ is indeed invertible with an inverse of $\mathbf{I} + \mathbf{A}$.

6.2 Question 1(b)

Remember this polynomial identity!

$$1 - x^n = (1 - x)(1 + x + x^2 + ... + x^{n-1})$$

Let's observe how if this polynomial identity is applied to matrix equations!

$$\mathbf{I} - \mathbf{A}^n = (\mathbf{I} - \mathbf{A})(\mathbf{I} + \mathbf{A} + \mathbf{A}^2 + ... + \mathbf{A}^{n-1})$$

Let $n = 3$, the equation will be:

$$\mathbf{I} - \mathbf{A}^3 = (\mathbf{I} - \mathbf{A})(\mathbf{I} + \mathbf{A} + \mathbf{A}^2)$$

Substituting $\mathbf{A}^3 = 0$, we will then get:

$$\mathbf{I} = (\mathbf{I} - \mathbf{A})(\mathbf{I} + \mathbf{A} + \mathbf{A}^2)$$

Observe that the order of $\mathbf{I} - \mathbf{A}$ and $\mathbf{I} + \mathbf{A} + \mathbf{A^2}$ does not matter since $\mathbf{AI} = \mathbf{IA}$. Thus, following the definition of inverse matrices, we can conclude that $\mathbf{I} - \mathbf{A}$ is invertible, with the inverse equal to $\mathbf{I} + \mathbf{A} + \mathbf{A^2}$.

6.3 Question 1(c)

Use the same identity as Question 1(b)!

$$\mathbf{I} - \mathbf{A}^n = (\mathbf{I} - \mathbf{A})(\mathbf{I} + \mathbf{A} + \mathbf{A}^2 + ... + \mathbf{A}^{n-1})$$

When $\mathbf{A}$ is *nilpotent*, this means $\mathbf{A^n} = 0$. Thus,

$$\begin{aligned} \mathbf{I} - \mathbf{A^n} &= (\mathbf{I} - \mathbf{A})(\mathbf{I} + \mathbf{A} + \mathbf{A}^2 + ... + \mathbf{A}^{n-1}) \\ \mathbf{I} &= (\mathbf{I} - \mathbf{A})(\mathbf{I} + \mathbf{A} + \mathbf{A}^2 + ... + \mathbf{A}^{n-1}) \end{aligned}$$

Since $\mathbf{IA} = \mathbf{AI}$, we can observe that $\mathbf{I} - \mathbf{A}$ is invertible with the inverse equal to $\mathbf{I} + \mathbf{A} + \mathbf{A^2} + ... + \mathbf{A^{n-1}}$.

6.4 Question 3(a)

Substitute, $n = 4$, $(x_1, y_1) = (1, 3)$, $(x_2, y_2) = (2, -2)$, $(x_3, y_3) = (3, -5)$, and $(x_4, y_4) = (4, 0)$ to the augmented matrix. The augmented matrix will be in this following form:

$$\left(\begin{array}{cccc|c} 1 & 1 & 1^2 & 1^3 & 3 \\ 1 & 2 & 2^2 & 2^3 & -2 \\ 1 & 3 & 3^2 & 3^3 & -5 \\ 1 & 4 & 4^2 & 4^3 & 0 \end{array}\right) = \left(\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 3 \\ 1 & 2 & 4 & 8 & -2 \\ 1 & 3 & 9 & 27 & -5 \\ 1 & 4 & 16 & 64 & 0 \end{array}\right) \xrightarrow{RREF} \left(\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 4 \\ 0 & 1 & 0 & 0 & 3 \\ 0 & 0 & 1 & 0 & -5 \\ 0 & 0 & 0 & 1 & 1 \end{array}\right)$$

6.5 Question 3(b)

Here is the Matlab code to solve the problem!

```
v=[1;2;3;4;5;6;7;8;];
A=fliplr(vander(v));
Newcol=[12;70;1244;10500;54268;205682;630540;1657024];
A=[A Newcol];
display(A);
B=rref(A);
display(B);
```

6.6 Question 4(a)

Using KCL, we can get that $I_1 + I_2 = I_3$. Then, analyzing the left loop, using KVL, we can get that $-20I_3 + 50 - 5I_1 = 0$. Next, analyzing the right loop, using KVL, we can also get that $20I_3 + 10I_2 + 30 = 0$. Let's rearrange this!

$$\begin{cases} I_1 + I_2 - I_3 = 0 \\ 5I_1 + 20I_3 = 50 \\ 10I_2 + 20I_3 = -30 \end{cases}$$

Now, let's represent this as an augmented matrix.

$$\left(\begin{array}{ccc|c} 1 & 1 & -1 & 0 \\ 5 & 0 & 20 & 50 \\ 0 & 10 & 20 & -30 \end{array}\right) \xrightarrow{RREF} \left(\begin{array}{ccc|c} 1 & 0 & 0 & 6 \\ 0 & 1 & 0 & -5 \\ 0 & 0 & 1 & 1 \end{array}\right)$$

Thus, $I_1 = 6A$, $I_2 = -5A$, and $I_3 = 1A$.

6.7 Question 4(b)

Using KCL, we can get that $I_1 + I_3 = I_2$. Then, analyzing the top loop, using KVL, we can get the equation $-5 + I_1 + 2I_2 = 0$. Next, analyzing the bottom loop, using KVL, we can get the equation $-2I_2 + 8 - 4I_3 = 0$. Let's rearrange this!

$$\begin{cases} I_1 - I_2 + I_3 = 0 \\ I_1 + 2I_2 = 5 \\ 2I_2 + 4I_3 = 8 \end{cases}$$

Let's represent this as an augmented matrix!

$$\left(\begin{array}{ccc|c} 1 & -1 & 1 & 0 \\ 1 & 2 & 0 & 5 \\ 0 & 2 & 4 & 8 \end{array}\right) \xrightarrow{RREF} \left(\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 1 \end{array}\right)$$

Thus, we can conclude that $I_1 = 1A$, $I_2 = 2A$, $I_3 = 1A$.

6.8 Question 5

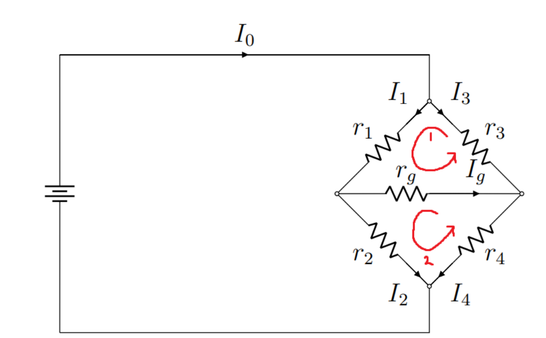

Figure 2: Wheatstone Bridge

Using KCL, we can get:

1. $I_0 = I_1 + I_3$
2. $I_1 = I_g + I_2$

3. $I_3 + I_g = I_4$
4. $I_2 + I_4 = I_0$

Then, using KVL, we can get:

1. $-I_1 r_1 - I_g r_g + I_3 r_3 = 0$
2. $-I_2 r_2 + I_4 r_4 + I_g r_g = 0$

6.8.1 Part a

Using the information that $I_g = 0$, we will simplify the equations above!

1. $I_0 = I_1 + I_3$
2. $I_1 = I_g + I_2 \rightarrow I_1 = I_2$
3. $I_3 + I_g = I_4 \rightarrow I_3 = I_4$
4. $I_2 + I_4 = I_0$
5. $-I_1 r_1 - I_g r_g + I_3 r_3 = 0 \rightarrow I_3 r_3 = I_1 r_1$
6. $-I_2 r_2 + I_4 r_4 + I_g r_g = 0 \rightarrow I_2 r_2 = I_4 r_4$

Let's rearrange this clearly!

$$\begin{cases} I_0 - I_1 - I_3 = 0 & ...(1) \\ I_1 - I_2 = 0 & ...(2) \\ I_3 - I_4 = 0 & ...(3) \\ I_0 - I_2 - I_4 = 0 & ...(4) \\ I_1 r_1 - I_3 r_3 = 0 & ...(5) \\ I_2 r_2 - I_4 r_4 = 0 & ...(6) \end{cases}$$

For Part a, we are tasked to express r_4 in terms of r_1, r_2, r_3.

$$\begin{aligned} I_2 r_2 - I_4 r_4 &= 0 && \text{from (6)} \\ I_4 r_4 &= I_2 r_2 \\ r_4 &= \frac{I_2}{I_4} r_2 \\ r_4 &= \frac{I_1}{I_3} r_2 && \text{from (2) and (3)} \end{aligned}$$

Let's observe what we can get from (5)!

$$\begin{aligned} I_1 r_1 &= I_3 r_3 \\ \frac{I_1}{I_3} &= \frac{r_3}{r_1} && ...(7) \end{aligned}$$

Let's substitute this new information to what we have derived previously!

$$r_4 = \frac{I_1}{I_3} r_2 = \frac{r_3}{r_1} r_2 = \frac{r_3 r_2}{r_1}$$

6.8.2 Part b

In part b, we don't know the value of I_g! Let's rearrange what we get from KCL and KVL!

1. $I_0 - I_1 - I_3 = 0$
2. $I_1 - I_2 - I_g = 0$
3. $I_3 - I_4 + I_g = 0$
4. $I_0 - I_2 - I_4 = 0$
5. $I_1 r_1 - I_3 r_3 + I_g r_g = 0$
6. $I_2 r_2 - I_4 r_4 - I_g r_g = 0$

In addition, in this part b, we are given new information on the battery voltage. Thus, let's add a new cycle on the circuit.

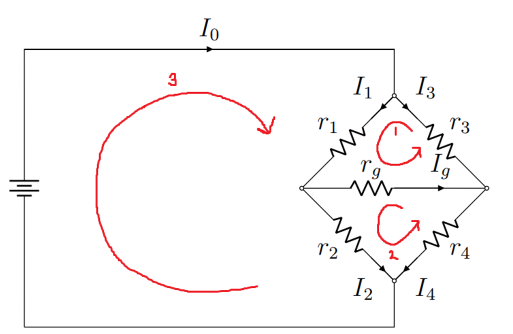

Figure 3: Wheatstone Bridge

From the third cycle, we can formulate an equation using KVL, $10 - I_1 r_1 - I_2 r_2 = 0 \rightarrow I_1 r_1 + I_2 r_2 = 10$.

Then, let's add this new equation and represent the linear system as an augmented matrix where the columns indicate $I_0, I_1, I_2, I_3, I_4, I_g$, respectively!

$$\left(\begin{array}{cccccc|c} 1 & -1 & 0 & -1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & -1 & 1 & 0 \\ 1 & 0 & -1 & 0 & -1 & 0 & 0 \\ 0 & r_1 & 0 & -r_3 & 0 & r_g & 0 \\ 0 & 0 & r_2 & 0 & -r_4 & -r_g & 0 \\ 0 & r_1 & r_2 & 0 & 0 & 0 & 10 \end{array}\right)$$

We know that $r_1 = 5\Omega$, $r_2 = 10\Omega$, $r_3 = 2\Omega$, $r_4 = 4\Omega$, $r_g = 50\Omega$. Let's substitute this into our augmented matrix!

$$\left(\begin{array}{cccccc|c} 1 & -1 & 0 & -1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & -1 & 1 & 0 \\ 1 & 0 & -1 & 0 & -1 & 0 & 0 \\ 0 & 5 & 0 & -2 & 0 & 50 & 0 \\ 0 & 0 & 10 & 0 & -4 & -50 & 0 \\ 0 & 5 & 10 & 0 & 0 & 0 & 10 \end{array}\right) \xrightarrow{RREF} \left(\begin{array}{cccccc|c} 1 & 0 & 0 & 0 & 0 & 0 & \frac{7}{3} \\ 0 & 1 & 0 & 0 & 0 & 0 & \frac{2}{3} \\ 0 & 0 & 1 & 0 & 0 & 0 & \frac{2}{3} \\ 0 & 0 & 0 & 1 & 0 & 0 & \frac{5}{3} \\ 0 & 0 & 0 & 0 & 1 & 0 & \frac{5}{3} \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

Then, because $I_g = 0A$, the relationship between r_1, r_2, r_3, r_4 that we obtained in part a will still hold! Thus,

$$r_4 = \frac{r_3 r_2}{r_1}$$

7 Week 4

7.1 Equivalent Statements for Invertibility

Let $\mathbf{A}$ be a square matrix of order n. The following statements are equivalent.

1. $\mathbf{A}$ is invertible
2. $\mathbf{A}^T$ is invertible
3. $\mathbf{A}$ has a left-inverse, that is, there is a matrix $\mathbf{B}$ such that $\mathbf{BA} = \mathbf{I}$.
4. $\mathbf{A}$ has a right-inverse, that is, there is a matrix $\mathbf{B}$ such that $\mathbf{AB} = \mathbf{I}$.
5. The reduced row-echelon form of $\mathbf{A}$ is the identity matrix.
6. $\mathbf{A}$ can be expressed as a product of elementary matrices.
7. The homogeneous system $\mathbf{Ax} = \mathbf{0}$ has only trivial solution.
8. For any $\mathbf{b}$, the system $\mathbf{Ax} = \mathbf{b}$ is consistent.

7.2 Determinant

7.2.1 Definition

The determinant of $\mathbf{A}$ is defined to be

$$\begin{aligned} \det(\mathbf{A}) &= a_{i1}A_{i1} + a_{i2}A_{i2} + ... + a_{in}A_{in} = \sum_{k=1}^{n} a_{ik}A_{ik} \\ &= a_{1j}A_{1j} + a_{2j}A_{2j} + ... + a_{nj}A_{nj} = \sum_{k=1}^{n} a_{kj}A_{kj} \end{aligned}$$

where $A_{ij} = (-1)^{i+j} \det(\mathbf{M}_{ij})$.

Note: A_{ij} is the (i, j) cofactor of $\mathbf{A}$, and $\mathbf{M}_{ij}$ is the (i, j) matrix minor of $\mathbf{A}$, obtained from $\mathbf{A}$ be deleting the i-th row and j-th column.

Trick: Choose the column or row that has the greatest numbers of zeros. With this, you don't need to calculate the value of A_{ij} because a_{ij} is zero already.

7.2.2 Determinant and Transpose

$$\det(\mathbf{A}) = \det(\mathbf{A}^T)$$

Proof: As we choose row i for computing $\det(\mathbf{A})$, we can always choose column i to compute $\det(\mathbf{A}^T)$ which is equivalent to row i after the transpose.

7.2.3 A Corollary

The determinant of a triangular matrix is the multiplication of the diagonal entries. That is, $\mathbf{A} = (a_{ij})_n$ is a triangular matrix, then

$$\det(\mathbf{A}) = a_{11}a_{22}...a_{nn} = \prod_{k=1}^{n} a_{kk}$$

Proof: Choose row $1, 2, ..., n$ sequentially. Notice that other values except a_{ii}, where i is the chosen row, is 0, which is irrelevant. Thus,

$$\det(\mathbf{A}) = a_{11}A_{11} = a_{11}a_{22}(A_{22})_{11} = \; ... \; = a_{11}a_{22}a_{33}...a_{nn}$$

Note: $(A_{22})_{11}$ means that row 1,2 and column 1,2 are removed.

7.2.4 Determinant and Elementary Row Operation

Suppose $\mathbf{B}$ is obtained from $\mathbf{A}$ by a single elementary row operation $\mathbf{A} \xrightarrow{r} \mathbf{B}$. Then, the determinant of $\mathbf{B}$ is obtained from the determinant of $\mathbf{A}$ as such:

- If $r = R_i + aR_j$, then $\det(\mathbf{B}) = \det(\mathbf{A})$
- If $r = cR_i$, then $\det(\mathbf{B}) = c \det(\mathbf{A})$
- If $r = R_i \leftrightarrow R_j$, then $\det(\mathbf{B}) = -\det(\mathbf{A})$

7.2.5 A Corollary

The determinant of an elementary matrix $\mathbf{E}$ is given as such:

1. If $\mathbf{E}$ corresponds to $R_i + aR_j$, then $\det(\mathbf{E}) = 1$.
2. If $\mathbf{E}$ corresponds to cR_i, then $\det(\mathbf{E}) = c$.
3. If $\mathbf{E}$ corresponds to $R_i \leftrightarrow R_j$, then $\det(\mathbf{E}) = -1$

7.2.6 A Theorem

Let $\mathbf{A}$ and $\mathbf{R}$ be square matrices such that

$$\mathbf{R} = \mathbf{E}_k \cdots \mathbf{E}_2 \mathbf{E}_1 \mathbf{A}$$

for some elementary matrices $\mathbf{E}_1, \mathbf{E}_2, \ldots, \mathbf{E}_k$. Then

$$\det(\mathbf{R}) = \det(\mathbf{E}_k) \cdots \det(\mathbf{E}_2) \det(\mathbf{E}_1) \det(\mathbf{A}).$$

7.2.7 A Corollary

Let $\mathbf{A}$ be a $n \times n$ square matrix. Suppose

$$\mathbf{A} \xrightarrow{r_1} \xrightarrow{r_2} \cdots \xrightarrow{r_k} \mathbf{R} = \begin{pmatrix} d_1 & * & \cdots & * \\ 0 & d_2 & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n \end{pmatrix},$$

where $\mathbf{R}$ is the reduced row-echelon form of $\mathbf{A}$. Let $\mathbf{E}_i$ be the elementary matrix corresponding to the elementary row operation r_i, for $i = 1, \ldots, k$. Then

$$\det(\mathbf{A}) = \frac{d_1 d_2 \cdots d_n}{\det(\mathbf{E}_k) \cdots \det(\mathbf{E}_2) \det(\mathbf{E}_1)}.$$

7.2.8 Determinant of Products

Let $\mathbf{A}_1, \mathbf{A}_2, ..., \mathbf{A}_k$ be square matrices of the same size. Thus,

$$\det(\mathbf{A_1 A_2 A_3} \cdots \mathbf{A_k}) = \det(\mathbf{A_1}) \det(\mathbf{A_2}) \cdots \det(\mathbf{A_k})$$

7.2.9 Determinant of Inverse

If $\mathbf{A}$ is invertible, then

$$\det(\mathbf{A}^{-1}) = \det(\mathbf{A})^{-1}$$

7.2.10 Determinant of Scalar Multiplication

For any square matrix $\mathbf{A}$ of order n and scalar c,

$$\det(c\mathbf{A}) = c^n \det(\mathbf{A})$$

8 Week 5

8.1 Euclidean Vector Spaces

8.1.1 What is Vector?

A vector is a collection of n ordered real numbers,

$$n = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} \text{ where } v_i \in \mathbb{R} \text{ for } i = 1, 2, ..., n.$$

The real numbers v_i is called the i-th coordinate of vector v.

8.1.2 What is Euclidean n-space?

The Euclidean n-space is denoted as $\mathbb{R}^n$. It is the collection of all n-vectors.

$$\mathbb{R}^n = \left\{ \mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} \middle| \; v_i \in \mathbb{R} \text{ for } i = 1, \ldots, n. \right\}$$

8.2 Dot Product, Norm, Distance

8.2.1 Dot Product

The dot product of vectors $\mathbf{u} = (u_i)$ and $\mathbf{v} = (v_i)$ in $\mathbb{R}^n$ is defined to be

$$\mathbf{u} \cdot \mathbf{v} = u_1 v_1 + u_2 v_2 + ... + u_n v_n$$

Sometimes, people call dot product as inner product.

8.2.2 Norm

The norm of a vector $\mathbf{u} \in \mathbb{R}^n$ is defined to be the square root of the dot product of $\mathbf{u}$ with itself. It is denoted by $||\mathbf{u}||$.

$$||\mathbf{u}|| = \sqrt{\mathbf{u} \cdot \mathbf{u}} = \sqrt{u_1^2 + u_2^2 + ... + u_n^2}$$

This is also known as the length or magnitude of a vector.

8.2.3 Properties of inner product and norm

Let $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ be vectors and $a, b, c \in \mathbb{R}$ be real numbers. Here are the properties that hold:

1. Inner product is symmetric,
$$\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}.$$
2. Inner product commutes with scalar multiple,
$$c\mathbf{u} \cdot \mathbf{v} = (c\mathbf{u}) \cdot \mathbf{v} = \mathbf{u} \cdot (c\mathbf{v}).$$
3. Inner product is distributive,
$$\mathbf{u} \cdot (a\mathbf{v} + b\mathbf{w}) = a\mathbf{u} \cdot \mathbf{v} + b\mathbf{u} \cdot \mathbf{w}.$$
4. Inner product is positive definite, $\mathbf{u} \cdot \mathbf{u} \geq 0$ with equality if and only if $\mathbf{u} = \mathbf{0}$.
5. $\|\mathbf{cu}\| = |c|\|\mathbf{u}\|$.

8.2.4 Unit vector

A vector $\mathbf{u}$ in $\mathbb{R}^n$ is a unit vector if its norm is 1. For a nonzero vector $\mathbf{u}$, we can multiply the vector by **the reciprocal of the norm**; this technique is called normalizing a vector.

8.2.5 Distance between vectors

The distance between two vectors $\mathbf{u}$ and $\mathbf{v}$ is denoted by $d(\mathbf{u}, \mathbf{v})$ and is defined to be:
$$d(\mathbf{u}, \mathbf{v}) = ||\mathbf{u} - \mathbf{v}||$$

8.2.6 Angle between two nonzero vectors

The angle between two nonzero vectors, $\mathbf{u}$ and $\mathbf{v}$ can be calculated with the following formula:
$$cos(\theta) = \frac{\mathbf{u} \cdot \mathbf{v}}{||\mathbf{u}||\ ||\mathbf{v}||}$$

8.3 Linear Combination and Linear Spans

8.3.1 Linear Combination

A linear combination of $\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k \in \mathbb{R}^n$ is
$$c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + ... + c_k\mathbf{u}_k, \text{for some } c_1, c_2, c_3, ..., c_k \in \mathbb{R}$$
The scalars $c_1, c_2, ..., c_k$ are called coefficients.

8.3.2 Linear Span

Let $\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k$ be vectors in $\mathbb{R}^n$. The linear span of $\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k$ is the subset of $\mathbb{R}^n$ containing all the linear combinations of $\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k$.
$$\text{span}\,\{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k\} = \{c_1\mathbf{u_1} + c_1\mathbf{u_1} + ... + c_1\mathbf{u_1} \mid c_1, c_2, ..., c_k \in \mathbb{R}\}$$

8.3.3 How to check if the vector v is in a span S?

Let $S = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k\}$ be a set of vectors in $\mathbb{R}^n$.

1. Form the $n \times k$ matrix $\mathbf{A} = (\mathbf{u}_1 \quad \mathbf{u}_2 \quad \mathbf{u}_3 \quad ... \quad \mathbf{u}_k)$ whose columns are the vectors in S.

2. Then, a vector $\mathbf{v}$ in $\mathbb{R}^n$ is in $span\{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k\}$ if and only if the system $\mathbf{Ax} = \mathbf{v}$ is consistent. To check if a system is consistent or not, we will represent the linear systems as an augmented matrix

$$(\mathbf{u}_1 \quad \mathbf{u}_2 \quad \mathbf{u}_3 \quad \ldots \quad \mathbf{u}_k \mid \mathbf{v})$$

   Then, find the RREF! If the system is consistent, then we can directly find the coefficients for the solution.

8.3.4 How to check if a span $S = \mathbb{R}^n$?

Let $S = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k\}$ be a set of vectors in $\mathbb{R}^n$.

1. Form the $n \times k$ matrix $\mathbf{A} = (\mathbf{u}_1 \quad \mathbf{u}_2 \quad \mathbf{u}_3 \quad ... \quad \mathbf{u}_k)$ whose columns are the vectors in S.

2. Look at the reduced row-echelon form of the matrix $(\mathbf{u}_1 \quad \mathbf{u}_2 \quad \mathbf{u}_3 \quad \ldots \quad \mathbf{u}_k)$. If there is no zero rows, $S = span(\mathbb{R}^n)$. Otherwise, it doesn't.

8.3.5 Properties of Linear Span

Let $S = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k\}$ be a finite set of vectors. The span of S, span(S), has the following properties.

1. The span of S contains the origin,

$$\mathbf{0} \in \text{span}(S).$$

2. The span of S is closed under vector addition, for any $\mathbf{u}, \mathbf{v} \in \text{span}(S)$,

$$\mathbf{u} + \mathbf{v} \in \text{span}(S).$$

3. The span S is closed under scalar multiplication, for any $\mathbf{u} \in \text{span}(S)$ and real number $\alpha \in \mathbb{R}$,

$$\alpha\mathbf{u} \in \text{span}(S).$$

Extra notes:

Properties (ii) and (iii) can be combined together into one property (ii'):

- The span is closed under linear combinations, that is, if $\mathbf{u}, \mathbf{v}$ are vectors in span(S) and α, β are any scalars, then the linear combination

$$\alpha\mathbf{u} + \beta\mathbf{v} \text{ is a vector in span}(S).$$

8.3.6 A Theorem

Let $S = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k\}$ be a set of vectors in $\mathbb{R}^n$. For any vectors $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, ..., \mathbf{v}_m$ in $span(S)$, the span of $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, ..., \mathbf{v}_m$ is a subset of $span(S)$,

$$span\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, ..., \mathbf{v}_m\} \subseteq span(S)$$

8.3.7 How to check if span T is a subset of span S?

Suppose the set of vectors $T = \{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_m\}$ and $S = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k\}$. We can check whether $span(T)$ is a subset of $span(S)$ by checking whether the augmented matrix

$$\begin{pmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \mathbf{u}_3 & \dots & \mathbf{u}_k \mid \mathbf{v}_1 \mid \mathbf{v}_2 \mid ... \mid \mathbf{v}_m \end{pmatrix}$$

is consistent or not. If it's consistent, then $span(T)$ is indeed the subset of $span(S)$.

8.4 Subspaces

8.4.1 Solution Sets to a Linear Systems

Recall that the set of solutions to a linear system $\mathbf{Ax} = \mathbf{b}$ can be expressed implicitly as

$$V = \{\mathbf{u} \in \mathbb{R}^n \mid \mathbf{Au} = \mathbf{b}\}$$

or explicitly as

$$V = \{\mathbf{u} + s_1\mathbf{v}_1 + s_2\mathbf{v}_2 + ... + s_k\mathbf{v}_k \mid s_1, s_2, ..., s_k \in \mathbb{R}\}$$

where $\mathbf{u} + s_1\mathbf{v}_1 + s_2\mathbf{v}_2 + ... + s_k\mathbf{v}_k$ where $s_1, s_2, ..., s_k \in \mathbb{R}$ is the general solution.

8.4.2 Solution Sets to a Homogeneous Systems

Recall that the set of solutions to a homogeneous system $\mathbf{Ax} = \mathbf{0}$ has the form

$$s_1\mathbf{v}_1 + s_2\mathbf{v}_2 + ... + s_k\mathbf{v}_k, \text{where } s_1, s_2, ..., s_k \in \mathbb{R}$$

Explicitly, the solution set is

$$V = \{s_1\mathbf{v}_1 + s_2\mathbf{v}_2 + ... + s_k\mathbf{v}_k \mid s_1, s_2, ..., s_k \in \mathbb{R}\}$$

Notice that this is just the span$\{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_k\}$.

By Theorem 8.3.6, the solution set to a homogeneous system is a vector space that is a subset of the Euclidean vector space. This kind of scenario is called subspace; we will define the relationship between subspace and solution set to a homogeneous system in Chapter 8.4.4

8.4.3 Definition of Subspace

A subset V of $\mathbb{R}^n$ is a subspace if it satisfies the following properties:

1. V contains the zero vector, $0 \in V$
2. V is closed under scalar multiplication. For any vector $\mathbf{v}$ in V and scalar α, the vector $\alpha\mathbf{v}$ is in V.
3. V is closed under addition. For any vectors $\mathbf{u}$, $\mathbf{v}$ in V, the sum $\mathbf{u}+\mathbf{v}$ is in V.

Note 1: Property 1 can be replaced with the property saying that V is nonempty

Note 2: Property 2 and 3 is equivalent to the property saying that V is closed under linear combination. For any $\mathbf{u}$ and $\mathbf{v}$ in V, and scalars α and β, the linear combination $\alpha\mathbf{u}+\beta\mathbf{v}$ is in V.

8.4.4 A Theorem Connecting 8.4.2 and 8.4.3

The solution set $V = \{\mathbf{u} \mid \mathbf{Au} = \mathbf{b}\}$ to a linear system $\mathbf{Ax} = \mathbf{b}$ is a **subspace** if and only if $\mathbf{b} = \mathbf{0}$, that is, the system is **homogeneous**.

Proof:

1. If V is a subspace, then $0 \in V$. Thus, $A \cdot 0 = b \rightarrow b = 0$. Indeed, the system is homogeneous.
2. If $b = 0$, we will prove that $V = \{\mathbf{u} \mid \mathbf{Au} = \mathbf{0}\}$ is a subspace, satisfying the three properties of a subspace.
   (a) It indeed contains 0; a solution to a homogeneous system is the trivial solution.
   (b) For $v \in V$ and $Av = 0$, $A(\alpha v) = \alpha(Av) = \alpha 0 = 0$. Thus, αv is $\in V$.
   (c) For $u, v \in V$ and $Au = Av = 0$, $A(u+v) = Au + Av = 0 + 0 = 0$. Thus, $u + v$ is $\in V$.

8.4.5 A Theorem

A subset $V \subseteq \mathbb{R}^n$ is a subspace if and only if it is a linear span, $V = span(S)$ for some finite set $S = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k\}$.

8.4.6 How to Check Whether A Subset is A Subspace or Not?

To show that a set V is a subspace, we can either:

1. Find a spanning set, i.e., find a set S such that $V = span(S)$
2. Show that V satisfies the 3 conditions of being a subspace.

To show that a subset V is not a subspace, we can either:

1. Show that it does not contain the zero vector, $0 \notin V$.
2. Find a vector $\mathbf{v} \in V$ and a scalar $\alpha \in \mathbb{R}$ such that $\alpha\mathbf{v} \notin V$
3. Find vectors $\mathbf{u}$ and $\mathbf{v} \in \mathbb{R}$ such that the sum is not in V, $\mathbf{u} + \mathbf{v} \notin V$.

8.4.7 Affine Spaces

The solution set of a homogeneous system is indeed a subspace, but then how about the solution set of a non-homogeneous system?

Similar to Chapter 8.4.1, recall that the general solution to a linear system $\mathbf{Ax} = \mathbf{b}$ where $\mathbf{b} \neq 0$ can be expressed explicitly as

$$V = \{\mathbf{u} + s_1\mathbf{v}_1 + s_2\mathbf{v}_2 + ... + s_k\mathbf{v}_k \mid s_1, s_2, ..., s_k \in \mathbb{R}\}$$

From Chapter 8.4.2, recall that the general solution to a homogeneous system $\mathbf{Ax} = \mathbf{0}$ can be expressed explicitly as

$$V = \{s_1\mathbf{v}_1 + s_2\mathbf{v}_2 + ... + s_k\mathbf{v}_k \mid s_1, s_2, ..., s_k \in \mathbb{R}\}$$

Notice that $\mathbf{u}$ is a particular solution to the non-homogeneous system $\mathbf{Ax} = \mathbf{b}$. In my imagination, this $\mathbf{u}$ acts a translation factor that shift a subspace. Thus, in general, you can say that an affine space is a "translated" version if a vector subspace.

9 Week 6

9.1 Linear Independence

A set of vector $\{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k\}$ in $\mathbb{R}^n$ is linearly independent if and only if

1. The homogeneous system $(\mathbf{u}_1 \quad \mathbf{u}_2 \quad \mathbf{u}_3 \quad ... \quad \mathbf{u}_k)\, \mathbf{x} = 0$ has only the trivial solution.
2. The reduced row-echelon form of $(\mathbf{u}_1 \quad \mathbf{u}_2 \quad \mathbf{u}_3 \quad ... \quad \mathbf{u}_k)$ has no non-pivot column.

Proof for Condition 1: Checking for linear independence is equivalent to checking whether there exists $\mathbf{u}_i \in \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k\}$ where

$$\mathbf{u}_i = c_1\mathbf{u}_1 + ... + c_{i-1}\mathbf{u}_{i-1} + c_{i+1}\mathbf{u}_{i+1} + ... + c_k\mathbf{u}_k \tag{1}$$

For the same goal, we could address this problem with a different approach.

Suppose we are able to find some $c_1, c_2, ..., c_k$ not all zero such that

$$c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + ... + c_k\mathbf{u}_k = 0$$

Without lost of generality, let say $\mathbf{u}_k \neq 0$. Observe that we can manipulate the equation, such that

$$\frac{c_1}{-c_k}\mathbf{u}_1 + \frac{c_2}{-c_k}\mathbf{u}_2 + ... + \frac{c_{k-1}}{-c_k}\mathbf{u}_{k-1} = \mathbf{u}_k$$

This form is similar to form (1).

Thus, checking whether there exists $\mathbf{u}_i \in \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k\}$ where

$$\mathbf{u}_i = c_1\mathbf{u}_1 + ... + c_{i-1}\mathbf{u}_{i-1} + c_{i+1}\mathbf{u}_{i+1} + ... + c_k\mathbf{u}_k$$

is the same as checking whether there exists a non-trivial solution for this system:

$$c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + ... + c_k\mathbf{u}_k = 0$$

If there exists a non-trivial solution for the system, the system is linearly dependent. Otherwise, the system is linearly independent.

9.2 Basis and Coordinates

9.2.1 Definition of Basis

Let V be a subspace of $\mathbb{R}^n$. A set $S = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_3\}$ is a **basis** for V if

1. $span(S) = V$
2. S is linearly independent

9.2.2 A Theorem

Suppose $S = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k\}$ is a basis for V. Then, every vector $\mathbf{v}$ in the subspace V can be written as a linear combination of vectors in S uniquely.

Proof of Existence: If $span(S) = V$, by the definition of a span, every $v \in V$ is a linear combination of $\{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k\}$.

Proof of Uniqueness: Suppose v in V, and v can be expressed both as $c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + ... + c_k\mathbf{u}_k$ and $d_1\mathbf{u}_1 + d_2\mathbf{u}_2 + ... + d_k\mathbf{u}_k$. This means:

$$v = c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + ... + c_k\mathbf{u}_k = d_1\mathbf{u}_1 + d_2\mathbf{u}_2 + ... + d_k\mathbf{u}_k$$
$$(c_1 - d_1)\mathbf{u}_1 + (c_2 - d_2)\mathbf{u}_2 + ... + (c_k - d_k)\mathbf{u}_k = 0$$

By the definition of a basis, $\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u_k}$ is linearly independent. Then, by the definition of linear independence, the only solution to the equation $(c_1 - d_1)\mathbf{u}_1 + (c_2 - d_2)\mathbf{u}_2 + ... + (c_k - d_k)\mathbf{u}_k = 0$ is the trivial solution. Thus, this means, $c_1 = d_1$, $c_2 = d_2$, ..., $c_k = d_k$ implying that $v \in V$ can be uniquely expressed as a linear combination of vectors in S.

9.2.3 Yet Another Theorem

Let $V = \{\mathbf{u} \mid \mathbf{Au} = \mathbf{0}\}$ be the solution space to some homogeneous system. Suppose

$$s_1\mathbf{u}_1 + s_2\mathbf{u}_2 + ... + s_k\mathbf{u}_k \text{ , where } s_1, s_2, ..., s_k \in \mathbb{R}$$

is a general solution to the homogeneous system $\mathbf{Ax} = \mathbf{0}$. Then, $\{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k\}$ is a basis for the subspace $V = \{\mathbf{u} \mid \mathbf{Au} = \mathbf{0}\}$.

9.2.4 A Theorem

Basis for the zero space $\{\mathbf{0}\}$ of $\mathbb{R}^n$ is the empty set $\{\}$ or $\varnothing$.

Note: $\{\mathbf{0}\}$ is not linearly independent, so it cannot be the basis for $\{\mathbf{0}\}$.

9.3 Basis and Invertibility

9.3.1 A Theorem

A $n \times n$ square matrix A is invertible if and only if the columns are linearly independent.

Proof: Recall Chapter 7.1 on the equivalent statement of invertibility number 5! If A is invertible, the reduced row-echelon form of A is the identity matrix.

Now, express A as $(\mathbf{u}_1 \quad \mathbf{u}_2 \quad ... \quad \mathbf{u}_n)$. Let $S = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_n\}$. Recall from Chapter 9.1 on the criteria for a vector space to be linearly independent! S is linearly independent when the reduced row-echelon form of A is the identity matrix.

Notice that these two are equivalent, both of them are intertwined with the condition that the reduced row-echelon form of A is the identity matrix.

9.3.2 Another Theorem

A $n \times n$ square matrix A is invertible if and only if the columns spans $\mathbb{R}^n$.

9.3.3 A Corollary

Let $S = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_n\}$ be a subset of $\mathbb{R}^n$ containing n vectors. Then, S is linearly independent if and only if S spans $\mathbb{R}^n$.

9.3.4 Another Corollary

Let $S = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k\}$ be a subset of $\mathbb{R}^n$ and $A = (\mathbf{u}_1 \quad \mathbf{u}_2 \quad ...\mathbf{u}_k)$ be the matrix whose columns are vectors in S. Then, S is a basis for $\mathbb{R}^n$ if and only if $k = n$ and A is an invertible matrix.

Proof:

1. If $k < n$, $span(S) \subsetneq \mathbb{R}^n$
2. If $k > n$, S must be linearly dependent

9.3.5 An Equivalent Theorem

A $n \times n$ square matrix A is invertible if and only if the rows of A form a basis for $\mathbb{R}^n$.

Proof: Change A in Theorem 9.3.1 to A^T!

9.3.6 Another Equivalent Theorem

A square matrix A of order n is invertible if and only if the rows of A are linearly independent.

Proof: Change A in Theorem 9.3.1 to A^T!

9.3.7 The Updated Version of the Equivalent Statements for Invertibility

Let $\mathbf{A}$ be a square matrix of order n. The following statements are equivalent.

1. $\mathbf{A}$ is invertible
2. $\mathbf{A}^T$ is invertible
3. $\mathbf{A}$ has a left-inverse, that is, there is a matrix $\mathbf{B}$ such that $\mathbf{BA} = \mathbf{I}$.
4. $\mathbf{A}$ has a right-inverse, that is, there is a matrix $\mathbf{B}$ such that $\mathbf{AB} = \mathbf{I}$.
5. The reduced row-echelon form of $\mathbf{A}$ is the identity matrix.
6. $\mathbf{A}$ can be expressed as a product of elementary matrices.
7. The homogeneous system $\mathbf{Ax} = \mathbf{0}$ has only trivial solution.
8. For any $\mathbf{b}$, the system $\mathbf{Ax} = \mathbf{b}$ is consistent.
9. The determinant of $\mathbf{A}$ is nonzero, $\det(A) \neq 0$.
10. The columns/rows of $\mathbf{A}$ are linearly independent for $\mathbb{R}^n$
11. The columns/rows of $\mathbf{A}$ spans $\mathbb{R}^n$

9.4 Coordinates Relative to a Basis

9.4.1 Definition

Let $S = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k\}$ be a basis for a subspace V of $\mathbb{R}^n$. Then, given any vector $\mathbf{v} \in V$, we can write $\mathbf{v}$ unique as

$$c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + ... + c_k\mathbf{u}_k$$

The coordinates of $\mathbf{v}$ relative to the basis S is defined to be the vector

$$[\mathbf{v}]_S = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_k \end{pmatrix}$$

9.4.2 An Important Note

Notice that it is declared that $\mathbf{v}$ can be written uniquely as

$$c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + ... + c_k\mathbf{u}_k$$

This uniqueness property can be satisfied only if S is a basis. If this condition is not satisfied:

1. If S is not linearly independent, a few vectors in $\mathbb{R}^k$ can map to the same $\mathbf{v} \in V$.
2. If S does not span V, there are a few vectors $\mathbf{v} \in V$ that we cannot represent.

9.4.3 How Can We Compute Relative Coordinate?

Let $S = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k\}$ be a basis for a subspace V of $\mathbb{R}^n$. Our target is to find real numbers $c_1, c_2, ..., c_k \in \mathbb{R}$ such that for $\mathbf{v} \in V$, $c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + ... + c_k\mathbf{u}_k = \mathbf{v}$. This is equivalent to solving for the augmented matrix

$$\left(\begin{array}{cccc|c} \mathbf{u}_1 & \mathbf{u}_2 & ... & \mathbf{u}_k & \mathbf{v} \end{array}\right)$$

9.5 Dimensions

9.5.1 A Theorem

Let V be a subspace of $\mathbb{R}^n$ and B be a basis for V. Suppose B contains k vectors, $|B| = k$. Let $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_m$ be vectors in V. Then,

1. $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_m$ is linearly independent if and only if $[\mathbf{v}_1]_B, [\mathbf{v}_2]_B, ..., [\mathbf{v}_m]_B$ is linearly independent in $\mathbb{R}^k$
2. $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_m$ is linearly dependent if and only if $[\mathbf{v}_1]_B, [\mathbf{v}_2]_B, ..., [\mathbf{v}_m]_B$ is linearly dependent in $\mathbb{R}^k$
3. $\{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_m\}$ spans V if and only if $\{[\mathbf{v}_1]_B, [\mathbf{v}_2]_B, ..., [\mathbf{v}_m]_B\}$ spans $\mathbb{R}^k$

9.5.2 A Corollary

Let V be a subspace of $\mathbb{R}^n$ and B be a basis for V. Suppose B contains k vectors, $|B| = k$.

1. If $S = \{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_p\}$ is a subset of V with $p > k$, then S is linearly dependent.
2. If $S = \{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_p\}$ is a subset of V with $p < k$, then S is cannot span V.

9.5.3 A Follow-Up Corollary

From Corollary 9.5.2, we can deduce that the number of vectors in basis B must be unique, cannot be bigger, cannot be smaller. Thus, we can have a follow-up corollary: "Suppose $S = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k\}$ and $T = \{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_m\}$ are bases for a subspace $V \subseteq \mathbb{R}^n$. Then, $k = m$."

9.5.4 Definition

From Corollary 9.5.3, we know that the number of vectors in a basis must be unique. Let V be a subspace of $\mathbb{R}^n$. Thus, we can define the dimension of V as the number of vectors in any basis of V.

9.5.5 An Intuition

More intuitively, we can imagine dimension as the independent degree of freedom of movement – in how many dimensions can we move around.

9.5.6 A Theorem

Let $\mathbf{A}$ be a $m \times n$ matrix. The number of non-pivot columns in the reduced row-echelon form of $\mathbf{A}$ is the dimension of the solution space

$$V = \{\mathbf{u} \in \mathbb{R}^n \mid \mathbf{Au} = \mathbf{0}\}$$

Intuitively speaking, the independent degree of freedom of movement is determined by the number of parameters in the general solution. Then, recall that the number of parameters is equal to the number of non-pivot columns.

9.5.7 Subset and Dimensions

Let U and V be subspaces of $\mathbb{R}^n$!

1. If $U \subseteq V$, then $dim(U) \leq dim(V)$
2. If $U \subseteq V$ and $U \neq V$, then $dim(U) < dim(V)$
3. Given that $U \subseteq V$, $dim(U) = dim(V)$ if and only if $U = V$

9.5.8 Alternative Way 1 to Check for Basis

Let V be a k-dimensional subspace for $\mathbb{R}^n$, $dim(V) = k$. Suppose $S \subseteq V$ is linearly independent subset containing k vectors, $|S| = k$. Then, S is a basis for V.

Proof:

1. Let U be the $span(S)$. Because S is linearly independent, according to the definition of basis in Chapter 9.2.1, S satisfies the two criteria to be the basis for U.

2. According to the definition of dimensions in Chapter 9.5.4, $dim(U) = |S| = k$.

3. Then, we know that S is a linearly independent subset of subspace V, so $span(S) \subseteq V$. Because $U = span(S)$, then $U \subseteq V$.

4. We are given that $U \subseteq V$ and $dim(U) = dim(V) = k$, according to Chapter 9.5.7, we can deduce that $U = V$.

5. Because S is the basis for U and $U = V$, this means that S is the basis for V.

9.5.9 Alternative Way 2 to Check for Basis

Let V be a k dimensional subspace of $\mathbb{R}^n$, $dim(V) = k$. Suppose S is a set containing k vectors, $|S| = k$, such that $V \subseteq span(S)$. Then, S is a basis for V.

9.5.10 A Summary

Definition	**(B1)**	**(B2)**
1. $\text{span}(S) = V$ 2. S is linearly independent	1. $\lvert S\rvert = \dim(V)$ 2. $S \subseteq V$ 3. S is linearly independent	1. $\lvert S\rvert = \dim(V)$ 2. $V \subseteq \text{span}(S)$

Figure 4: Checking for Basis

10 Tutorial 3

10.0.1 Question 1(a)

$$\left(\begin{array}{cc|cc} -1 & 3 & 1 & 0 \\ 3 & -2 & 0 & 1 \end{array}\right) \xrightarrow{R_2+3R_1} \left(\begin{array}{cc|cc} -1 & 3 & 1 & 0 \\ 0 & 7 & 3 & 1 \end{array}\right) \xrightarrow{\frac{1}{7}R_2} \left(\begin{array}{cc|cc} -1 & 3 & 1 & 0 \\ 0 & 1 & \frac{3}{7} & \frac{1}{7} \end{array}\right)$$
$$\xrightarrow{R_1-3R_2} \left(\begin{array}{cc|cc} -1 & 0 & -\frac{2}{7} & -\frac{3}{7} \\ 0 & 1 & \frac{3}{7} & \frac{1}{7} \end{array}\right) \xrightarrow{-R_1} \left(\begin{array}{cc|cc} 1 & 0 & \frac{2}{7} & \frac{3}{7} \\ 0 & 1 & \frac{3}{7} & \frac{1}{7} \end{array}\right)$$

The matrix is invertible with the inverse is $\begin{pmatrix} \frac{2}{7} & \frac{3}{7} \\ \frac{3}{7} & \frac{1}{7} \end{pmatrix}$.

10.0.2 Question 2(a)

$$\left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ a & b & c & 0 & 1 & 0 \\ a^2 & b^2 & c^2 & 0 & 0 & 1 \end{array}\right) \xrightarrow{R_2-aR_1} \left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & b-a & c-a & -a & 1 & 0 \\ a^2 & b^2 & c^2 & 0 & 0 & 1 \end{array}\right)$$
$$\xrightarrow{R_3-a^2R_1} \left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & b-a & c-a & -a & 1 & 0 \\ 0 & b^2-a^2 & c^2-a^2 & -a^2 & 0 & 1 \end{array}\right)$$
$$\xrightarrow{R_3-(b+a)R_2} \left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & b-a & c-a & -a & 1 & 0 \\ 0 & 0 & (c-a)(c+a)-(c-a)(b+a) & -a^2 & -(b+a) & 1 \end{array}\right)$$
$$= \left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & b-a & c-a & -a & 1 & 0 \\ 0 & 0 & (c-a)(c+a-b-a) & -a^2 & -(b+a) & 1 \end{array}\right)$$
$$= \left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & b-a & c-a & -a & 1 & 0 \\ 0 & 0 & (c-a)(c-b) & -a^2 & -(b+a) & 1 \end{array}\right)$$

From Chapter 7.1, remember that $\mathbf{A}$ is invertible if and only if the determinant of $\mathbf{A}$ is nonzero. Recall that the determinant of an upper triangular matrix is the multiplication of the values of the diagonal. Therefore,

$$\begin{aligned} 1 \cdot (b-a) \cdot (c-a)(c-b) &\neq 0 \\ (b-a)(c-a)(c-b) &\neq 0 \end{aligned}$$

From here, we can write down the conditions so that the matrix is invertible: $b \neq a$ and $c \neq a$ and $c \neq b$.

10.1 Question 3(a)

Write down the equation as an augmented matrix, and solve the equation by finding the RREF!

$$\left(\begin{array}{ccc|cccc} 2 & 1 & 1 & 2 & 3 & 4 & 1 \\ 0 & 1 & 2 & 1 & 0 & 3 & 7 \\ 1 & 3 & 2 & 2 & 1 & 1 & 2 \end{array}\right) \xrightarrow{RREF} \left(\begin{array}{ccc|cccc} 1 & 0 & 0 & \frac{5}{7} & \frac{11}{7} & \frac{12}{7} & -\frac{5}{7} \\ 0 & 1 & 0 & \frac{1}{7} & -\frac{2}{7} & -\frac{13}{7} & -\frac{15}{7} \\ 0 & 0 & 1 & \frac{3}{7} & \frac{1}{7} & \frac{17}{7} & \frac{32}{7} \end{array}\right)$$

From here, we can find that $\mathbf{X} = \begin{pmatrix} \frac{5}{7} & \frac{11}{7} & \frac{12}{7} & -\frac{5}{7} \\ \frac{1}{7} & -\frac{2}{7} & -\frac{13}{7} & -\frac{15}{7} \\ \frac{3}{7} & \frac{1}{7} & \frac{17}{7} & \frac{32}{7} \end{pmatrix}$.

10.2 Question 3(b)

Write down the equation as an augmented matrix, and solve the equation by finding the RREF!

$$\left(\begin{array}{ccc|c} 2 & 1 & 1 & 1 \\ 0 & 1 & 2 & -1 \\ 1 & 3 & 2 & -1 \end{array}\right) \xrightarrow{RREF} \left(\begin{array}{ccc|c} 1 & 0 & 0 & \frac{6}{7} \\ 0 & 1 & 0 & -\frac{3}{7} \\ 0 & 0 & 1 & -\frac{2}{7} \end{array}\right)$$

From here, we can find that $\mathbf{X} = \begin{pmatrix} \frac{6}{7} \\ -\frac{3}{7} \\ -\frac{2}{7} \end{pmatrix}$

10.3 Question 4(a)(i)

$$\begin{aligned} |\mathbf{A}| &= 1 \cdot \begin{vmatrix} 2 & -2 \\ 1 & 3 \end{vmatrix} - 0 \cdot \begin{vmatrix} 5 & 3 \\ 1 & 3 \end{vmatrix} + 0 \cdot \begin{vmatrix} 5 & 3 \\ 2 & -2 \end{vmatrix} \\ &= 1 \cdot (2 \cdot 3 - 1 \cdot (-2)) = 6 + 2 = 8 \end{aligned}$$

10.4 Question 4(a)(ii)

$$\begin{aligned} |\mathbf{A}_1| &= 1 \cdot \begin{vmatrix} 2 & -2 \\ 1 & 3 \end{vmatrix} - 2 \cdot \begin{vmatrix} 5 & 3 \\ 1 & 3 \end{vmatrix} + 0 \cdot \begin{vmatrix} 5 & 3 \\ 2 & -2 \end{vmatrix} \\ &= 1 \cdot (2 \cdot 3 - 1 \cdot (-2)) - 2 \cdot (5 \cdot 3 - 1 \cdot 3) \\ &= (6 + 2) - 2 \cdot 12 = 8 - 24 = -16 \end{aligned}$$

10.5 Question 4(a)(iii)

$$\begin{aligned} |\mathbf{A}_2| &= 1 \cdot \begin{vmatrix} 2 & -2 \\ 0 & 3 \end{vmatrix} - 0 \cdot \begin{vmatrix} 1 & 3 \\ 0 & 3 \end{vmatrix} + 0 \cdot \begin{vmatrix} 1 & 3 \\ 2 & -2 \end{vmatrix} \\ &= 1 \cdot (2 \cdot 3 - 0 \cdot (-2)) = 6 \end{aligned}$$

10.6 Question 4(a)(iv)

$$\begin{aligned}|\mathbf{A}_3| &= 1 \cdot \begin{vmatrix} 2 & 2 \\ 1 & 0 \end{vmatrix} - 0 \cdot \begin{vmatrix} 5 & 1 \\ 1 & 0 \end{vmatrix} + 0 \cdot \begin{vmatrix} 5 & 1 \\ 2 & 2 \end{vmatrix} \\ &= 1 \cdot (2 \cdot 0 - 1 \cdot 2) = -2\end{aligned}$$

10.7 Question 4(b)

Write down the equation as an augmented matrix, and solve the equation by finding the RREF!

$$\left(\begin{array}{ccc|c} 1 & 5 & 3 & 1 \\ 0 & 2 & -2 & 2 \\ 0 & 1 & 3 & 0 \end{array}\right) \xrightarrow{RREF} \left(\begin{array}{ccc|c} 1 & 0 & 0 & -2 \\ 0 & 1 & 0 & \frac{3}{4} \\ 0 & 0 & 1 & -\frac{1}{4} \end{array}\right)$$

From here, we can find that $\mathbf{x}$ is $\begin{pmatrix} -2 \\ \frac{3}{4} \\ -\frac{1}{4} \end{pmatrix}$

10.8 Question 4(c)

$$\frac{1}{\det(\mathbf{A})} \begin{pmatrix} \det(\mathbf{A}_1) \\ \det(\mathbf{A}_2) \\ \det(\mathbf{A}_3) \end{pmatrix} = \frac{1}{8} \begin{pmatrix} -16 \\ 6 \\ -2 \end{pmatrix} = \begin{pmatrix} -2 \\ \frac{3}{4} \\ -\frac{1}{4} \end{pmatrix}$$

The answer that we get from 4(b) is indeed the same as we get here!

10.9 Question 5

$$\begin{aligned}\begin{vmatrix} -x & 1 & 0 \\ 0 & -x & 1 \\ 2 & -5 & 4-x \end{vmatrix} &= -x \cdot \begin{vmatrix} -x & 1 \\ -5 & 4-x \end{vmatrix} + 2 \cdot \begin{vmatrix} 1 & 0 \\ -x & 1 \end{vmatrix} \\ &= -x \cdot (-x \cdot (4-x) - 1 \cdot (-5)) + 2 \cdot 1 \\ &= x \cdot (x \cdot (4-x) - 5) + 2 \\ &= x \cdot (4x - x^2 - 5) + 2 \\ &= -x^3 + 4x^2 - 5x + 2 \\ &= (x-1)(-x^2 + 3x - 2) \\ &= -(x-1)(x^2 - 3x + 2) \\ &= -(x-1)(x-1)(x-2) \\ &= -(x-1)^2(x-2)\end{aligned}$$

When $\det(\mathbf{A}) = 0$, $-(x-1)^2(x-2) = 0 \rightarrow x = 1$ or $x = 2$.

When $x = 1$, let's express the homogeneous linear system $\mathbf{Ax} = \mathbf{0}$ as an augmented matrix.

$$\left(\begin{array}{ccc|c} -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 2 & -5 & 3 & 0 \end{array}\right) \xrightarrow{RREF} \left(\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

Assume $x_3 = p$. Using the information in the augmented matrix, we can get $x_1 = p$ and $x_2 = p$.

When $x = 2$, let's express the homogeneous linear system $\mathbf{Ax} = \mathbf{0}$ as an augmented matrix.

$$\left(\begin{array}{ccc|c} -2 & 1 & 0 & 0 \\ 0 & -2 & 1 & 0 \\ 2 & -5 & 2 & 0 \end{array}\right) \xrightarrow{RREF} \left(\begin{array}{ccc|c} 1 & 0 & -\frac{1}{4} & 0 \\ 0 & 1 & -\frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

Assume $x_3 = p$. Using the information in the augmented matrix, we can get $x_1 = \frac{1}{4}p$ and $x_2 = \frac{1}{2}p$.

10.10 Question 6(a)

An invertible stochastic matrix: $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

A singular stochastic matrix: $\begin{pmatrix} 1 & 0 & 0.8 \\ 0 & 1 & 0.2 \\ 0 & 0 & 0 \end{pmatrix}$

10.11 Question 6(b)

Let $\mathbf{P} = \begin{pmatrix} p_{11} & p_{12} & p_{13} & \cdots & p_{1m} \\ p_{21} & p_{22} & p_{23} & \cdots & p_{2m} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ p_{n1} & p_{n2} & p_{n3} & \cdots & p_{nm} \end{pmatrix}$.

Then, $\mathbf{I} - \mathbf{P} = \begin{pmatrix} 1 - p_{11} & -p_{12} & -p_{13} & \cdots & -p_{1m} \\ -p_{21} & 1 - p_{22} & -p_{23} & \cdots & -p_{2m} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -p_{n1} & -p_{n2} & -p_{n3} & \cdots & 1 - p_{nm} \end{pmatrix}$.

The main idea is to utilize the fact that the sum of every column in $\mathbf{I} - \mathbf{P}$ is always 0 – the property of stochastic matrix.

To really utilize this information, we will perform elementary row operations on $\mathbf{I} - \mathbf{P}$ by adding row n with row $1, 2, ..., n-1$. By doing this, we will get the matrix:

$$\begin{pmatrix} 1-p_{11} & -p_{12} & -p_{13} & \dots & -p_{1m} \\ -p_{21} & 1-p_{22} & -p_{23} & \dots & -p_{2m} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1-\sum_{i=1}^{n} p_{i1} & 1-\sum_{i=1}^{n} p_{i2} & 1-\sum_{i=1}^{n} p_{i3} & \cdots & 1-\sum_{i=1}^{n} p_{im} \end{pmatrix}$$

Observe the last row! We know that $\sum_{i=1}^{n} p_{ij}$ will always equal to 1 for $1 \leq j \leq m$. Therefore, the last row is all equal to 0.

$$\begin{pmatrix} 1-p_{11} & -p_{12} & -p_{13} & \dots & -p_{1m} \\ -p_{21} & 1-p_{22} & -p_{23} & \dots & -p_{2m} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}$$

Because the matrix contains a zero row, we know that it will never be invertible. A non invertible matrix will always have a determinant of zero. Therefore, $\mathbf{I} - \mathbf{P}$ is singular.

10.12 Question 6(c)

Let's sum up every column in the matrix!

$$\begin{aligned} 0.2 + 0.3 + 0.5 &= 1 \\ 0.8 + 0.2 + 0 &= 1 \\ 0.4 + 0.4 + 0.2 &= 1 \end{aligned}$$

All of them sums up to 1, so $\mathbf{P}$ is indeed a stochastic matrix.

$$\mathbf{I} - \mathbf{P} = \begin{pmatrix} 0.8 & -0.8 & -0.4 \\ -0.3 & 0.8 & -0.4 \\ -0.5 & 0 & 0.8 \end{pmatrix}$$

Then, let's express the homogeneous system $(\mathbf{I} - \mathbf{P})\mathbf{x} = \mathbf{0}$ as an augmented matrix.

$$\left(\begin{array}{ccc|c} 0.8 & -0.8 & -0.4 & 0 \\ -0.3 & 0.8 & -0.4 & 0 \\ -0.5 & 0 & 0.8 & 0 \end{array}\right) \xrightarrow{RREF} \left(\begin{array}{ccc|c} 1 & 0 & -\frac{8}{5} & 0 \\ 0 & 1 & -\frac{11}{10} & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

Assume $x_3 = p$! We will get the solution to this homogeneous system as $x_1 = \frac{8}{5}p$ and $x_2 = \frac{11}{10}p$.

10.13 Question 7

Focus on the first row! We want to remove the occurrence of p, q, r! Apply the elementary row operation!

$$\begin{vmatrix} a+px & b+qx & c+rx \\ p+ux & q+vx & r+wx \\ u+ax & v+bx & w+cx \end{vmatrix}$$

$$\xrightarrow{R_1-R_2x} \begin{vmatrix} a+px-px-ux^2 & b+qx-qx-vx^2 & c+rx-rx-wx^2 \\ p+ux & q+vx & r+wx \\ u+ax & v+bx & w+cx \end{vmatrix}$$

$$= \begin{vmatrix} a-ux^2 & b-vx^2 & c-wx^2 \\ p+ux & q+vx & r+wx \\ u+ax & v+bx & w+cx \end{vmatrix}$$

$$\xrightarrow{R_1+R_3x^2} \begin{vmatrix} a-ux^2+ux^2+ax^3 & b-vx^2+vx^2+bx^3 & c-wx^2+wx^2+cx^3 \\ p+ux & q+vx & r+wx \\ u+ax & v+bx & w+cx \end{vmatrix}$$

$$= \begin{vmatrix} a+ax^3 & b+bx^3 & c+cx^3 \\ p+ux & q+vx & r+wx \\ u+ax & v+bx & w+cx \end{vmatrix} = \begin{vmatrix} a(1+x^3) & b(1+x^3) & c(1+x^3) \\ p+ux & q+vx & r+wx \\ u+ax & v+bx & w+cx \end{vmatrix}$$

Note that even though we perform these elementary row operations, the determinant won't change because the elementary matrix corresponds to $R_i + aR_j$ has a determinant of 1.

Now, let's further simplify the matrix by multiplying the first row with $\frac{1}{1+x^3}$. This elementary row operation has a determinant of $\frac{1}{1+x^3}$.

$$\begin{vmatrix} a(1+x^3) & b(1+x^3) & c(1+x^3) \\ p+ux & q+vx & r+wx \\ u+ax & v+bx & w+cx \end{vmatrix} \xrightarrow{\frac{1}{1+x^3}R_1} \begin{vmatrix} a & b & c \\ p+ux & q+vx & r+wx \\ u+ax & v+bx & w+cx \end{vmatrix}$$

Next, we will reduce the matrix with the elementary row operation $R_i + aR_j$, which has a determinant of 1!

$$\begin{vmatrix} a & b & c \\ p+ux & q+vx & r+wx \\ u+ax & v+bx & w+cx \end{vmatrix} \xrightarrow{R_3-xR_1} \begin{vmatrix} a & b & c \\ p+ux & q+vx & r+wx \\ u & v & w \end{vmatrix}$$

$$\xrightarrow{R_2-xR_3} \begin{vmatrix} a & b & c \\ p & q & r \\ u & v & w \end{vmatrix}$$

So far, to finally reduce the matrix to $\begin{vmatrix} a & b & c \\ p & q & r \\ u & v & w \end{vmatrix}$, we use multiple $R_i + aR_j$ operations and one $\frac{1}{1+x^3}R_1$. Because the elementary matrix corresponding to

$R_i + aR_j$ has a determinant of 1, we don't need to care that much.
Now, let's formulate the determinant after applying these elementary row operations.

$$\begin{vmatrix} a+px & b+qx & c+rx \\ p+ux & q+vx & r+wx \\ u+ax & v+bx & w+cx \end{vmatrix} = \frac{\begin{vmatrix} a & b & c \\ p & q & r \\ u & v & w \end{vmatrix}}{\frac{1}{1+x^3}x^{‘}} = (1+x^3)\begin{vmatrix} a & b & c \\ p & q & r \\ u & v & w \end{vmatrix}$$

11 Tutorial 4

11.1 Question 1a

Because $\mathbf{A}$ and $\mathbf{PDP}^{-1}$ are equal, their determinant must also be equal, i.e., $\det(\mathbf{A}) = \det(\mathbf{PDP}^{-1})$. Then, by the property of products in determinant, we will have $\det(\mathbf{A}) = \det(\mathbf{P}) \cdot \det(\mathbf{D}) \cdot \det(\mathbf{P^{-1}})$.

Then, as we know that $\mathbf{P}$ is invertible, by the property of inverse in determinant, $\det(\mathbf{P}^{-1}) = \det(\mathbf{P})^{-1}$. Using these ideas:

$$\det(\mathbf{A}) = \det(\mathbf{P}) \cdot \det(\mathbf{D}) \cdot \det(\mathbf{P}^{-1}) = \det(\mathbf{P}) \cdot \det(\mathbf{D}) \cdot \det(\mathbf{P})^{-1} = \det(\mathbf{D})$$

11.2 Question 1b

From Question 1(a), we have obtained that given $\mathbf{A} = \mathbf{PDP}^{-1}$ for some invertible matrix $\mathbf{P}$ and $\mathbf{D}$ is a diagonal matrix, $\det(\mathbf{A}) = \det(\mathbf{D})$.

$$\begin{aligned} \mathbf{A} \text{ is invertible} &\Leftrightarrow \det(\mathbf{A}) \neq 0 && \text{equivalent statements of invertibility} \\ &\Leftrightarrow \det(\mathbf{D}) \neq 0 && \text{given } \det(\mathbf{A}) = \det(\mathbf{D}) \\ &\Leftrightarrow \prod \text{diagonal entries} \neq 0 && \text{by Corollary 7.2.3} \\ &\Leftrightarrow \text{Each diagonal entry} \neq 0 && \text{by the property of multiplication} \end{aligned}$$

11.3 Question 1c

By the property of products in determinant, for any positive integer $k > 0$, this property will hold. Call this property the property of power in determinant.

$$\det(\mathbf{A}^k) = \det(\mathbf{A}) \cdot \det(\mathbf{A}) \cdot \ldots \cdot \det(\mathbf{A}) \text{ (for } k \text{ times)} = \det(\mathbf{A})^k$$

We will apply this property to prove the statement in the problem.

$$\begin{aligned} \mathbf{A} \text{ is nilpotent} &\Leftrightarrow \mathbf{A}^k = \mathbf{0} && \text{by the definition of nilpotent} \\ &\Leftrightarrow \det(\mathbf{A}^k) = \det(\mathbf{0}) \\ &\Leftrightarrow \det(\mathbf{A}^k) = 0 \\ &\Leftrightarrow \det(\mathbf{A})^k = 0 && \text{by the property of power in determinant} \\ &\Leftrightarrow \det(\mathbf{A}) = 0 \end{aligned}$$

11.4 Question 1d

$$\begin{aligned}
\mathbf{A} \text{ is orthogonal} &\Leftrightarrow \mathbf{A}^T = \mathbf{A}^{-1} && \text{by the definition of orthogonal} \\
&\Leftrightarrow \det(\mathbf{A}^T) = \det(\mathbf{A}^{-1}) \\
&\Leftrightarrow \det(\mathbf{A}) = \det(\mathbf{A}^{-1}) && \text{by Chapter 7.2.2} \\
&\Leftrightarrow \det(\mathbf{A}) = \det(\mathbf{A})^{-1} && \text{by Chapter 7.2.9} \\
&\Leftrightarrow \det(\mathbf{A}) = \frac{1}{\det(\mathbf{A})} \\
&\Leftrightarrow \det(\mathbf{A})^2 = 1 \\
&\Leftrightarrow \det(\mathbf{A}) = \pm 1
\end{aligned}$$

11.5 Question 2a

$$\begin{aligned}
\begin{vmatrix} -1 & 3 & -4 \\ 2 & 4 & 1 \\ -4 & 2 & -9 \end{vmatrix} &= -1 \cdot \begin{vmatrix} 4 & 1 \\ 2 & -9 \end{vmatrix} - 2 \cdot \begin{vmatrix} 3 & -4 \\ 2 & -9 \end{vmatrix} - 4 \cdot \begin{vmatrix} 3 & -4 \\ 4 & 1 \end{vmatrix} \\
&= -1 \cdot (-36 - 2) - 2 \cdot (-27 + 8) - 4 \cdot (3 + 16) \\
&= 38 + 38 - 76 = 0
\end{aligned}$$

11.6 Question 2b

Recall that in the equivalent statements of invertibility, a square matrix $\mathbf{P}$ of order n is invertible if and only if the only solution to the homogeneous linear system $\mathbf{Px} = \mathbf{0}$ is the trivial solution.

Negating both sides of this statement, a square matrix $\mathbf{P}$ of order n is singular if and only if the homogeneous linear system $\mathbf{Px} = \mathbf{0}$ has infinitely many solutions.

Therefore, with this information in mind, showing that the homogeneous linear system $\mathbf{ABx} = 0$ have infinitely many solutions is equivalent to showing that $\mathbf{AB}$ is a singular matrix.

Recall also in the equivalent statements of invertibility, a square matrix $\mathbf{P}$ of order n is invertible if and only if $\det(\mathbf{P}) \neq 0$.

Negating both sides of this statement, a square matrix $\mathbf{P}$ of order n is singular if and only if $\det(\mathbf{P}) = 0$.

Following up, this means that to show $\mathbf{AB}$ is a singular matrix, we will need to show that $\det(\mathbf{AB}) = 0$. By the property of products in determinant, $\det(\mathbf{AB}) = \det(\mathbf{A}) \cdot \det(\mathbf{B})$. Then, from Question 2a, we know that $\det(\mathbf{A}) = 0 \to \det(\mathbf{AB}) = \det(\mathbf{A}) \cdot \det(\mathbf{B}) = 0$, implying that $\mathbf{AB}$ is a singular matrix causing the homogeneous linear system $\mathbf{ABx} = 0$ have infinitely many solutions.

11.7 Question 3a

Let's express the linear system as an augmented matrix.

$$\left(\begin{array}{cccc|c} 1 & 1 & -1 & -2 & 0 \\ 2 & 1 & -1 & 1 & -2 \\ -1 & 1 & -3 & 1 & 4 \end{array}\right) \xrightarrow{RREF} \left(\begin{array}{cccc|c} 1 & 0 & 0 & 3 & -2 \\ 0 & 1 & 0 & -\frac{19}{2} & 2 \\ 0 & 0 & 1 & -\frac{9}{2} & 0 \end{array}\right)$$

Let $d = p$. Then, from the augmented matrix, we can obtain $a = -2 - 3p$, $b = 2 + \frac{19}{2}p$, and $c = \frac{9}{2}p$. By this, we can construct:

$$\begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = \begin{pmatrix} -2-3p \\ 2+\frac{19}{2}p \\ \frac{9}{2}p \\ p \end{pmatrix} = \begin{pmatrix} -2 \\ 2 \\ 0 \\ 0 \end{pmatrix} + p \begin{pmatrix} -3 \\ \frac{19}{2} \\ \frac{9}{2} \\ 1 \end{pmatrix}$$

Therefore, the solution set for this linear system can be expressed as

$$\left\{ \begin{pmatrix} -2 \\ 2 \\ 0 \\ 0 \end{pmatrix} + p \begin{pmatrix} -3 \\ \frac{19}{2} \\ \frac{9}{2} \\ 1 \end{pmatrix},\ p \in \mathbb{R} \right\}$$

11.8 Question 4a(i)

To express the vector as a linear combination of $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$, we will need to solve this equation:

$$c_1 \begin{pmatrix} 2 \\ 1 \\ 0 \\ 3 \end{pmatrix} + c_2 \begin{pmatrix} 3 \\ -1 \\ 5 \\ 2 \end{pmatrix} + c_3 \begin{pmatrix} -1 \\ 0 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \\ -7 \\ 3 \end{pmatrix}$$

Let's express this as an augmented matrix.

$$\left(\begin{array}{ccc|c} 2 & 3 & -1 & 2 \\ 1 & -1 & 0 & 3 \\ 0 & 5 & 2 & -7 \\ 3 & 2 & 1 & 3 \end{array}\right) \xrightarrow{RREF} \left(\begin{array}{ccc|c} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

By this augmented matrix, we can deduce that the vector can be expressed as a linear combination of $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$ as follows:

$$2 \begin{pmatrix} 2 \\ 1 \\ 0 \\ 3 \end{pmatrix} - 1 \begin{pmatrix} 3 \\ -1 \\ 5 \\ 2 \end{pmatrix} - 1 \begin{pmatrix} -1 \\ 0 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \\ -7 \\ 3 \end{pmatrix}$$

11.9 Question 4b

Only three vectors cannot span $\mathbb{R}^4$. Just pick any two vectors $\in \mathbb{R}^4$ which are not a multiple of each other and not in the span of $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$.

11.10 Question 5a

Express the equation as an augmented matrix.

$$\left(\begin{array}{ccc|c} 1 & -1 & -1 & 0 \end{array}\right)$$

Let $y = p$ (the second column) and $z = q$ (the third column), then from the augmented matrix, we can get $x = p + q$. By this, we can construct:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} p+q \\ p \\ q \end{pmatrix} = p\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + q\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$$

By this, we can express $V = \left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \right\}$.

Then, to prove that $\text{span}(S) = V$, we will prove that $S \subseteq V$ and $V \subseteq S$.

To prove that $S \subseteq V$, we will check for the consistency of this augmented matrix!

$$\left(\begin{array}{cc|cc} 1 & 5 & 1 & 1 \\ 1 & 2 & 1 & 0 \\ 0 & 3 & 0 & 1 \end{array}\right) \xrightarrow{RREF} \left(\begin{array}{cc|cc} 1 & 0 & 1 & -\frac{2}{3} \\ 0 & 1 & 0 & \frac{1}{3} \\ 0 & 0 & 0 & 0 \end{array}\right)$$

This augmented matrix is indeed consistent; therefore, we can conclude that S is indeed $\subseteq V$.

Next, to prove that $V \subseteq S$, we will check the consistency of this augmented matrix!

$$\left(\begin{array}{cc|cc} 1 & 1 & 1 & 5 \\ 1 & 0 & 1 & 2 \\ 0 & 1 & 0 & 3 \end{array}\right) \xrightarrow{RREF} \left(\begin{array}{cc|cc} 1 & 0 & 1 & 2 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

This augmented matrix is also consistent; therefore, we can conclude that V is indeed $\subseteq S$. Combining these arguments up, we can deduce that $span(S) = V$.

11.11 Question 5b

$$T = \left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 5 \\ 2 \\ 3 \end{pmatrix} \right\} \bigcup \left\{ \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\} = \left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 5 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$$

To check that $span(T) = \mathbb{R}^3$, we will check the existence of zero rows from this matrix.

$$\begin{pmatrix} 1 & 5 & 0 \\ 1 & 2 & 0 \\ 0 & 3 & 1 \end{pmatrix} \xrightarrow{RREF} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

There are no zero rows, $span(T) = \mathbb{R}^3$.

11.12 Question 6(iii)

To check that $span(S) = \mathbb{R}^4$, let's check the existence of non zero rows from this matrix.

$$\begin{pmatrix} 6 & 2 & 3 & 5 & 0 \\ 4 & 0 & 2 & 6 & 4 \\ -2 & 0 & -1 & -3 & -2 \\ 4 & 1 & 2 & 2 & -1 \end{pmatrix} \xrightarrow{RREF} \begin{pmatrix} 1 & 0 & \frac{1}{2} & 0 & -\frac{1}{2} \\ 0 & 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

There are zero rows, so $span(S) \neq \mathbb{R}^4$.

12 MA1508E Midterm AY2023/2024 Sem 2

12.1 Question 1

Express the linear system as an augmented matrix!

$$\left(\begin{array}{ccc|c} 1 & a & a & 1 \\ 3 & 3 & 3a & 3-3a \\ 1 & 1 & 2 & 3-2a \\ 2 & a+1 & a+2 & 4-2a \end{array}\right)$$

Perform these Matlab operations on this augmented matrix!

```
syms a
A = [1 a a 1; 3 3 3*a 3-3*a; 1 1 2 3-2*a; 2 a+1 a+2 4-2*a];
A(2,:) = A(2,:)-3*A(1,:);
A(3,:) = A(3,:)-A(1,:);
A(4,:) = A(4,:)-2*A(1,:);
A(4,:) = A(4,:)-A(3,:);
A(2,:) = A(2,:)-3*A(3,:);
A([2,3], :) = A([3, 2], :);
display(A)
```

We will get the augmented matrix:

$$\left(\begin{array}{ccc|c} 1 & a & a & 1 \\ 0 & 1-a & 2-a & 2-2a \\ 0 & 0 & 3a-6 & 3a-6 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

If $3a-6 \neq 0$, we can divide the third row by $3a-6$ and find that $z = 1$. However, it is given in the question that $z = 3$. This is a contradiction. Thus, $3a - 6 = 0 \to a = 2$.

Substitute the value of $a = 2$ into the augmented matrix; then, find the RREF of the augmented matrix.

```
A_subs = subs(A, a, 2);
display(rref(A_subs));
```

We will get the augmented matrix to be

$$\left(\begin{array}{ccc|c} 1 & 0 & 2 & -3 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

We know the value of the non-pivot column 3, z, is equal to 3. Therefore, we can get $x = -3 - 2z = -3 - 6 = -9$ and $y = 2$.

Correct MCQ Option: **Option A**

12.2 Question 2

Let's express the system into an augmented matrix, and find the RREF!

$$\left(\begin{array}{ccccc|c} 1 & 2 & -1 & 1 & -1 & 1 \\ 1 & 1 & 2 & 1 & 0 & 11 \\ 1 & 2 & 0 & -1 & -1 & 9 \\ 2 & 1 & 2 & 1 & -1 & 10 \\ 2 & -1 & -1 & 0 & -1 & -8 \end{array}\right) \xrightarrow{RREF} \left(\begin{array}{ccccc|c} 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 4 \\ 0 & 0 & 1 & 0 & 0 & 4 \\ 0 & 0 & 0 & 1 & 0 & -2 \\ 0 & 0 & 0 & 0 & 1 & 2 \end{array}\right)$$

Let's extract the information from this augmented matrix!

$$\begin{aligned} x^2 &= 1 \\ y^2 &= 4 \\ z^2 &= 4 \\ xy &= -2 \\ xz &= 2 \end{aligned}$$

Because $z > 1$ and $xz = 2$, $x > 0$. Then, divide $xy = -2$ with $xz = 2$, we will get $\frac{y}{z} = -1 \to y = -z$. Then, because $z > 1$, then $y < -1$.

Reasonable MCQ Options: **Option A**

12.3 Question 3

Assume the number of 10 cent coins is x, 50 cent coins is y, and \$1 coins is z. We can formulate the condition in the problem to be:

$$\begin{aligned} 0.1x + 0.5y + z &= 10 \\ x + y + z &= 32 \end{aligned}$$

Next, let's express this linear system as an augmented matrix!

$$\left(\begin{array}{ccc|c} 0.1 & 0.5 & 1 & 10 \\ 1 & 1 & 1 & 32 \end{array}\right) \xrightarrow{RREF} \left(\begin{array}{ccc|c} 1 & 0 & -1.25 & 15 \\ 0 & 1 & 2.25 & 17 \end{array}\right)$$

Assume $z = p$! From the augmented matrix, we can get that $x = 15 + 1.25p$ and $y = 17 - 2.25p$.

It is a common sense to know that number of coins must be an integer. Thus, p must be a multiple of 4. Then, it is also a common sense to know that number of coins must be positive. Thus, $17 - 2.25p \geq 0 \to 17 \geq 2.25p \to p \leq \frac{17}{2.25} \approx 7.56$.

Let's list up the condition:

1. $p > 0$
2. $p \leq 7.56$
3. p must be a multiple of 4

The only value of p that satisfies these conditions are 4. Thus, the number of 50 cents in the bag must be 8.

Correct MCQ Option: **Option A**

12.4 Question 5

Let's express the linear system as an augmented matrix!

$$\left(\begin{array}{ccc|c} -1 & 1 & -(1-\alpha) & 0 \\ \alpha & 1-\alpha & -(1-\alpha)^2 & 0 \\ 1 & -1 & \frac{-(1-\alpha)^2}{\alpha} & \beta \end{array}\right)$$

Perform these Matlab operations on this augmented matrix!

```
syms a b
A = [-1 1 -(1-a) 0; a 1-a -(1-a)^2 0; 1 -1 -(1-a)^2/a b;];
A(3,:) = A(3,:) + A(1,:);
A(2,:) = A(2,:) + a*A(1,:);
A(1,:) = -A(1,:);
A = simplify(A);
display(A)
```

Note: We need to make sure if we use variables in our matrix, our elementary row operations cannot break the scope of the variable. For example, if we want to perform row multiplication with α, α must be > 0.

We will get an updated augmented matrix, which is

$$\left(\begin{array}{ccc|c} 1 & -1 & 1-\alpha & 0 \\ 0 & 1 & \alpha-1 & 0 \\ 0 & 0 & \frac{\alpha-1}{\alpha} & \beta \end{array}\right)$$

Because $\alpha \neq 0$, we can multiply the last row with α!

$$\left(\begin{array}{ccc|c} 1 & -1 & 1-\alpha & 0 \\ 0 & 1 & \alpha-1 & 0 \\ 0 & 0 & \frac{\alpha-1}{\alpha} & \beta \end{array}\right) \xrightarrow{\alpha R_3} \left(\begin{array}{ccc|c} 1 & -1 & 1-\alpha & 0 \\ 0 & 1 & \alpha-1 & 0 \\ 0 & 0 & \alpha-1 & \alpha\beta \end{array}\right)$$
$$\xrightarrow{R_2-R_3} \left(\begin{array}{ccc|c} 1 & -1 & 1-\alpha & 0 \\ 0 & 1 & 0 & -\alpha\beta \\ 0 & 0 & \alpha-1 & \alpha\beta \end{array}\right)$$

Look! Because $0 < \alpha < 1$, the system is always consistent. Now, look at row 2 to observe the value of y! We can see that $y = -\alpha\beta$. As we know that $\alpha > 0$ and $\beta > 0$, we can conclude that y will always < 0.

Appropriate MCQ Choice: **Option B**

Note: Another approach for this problem is to use the Cramer's Rule discussed in Tutorial 3.

12.5 Question 6

First, let's find **A**! We can find that by inversing the elementary row operation that we done:

1. Inverse of $R_4 + \frac{9}{31}R_3$: $R_4 - \frac{9}{31}R_3$
2. Inverse of $R_4 + 2R_2$: $R_4 - 2R_2$
3. Inverse of $R_3 - 7R_2$: $R_3 + 7R_2$
4. Inverse of $R_3 - 2R_1$: $R_3 + 2R_1$

After finding A, find the RREF for each of the linear system. Here is the matlab code to achieve this:

```
syms a b c
A = [1 a 3 3; 0 -1 b -5; 0 0 -31 c; 0 0 0 0;];
A(4,:) = A(4,:) - 9/31 * A(3,:);
A(4,:) = A(4,:) - 2 * A(2,:);
A(3,:) = A(3,:) + 7 * A(2,:);
A(3,:) = A(3,:) + 2 * A(1,:);
A = simplify(A);

OA = [8; 3; 6; 3];
OAM = [A OA];
display(rref(OAM));

OB = [8; 3; 6; 4];
OBM = [A OB];
display(rref(OBM));

OC = [8; 4; 6; 3];
OCM = [A OC];
```

```
display(rref(OCM));

OD = [8; 4; 6; 4];
ODM = [A OD];
display(rref(ODM));
```

When running this code, we will get this output:

```
ans =

[1, 0, 0, (3*c)/31 - 5*a + (a*b*c)/31 + 3, 3*a - a*b + 5]
[0, 1, 0,                    5 - (b*c)/31,         b - 3]
[0, 0, 1,                           -c/31,             1]
[0, 0, 0,                               0,             0]

ans =

[1, 0, 0, (3*c)/31 - 5*a + (a*b*c)/31 + 3, 0]
[0, 1, 0,                    5 - (b*c)/31, 0]
[0, 0, 1,                           -c/31, 0]
[0, 0, 0,                               0, 1]

ans =

[1, 0, 0, (3*c)/31 - 5*a + (a*b*c)/31 + 3, 0]
[0, 1, 0,                    5 - (b*c)/31, 0]
[0, 0, 1,                           -c/31, 0]
[0, 0, 0,                               0, 1]

ans =

[1, 0, 0, (3*c)/31 - 5*a + (a*b*c)/31 + 3, 0]
[0, 1, 0,                    5 - (b*c)/31, 0]
[0, 0, 1,                           -c/31, 0]
[0, 0, 0,                               0, 1]
```

Observe that the linear system is consistent is when $\mathbf{b} = \begin{pmatrix} 8 \\ 3 \\ 6 \\ 3 \end{pmatrix}$

Correct MCQ Option: **Option A**

Note: Another approach to solve this problem is to apply the same elementary row operations to each of the options. Then, we will check whether the last row will be 0 or not. If it is 0, then the system is consistent. Otherwise, it is not consistent.

12.6 Question 7

From matrix block multiplication, we know that

$$\mathbf{AB} = \mathbf{A}\begin{pmatrix} \mathbf{b}_1 & \mathbf{b}_2 & ... & \mathbf{b}_n \end{pmatrix} = \begin{pmatrix} \mathbf{Ab}_1 & \mathbf{Ab}_2 & ... & \mathbf{Ab}_n \end{pmatrix}$$

where $\mathbf{b}_1, \mathbf{b}_2, ..., \mathbf{b}_n$ is the columns matrix of $\mathbf{B}$.

Using this idea, we can extract the information from the matrix equation:

1. $\mathbf{A} \cdot \begin{pmatrix} 0 \\ -1 \\ -1 \end{pmatrix} = \begin{pmatrix} 1 \\ -3 \\ -4 \\ -8 \end{pmatrix}$

2. $\mathbf{A} \cdot \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} = \begin{pmatrix} -1 \\ 2 \\ 0 \\ 8 \end{pmatrix}$

3. $\mathbf{A} \cdot \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix} = \begin{pmatrix} -1 \\ 7 \\ 9 \\ 13 \end{pmatrix}$

4. $\mathbf{A} \cdot \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -2 \\ -1 \\ 2 \\ 6 \end{pmatrix}$

The question asks us to find what is the value of $\mathbf{A} \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$. In order to find this, we need to search for combination of these 4 matrix, that is $\begin{pmatrix} 0 \\ -1 \\ -1 \end{pmatrix}$, $\begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}$, $\begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}$, $\begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}$ that can evaluates to $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$.

Let's formulate it into an equation! Assume p, q, r, s as the coefficient of each of the matrix respectively. Therefore, the matrix equation will be:

$$p\begin{pmatrix} 0 \\ -1 \\ -1 \end{pmatrix} + q\begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} + r\begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix} + s\begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

Then, express this equation as an augmented matrix and find the RREF!

$$\left(\begin{array}{cccc|c} 0 & 1 & 0 & -1 & 1 \\ -1 & 2 & 1 & 1 & 1 \\ -1 & 0 & 2 & 1 & 1 \end{array}\right) \xrightarrow{RREF} \left(\begin{array}{cccc|c} 1 & 0 & 0 & -5 & 3 \\ 0 & 1 & 0 & -1 & 1 \\ 0 & 0 & 1 & -2 & 2 \end{array}\right)$$

Assume $s = t$. Then, from the augmented matrix, we can deduce that $p = 3+5t$, $q = 1 + t$, $r = 2 + 2t$. To make it simple, we will choose $t = -1$, so we get the value $p = -2$, $q = 0$, $r = 0$, $s = -1$. With this coefficient, we can find what we want!

$$\begin{aligned}
-2 \cdot \mathbf{A} \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} &= -2 \cdot \mathbf{A} \cdot \begin{pmatrix} 0 \\ -1 \\ -1 \end{pmatrix} - 1 \cdot \mathbf{A} \cdot \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} \\
&= -2 \cdot \begin{pmatrix} 1 \\ -3 \\ -4 \\ -8 \end{pmatrix} - 1 \cdot \begin{pmatrix} -2 \\ -1 \\ 2 \\ 6 \end{pmatrix} = \begin{pmatrix} -2+2 \\ 6+1 \\ 8-2 \\ 16-6 \end{pmatrix} = \begin{pmatrix} 0 \\ 7 \\ 6 \\ 10 \end{pmatrix}
\end{aligned}$$

Correct MCQ Option: **Option A**

12.7 Question 8

$$\det(\mathbf{A}) = \frac{\begin{vmatrix} 1 & 3 & 5 & -1 & 2 \\ 0 & 2 & -1 & 3 & 0 \\ 0 & 0 & -1 & 2 & 9 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{vmatrix}}{-4 \cdot (-1) \cdot \frac{5}{3}}$$

The division by -4 comes from $(-4)R_5$; the division by -1 comes from $R_1 \leftrightarrow R_2$; the division by $\frac{5}{3}$ comes from $\frac{5}{3}R_3$.

Then, find the determinant by using the last row. The thing that matter then is only the $(5, 4)$-th entries.

$$\det(\mathbf{A}) = (-1)^9 \cdot \begin{vmatrix} 1 & 3 & 5 & 2 \\ 0 & 2 & -1 & 0 \\ 0 & 0 & -1 & 9 \\ 0 & 0 & 0 & 1 \end{vmatrix} \cdot \frac{3}{20} = (-1) \cdot (-2) \cdot \frac{3}{20} = \frac{3}{10}$$

Correct MCQ Option: **Option A**

12.8 Question 12

Let's assess the consistency of these augmented matrix corresponding to each of vectors!

1. For $\mathbf{v}_1$, the augmented matrix is $\left(\begin{array}{ccc|c} 2 & 2 & 1 & 1 \\ -1 & 2 & 2 & 2 \\ 0 & 2 & -1 & -1 \\ 1 & 1 & 2 & 1 \end{array}\right)$. Reducing this to its RREF form, we would get $\left(\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right)$ – inconsistent

2. For $\mathbf{v}_2$, the augmented matrix is $\left(\begin{array}{ccc|c} 2 & 2 & 1 & 2 \\ -1 & 2 & 2 & 1 \\ 0 & 2 & -1 & -4 \\ 1 & 1 & 2 & 4 \end{array}\right)$. Reducing this to its RREF form, we would get $\left(\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{array}\right)$ – consistent

3. For $\mathbf{v}_3$, the augmented matrix is $\left(\begin{array}{ccc|c} 2 & 2 & 1 & 4 \\ -1 & 2 & 2 & 3 \\ 0 & 2 & -1 & -2 \\ 1 & 1 & 2 & 2 \end{array}\right)$. Reducing this to its RREF form, we would get $\left(\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right)$ – inconsistent

4. For $\mathbf{v}_4$, the augmented matrix is $\left(\begin{array}{ccc|c} 2 & 2 & 1 & -2 \\ -1 & 2 & 2 & 3 \\ 0 & 2 & -1 & -4 \\ 1 & 1 & 2 & 2 \end{array}\right)$. Reducing this to its RREF form, we would get $\left(\begin{array}{ccc|c} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{array}\right)$ – consistent

The augmented matrix that is consistent is the one for $\mathbf{v}_2, \mathbf{v}_4$. Therefore, the correct MCQ options to choose is **Option A**.

12.9 Question 13

V is an affine space. V is a subspace if the particular solution is the zero vector. However, no matter what the value of a you plugged in, we will never get the zero vector. Therefore, V is not a subspace.

12.9.1 Why is the Answer Key Wrong?

The Answer Key says that V is a subspace for any value $a \in R$. Here is the explanation it offers:

In fact, $V = \mathbb{R}^2$. For any $\begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2$, $\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} a \\ 1 \end{pmatrix} + s\begin{pmatrix} 1 \\ 1 \end{pmatrix} + t\begin{pmatrix} 1 \\ 2 \end{pmatrix}$ if and only if $s\begin{pmatrix} 1 \\ 1 \end{pmatrix} + t\begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} x-a \\ y-1 \end{pmatrix}$,

$$\left(\begin{array}{cc|c} 1 & 1 & x-a \\ 1 & 2 & y-1 \end{array}\right) \xrightarrow{RREF} \left(\begin{array}{cc|c} 1 & 0 & 2x-y-1 \\ 0 & 1 & y-x \end{array}\right),$$

which is consistent for any x, y, a.

Figure 5: Answer Key Argument

If the RREF of the augmented matrix is $\left(\begin{array}{cc|c} 1 & 0 & 2x-y-1 \\ 0 & 1 & y-x \end{array}\right)$, V is indeed a subspace because s and t can be expressed just by x and y, not dependent to the value of a.

However, this is not the case. The actual RREF of the augmented matrix is $\left(\begin{array}{cc|c} 1 & 0 & 2x-2a-y+1 \\ 0 & 1 & a-x+y-1 \end{array}\right)$ – the Answer Key calculate it wrongly. As we can see, s and t is dependent to the value of a. A subspace should be independent of external parameters—its structure should not change based on the choice of a. If s and t depend on a, then the span of the basis vectors does not consistently form a subspace for all values of a, violating the closure properties required for subspaces.

12.10 Question 14

For V to be a subspace, this property needs to be satisfied: for any $\mathbf{u}, \mathbf{v} \in V$, $\mathbf{u} + \mathbf{v} \in V$.

Now, for whatever value of c_1 and c_2, take $\mathbf{u} = \begin{pmatrix} 1 \\ 1 \\ c_1 \end{pmatrix}$ and $\mathbf{v} = \begin{pmatrix} 1 \\ -1 \\ c_2 \end{pmatrix}$, so $\mathbf{u} + \mathbf{v} = \begin{pmatrix} 2 \\ 0 \\ c_1 + c_2 \end{pmatrix}$. However, $\mathbf{u} + \mathbf{v}$ is not $\in V$ as $2^2 \neq 0^2$ no matter what the value of c_1 and c_2 is.

Because of this, we can conclude that V is not a subspace.

Correct MCQ Options: **Option A**

12.11 Question 15

According to the definition of subspace, a subspace always contains the zero vector. This means $\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ is in V. Substituting this into the equation, we will get $a(0) + b(0) + c(0) = d \to d = 0$.

Then, for the other vectors, $\begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}$ and $\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$, when we substitute these in to the equation we will get $a+b-c = 0$ and $a-b = 0$. From the problem, we know that $c = 2$. After substituting this, we have $a + b = 2$ and $a = b$. Solving this, we have $a = b = 1$.

Correct MCQ Options: **Option A**

12.12 Question 16

Use this code to solve this problem:

```
syms t

% Compute the coefficient matrix
f = [cos(2*pi*t) sin(2*pi*t) cos(4*pi*t) sin(4*pi*t) cos(6*pi
    *t) sin(6*pi*t)];
A = [subs(f,t,0); subs(f,t,1/6); subs(f,t,1/4); subs(f,t,1/3)
    ; subs(f,t,1/2); subs(f,t,2/3)];

% Compute the RHS column
res = [-2; 2+7*sqrt(3)/2; 7; 1-3*sqrt(3)/2; -4; 1+3*sqrt(3)
    /2];

% Combine the coefficient matrix with the RHS column
combined = [A res];

% Display the RREF of the augmented matrix
display(eval(rref(combined)))
```

Note:
The *subs* function in Matlab only substitutes a numerical value for the *syms* you declared, so the results will remain in exact form.

This also happens for *rref*! *rref* function preserves exact symbolic values when working with symbolic inputs.

If we want to calculate the real decimal format, we can use one of these functions: *eval*, *double*, *vpa*. In this code, I use *eval*.

12.13 Question 17

The main idea is to utilize:

$$\begin{pmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n \end{pmatrix}^p = \begin{pmatrix} d_1^p & 0 & \cdots & 0 \\ 0 & d_2^p & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n^p \end{pmatrix}$$

To utilize this information, we will first assume $\mathbf{B} = \begin{pmatrix} \alpha & \beta \\ \beta & -\alpha \end{pmatrix}$, the top-leftmost 2×2 block entries and $\mathbf{C} = \begin{pmatrix} 1 & \gamma \\ 0 & 1 \end{pmatrix}$. By this assumption, we can simplify the matrix to be:

$$\mathbf{A} = \begin{pmatrix} \mathbf{B} & 0 \\ 0 & \mathbf{C} \end{pmatrix}$$

Then, because $\mathbf{A}$ is a diagonal matrix, $\mathbf{A}^{1508E} = \begin{pmatrix} \mathbf{B}^{1508E} & 0 \\ 0 & \mathbf{C}^{1508E} \end{pmatrix}$

Now, we are left with $\mathbf{B}^{1508E}$ and $\mathbf{C}^{1508E}$.

First, let's try to square $\mathbf{B}$!

$\mathbf{B}^2 = \begin{pmatrix} \alpha & \beta \\ \beta & -\alpha \end{pmatrix} \begin{pmatrix} \alpha & \beta \\ \beta & -\alpha \end{pmatrix} = \begin{pmatrix} \alpha^2 + \beta^2 & \beta\alpha - \alpha\beta \\ \alpha\beta - \beta\alpha & \beta^2 + (-\alpha)^2 \end{pmatrix} = \begin{pmatrix} \alpha^2 + \beta^2 & 0 \\ 0 & \alpha^2 + \beta^2 \end{pmatrix}$

Observe that $\mathbf{B}^2$ is a diagonal matrix. Therefore,

$$(\mathbf{B}^2)^{754E} = \begin{pmatrix} \alpha^2 + \beta^2 & 0 \\ 0 & \alpha^2 + \beta^2 \end{pmatrix}^{754E} = \begin{pmatrix} (\alpha^2 + \beta^2)^{754E} & 0 \\ 0 & (\alpha^2 + \beta^2)^{754E} \end{pmatrix}$$

Next, let's observe what happen when we power $\mathbf{C}$!

$$\mathbf{C}^2 = \begin{pmatrix} 1 & \gamma \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & \gamma \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & \gamma + \gamma \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2\gamma \\ 0 & 1 \end{pmatrix}$$

$$\mathbf{C}^3 = \mathbf{C}^2\mathbf{C} = \begin{pmatrix} 1 & 2\gamma \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & \gamma \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & \gamma + 2\gamma \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 3\gamma \\ 0 & 1 \end{pmatrix}$$

$$\mathbf{C}^4 = \mathbf{C}^3\mathbf{C} = \begin{pmatrix} 1 & 3\gamma \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & \gamma \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & \gamma + 3\gamma \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 4\gamma \\ 0 & 1 \end{pmatrix}$$

Observe that there is a pattern that we can generalize to:

$$\mathbf{C}^n = \begin{pmatrix} 1 & n\gamma \\ 0 & 1 \end{pmatrix}$$

By this pattern, we can find that

$$\mathbf{C}^{1508E} = \begin{pmatrix} 1 & 1508E\gamma \\ 0 & 1 \end{pmatrix}$$

Plugging $\mathbf{B}^{1508E}$ and $\mathbf{C}^{1508E}$ back to $\mathbf{A}^{1508E}$, we will get the matrix:

$$\begin{pmatrix} (\alpha^2+\beta^2)^{754E} & 0 & 0 & 0 \\ 0 & (\alpha^2+\beta^2)^{754E} & 0 & 0 \\ 0 & 0 & 1 & 1508E\gamma \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

12.14 Question 18

First of all, because $\mathbf{u}$ and $\mathbf{v}$ are unit vectors, $||\mathbf{u}|| = ||\mathbf{v}|| = 1$. Then, according to the property of norm, as $\mathbf{w} = c\mathbf{v}$, $||\mathbf{w}|| = |c| \cdot ||\mathbf{v}||$. Because $c > 0$ and $||\mathbf{v}|| = 1$, $||\mathbf{w}|| = c$

Then, the angle of $\mathbf{u}$ and $\mathbf{v}$ is the same to the angle of $\mathbf{u}$ and $\mathbf{w}$ as $\mathbf{v}$ and $\mathbf{w}$ have the same direction. Using the formula of dot product, we will get:

$$\frac{\mathbf{u} \cdot \mathbf{w}}{||\mathbf{u}|| \cdot ||\mathbf{w}||} = \cos(120) = -\frac{1}{2}$$

Substituting the value of $||\mathbf{u}||$ and $||\mathbf{w}||$, we will get:

$$\mathbf{u} \cdot \mathbf{w} = -\frac{c}{2}$$

Then, we are given that the distance between $\mathbf{u}$ and $\mathbf{w}$ is $\sqrt{7}$.

$$||\mathbf{u} - \mathbf{w}|| = \sqrt{7}$$

Then, according to the definition of norm, we can extend this to be:

$$\sqrt{(\mathbf{u} - \mathbf{w}) \cdot (\mathbf{u} - \mathbf{w})} = \sqrt{7}$$
$$(\mathbf{u} - \mathbf{w}) \cdot (\mathbf{u} - \mathbf{w}) = 7$$

Then, by the distributivity and symmetric property of dot product, we can simplify this further to:

$$\mathbf{u} \cdot \mathbf{u} - \mathbf{w} \cdot \mathbf{u} - \mathbf{u} \cdot \mathbf{w} + \mathbf{w} \cdot \mathbf{w} = 7$$

Next, we will reverse the definition of norm! We know that $||\mathbf{u}|| = \sqrt{\mathbf{u} \cdot \mathbf{u}}$. This means that $\mathbf{u} \cdot \mathbf{u} = ||\mathbf{u}||^2$. Similarly, $\mathbf{w} \cdot \mathbf{w} = ||\mathbf{w}||^2$.

Let's apply this to the last equation that we have!

$$\begin{aligned} &||\mathbf{u}||^2 - 2(\mathbf{u} \cdot \mathbf{w}) + ||\mathbf{w}||^2 = 7 \\ &1^2 - 2 \cdot (-\frac{c}{2}) + c^2 = 7 \\ &1 + c + c^2 = 7 \\ &c^2 + c - 6 = 0 \\ &(c+3)(c-2) = 0 \\ &c = -3 \vee c = 2 \end{aligned}$$

We know that $c > 0$, so $c = 2 = ||\mathbf{w}||$.

13 Week 7

13.1 Row Space and Column Space

Let $\mathbf{A}$ be an $m \times n$ matrix,

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

The **row space** of $\mathbf{A}$ is the subspace of $\mathbb{R}^n$ spanned by the rows of $\mathbf{A}$. We usually denote a row space of $\mathbf{A}$ as $\text{Row}(\mathbf{A})$.

$$\text{Row}(\mathbf{A}) = \text{span}\{\begin{pmatrix} a_{11} & a_{12} & ... & a_{1n} \end{pmatrix}, \begin{pmatrix} a_{21} & a_{22} & ... & a_{2n} \end{pmatrix}, ..., \begin{pmatrix} a_{m1} & a_{m2} & ... & a_{mn} \end{pmatrix}\}$$

The **column space** of $\mathbf{A}$ is the subspace of $\mathbb{R}^m$ spanned by the columns of $\mathbf{A}$. We usually denote a column space of $\mathbf{A}$ as $\text{Column}(\mathbf{A})$.

$$\text{Column}(\mathbf{A}) = \text{span}\left\{ \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix}, \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix}, ..., \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix} \right\}$$

13.2 A Theorem about Row Space

Suppose $\mathbf{A}$ and $\mathbf{B}$ are row equivalent matrices. Then, $\text{Row}(\mathbf{A}) = \text{Row}(\mathbf{B})$.

Intuition: Recall that linear span of $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, ..., \mathbf{u}_k$ contains the linear combinations of $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, ..., \mathbf{u}_k$. Then, observe that applying elementary row operations to a set of vectors do not do anything towards their linear combinations – their linear combination will stay the same, implying that the linear span is equal.

13.3 Another Theorem about Row Space

For any matrix $\mathbf{A}$, the nonzero rows of the reduced row echelon form (RREF) of $\mathbf{A}$ form a basis for the row space of $\mathbf{A}$.

13.4 A Theorem about Column Space

Let $\mathbf{A} = \begin{pmatrix} \mathbf{a}_1 & \mathbf{a}_2 & ... & \mathbf{a}_n \end{pmatrix}$ and $\mathbf{B} = \begin{pmatrix} \mathbf{b}_1 & \mathbf{b}_2 & ... & \mathbf{b}_n \end{pmatrix}$ be row equivalent $m \times n$ matrices, where $\mathbf{a}_i$ and $\mathbf{b}_i$ are the i-th column of $\mathbf{A}$ and $\mathbf{B}$ for $i = 1, 2, ..., n$.

$$c_1\mathbf{a}_1 + c_2\mathbf{a}_2 + ... + c_n\mathbf{a}_n = 0$$

if and only if

$$c_1\mathbf{b}_1 + c_2\mathbf{b}_2 + ... + c_n\mathbf{b}_n = 0$$

Simply speaking, row operations preserve linear relationship between columns.

Proof:

We will divide this proof into two parts:

1. If $c_1\mathbf{b}_1 + c_2\mathbf{b}_2 + c_3\mathbf{b}_3 + ... + c_n\mathbf{b}_n = 0$, $c_1\mathbf{a}_1 + c_2\mathbf{a}_2 + c_3\mathbf{a}_3 + ... + c_n\mathbf{a}_n = 0$

   Since $\mathbf{A}$ and $\mathbf{B}$ are row equivalent, the equation $\mathbf{A} = \mathbf{PB}$ holds for an invertible matrix $\mathbf{P}$ of size $m \times m$. This comes from the idea that performing elementary row operation is equivalent to multiplying an elementary matrix to the matrix (Chapter 5.6).

   As $\mathbf{A} = \mathbf{PB}$, by block multiplication, we can say that $\mathbf{a}_i = \mathbf{Pb}_i$ for $i = 1, 2, ..., n$.

   $$\begin{aligned} \mathbf{P} \cdot 0 &= 0 \\ \mathbf{P} \cdot (c_1\mathbf{b}_1 + c_2\mathbf{b}_2 + ... + c_n\mathbf{b}_n) &= 0 \\ c_1\mathbf{Pb}_1 + c_2\mathbf{Pb}_2 + ... + c_n\mathbf{Pb}_n &= 0 \\ c_1\mathbf{a}_1 + c_2\mathbf{a}_2 + ... + c_n\mathbf{a}_n &= 0 \qquad \text{By } \mathbf{Pb}_i = \mathbf{a}_i \end{aligned}$$

2. If $c_1\mathbf{a}_1 + c_2\mathbf{a}_2 + c_3\mathbf{a}_3 + ... + c_n\mathbf{a}_n = 0$, $c_1\mathbf{b}_1 + c_2\mathbf{b}_2 + c_3\mathbf{b}_3 + ... + c_n\mathbf{b}_n = 0$

   By the same argument as number 1, since $\mathbf{A}$ and $\mathbf{B}$ are row equivalent, the equation $\mathbf{B} = \mathbf{QA}$ holds for an invertible matrix $\mathbf{Q}$ of size $m \times m$.

   As $\mathbf{B} = \mathbf{QA}$, by block multiplication, we can say that $\mathbf{b}_i = \mathbf{Qa}_i$ for $i = 1, 2, ..., n$.

   $$\begin{aligned} \mathbf{Q} \cdot 0 &= 0 \\ \mathbf{Q} \cdot (c_1\mathbf{a}_1 + c_2\mathbf{a}_2 + ... + c_n\mathbf{a}_n) &= 0 \\ c_1\mathbf{Qa}_1 + c_2\mathbf{Qa}_2 + ... + c_n\mathbf{Qa}_n &= 0 \\ c_1\mathbf{b}_1 + c_2 \cdot \mathbf{b}_2 + ... + c_n\mathbf{b}_n &= 0 \qquad \text{By } \mathbf{Qa}_i = \mathbf{b}_i \end{aligned}$$

13.5 Another Theorem about Column Space

Suppose $\mathbf{R}$ is the reduced row-echelon form of $\mathbf{A}$. Then, the columns of $\mathbf{A}$ that corresponds to the pivot columns of $\mathbf{R}$ form a basis for the column space of $\mathbf{A}$.

13.6 A Warning

1. Row operations do not preserve column space. Consider the row operation $R_i \leftrightarrow R_j$! When performing this, all the values in row 1 are swapped with those in row 2 – the columns no longer preserved.

2. Row operations do not preserve linear relationship between the rows. Recall the step by step solutions of Gaussian Elimination:

   Step 1: Locate the leftmost column that does not consist entirely of zeros

   Step 2: Interchange the top row with another row, if necessary, to bring a nonzero entry to the top of the column found in step 1

   Step 3: For each row below the top row, add a suitable multiple of the top row to it so that the entry below the leading entry of the top row becomes 0

   Step 4: Now cover the top row in the augmented matrix and begin again with Step 1 applied to the submatrix that remains

   Observe that after performing step 3, the linear relationship between columns are no longer preserved – we can no longer see the relationship between the columns as the value already becomes 0.

13.7 Another Perspective

Let $\mathbf{v} \in \mathbb{R}^m$ be a vector and $\mathbf{A}$ be a $m \times n$ matrix. By block multiplication, express $\mathbf{A}$ as $\begin{pmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \dots & \mathbf{a}_n \end{pmatrix}$. Observe that to check whether a vector $\mathbf{v}$ is in the column space of $\mathbf{A}$ or not, we need to check whether there exists $\mathbf{x} = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix}$ where $c_1, c_2, \dots, c_n \in \mathbb{R}$ such that

$$c_1\mathbf{a}_1 + c_2\mathbf{a}_2 + \dots + c_n\mathbf{a}_n = \mathbf{v}$$
$$\begin{pmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \dots & \mathbf{a}_n \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} = \mathbf{v}$$
$$\mathbf{A}\mathbf{x} = \mathbf{v}$$

Notice that this is equivalent to checking whether $\mathbf{A}\mathbf{x} = \mathbf{v}$ is consistent or not. Then, for every $\mathbf{v}$ where $\mathbf{A}\mathbf{x} = \mathbf{v}$ is consistent, $\mathbf{v}$ is in the column space of $\mathbf{A}$. By this observation, the column space can be characterized by:

$$\text{Col}(\mathbf{A}) = \{\mathbf{v} \mid \mathbf{A}\mathbf{x} = \mathbf{v} \text{ is consistent}\}$$

or

$$\text{Col}(\mathbf{A}) = \{\mathbf{v} = \mathbf{A}\mathbf{u} \mid \mathbf{u} \in \mathbb{R}^n\}$$

13.8 Definition of Nullspace

The **nullspace** of a $m \times n$ matrix $\mathbf{A}$ is the solution space to the homogeneous system $\mathbf{Ax} = \mathbf{0}$ with coefficient matrix $\mathbf{A}$. It is denoted as

$$\text{Null}(\mathbf{A}) = \{\mathbf{v} \in \mathbb{R}^n \mid \mathbf{Av} = \mathbf{0}\}$$

13.9 Definition of Nullity

The **nullity** of $\mathbf{A}$ is the dimension of the nullspace of $\mathbf{A}$, denoted as

$$\text{nullity}(\mathbf{A}) = \dim(\text{Null}(\mathbf{A}))$$

Note: Recall that the dimension of a subspace V is the number of vectors in the basis of V. So, the dimension of the nullspace of $\mathbf{A}$ corresponds to the number of vectors in the basis of the nullspace of $\mathbf{A}$. By the definition of nullspace, finding the number of vectors in the basis of the nullspace of $\mathbf{A}$ is equivalent to finding the number of vectors in the basis of $\{\mathbf{v} \in \mathbb{R}^n \mid \mathbf{Av} = \mathbf{0}\}$. This is the same as finding the number of nonpivot columns of the augmented matrix $\mathbf{Av} = \mathbf{0}$.

13.10 Rank

13.10.1 A Theorem

Let $\mathbf{A}$ be a $m \times n$ matrix! The equation $\dim(\text{Col}(\mathbf{A})) = \dim(\text{Row}(\mathbf{A}))$ will always hold true.

Proof:

$$\begin{aligned}
\dim(\text{Col}(\mathbf{A})) &= \#\text{ of pivot columns in RREF of }\mathbf{A} && \text{by Chapter 13.5} \\
&= \#\text{ of leading entries in RREF of }\mathbf{A} && \text{by the definition of pivot columns} \\
&= \#\text{ of nonzero rows in RREF of }\mathbf{A} && \\
&= \dim(\text{Row}(\mathbf{A})) && \text{by Chapter 13.3}
\end{aligned}$$

13.10.2 The Definition

By Theorem 13.10.1, we know that the dimension of the column space of $\mathbf{A}$ is always equal to the dimension of the row space of $\mathbf{A}$. We define this value as the **rank** of $\mathbf{A}$.

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

13.10.3 A Corollary

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

Proof:

Definition-wise, transpose means to swap rows and columns.

$$\begin{aligned}
\text{rank}(\mathbf{A}) &= \dim(\text{Col}(\mathbf{A})) && \text{By the definition of rank} \\
&= \dim(\text{Row}(\mathbf{A}^T)) && \text{By the definition of transpose} \\
&= \text{rank}(\mathbf{A}^T) && \text{By the definition of rank}
\end{aligned}$$

13.10.4 A Lemma

Let $\mathbf{A}$ be a $m \times n$ matrix and $\mathbf{B}$ be a $n \times p$ matrix. The column space of the product $\mathbf{AB}$ is a subspace of the column space of $\mathbf{A}$. In other words, $\text{Col}(\mathbf{AB}) \subseteq \text{Col}(\mathbf{A})$.

Proof:

Express $\mathbf{B}$ as $\begin{pmatrix} \mathbf{b_1} & \mathbf{b}_2 & ... & \mathbf{b}_p \end{pmatrix}$ where $\mathbf{b}_i \in \mathbb{R}^n$ for $i = 1, 2, ..., p$! Using this expression, let's take a look what $\mathbf{AB}$ will look like!

$$\mathbf{A} \begin{pmatrix} \mathbf{b_1} & \mathbf{b}_2 & ... & \mathbf{b}_p \end{pmatrix} = \begin{pmatrix} \mathbf{Ab_1} & \mathbf{Ab}_2 & ... & \mathbf{Ab}_p \end{pmatrix}$$

Recall from the definition of column space – Chapter 13.7 – we know that $\text{Col}(\mathbf{A})$ $= \{\mathbf{v} = \mathbf{Au} \mid \mathbf{u} \in \mathbb{R}^n\}$. In other words, $\text{Col}(\mathbf{A})$ is the set of all $\mathbf{Au}$ for $\mathbf{u} \in \mathbb{R}^n$.

As $\mathbf{b}_i \in \mathbb{R}^n$ for $i = 1, 2, ..., p$, by the definition of column space above, $\mathbf{Ab}_i \in \text{Col}(\mathbf{A})$ for all $i = 1, 2, ..., p$.

Then, by the definition of column space, we know that

$$\text{Col}(\mathbf{AB}) = \text{span}\{\mathbf{Ab}_1, \mathbf{Ab}_2, ..., \mathbf{Ab}_p\}$$

Then, according to definition of subset, as we know that $\mathbf{Ab}_i \in \text{Col}(\mathbf{A})$ for all $i = 1, 2, ..., p$,

$$\text{span}\{\mathbf{Ab}_1, \mathbf{Ab}_2, ..., \mathbf{Ab}_p\} \subseteq \text{Col}(\mathbf{A})$$

By this, we can conclude that $\text{Col}(\mathbf{AB}) \subseteq \text{Col}(\mathbf{A})$.

13.10.5 A Theorem

Let $\mathbf{A}$ be a $m \times n$ matrix and $\mathbf{B}$ be a $n \times p$ matrix. Then,

$$\text{rank}(\mathbf{AB}) \leq \min\{\text{rank}(\mathbf{A}), \text{rank}(\mathbf{B})\}$$

Proof:

First thing, from Lemma 13.10.4, we know that $\text{Col}(\mathbf{AB}) \subseteq \text{Col}(\mathbf{A})$. Then, from the property of subsets and dimensions – Chapter 9.5.7 No. 1 – we know that if $\text{Col}(\mathbf{AB}) \subseteq \text{Col}(\mathbf{A})$, $\dim(\text{Col}(\mathbf{AB})) \leq \dim(\text{Col}(\mathbf{A}))$.

Then, by the definition of rank – Chapter 13.10.2 – we can deduce that

$$\text{rank}(\mathbf{AB}) = \dim(\text{Col}(\mathbf{AB})) \leq \dim(\text{Col}(\mathbf{A})) = \text{rank}(\mathbf{A})$$

In short, $\text{rank}(\mathbf{AB}) \leq \text{rank}(\mathbf{A})$.

Then, we will use what we have known in Corollary 13.10.3 and apply it to $\text{rank}(\mathbf{AB})$. We will have $\text{rank}(\mathbf{AB}) = \text{rank}((\mathbf{AB})^T) = \text{rank}(\mathbf{B}^T\mathbf{A}^T)$ – Claim 1

Next, observe that we can still apply Lemma 13.10.4, just in with a different matrix size. Therefore, we can say that

$$\text{Col}(\mathbf{B}^T\mathbf{A}^T) \subseteq \text{Col}(\mathbf{B}^T)$$

By the property of subsets and dimensions, we can say that

$$\dim(\text{Col}(\mathbf{B}^T\mathbf{A}^T)) \leq \dim(\text{Col}(\mathbf{B}^T)) \text{ - Claim 2}$$

Let's arrange what we have known so far!

$$\begin{aligned}
\dim(\text{Col}(\mathbf{B}^T\mathbf{A}^T)) &\leq \dim(\text{Col}(\mathbf{B}^T)) && \text{By Claim 2} \\
\text{rank}(\mathbf{B}^T\mathbf{A}^T) = \dim(\text{Col}(\mathbf{B}^T\mathbf{A}^T)) &\leq \dim(\text{Col}(\mathbf{B}^T)) = \text{rank}(\mathbf{B}^T) && \text{By the definition of rank} \\
\text{rank}(\mathbf{B}^T\mathbf{A}^T) &\leq \text{rank}(\mathbf{B}^T) && \text{Write shorter} \\
\text{rank}(\mathbf{AB}) = \text{rank}(\mathbf{B}^T\mathbf{A}^T) &\leq \text{rank}(\mathbf{B}^T) && \text{By Claim 1} \\
\text{rank}(\mathbf{AB}) = \text{rank}(\mathbf{B}^T\mathbf{A}^T) &\leq \text{rank}(\mathbf{B}^T) = \text{rank}(\mathbf{B}) && \text{By Corollary 13.10.3}
\end{aligned}$$

In short, $\text{rank}(\mathbf{AB}) \leq \text{rank}(\mathbf{B})$.

In summary, as we know that $\text{rank}(\mathbf{AB}) \leq \text{rank}(\mathbf{A})$ and $\text{rank}(\mathbf{AB}) \leq \text{rank}(\mathbf{B})$, we can conclude that $\text{rank}(\mathbf{AB}) \leq \min(\text{rank}(\mathbf{A}), \text{rank}(\mathbf{B}))$.

13.10.6 Rank-Nullity Theorem

Let $\mathbf{A}$ be a $m \times n$ matrix. The sum of its rank and nullity is equal to the number of columns.

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

Proof:

By the definition of rank at Chapter 13.10.1, we know that $\text{rank}(\mathbf{A})$ is equal to the number of pivot columns in the RREF of $\mathbf{A}$.

By the definition of nullity at Chapter 13.9, we know that $\text{nullity}(\mathbf{A})$ is equal to the number of nonpivot columns in the RREF of $\mathbf{A}$.

Then, because a column can be either a pivot column or nonpivot column (cannot be both), the number of pivot columns plus the number of nonpivot columns is equal to the total number of columns, which is n.

13.10.7 A Summary

Summary

Let $\mathbf{A}$ be a $m \times n$ matrix.

Subspace	Subspace of	Basis	Dimension
$Col(\mathbf{A})$	$\mathbb{R}^m$	Columns of $\mathbf{A}$ corresponding to pivot columns in RREF.	$rank(\mathbf{A})$=number of pivot columns in RREF.
$Row(\mathbf{A})$	$\mathbb{R}^n$	Nonzero rows of RREF.	$rank(\mathbf{A})$=number of nonzero rows of RREF
$Null(\mathbf{A})$	$\mathbb{R}^n$	Vectors in general solution to $\mathbf{Ax} = \mathbf{0}$.	$nullity(\mathbf{A})$=number of non-pivot columns in RREF.

Figure 6: Column, Row, and Null Space

13.11 Full Rank Matrices

13.11.1 An Observation

Let $\mathbf{A}$ be a $m \times n$ matrix.

1. By the definition of rank in Chapter 13.10.2, it is known that rank($\mathbf{A}$) is equal to dim(Col($\mathbf{A}$)). Also, by the property of column space in Chapter 13.5, we know that the pivot columns in RREF of $\mathbf{A}$ forms a basis for $\mathbf{A}$. By the definition of dimension, we can therefore conclude that the rank($\mathbf{A}$) is equal to the number of pivot columns in RREF of $\mathbf{A}$.

2. By the definition of rank in Chapter 13.10.2, it known that rank($\mathbf{A}$) is equal to dim(Row($\mathbf{A}$)). Then, by the property of row space in Chapter 13.3, we know that the nonzero rows of the RREF of $\mathbf{A}$ forms a basis for $\mathbf{A}$. By the definition of dimension, we can therefore conclude that rank($\mathbf{A}$) is equal to the number of nonzero rows in RREF of $\mathbf{A}$.

It is intuituive to observe that the number of pivot columns in the RREF of $\mathbf{A}$ is $\leq$ the number of columns and the number of nonzero rows in the RREF of $\mathbf{A}$ is $\leq$ the number of rows.

Therefore, by this observation, it can be concluded that the rank of $\mathbf{A}$ is no greater than the number of rows or columns.

$$\text{rank}(\mathbf{A}) \leq \min\{m, n\}$$

13.11.2 The Definition

A $m \times n$ matrix $\mathbf{A}$ is said to be of **full rank** if its rank is equal to either the number of rows or columns.

$$\text{rank}(\mathbf{A}) = \min\{m, n\}$$

13.11.3 Equivalent Statement of Invertibility

Consider the case when $\mathbf{A}$ is a full rank matrix where the number of rows is equal to columns, call it n. Thus, this means that $\mathbf{A}$ is a full rank matrix if and only if $\text{rank}(\mathbf{A}) = n$.

Now, let's observe how invertibility is related to the fact that $\mathbf{A}$ is a full rank matrix.

$$
\begin{aligned}
\mathbf{A} \text{ is invertible} &\Leftrightarrow \text{The RREF of } \mathbf{A} \text{ is the identity matrix} \\
&\quad \text{by the equivalent statements of invertibility} \\
&\Leftrightarrow \dim(\text{Col}(\mathbf{A})) = \dim(\text{Row}(\mathbf{A})) = n \\
&\quad \text{by the definition of dimension} \\
&\Leftrightarrow \text{rank}(\mathbf{A}) = n \\
&\quad \text{by the definition of rank} \\
&\Leftrightarrow \mathbf{A} \text{ is a full rank matrix} \\
&\quad \text{by the property of full rank}
\end{aligned}
$$

By this, we can observe that having $\mathbf{A}$ invertible is equivalent to having $\mathbf{A}$ being a full rank matrix.

Then, by the Rank-Nullity Theorem, we know that

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

Because $\mathbf{A}$ is a full rank matrix, $\text{rank}(\mathbf{A}) = n$. By substituting this in, we will have $\text{nullity}(\mathbf{A}) = 0$. With this, we have that having $\mathbf{A}$ invertible is also equivalent to having the nullity of $\mathbf{A}$ equals to 0.

13.11.4 The Updated Version of the Equivalent Statements for Invertibility

Let $\mathbf{A}$ be a square matrix of order n. The following statements are equivalent.

1. $\mathbf{A}$ is invertible
2. $\mathbf{A}^T$ is invertible
3. $\mathbf{A}$ has a left-inverse, that is, there is a matrix $\mathbf{B}$ such that $\mathbf{BA} = \mathbf{I}$.
4. $\mathbf{A}$ has a right-inverse, that is, there is a matrix $\mathbf{B}$ such that $\mathbf{AB} = \mathbf{I}$.
5. The reduced row-echelon form of $\mathbf{A}$ is the identity matrix.
6. $\mathbf{A}$ can be expressed as a product of elementary matrices.
7. The homogeneous system $\mathbf{Ax} = \mathbf{0}$ has only trivial solution.
8. For any $\mathbf{b}$, the system $\mathbf{Ax} = \mathbf{b}$ is consistent.
9. The determinant of $\mathbf{A}$ is nonzero, $\det(A) \neq 0$.
10. The columns/rows of $\mathbf{A}$ are linearly independent for $\mathbb{R}^n$

11. The columns/rows of $\mathbf{A}$ spans $\mathbb{R}^n$
12. $\mathbf{A}$ is of full rank, $\text{rank}(\mathbf{A}) = n$
13. $\text{nullity}(\mathbf{A}) = 0$

13.11.5 Full Rank Equals Number of Columns

Suppose $\mathbf{A}$ is a $m \times n$ matrix. The following statements are equivalent.

1. $\mathbf{A}$ is full rank, where rank is equal to the number of columns, rank($\mathbf{A}$)= n.
2. The rows of $\mathbf{A}$ spans $\mathbb{R}^n$, $\text{Row}(\mathbf{A}) = \mathbb{R}^n$
3. The columns of $\mathbf{A}$ are linearly independent
4. The homogeneous system $\mathbf{Ax} = \mathbf{0}$ has only the trivial solution, that is Null($\mathbf{A}$)= $\{0\}$
5. $\mathbf{A}^T\mathbf{A}$ is an invertible matrix of order n
6. $\mathbf{A}$ has a left inverse

Proof:

Assume that statement 1 is true! By the definition of rank, we could deduce that $\text{rank}(\mathbf{A}) = n = \dim(\text{Col}(\mathbf{A}))$, implying that all columns of $\mathbf{A}$ are pivot columns. Then, because all columns of $\mathbf{A}$ are pivot columns, the reduced row-echelon form of $\mathbf{A}$ can be expressed as:

$$\begin{pmatrix} \mathbf{I}_n \\ \mathbf{0}_{(m-n)\times n} \end{pmatrix}$$

By having this RREF, it is clear to observe that the rows of $\mathbf{A}$ indeed span $\mathbb{R}^n$, the second statement, because all of the remaining rows which are not included in the identity matrix are just zero rows.

Recall that all columns of $\mathbf{A}$ are pivot columns, by the definition of rank. Therefore, they must be linearly independent, the third statement.

By the Rank-Nullity Theorem, we know that $\text{rank}(\mathbf{A}) + \text{nullity}(\mathbf{A}) = n$. Then, by our assumption, we know that rank($\mathbf{A}$) is n. Thus, substituting this into the equation, we will have the nullity of $\mathbf{A}$ equals 0, the fourth statement.

We will prove the fourth statement by contradiction. Let $\mathbf{A}^T\mathbf{A}$ be a non-invertible matrix. By the equivalent statement of invertibility, this means that there exists a nontrivial solution x such that $\mathbf{A}^T\mathbf{A}x = 0$. Next, multiply both sides by x^T! This means that there exists a nontrivial solution x such that $x^T\mathbf{A}^T\mathbf{A}x = 0$.

By the property of transpose, $x^T\mathbf{A}^T\mathbf{A}x = 0$ is equivalent to $(\mathbf{A}x)^T\mathbf{A}x = 0$. Then, by the property of inner product, this is equivalent to $||\mathbf{A}x||^2 = 0 \rightarrow \mathbf{A}x = 0$. However, by the fourth statement, we know that the only solution for x is the trivial solution. Therefore, this is a contradiction! $\mathbf{A}^T\mathbf{A}$ must be an invertible matrix; the fifth statement is proven.

By the fifth statement, we already know that $\mathbf{A}^T\mathbf{A}$ is an invertible matrix of order n. By the definition of invertibility, $(\mathbf{A}^T\mathbf{A})^{-1}(\mathbf{A}^T\mathbf{A}) = \mathbf{I}$. Observe that this is equivalent to $((\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T)\mathbf{A} = \mathbf{I}$, implying that $(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T$ is the left inverse of $\mathbf{A}$, i.e., $\mathbf{A}$ has a left inverse.

13.11.6 Full Rank Equals Number of Rows

Suppose $\mathbf{A}$ is a $m \times n$ matrix. The following statements are equivalent.

1. $\mathbf{A}$ is full rank, where rank is equal to the number of rows, $\text{rank}(\mathbf{A}) = m$
2. The columns of $\mathbf{A}$ spans $\mathbb{R}^m$, $\text{Col}(\mathbf{A}) = \mathbb{R}^m$
3. The rows of $\mathbf{A}$ are linearly independent
4. The linear system $\mathbf{Ax} = \mathbf{b}$ is consistent for every $\mathbf{b} \in \mathbb{R}^m$
5. $\mathbf{AA}^T$ is an invertible matrix of order m
6. $\mathbf{A}$ has a right inverse

14 Week 8

14.1 Definition of Orthogonality

Two vectors $\mathbf{u}, \mathbf{v}$ in $\mathbb{R}^n$ are *orthogonal* if $\mathbf{u} \cdot \mathbf{v} = 0$. Let's observe the two cases when will this equation holds:

1. Either $\mathbf{u} = \mathbf{0}$ or $\mathbf{v} = \mathbf{0}$
2. Recall the formula to calculate the angle between two vectors!
$$\cos(\theta) = \frac{\mathbf{u} \cdot \mathbf{v}}{||\mathbf{u}||\ ||\mathbf{v}||}$$
Then, because $\mathbf{u} \cdot \mathbf{v} = 0$, we can deduce that $\cos(\theta) = 0$, which tells us that $\theta = \frac{\pi}{2}$.

By this, we can say that $\mathbf{u}, \mathbf{v}$ are orthogonal if and only if one of them is the zero vector or they are perpendicular to each other.

14.2 Definition of Orthogonal Sets

A set $S = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, ..., \mathbf{v}_k\}$ of vectors is **orthogonal** if $\mathbf{v}_i \cdot \mathbf{v}_j = 0$ for every $i \neq j$. In other words, the set S is orthogonal if vectors in S are pairwise orthogonal.

14.3 Definition of Orthonormal

A set of $S = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, ..., \mathbf{v}_k\}$ of vectors is **orthonormal** if for all $i, j = 1, 2, ..., k$,

$$\mathbf{v}_i \cdot \mathbf{v}_j = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j \end{cases}$$

In other words, the set S is orthonormal if S is orthogonal and all the vectors in S are unit vectors.

Note: Orthogonal sets may contain zero vectors, but orthonormal sets cannot contain zero vectors.

14.4 Orthogonal to Subspace

14.4.1 The Definition

Let V be a subspace of $\mathbb{R}^n$. A vector $\mathbf{n} \in \mathbb{R}^n$ is orthogonal to V if for every $\mathbf{v}$ in V, $\mathbf{n} \cdot \mathbf{v} = 0$ – $\mathbf{n}$ is orthogonal to every vector in V. We denote this type of relationship as $\mathbf{n} \perp V$.

14.4.2 A Theorem

Let V be a subspace of $\mathbb{R}^n$ and $S = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k\}$ be a spanning set for V. Simply speaking, $\text{span}(S) = V$. Then, a vector $\mathbf{w}$ is orthogonal to V if and only if $\mathbf{w} \cdot \mathbf{u}_i = 0$ for all $i = 1, 2, ..., k$.

14.4.3 Another Theorem

Let V be a subspace of $\mathbb{R}^n$ and $S = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k\}$ be a spanning set for V. Then, $\mathbf{w}$ is orthogonal to V if and only if $\mathbf{w}$ is in the nullspace of $\mathbf{A}^T$, where $\mathbf{A} = \begin{pmatrix} \mathbf{u}_1 & \mathbf{u}_2 & ... & \mathbf{u}_k \end{pmatrix}$.

Proof:

$$\begin{aligned}
\mathbf{w} \text{ is orthogonal to } V &\Leftrightarrow \mathbf{w} \cdot \mathbf{u}_i = 0 \text{ for } i = 1, 2, ..., k && \text{By Theorem 14.4.2} \\
&\Leftrightarrow \mathbf{u}_i \cdot \mathbf{w} = 0 \text{ for } i = 1, 2, ..., k && \text{By commutativity of dot product} \\
&\Leftrightarrow \mathbf{u}_i^T \mathbf{w} = 0 \text{ for } i = 1, 2, ..., k && \text{By the definition of dot product} \\
&\Leftrightarrow \begin{pmatrix} -\mathbf{u}_1^T- \\ -\mathbf{u}_2^T- \\ \vdots \\ -\mathbf{u}_k^T- \end{pmatrix} \mathbf{w} = 0 && \text{By block multiplication} \\
&\Leftrightarrow \mathbf{A}^T \mathbf{w} = 0 && \text{By the definition of } \mathbf{A} \\
&\Leftrightarrow \mathbf{w} \in \{\mathbf{v} \in \mathbb{R}^n \mid \mathbf{A}^T \mathbf{v} = 0\} && \text{By the definition of subspace} \\
&\Leftrightarrow \mathbf{w} \in Null(A) && \text{By the definition of nullspace}
\end{aligned}$$

14.5 Orthogonal Complement

14.5.1 The Definition

For V be a subspace of $\mathbb{R}^n$, observe that the set of vectors that are orthogonal to a subspace V is a subspace. We define this subspace as the **orthogonal complement** of V.

Formally speaking, the **orthogonal complement** of V is the set of all vectors that are orthogonal to V, and is denoted as

$$V^{\perp} = \{\mathbf{w} \in \mathbb{R}^n \mid \mathbf{w} \cdot \mathbf{v} = \mathbf{0} \text{ for all } \mathbf{v} \text{ in } V\}$$

14.5.2 A Theorem

Let V be a subspace of $\mathbb{R}^n$ and $S = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k\}$ be a spanning set for V. Let $\mathbf{A} = \begin{pmatrix}\mathbf{u}_1 & \mathbf{u}_2 & ... & \mathbf{u}_k\end{pmatrix}$. As a result of Theorem 14.4.3 and the definition of orthogonal complement, it is intuitive to deduce that the orthogonal complement of V is the nullspace of $\mathbf{A}^T$. In other words, $V^{\perp} = \text{Null}(\mathbf{A}^T)$.

14.6 Orthogonal and Orthonormal Bases

14.6.1 A Theorem

Suppose $S = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k\}$ is an orthogonal set of nonzero vectors. Then, S is linearly independent.

Proof:

From Chapter 9.1, recall that a set of vectors $\{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k\}$ is linear independent if and only if the only solution for $\begin{pmatrix}\mathbf{u}_1 & \mathbf{u}_2 & ... & \mathbf{u}_k\end{pmatrix}\mathbf{x} = 0$ is the trivial solution. This is equivalent to saying that the only solution for $c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + ... + c_k\mathbf{u}_k = 0$ is $c_i = 0$ for all $1 \leq i \leq k$.

To prove this, multiply $\mathbf{u}_i$ to both sides of $c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + ... + c_k\mathbf{u}_k = 0$ for any $1 \leq i \leq k$.

$$\mathbf{u}_i \cdot (c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + ... + c_k\mathbf{u}_k) = 0$$
$$c_1\mathbf{u}_1 \cdot \mathbf{u}_i + c_2\mathbf{u}_2 \cdot \mathbf{u}_i + ... + c_k\mathbf{u}_k \cdot \mathbf{u}_i = 0$$

By the property of orthogonal set, we know that for any i, j where $i \neq j$, $\mathbf{u}_i \cdot \mathbf{u}_j = 0$. By this, the equation above can be simplified to

$$c_i\mathbf{u}_i \cdot \mathbf{u}_i = 0$$
$$c_i||\mathbf{u}_i|| = 0$$

We know that S is an orthogonal set of nonzero vectors. This means $\mathbf{u}_i$ is always non-zero, so c_i is the one that must be equal to 0.

14.6.2 A Corollary

Every orthonormal set is linearly independent.

Proof: Orthonormal set is a normalized orthogonal set of nonzero vectors.

14.6.3 The Definition

Let V be a subspace of $\mathbb{R}^n$. A set $S \subseteq V$ is an orthogonal/orthonormal basis of V if S is a basis of V and S is an orthogonal/orthonormal set.

14.7 Coordinates Relative to an Orthogonal Basis

Let $S = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k\}$ be an orthogonal basis for a subspace V of $\mathbb{R}^n$.

Let's express $\mathbf{v} \in V$, $\mathbf{v} = c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + ... + c_k\mathbf{u}_k$ as a linear combination of vectors in the basis S.

Recall that as $S = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k\}$ is an orthogonal set, for any i, j where $i \neq j$, $\mathbf{u}_i \cdot \mathbf{u}_j = 0$! We will utilize this fact and find the formula for c_i.

Multiply $\mathbf{u}_i$ to both sides of equation $\mathbf{v} = c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + ... + c_k\mathbf{u}_k$. By doing this, we will have:

$$\begin{aligned}
\mathbf{u}_i \cdot \mathbf{v} &= \mathbf{u}_i \cdot (c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + ... + c_k\mathbf{u}_k) \\
\mathbf{u}_i \cdot \mathbf{v} &= c_1\mathbf{u}_1 \cdot \mathbf{u}_i + c_2\mathbf{u}_2 \cdot \mathbf{u}_i + ... + c_k\mathbf{u}_k \cdot \mathbf{u}_i
\end{aligned}$$

Utilizing the fact that we recall earlier, we will have:

$$\begin{aligned}
\mathbf{u}_i \cdot \mathbf{v} &= c_i\mathbf{u}_i \cdot \mathbf{u}_i \\
c_i &= \frac{\mathbf{u}_i \cdot \mathbf{v}}{\mathbf{u}_i \cdot \mathbf{u}_i} = \frac{\mathbf{u}_i \cdot \mathbf{v}}{||\mathbf{u}_i||^2}
\end{aligned}$$

By the definition of coordinates relative we have defined before in Chapter 9.4.1, we know that

$$[\mathbf{v}]_S = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_k \end{pmatrix}$$

By the formula of c_i we have obtained, we can define $[\mathbf{v}]_S$ to be:

$$[\mathbf{v}]_S = \begin{pmatrix} \frac{\mathbf{u}_1 \cdot \mathbf{v}}{||\mathbf{u}_1||^2} \\ \frac{\mathbf{u}_2 \cdot \mathbf{v}}{||\mathbf{u}_2||^2} \\ \vdots \\ \frac{\mathbf{u}_k \cdot \mathbf{v}}{||\mathbf{u}_k||^2} \end{pmatrix}$$

Special case when S is orthonormal, we know that $||\mathbf{u}_i||$ is always 1 for $1 \leq i \leq k$. Therefore,

$$[\mathbf{v}]_S = \begin{pmatrix} \mathbf{u}_1 \cdot \mathbf{v} \\ \mathbf{u}_2 \cdot \mathbf{v} \\ \vdots \\ \mathbf{u}_k \cdot \mathbf{v} \end{pmatrix}$$

14.8 Orthogonal Matrices

14.8.1 An Observation

Let $S = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k\}$ be set of vectors in $\mathbb{R}^n$. Construct the $n \times k$ matrix $\mathbf{Q} = \begin{pmatrix} \mathbf{u}_1 & \mathbf{u}_2 & ... & \mathbf{u}_k \end{pmatrix}$ whose columns are the vectors in S. If we multiply $\mathbf{Q}^T$ and $\mathbf{Q}$, here is the result!

$$\mathbf{Q}^T\mathbf{Q} = \begin{pmatrix} \mathbf{u}_1^T \\ \mathbf{u}_2^T \\ \vdots \\ \mathbf{u}_k^T \end{pmatrix} \begin{pmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_k \end{pmatrix}$$

By multiplying the matrices together, we will get

$$\mathbf{Q}^T\mathbf{Q} = \begin{pmatrix} \mathbf{u}_1^T\mathbf{u}_1 & \mathbf{u}_1^T\mathbf{u}_2 & \cdots & \mathbf{u}_1^T\mathbf{u}_k \\ \mathbf{u}_2^T\mathbf{u}_1 & \mathbf{u}_2^T\mathbf{u}_2 & \cdots & \mathbf{u}_2^T\mathbf{u}_k \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{u}_k^T\mathbf{u}_1 & \mathbf{u}_k^T\mathbf{u}_2 & \cdots & \mathbf{u}_k^T\mathbf{u}_k \end{pmatrix}$$

Notice that this form is just the inner product!

$$\mathbf{Q}^T\mathbf{Q} = \begin{pmatrix} \mathbf{u}_1 \cdot \mathbf{u}_1 & \mathbf{u}_1 \cdot \mathbf{u}_2 & \cdots & \mathbf{u}_1 \cdot \mathbf{u}_k \\ \mathbf{u}_2 \cdot \mathbf{u}_1 & \mathbf{u}_2 \cdot \mathbf{u}_2 & \cdots & \mathbf{u}_2 \cdot \mathbf{u}_k \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{u}_k \cdot \mathbf{u}_1 & \mathbf{u}_k \cdot \mathbf{u}_2 & \cdots & \mathbf{u}_k \cdot \mathbf{u}_k \end{pmatrix}$$

Recall that as $S = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k\}$ is an orthogonal set, for any i, j where $i \neq j$, $\mathbf{u}_i \cdot \mathbf{u}_j = 0$! Applying this information to the matrix, the matrix will be:

$$\mathbf{Q}^T\mathbf{Q} = \begin{pmatrix} \mathbf{u}_1 \cdot \mathbf{u}_1 & 0 & \cdots & 0 \\ 0 & \mathbf{u}_2 \cdot \mathbf{u}_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{u}_k \cdot \mathbf{u}_k \end{pmatrix}$$

Notice that $\mathbf{u}_i \cdot \mathbf{u}_i = ||\mathbf{u}_i||^2$ plays an important role on deciding what $\mathbf{Q}^T\mathbf{Q}$ will be. By the definition of orthogonal and orthonormal, S is orthogonal if and only if $\mathbf{Q}^T\mathbf{Q}$ is a diagonal matrix, and S is orthonormal if and only if $\mathbf{Q}^T\mathbf{Q}$ is the identity matrix.

When $n \neq k$, even though $\mathbf{Q}^T\mathbf{Q} = \mathbf{I}_k$, we cannot say that $\mathbf{Q}^T = \mathbf{Q}^{-1}$. This is because $\mathbf{Q}^T$ is only the left inverse of $\mathbf{Q}$.

When $n = k$, we can say $\mathbf{Q}^T = \mathbf{Q}^{-1}$. This is because $\mathbf{Q}$ is a square matrix. As $\mathbf{Q}$ is a square matrix and $\mathbf{Q}^T$ is the left inverse of $\mathbf{Q}$, by the equivalent statements for invertibility, $\mathbf{Q}$ is indeed invertible, implying that $\mathbf{Q}^{-1}$ always exists!

14.8.2 A Corollary

Let $S = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_n\}$ be a set of vectors in $\mathbb{R}^n$. Let $\mathbf{Q}$ be a $n \times n$ matrix where $\mathbf{Q} = \begin{pmatrix}\mathbf{u}_1 & \mathbf{u}_2 & ... & \mathbf{u}_n\end{pmatrix}$. Then, $\mathbf{Q}^T = \mathbf{Q}^{-1}$ if and only if S is orthonormal.

14.8.3 The Definition

A $n \times n$ square matrix $\mathbf{A}$ is *orthogonal* if $\mathbf{A}^T = \mathbf{A}^{-1}$, equivalently $\mathbf{A}^T\mathbf{A} = \mathbf{I} = \mathbf{A}\mathbf{A}^T$.

14.8.4 A Theorem

Let $\mathbf{A}$ be a square matrix of order n. The following statements are equivalent.

1. $\mathbf{A}$ is an orthogonal matrix
2. The columns of $\mathbf{A}$ form an orthonormal basis for $\mathbb{R}^n$
3. The rows of $\mathbf{A}$ form an orthonormal basis for $\mathbb{R}^n$

Proof:

By block multiplication, let $\mathbf{A} = \begin{pmatrix}\mathbf{a}_1 & \mathbf{a}_2 & ... & \mathbf{a}_n\end{pmatrix}$. By the definition of an orthogonal matrix, $\mathbf{A}$ is orthogonal if and only if $\mathbf{A}^T = \mathbf{A}^{-1}$. By Corollary 14.8.2, $\mathbf{A}^T = \mathbf{A}^{-1}$ if and only if $S = \{\mathbf{a}_1, \mathbf{a}_2, ..., \mathbf{a}_n\}$ is orthonormal. Notice that S is just the columns of $\mathbf{A}$, and S forms a basis for $\mathbf{A}$ since $\mathbf{A}$ is invertible.

By block multiplication, let $\mathbf{A}^T = \begin{pmatrix}\mathbf{a}_1 & \mathbf{a}_2 & ... & \mathbf{a}_n\end{pmatrix}$. Assume $S = \{\mathbf{a}_1, \mathbf{a}_2, ..., \mathbf{a}_n\}$! Notice that S is just the rows of $\mathbf{A}$.

$$\begin{aligned}
\mathbf{A}^T \text{ is orthogonal} &\Leftrightarrow (\mathbf{A}^T)^T = (\mathbf{A}^T)^{-1} && \text{by definition of orthogonal matrix} \\
&\Leftrightarrow \mathbf{A} = (\mathbf{A}^T)^{-1} && \text{by the property of transpose} \\
&\Leftrightarrow \mathbf{A}^{-1} = ((\mathbf{A}^T)^{-1})^{-1} && \text{inversing both sides} \\
&\Leftrightarrow \mathbf{A}^{-1} = \mathbf{A}^T && \text{by the property of inverse} \\
&\Leftrightarrow S \text{ is orthonormal} && \text{by Corollary 14.8.2} \\
&\Leftrightarrow \text{Rows of } \mathbf{A} \text{ is orthonormal} && \text{by the definition of } S
\end{aligned}$$

Then, because $\mathbf{A}$ is invertible, S is not only orthonormal but also an orthonormal basis for $\mathbb{R}^n$.

15 Week 9

15.1 Orthogonal Projection

15.1.1 The Motivation

Recall Chapter 14.7! Let $S = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k\}$ be an orthogonal basis for a subspace V of $\mathbb{R}^n$, $\mathbf{v} \in V$ can be represented as

$$\mathbf{v} = \frac{\mathbf{v} \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}\mathbf{u}_1 + \frac{\mathbf{v} \cdot \mathbf{u}_2}{\mathbf{u}_2 \cdot \mathbf{u}_2}\mathbf{u}_2 + ... + \frac{\mathbf{v} \cdot \mathbf{u}_k}{\mathbf{u}_k \cdot \mathbf{u}_k}\mathbf{u}_k$$

However, how if $\mathbf{w} \notin V$? What is the value of $\mathbf{w}'$?

$$\mathbf{w}' = \frac{\mathbf{w} \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}\mathbf{u}_1 + \frac{\mathbf{w} \cdot \mathbf{u}_2}{\mathbf{u}_2 \cdot \mathbf{u}_2}\mathbf{u}_2 + ... + \frac{\mathbf{w} \cdot \mathbf{u}_k}{\mathbf{u}_k \cdot \mathbf{u}_k}\mathbf{u}_k$$

15.1.2 Vector Projection

Let say we have a vector $\mathbf{a}$ and a vector $\mathbf{b}$ that has an angle between them equals to θ! We want to project $\mathbf{a}$ to $\mathbf{b}$.

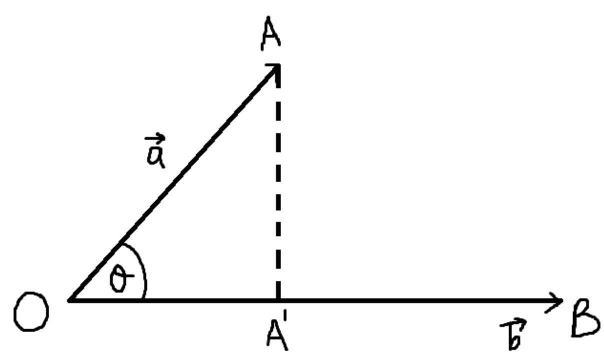

Figure 7: Vector Projection

The norm of $\mathbf{a}$ can be illustrated as OA, which is equal to $||\mathbf{a}||$. By trigonometry, we will have the norm of the projection vector to be $OA' = ||\mathbf{a}|| \cos\theta$. To change it into a vector, we will multiply the norm of this projection vector by the normalized vector $\mathbf{b}$. By this we have the projection vector to be

$$||\mathbf{a}|| \cos\theta \cdot \frac{\mathbf{b}}{||\mathbf{b}||}$$

Recall the formula for calculating the angle between two vectors $\mathbf{a}$ and $\mathbf{b}$ to be $\cos\theta = \frac{\mathbf{a}\cdot\mathbf{b}}{||\mathbf{a}||\cdot||\mathbf{b}||}$. Substituting this in to the projection vector, we will have:

$$||\mathbf{a}|| \cos\theta \cdot \frac{\mathbf{b}}{||\mathbf{b}||} = ||\mathbf{a}|| \cdot \frac{\mathbf{a} \cdot \mathbf{b}}{||\mathbf{a}|| \cdot ||\mathbf{b}||} \cdot \frac{\mathbf{b}}{||\mathbf{b}||} = \frac{\mathbf{a} \cdot \mathbf{b} \cdot \mathbf{b}}{||\mathbf{b}||^2} = \frac{\mathbf{a} \cdot \mathbf{b}}{\mathbf{b} \cdot \mathbf{b}}\mathbf{b}$$

15.1.3 Continuing the Motivation

Observe that the formula for vector projection is exactly the same as the one in our motivation. Yes, $\mathbf{w}'$ is the projection of $\mathbf{w}$ onto span(S)$= V$.

15.1.4 The Orthogonal Projection Theorem

Let V be a subspace of $\mathbb{R}^n$. Every vector $\mathbf{w}$ in $\mathbb{R}^n$ can be decomposed uniquely as a sum $\mathbf{w} = \mathbf{w}_p + \mathbf{w}_n$ where $\mathbf{w}_n$ is orthogonal to V and $\mathbf{w}_p$ is in V. Moreover, if $S = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k\}$ is an orthogonal basis for V, then

$$\mathbf{w}_p = \frac{\mathbf{w} \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}\mathbf{u}_1 + \frac{\mathbf{w} \cdot \mathbf{u}_2}{\mathbf{u}_2 \cdot \mathbf{u}_2}\mathbf{u}_2 + ... + \frac{\mathbf{w} \cdot \mathbf{u}_k}{\mathbf{u}_k \cdot \mathbf{u}_k}\mathbf{u}_k$$

An Illustration:

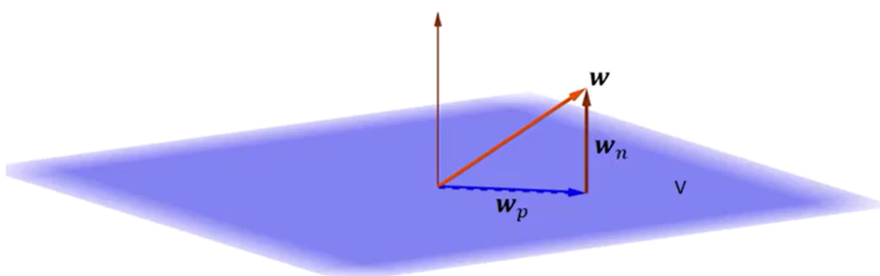

Figure 8: Illustration

15.1.5 A Definition

The vector $\mathbf{w}_p$ mentioned in Theorem 15.1.4 is called the orthogonal projection / projection of $\mathbf{w}$ onto the subspace V.

15.2 Best Approximation Theorem

Let V be a subspace of $\mathbb{R}^n$ and $\mathbf{w}$ a vector in $\mathbb{R}^n$. Let $\mathbf{w}_p$ be the projection of $\mathbf{w}$ onto V. Then, $\mathbf{w}_p$ is vector in V closest to $\mathbf{w}$. Defining it mathematically, for all $\mathbf{v}$ in V,

$$||\mathbf{w} - \mathbf{w}_p|| \leq ||\mathbf{w} - \mathbf{v}||$$

Proof:

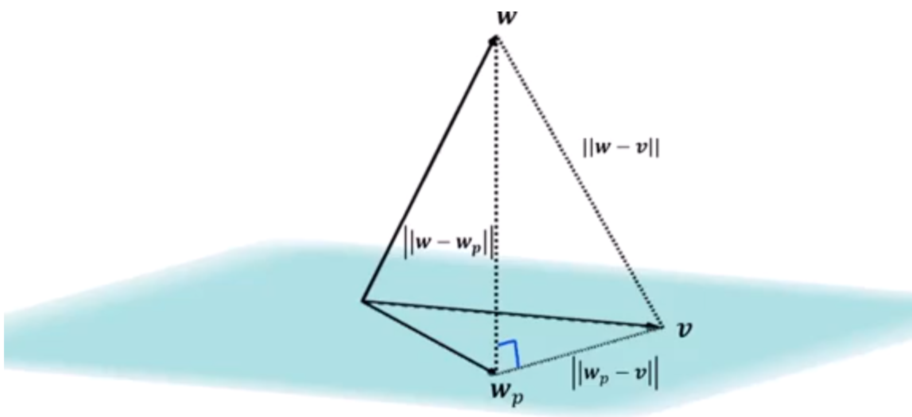

Figure 9: Geometrical Representation

By the Orthogonal Projection Theorem, we know that $\mathbf{w} - \mathbf{w}_p = \mathbf{w}_n$ which is orthogonal to V. Because $\mathbf{w}_p - \mathbf{v}$ is just a vector in V, we know that $\mathbf{w} - \mathbf{w}_p$ is orthogonal to $\mathbf{w}_p - \mathbf{v}$.

Because of this, now we can apply the Pythagorean Theorem!

$$||\mathbf{w} - \mathbf{v}||^2 = ||\mathbf{w} - \mathbf{w}_p||^2 + ||\mathbf{v} - \mathbf{w}_p||^2$$

Since $||\mathbf{v} - \mathbf{w}_p||^2$ is always ≥ 0, we could therefore conclude that $||\mathbf{w} - \mathbf{w}_p|| \leq ||\mathbf{w} - \mathbf{v}||$.

15.3 Gram-Schmidt Orthogonalization

15.3.1 Why we need this?

Recall the Orthogonal Projection Theorem in Chapter 15.1.4! In order for us to find $\mathbf{w}_p$, we need to have S that is an orthogonal basis for V. The problem is, how if we are only given a basis that are not orthogonal? This is where Gram-Schmidt Orthogonalization comes in!

15.3.2 The Process

Let $S = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k\}$ be a linearly independent set. Let

$$
\begin{aligned}
\mathbf{v}_1 &= \mathbf{u}_1 \\
\mathbf{v}_2 &= \mathbf{u}_2 - \left(\frac{\mathbf{v}_1 \cdot \mathbf{u}_2}{||\mathbf{v}_1||^2}\right)\mathbf{v}_1 \\
\mathbf{v}_3 &= \mathbf{u}_3 - \left(\frac{\mathbf{v}_1 \cdot \mathbf{u}_3}{||\mathbf{v}_1||^2}\right)\mathbf{v}_1 - \left(\frac{\mathbf{v}_2 \cdot \mathbf{u}_3}{||\mathbf{v}_2||^2}\right)\mathbf{v}_2 \\
&\vdots \\
\mathbf{v}_k &= \mathbf{u}_k - \left(\frac{\mathbf{v}_1 \cdot \mathbf{u}_k}{||\mathbf{v}_1||^2}\right)\mathbf{v}_1 - \left(\frac{\mathbf{v}_2 \cdot \mathbf{u}_k}{||\mathbf{v}_2||^2}\right)\mathbf{v}_2 - \cdots - \left(\frac{\mathbf{v}_{k-1} \cdot \mathbf{u}_k}{||\mathbf{v}_{k-1}||^2}\right)\mathbf{v}_{k-1}
\end{aligned}
$$

The intuition: For every new vector we construct, we will extract "the part" of the vector which is orthogonal to all previous vectors we have added.

After doing this process, we will have the set $\{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_k\}$ to be an orthogonal set such that $\text{span}(\mathbf{v}_1, ..., \mathbf{v}_k\} = \text{span}\{\mathbf{u}_1, ..., \mathbf{u}_k\}$! Then, by normalizing each of them, the set $\{\frac{\mathbf{v}_1}{||\mathbf{v}_1||^2}, \frac{\mathbf{v}_2}{||\mathbf{v}_2||^2}, \cdots, \frac{\mathbf{v}_k}{||\mathbf{v}_k||^2}\}$ is an orthonormal set.

15.4 Least Square Approximation

15.4.1 A Motivation

In a real-world case, a linear system is usually inconsistent. This is because we are always trying to get as much data as we can, so the number of rows in the coefficient matrix tend to be significantly greater than the number of columns – making the system tends to be inconsistent.

Let $\mathbf{A}$ be the $m \times n$ coefficient matrix and $\mathbf{Ax} = \mathbf{b}$ is the equation that we are trying to solve. In most cases, $m > n$ and the system is inconsistent. The idea of approximation is to find $\mathbf{b}'$ such that $\mathbf{Ax} = \mathbf{b}'$ is consistent and $\mathbf{b}'$ is closest to $\mathbf{b}$, or equivalently finding $\mathbf{u}$ where $\mathbf{Au} = \mathbf{b}' \in \text{Col}(\mathbf{A})$ such that $||\mathbf{Au} - \mathbf{b}||$ is minimized.

15.4.2 A Definition of Least Square Solution

Let $\mathbf{A}$ be a $m \times n$ matrix and $\mathbf{b}$ be a vector in $\in \mathbb{R}^m$. A vector $\mathbf{u}$ in $\mathbb{R}^n$ is a **least square solution** of $\mathbf{Ax} = \mathbf{b}$ if for every vector $\mathbf{v} \in \mathbb{R}^n$,

$$||\mathbf{Au} - \mathbf{b}|| \leq ||\mathbf{Av} - \mathbf{b}||$$

15.4.3 A Theorem

Let $\mathbf{A}$ be a $m \times n$ matrix and $\mathbf{b}$ be a vector in $\mathbb{R}^m$. A vector $\mathbf{u}$ in $\mathbb{R}^n$ is a **least square solution** to $\mathbf{Ax} = \mathbf{b}$ if and only if $\mathbf{Au}$ is the projection of $\mathbf{b}$ onto the column space of Col($\mathbf{A}$).

Proof: Combining the definition of Least Square Solution in Chapter 15.4.2 and the Best Approximation Theorem in Chapter 15.2.

15.4.4 Another Theorem

Let $\mathbf{A}$ be a $m \times n$ matrix and $\mathbf{b}$ be a vector in $\mathbb{R}^m$. A vector $\mathbf{u}$ in $\mathbb{R}^n$ is a least square solution to $\mathbf{Ax} = \mathbf{b}$ if and only if $\mathbf{u}$ is a solution to $\mathbf{A}^T\mathbf{Ax} = \mathbf{A}^T\mathbf{b}$.

Proof:

By Theorem 15.4.3, we know that a vector $\mathbf{u}$ in $\mathbb{R}^n$ is a least square solution to $\mathbf{Ax} = \mathbf{b}$ if and only if $\mathbf{Au}$ is the projection of $\mathbf{b}$ onto the column space of Col($\mathbf{A}$). Then, by the Orthogonal Projection Theorem in Theorem 15.1.4, because $\mathbf{Au}$ is the projection of $\mathbf{b}$, $\mathbf{Au} - \mathbf{b}$ is orthogonal to the column space of $\mathbf{A}$.

Let Col($\mathbf{A}$) be $\{\mathbf{a}_1, \mathbf{a}_2, ..., \mathbf{a}_n\}$. By the property of orthogonality, this means $\mathbf{Au} - \mathbf{b}$ is orthogonal to every $\mathbf{a}_i$ for $1 \leq i \leq n$. By say, for $1 \leq i \leq n$, $\mathbf{a}_i \cdot (\mathbf{Au} - \mathbf{b}) = 0 \to \mathbf{a}_i^T(\mathbf{Au} - \mathbf{b}) = 0$. This is basically equivalent to saying that $\mathbf{A}^T(\mathbf{Au} - \mathbf{b}) = 0 \to \mathbf{A}^T\mathbf{Au} - \mathbf{A}^T\mathbf{b} = 0 \to \mathbf{A}^T\mathbf{Au} = \mathbf{A}^T\mathbf{b}$.

Note:

The value of $\mathbf{u}$ may not be unique, but the value of $\mathbf{Au}$, the projection, is always unique.

15.4.5 Another Way to find Orthogonal Projection

The Motivation:

Our previous method on finding Orthogonal Projection, Chapter 15.1.4, requires us to have an orthogonal basis S. However, in many cases, we are not given the orthogonal basis directly. Therefore, we need to perform the Gram-Schmidt Orthogonalization to convert the spanning set/basis to orthogonal basis. This is tedious!

In fact, there is another way to find orthogonal projection! We will utilize what we have learnt about the least square solution!

Another Method:

Let V be a subspace of $\mathbb{R}^n$ and $S = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k\}$ be a **spanning set (not necessarily a basis)** for V. Let $\mathbf{A} = \begin{pmatrix}\mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_k\end{pmatrix}$, $\mathbf{w}$ be a vector in $\mathbb{R}^n$, and $\mathbf{u}$ be a least square solution to $\mathbf{Ax} = \mathbf{w}$. Then, $\mathbf{w}_p = \mathbf{Au}$ is the orthogonal projection of a vector $\mathbf{w}$ onto V.

Note: We can find $\mathbf{u}$ by using Theorem 15.4.4.

Special Case:

Consider the case where S is a basis! If so, $\mathbf{A}$ will then be linearly independent. By the equivalent statement statement of invertibility, this means that $\mathbf{A}$ is invertible. Because of that, we can generalize the solution for $\mathbf{w}_p$!

First, we will find $\mathbf{u}$, the least square solution to $\mathbf{Ax} = \mathbf{w}$. By Theorem 15.4.4, $\mathbf{u}$ can be obtained by solving this equation $\mathbf{A}^T\mathbf{Ax} = \mathbf{A}^T\mathbf{b}$. Since $\mathbf{A}$ is invertible, then $\mathbf{u} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{b}$. Next, $\mathbf{w}_p = \mathbf{Au} = \mathbf{A}(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{b}$. We can state this as a theorem!

15.4.6 A Theorem

Let V be a subspace of $\mathbb{R}^n$ and $S = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k\}$ be a **basis** for V. Then, the orthogonal projection of a vector $\mathbf{w}$ onto V is

$$\mathbf{w}_p = \mathbf{A}(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{w}$$

where $\mathbf{A} = \begin{pmatrix}\mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_k\end{pmatrix}$.

16 Week 10

16.1 Eigenvalues and Eigenvectors

16.1.1 The Definition

Let $\mathbf{A}$ be a square matrix of order n. A real number λ is an **eigenvalue** of $\mathbf{A}$ if there is a **nonzero** vector $\mathbf{v}$ in $\mathbb{R}^n$, such that

$$\mathbf{Av} = \lambda\mathbf{v}$$

In this case, the nonzero vector $\mathbf{v}$ is called the **eigenvector** associated to λ.

16.1.2 Characteristic Polynomial

Let $\mathbf{A}$ be a square matrix of order n, the characteristic polynomial of $\mathbf{A}$, denoted as $\text{char}(\mathbf{A})$, is the degree n polynomial

$$\det(x\mathbf{I} - \mathbf{A})$$

16.1.3 A Theorem

Let $\mathbf{A}$ be a square matrix of order n. $\lambda \in \mathbb{R}$ is an eigenvalue of $\mathbf{A}$ if and only if the homogeneous system $(\lambda\mathbf{I} - \mathbf{A})\mathbf{x} = 0$ has nontrivial solutions.

Proof:

By the definition of eigenvalue, we know that $\mathbf{Av} = \lambda\mathbf{v}$ for a nonzero vector $\mathbf{v}$ in $\mathbb{R}^n$. Notice that the scalar multiplication $\lambda\mathbf{v}$ is equivalent to the matrix multiplication $\lambda\mathbf{Iv}$. Thus, by this, we can express the equation as $\mathbf{Av} = \lambda\mathbf{Iv}$. Manipulating the equation further, we will have:

$$\mathbf{Av} = \lambda\mathbf{Iv} \to \lambda\mathbf{Iv} - \mathbf{Av} = 0 \to \mathbf{v}(\lambda\mathbf{I} - \mathbf{A}) = 0 \to (\lambda\mathbf{I} - \mathbf{A})\mathbf{v} = 0$$

Because $\mathbf{v}$ is nonzero, this means that $\mathbf{v}$ is a nontrivial solution to the homogeneous system $(\lambda\mathbf{I} - \mathbf{A})\mathbf{v} = 0$.

16.1.4 A Follow-Up Theorem

Let $\mathbf{A}$ be a square matrix of order n. λ is an eigenvalue of $\mathbf{A}$ if and only if λ is a root of the characteristic polynomial $\det(x\mathbf{I} - \mathbf{A})$.

Proof:

By Theorem 16.1.3, we know that λ is an eigenvalue of $\mathbf{A}$ if and only if the homogeneous system $(\lambda\mathbf{I} - \mathbf{A})\mathbf{x} = 0$ has nontrivial solutions. By the equivalent statements of invertibility, we know that $\mathbf{Qx} = \mathbf{0}$ has only trivial solution if and only if $\mathbf{Q}$ is invertible. Equivalently, $\mathbf{Qx} = \mathbf{0}$ has nontrivial solutions if and only if $\mathbf{Q}$ is singular. Therefore, applying this property, we need to find λ such that $\lambda\mathbf{I} - \mathbf{A}$ is singular.

Recall that a singular matrix has a determinant of 0. Then, to find λ such that $\lambda\mathbf{I} - \mathbf{A}$ is singular, we will find the values of λ such that the determinant of $\lambda\mathbf{I} - \mathbf{A}$ is 0. In other words, by the definition of characteristic polynomial, we just need to find the root of the characteristic polynomial $\det(x\mathbf{I} - \mathbf{A})$.

16.1.5 A Theorem (An Equivalent Statement for Invertibility)

A square matrix $\mathbf{A}$ is invertible if and only if $\lambda = 0$ is not an eigenvalue of $\mathbf{A}$.

Proof:

Let's analyze the equivalence statement of 0 being an eigenvalue of $\mathbf{A}$.

$$\begin{aligned}
0 \text{ is the eigenvalue of } \mathbf{A} &\Leftrightarrow \exists \mathbf{v} \neq 0 \text{ such that } \mathbf{Av} = \lambda\mathbf{v} = 0 \\
&\quad \text{by the definition of eigenvalue} \\
&\Leftrightarrow \mathbf{v} \text{ is a non-trivial solution for } \mathbf{Ax} = 0 \\
&\quad \text{by the definition of non-trivial solution} \\
&\Leftrightarrow \mathbf{A} \text{ is singular} \\
&\quad \text{by the definition of singular matrix}
\end{aligned}$$

Next, let's negate these two! So,

$$\begin{aligned} 0 \text{ is not the eigenvalue of } \mathbf{A} &\Leftrightarrow \mathbf{A} \text{ is not singular} \\ &\Leftrightarrow \mathbf{A} \text{ is invertible} \end{aligned}$$

16.2 Algebraic Multiplicity

16.2.1 The Definition

Let λ be an eigenvalue of $\mathbf{A}$. The **algebraic multiplicity** of λ is the largest integer r_λ such that

$$\det(x\mathbf{I} - \mathbf{A}) = (x - \lambda)^{r_\lambda} p(x)$$

for some polynomial $p(x)$.

16.2.2 Another Definition

Let λ be an eigenvalue of $\mathbf{A}$. The **algebraic multiplicity** of λ is the positive integer such that in this equation:

$$\det(x\mathbf{I} - \mathbf{A}) = (x - \lambda)^{r_\lambda} p(x)$$

λ is not a root of $p(x)$.

16.2.3 A Theorem

The eigenvalues of a triangular matrix are the diagonal entries. The algebraic multiplicity of the eigenvalue is the number of times it appears as a diagonal entry of $\mathbf{A}$.

Proof:

Suppose $\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix}$.

By Theorem 16.1.4, we know that the eigenvalue of $\mathbf{A}$ is the root of the characteristic polynomial $\det(x\mathbf{I} - \mathbf{A})$. Let's express $x\mathbf{I} - \mathbf{A}$!

$$\begin{aligned} x\mathbf{I} - \mathbf{A} &= \begin{pmatrix} x & 0 & \cdots & 0 \\ 0 & x & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & x \end{pmatrix} - \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix} \\ &= \begin{pmatrix} x - a_{11} & -a_{12} & \cdots & -a_{1n} \\ 0 & x - a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & x - a_{nn} \end{pmatrix} \end{aligned}$$

Then, by Corollary 7.2.3, we know that the determinant of diagonal matrices are the multiplication of its diagonal entries. Therefore, we will know that

$$\det(x\mathbf{I} - \mathbf{A}) = (x - a_{11})(x - a_{22})...(x - a_{nn})$$

The root of the characteristic polynomial is $a_{11}, a_{22}, a_{33}, ..., a_{nn}$, which is the diagonal entries of $\mathbf{A}$.

16.3 Eigenspace

16.3.1 A Motivation

Recall that the eigenvectors of $\mathbf{A}$ associated to a certain eigenvalue λ are non-trivial solution to $(\lambda\mathbf{I} - \mathbf{A})\mathbf{x} = 0$. Because the system is homogeneous, the set of all solutions to this is a subspace. We'll define this subspace as eigenspace.

16.3.2 The Definition

More formally, the eigenspace associated to an eigenvalue λ of $\mathbf{A}$ is

$$E_\lambda = \{\mathbf{v} \in \mathbb{R}^n \mid \mathbf{A}\mathbf{v} = \lambda\mathbf{v}\} = Null(\lambda\mathbf{I} - \mathbf{A})$$

16.3.3 Geometric Multiplicity

The dimension of its associated eigenspace is the geometric multiplicity of an eigenvalue λ.

$$dim(E_\lambda) = nullity(\lambda\mathbf{I} - \mathbf{A})$$

16.4 Diagonalization

16.4.1 The Definition

A square matrix $\mathbf{A}$ of order n is **diagonalizable** if there exists an invertible matrix $\mathbf{P}$ and a diagonal matrix $\mathbf{D}$ such that

$$\mathbf{P^{-1}AP} = \mathbf{D}$$

Or, equivalently

$$\mathbf{A} = \mathbf{PDP^{-1}}$$

16.4.2 A Theorem

A $n \times n$ matrix $\mathbf{A}$ is diagonalizable if and only if $\mathbf{A}$ has n linearly independent eigenvectors.

Proof:

By the definition of diagonalization in Chapter 16.4.1, $\mathbf{A}$ is diagonalizable if and only if $\mathbf{A} = \mathbf{PDP^{-1}}$ for an invertible matrix $\mathbf{P}$ and a diagonal matrix $\mathbf{D}$.

By block multiplication, let $\mathbf{P} = \begin{pmatrix} \mathbf{u}_1 & \mathbf{u}_2 & ... & \mathbf{u}_n \end{pmatrix}$. Then, let $\mathbf{D} = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$.

According to the equivalent statements of invertibility, as $\mathbf{P}$ is invertible, $\{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_n\}$ are linearly independent for $\mathbb{R}^n$ (Claim 1).

Then, observe these equivalencies!

$\mathbf{A}$ is diagonalizable $\Leftrightarrow \mathbf{P^{-1}AP} = \mathbf{D}$

by the definition of diagonalizable

$\Leftrightarrow \mathbf{AP} = \mathbf{PD}$

by pre-multiplying $\mathbf{P}$ to both sides

$$\Leftrightarrow \mathbf{A} \begin{pmatrix} \mathbf{u}_1 & \mathbf{u}_2 & ... & \mathbf{u}_n \end{pmatrix} = \begin{pmatrix} \mathbf{u}_1 & \mathbf{u}_2 & ... & \mathbf{u}_n \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

$$\Leftrightarrow \begin{pmatrix} \mathbf{Au}_1 & \mathbf{Au}_2 & ... & \mathbf{Au}_n \end{pmatrix} = \begin{pmatrix} \lambda_1\mathbf{u}_1 & \lambda_2\mathbf{u}_2 & ... & \lambda_n\mathbf{u}_n \end{pmatrix}$$

by block multiplication

$\Leftrightarrow \mathbf{Au}_i = \lambda_i \mathbf{u}_i$ for $1 \leq i \leq n$

$\Leftrightarrow \mathbf{u}_i$ is the eigenvector of $\mathbf{A}$ for $1 \leq i \leq n$ (Claim 2)

Combining Claim 1 and Claim 2, we can deduce that $\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_n$ are eigenvectors of $\mathbf{A}$ that are linearly independent for $\mathbb{R}^n$.

16.4.3 Another Theorem

Let $\mathbf{A}$ be a $n \times n$ square matrix. Let λ_1 and λ_2 are distinct eigenvalues of $\mathbf{A}$, $\lambda_1 \neq \lambda_2$. Suppose $\{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k\}$ is a linearly independent subset of eigenspace associated to eigenvalue λ_1, and $\{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_m\}$ is a linearly independent subset of eigenspace associated to eigenvalue λ_2. Then, the union $\{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k, \mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_m\}$ is linearly independent.

16.5 Diagonalizability

16.5.1 A Theorem

The geometric multiplicity of an eigenvalue λ of a square matrix $\mathbf{A}$ is no greater than the algebraic multiplicity.

$$1 \leq \dim(E_\lambda) \leq r_\lambda$$

16.5.2 Equivalent Statements for Diagonalizability

Let $\mathbf{A}$ be a square matrix of order n. The following statements are equivalent.

1. $\mathbf{A}$ is diagonalizable
2. There exists a basis $\{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_n\}$ of $\mathbb{R}^n$ of eigenvectors of $\mathbf{A}$
3. The characteristic polynomial of $\mathbf{A}$ splits into linear factors,

   $$\det(x\mathbf{I} - \mathbf{A}) = (x - \lambda_1)^{r_{\lambda_1}}(x - \lambda_2)^{r_{\lambda_2}} \cdots (x - \lambda_k)^{r_{\lambda_k}}$$

   where r_{λ_i} is the algebraic multiciplity of λ_i for $i = 1, 2, ..., k$, and the eigenvalues are distinct, $\lambda_i \neq \lambda_j$ for all $i \neq j$, and the geometric multiplicity is equal to the algebraic multiplicity for each eigenvalue λ_i,

   $$\dim(E_{\lambda_i}) = r_{\lambda_i}$$

16.5.3 Not Diagonalizable

A square matrix $\mathbf{A}$ of order n is not diagonalizable if either

1. The characteristic polynomial $\det(x\mathbf{I} - \mathbf{A})$ does not split into linear factors
2. There exists an eigenvalue λ such that $\dim(E_\lambda) < r_\lambda$

16.5.4 A Theorem

If $\mathbf{A}$ is a square matrix of order n with n distinct eigenvalues, then $\mathbf{A}$ is diagonalizable.

16.5.5 Algorithm to Diagonalization

Let $\mathbf{A}$ be an order n square matrix.

1. Compute the characteristic polynomial

   $$\det(x\mathbf{I} - \mathbf{A}) = (x - \lambda_1)^{r_{\lambda_1}}(x - \lambda_2)^{r_{\lambda_2}} \cdots (x - \lambda_k)^{r_{\lambda_k}}$$

   If the characteristic polynomial do not split into linear factors, $\mathbf{A}$ is not diagonalizable.

2. For each eigenvalue λ_i of $\mathbf{A}$, $i = 1, 2, ..., k$, find a basis S_{λ_i} for the eigenspace. According to the definition of eigenspace, this is equivalent to finding S_{λ_i} for the solution space of the linear system $(\lambda_i\mathbf{I} - \mathbf{A})\mathbf{x} = 0$.

   Compute first the eigenspace associated to eigenvalues with algebraic multiplicity greater than 1. If $\dim(E_\lambda) < r_\lambda$, $\mathbf{A}$ is not diagonalizable.

3. Let $S = \bigcup_{i=1}^{k} S_{\lambda_i}$. Then, $S = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_n\}$ is a basis for $\mathbb{R}^n$ consisting of eigenvectors of $\mathbf{A}$.

4. Construct $\mathbf{P} = \begin{pmatrix}\mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_n\end{pmatrix}$ and $\mathbf{D} = \text{diag}(\mu_1, \mu_2, ..., \mu_n)$ from what we have computed. Here, μ_i is the eigenvalue associated to $\mathbf{u}_i$, $i = 1, 2, ..., n$. By the definition of eigenvalue and eigenvector, we know that $\mathbf{A}\mathbf{u}_i = \mu_i \mathbf{u}_i$ for $i = 1, 2, ..., n$. Therefore, by this, we can express $\mathbf{A}$ as

$$\mathbf{A} = \begin{pmatrix}\mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_n\end{pmatrix} \begin{pmatrix} \mu_1 & 0 & \cdots & 0 \\ 0 & \mu_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mu_n \end{pmatrix} \begin{pmatrix}\mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_n\end{pmatrix}^{-1}$$

16.5.6 How can the algorithm be true?

Recall $\mathbf{A}$, an order n square matrix, with μ_i as the eigenvalue associated to $\mathbf{u}_i$ for $i = 1, 2, ..., n$. By the definition of eigenvectors and eigenvalues, we have $\mathbf{A}\mathbf{u}_i = \mu_i \mathbf{u}_i$ for $i = 1, 2, ..., n$. Then, by block multiplication, we can represent it as

$$\begin{aligned}
\begin{pmatrix}\mathbf{A}\mathbf{u}_1 & \mathbf{A}\mathbf{u}_2 & \cdots & \mathbf{A}\mathbf{u}_n\end{pmatrix} &= \begin{pmatrix}\mu_1\mathbf{u}_1 & \mu_2\mathbf{u}_2 & \cdots & \mu_n\mathbf{u}_n\end{pmatrix} \\
\mathbf{A}\begin{pmatrix}\mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_n\end{pmatrix} &= \begin{pmatrix}\mu_1\mathbf{u}_1 & \mu_2\mathbf{u}_2 & \cdots & \mu_n\mathbf{u}_n\end{pmatrix} \\
\mathbf{A}\begin{pmatrix}\mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_n\end{pmatrix} &= \begin{pmatrix}\mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_n\end{pmatrix} \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}
\end{aligned}$$

By our assumption about $\mathbf{P}$ and $\mathbf{D}$, we can represent the equation above as

$$\begin{aligned}
\mathbf{AP} &= \mathbf{PD} \\
\mathbf{A} &= \mathbf{PDP}^{-1} && \text{the definition of diagonalization}
\end{aligned}$$

16.5.7 Another Perspective: Linear Map

Recall that $\mathbf{P} = \begin{pmatrix}\mathbf{u}_1 & \mathbf{u}_2 & ... & \mathbf{u}_n\end{pmatrix}$, where each $\mathbf{u}_i$ for $i = 1, 2, ..., n$ is the eigenvector of $\mathbf{A}$. View $\mathbf{P}$ as a function that transforms e_i, the standard basis at i, to $\mathbf{u}_i$, the eigenvector of $\mathbf{A}$ indexed at i, that is, $\mathbf{P}e_i = \mathbf{u}_i$ (Observation 1). Then, if we inverse $\mathbf{P}$, $\mathbf{P}^{-1}$ can be viewed as a function that transforms $\mathbf{u}_i$ to e_i (Observation 2).

Seeing the matrices as functions, we can say that two matrices $\mathbf{A}$ and $\mathbf{B}$ with size $m \times n$ are equivalent if and only if $\mathbf{Av} = \mathbf{Bv}$ for all $\mathbf{v} \in \mathbb{R}^n$ (Observation 3).

We will use these two observations to make an intuition for the diagonalization algorithm. First, observe that every matrix $\mathbf{v} \in \mathbb{R}^n$ can be written as linear combination of $\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_n$. Therefore, we can split $\mathbf{v}$ into sums of $\mathbf{u}_i$. Then, by observation 3, if $\mathbf{A} = \mathbf{PDP}^{-1}$, this equation holds true.

$$\begin{aligned}
\mathbf{A}\mathbf{u}_i &= \mathbf{PDP}^{-1}\mathbf{u_i} && \\
\mathbf{A}\mathbf{u}_i &= \mathbf{PD}e_i && \text{by Observation 2} \\
\mathbf{A}\mathbf{u}_i &= \mathbf{P}\mu_i && \text{by Matrix Multiplication} \\
\mathbf{A}\mathbf{u}_i &= \mu_i\mathbf{u}_i && \text{by Observation 1}
\end{aligned}$$

Observe that this equation comes from the definition of eigenvectors and eigenvalues. This is why we can diagonalize a matrix by computing its eigenvectors and eigenvalues.

16.6 Orthogonally Diagonalizable

16.6.1 The Definition

An order n square matrix $\mathbf{A}$ is orthogonally diagonalizable if $\mathbf{A} = \mathbf{PDP}^T$ for some orthogonal matrix $\mathbf{P}$ and diagonal matrix $\mathbf{D}$.

16.6.2 The Spectral Theorem

Let $\mathbf{A}$ be a $n \times n$ square matrix. $\mathbf{A}$ is orthogonally diagonalizable if and only if $\mathbf{A}$ is symmetric.

16.6.3 Equivalent statements for orthogonally diagonalizable

Let $\mathbf{A}$ be a square matrix of order n. The following statements are equivalent.

1. $\mathbf{A}$ is orthogonally diagonalizable
2. There exists orthonormal basis $\{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_n\}$ of $\mathbb{R}^n$ of eigenvectors of $\mathbf{A}$
3. $\mathbf{A}$ is a symmetric matrix

16.6.4 A Theorem

If $\mathbf{A}$ is a symmetric matrix, then the eigenspaces are orthogonal to each other. That is, suppose λ_1 and λ_2 are distinct eigenvalues of a symmetric matrix $\mathbf{A}$, $\lambda_1 \neq \lambda_2$, and $\mathbf{v}_i$ is an eigenvector associated to eigenvalue λ_i, for $i = 1, 2$. Then, $\mathbf{v}_1 \cdot \mathbf{v}_2 = 0$.

16.6.5 Algorithm to Orthogonal Diagonalization

Let $\mathbf{A}$ be an order n symmetric matrix. Since $\mathbf{A}$ is symmetric, it is orthogonally diagonalizable.

1. Compute the characteristic polynomial

$$\det(x\mathbf{I} - \mathbf{A}) = (x - \lambda_1)^{r_{\lambda_1}}(x - \lambda_2)^{r_{\lambda_2}} \cdots (x - \lambda_k)^{r_{\lambda_k}}$$

2. For each eigenvalue λ_i of $\mathbf{A}$, $i = 1, 2, ..., k$, find a basis S_{λ_i} for the eigenspace, that is, find a basis S_{λ_i} for the solution space of the following linear system,
$$(\lambda_i \mathbf{I} - \mathbf{A})\mathbf{x} = 0$$

3. Apply Gram-Schmidt process to each basis S_{λ_i} of the eigenspace E_{λ_i} to obtain an orthonormal basis T_{λ_i}. Let $T = \bigcup_{i=1}^{k} T_{\lambda_i}$. Then, $T = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_n\}$ is an orthonormal basis for $\mathbb{R}^n$.

   Note: By Theorem 16.6.4, it is guaranteed that for distinct eigenvector, the eigenspace is always orthogonal. Therefore, calculating $T = \bigcup_{i=1}^{k} T_{\lambda_i}$ do not break the orthogonality characteristic.

4. Let $\mathbf{P} = \begin{pmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_n \end{pmatrix}$, and $\mathbf{D} = diag(\mu_1, \mu_2, ..., \mu_n)$, where μ_i is the eigenvalue associated to $\mathbf{u}_i$, $i = 1, 2, ..., n$, $\mathbf{A}\mathbf{u}_i = \mu_i \mathbf{u}_i$. Then, $\mathbf{P}$ is an orthogonal matrix, and
$$\mathbf{A} = \mathbf{P}\mathbf{D}\mathbf{P}^T = \begin{pmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_n \end{pmatrix} \begin{pmatrix} \mu_1 & 0 & \cdots & 0 \\ 0 & \mu_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mu_n \end{pmatrix} \begin{pmatrix} \mathbf{u}_1^T \\ \mathbf{u}_2^T \\ \vdots \\ \mathbf{u}_n^T \end{pmatrix}$$

16.7 Markov Chain

16.7.1 A Theorem

Suppose $\mathbf{A} = \mathbf{P}\mathbf{D}\mathbf{P}^{-1}$. Then, for any positive integer m, $\mathbf{A}^m = \mathbf{P}\mathbf{D}^m\mathbf{P}^{-1}$.

Proof (by induction):

We will start by the base case when $m = 1$. So, $\mathbf{A}^1 = \mathbf{A} = \mathbf{P}\mathbf{D}\mathbf{P}^{-1}$, which holds by assumption. We will then continue with the inductive step. Assume that for some $m \geq 1$, $\mathbf{A}^m = \mathbf{P}\mathbf{D}^m\mathbf{P}^{-1}$. Now, we want to show that $\mathbf{A}^{m+1} = \mathbf{P}\mathbf{D}^{m+1}\mathbf{P}^{-1}$.

Starting from the inductive hypothesis $\mathbf{A}^m = \mathbf{P}\mathbf{D}^m\mathbf{P}^{-1}$, we will multiply both sides of the equation by $\mathbf{A}$. We will then obtain
$$\mathbf{A}^{m+1} = \mathbf{A}^m\mathbf{A} = (\mathbf{P}\mathbf{D}^m\mathbf{P}^{-1})(\mathbf{P}\mathbf{D}\mathbf{P}^{-1}) = \mathbf{P}\mathbf{D}^m(\mathbf{P}^{-1}\mathbf{P})\mathbf{D}\mathbf{P}^{-1}$$

Then, because $\mathbf{P}^{-1}\mathbf{P} = \mathbf{I}$, we can further simplify the expression into
$$\mathbf{A}^{m+1} = \mathbf{P}\mathbf{D}^m\mathbf{I}\mathbf{D}\mathbf{P}^{-1} = \mathbf{P}\mathbf{D}^m\mathbf{D}\mathbf{P}^{-1} = \mathbf{P}\mathbf{D}^{m+1}\mathbf{P}^{-1}$$

Thus, the statement holds for $m + 1$, completing the induction proof.

16.7.2 Another Theorem

Let $\mathbf{D} = \begin{pmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n \end{pmatrix}$ be a diagonal matrix. Then, for any positive integer m, $\mathbf{D}^m = \begin{pmatrix} d_1^m & 0 & \cdots & 0 \\ 0 & d_2^m & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n^m \end{pmatrix}$

Proof (by induction):

We will start by the base case when $m = 1$. So,

$$\mathbf{D}^1 = \mathbf{D} = \begin{pmatrix} d_1^1 & 0 & \cdots & 0 \\ 0 & d_2^1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n^1 \end{pmatrix} = \begin{pmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n \end{pmatrix}$$

which hold by assumption. We will then continue with the inductive step. Assume that for some $m \geq 1$,

$$\mathbf{D}^m = \begin{pmatrix} d_1^m & 0 & \cdots & 0 \\ 0 & d_2^m & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n^m \end{pmatrix}$$

Now, we want to show that

$$\mathbf{D}^{m+1} = \begin{pmatrix} d_1^{m+1} & 0 & \cdots & 0 \\ 0 & d_2^{m+1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n^{m+1} \end{pmatrix}$$

Starting from the inductive hypothesis, we will multiply both sides of the equation by $\mathbf{D}$. We will then obtain

$$\mathbf{D}^m\mathbf{D} = \begin{pmatrix} d_1^m & 0 & \cdots & 0 \\ 0 & d_2^m & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n^m \end{pmatrix} \begin{pmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n \end{pmatrix} = \begin{pmatrix} d_1^{m+1} & 0 & \cdots & 0 \\ 0 & d_2^{m+1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n^{m+1} \end{pmatrix}$$

Thus, the statement holds for $m + 1$, completing the induction proof.

16.7.3 A Corollary

Suppose $\mathbf{A}$ is diagonalizable. Write

$$\mathbf{A} = \mathbf{P}\begin{pmatrix} \mu_1 & 0 & \cdots & 0 \\ 0 & \mu_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mu_n \end{pmatrix}\mathbf{P}^{-1}$$

Then, for any positive integer $k > 0$,

$$\mathbf{A}^k = \mathbf{P}\begin{pmatrix} \mu_1^k & 0 & \cdots & 0 \\ 0 & \mu_2^k & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mu_n^k \end{pmatrix}\mathbf{P}^{-1}$$

Moreover, if $\mathbf{A}$ is invertible, then the identity above holds for any integer $k \in \mathbb{Z}$.

16.7.4 The Definition of Probability Vector

A probability vector is a vector $\mathbf{v} = (\mathbf{v}_i)_n$ with nonnegative coordinates that add up to 1, $\sum_{i=1}^{n} \mathbf{v}_i = 1$.

16.7.5 The Definition of Stochastic Matrix

A stochastic matrix is a square matrix whose columns are probability vectors.

16.7.6 The Definition of Markov Chain

A Markov chain is a sequence of probability vectors $\mathbf{x}_0, \mathbf{x}_1, ..., \mathbf{x}_k, ..., \mathbf{x}_\infty$, together with a stochastic matrix $\mathbf{P}$ such that

$$\mathbf{x}_1 = \mathbf{P}\mathbf{x}_0, \mathbf{x}_2 = \mathbf{P}\mathbf{x}_1, ..., \mathbf{x}_k = \mathbf{P}\mathbf{x}_{k-1}, ...$$

16.7.7 Equilibrium vector

A steady-state vector, or equilibrium vector, for stochastic vector $\mathbf{P}$ is a probability vector that is an eigenvector associated to eigenvalue 1.

Intuition: Use the definition of the Markov chain, and observe what happen when k approaches infinity; we will have

$$\mathbf{x}_\infty = \mathbf{P}\mathbf{x}_{\infty-1} = \mathbf{P}\mathbf{x}_\infty$$

Then, from the equation $\mathbf{P}\mathbf{x}_\infty = \mathbf{x}_\infty$, by the definition of eigenvalues and eigenvectors, observe that $\mathbf{x}_\infty$ is the eigenvector of $\mathbf{P}$ associated to eigenvalue 1.

$\mathbf{x}_\infty$, by definition, is the equilibrium vector. Therefore, finding equilibrium vector for stochastic vector $\mathbf{P}$, is equivalent to searching an eigenvector $\mathbf{u}$ associated to eigenvalue 1.

16.7.8 A Theorem

Let $\mathbf{P}$ be a $n \times n$ stochastic matrix and

$$\mathbf{x}_0, \mathbf{x}_1 = \mathbf{P}\mathbf{x}_0, \mathbf{x}_2 = \mathbf{P}\mathbf{x}_1, ..., \mathbf{x}_k = \mathbf{P}\mathbf{x}_{k-1}$$

be a Markov chain for some probability vector $\mathbf{x}_0$. If the Markov chain converges, it will converge to an equilibrium vector.

16.7.9 Regular Stochastic Matrix

A stochastic matrix is regular if for some positive integer $k > 0$, the matrix power $\mathbf{P}^k$ has positive entries,

$$\mathbf{P}^k = (a_{ij})_n, a_{ij} > 0 \text{ for all } i, j = 1, 2, ..., n$$

16.7.10 A Theorem

Suppose

$$\mathbf{x}_1 = \mathbf{P}\mathbf{x}_0, \mathbf{x}_2 = \mathbf{P}\mathbf{x}_1, ..., \mathbf{x}_k = \mathbf{P}\mathbf{x}_{k-1}, ...$$

is a Markov chain and $\mathbf{P}$ is a regular stochastic matrix. Then, the Markov chain will converge to the unique equilibrium vector.

16.7.11 Algorithm to Compute Equilibrium Vector

Let $\mathbf{P}$ be a $n \times n$ stochastic matrix.

1. By the intuition in Chapter 16.7.7, find an eigenvector $\mathbf{u}$ associated to eigenvalue $\lambda = 1$. Based on Theorem 16.1.3, equivalently, find a nontrivial solution to the homogeneous system $(\mathbf{I} - \mathbf{P})\mathbf{x} = \mathbf{0}$.
2. Write $\mathbf{u} = (\mathbf{u}_i)$. Then, divide each $\mathbf{u}_i$ with $\sum_{k=1}^{n} \mathbf{u}_k$ to obtain the equilibrium vector.

16.7.12 Alternative Algorithm

The equilibrium eigenvectors are solutions to the equation

$$\begin{pmatrix} & \mathbf{P} - \mathbf{I}_n & & \\ 1 & 1 & \cdots & 1 \end{pmatrix} \mathbf{x} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

where $\mathbf{I}_n$ is the identity $n \times n$ identity matrix.

16.7.13 Google PageRank Algorithm

Assume a set S of sites contains key words on a topic of common interest. Our goal is to design an algorithm to rank the sites, so that the sites with the highest rank appear first.

In 1996, a new search engine "Google" was developed by Larry Page and Sergey Brin. It is based on the PageRank algorithm that involve the use of dominant eigenvector of some matrix. This is one of the application of Markov chain that we are going to discuss.

The underlying assumption of this algorithm is that more important websites are likely to receive more links from other websites. With this assumption in mind, define the adjacency matrix for S be the order n square matrix $\mathbf{A} = (a_{ij})$ where

$$a_{ij} = \begin{cases} 1 & \text{if site } j \text{ has an outgoing link to site } i \\ 0 & \text{if site } j \text{ does not have an outgoing link to site } i \end{cases}$$

Good to observe:

1. The sum of the entries in the i-th row of $\mathbf{A}$ is the number of incoming links to site i from other sites
2. The sum of the entries in the j-th column of $\mathbf{A}$ is the number of outgoing links from site j to other sites.

From the adjacency matrix $\mathbf{A}$, define the probability transition matrix $\mathbf{P} = (p_{ij})$ by dividing each entry of $\mathbf{A}$ by the sum of the entries in the same column.

$$p_{ij} = \frac{a_{ij}}{\sum_{k=1}^{n} a_{kj}}$$

Observe that by doing this, we will obtain $\mathbf{P}$, a stochastic matrix. This matrix incorporates the probability information for advancing randomly from one site to the next with a mouse click. Define the value of $\mathbf{x}_0$ to be

$$\mathbf{x}_{0i} = \begin{cases} 1 & \text{if the surfer start at surfing from site } i \\ 0 & \text{otherwise} \end{cases}$$

Therefore,

$$\mathbf{x}_k = \mathbf{P}^k \mathbf{x}_0$$

will show the probability of landing on each site after k random clicks.

Then, if we calculate $\mathbf{x}_\infty$, we will have the probability that we will end up at each of the sites when surfing infinitely. The algorithm will use this information to rank the sites sorted based on the probability that users will visit them after infinitely surfing.

16.7.14 Other application of Markov Chain

Observe the figure below.

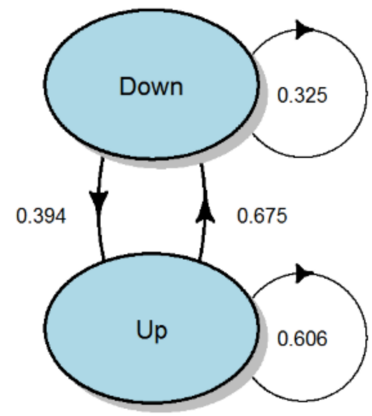

Figure 10: Application of Markov Chain on Analyzing Stocks

Markov chain can also be used to analyze stocks trend. By computing the probability transition matrix, we can get the probability of the stocks going up or down, which will be very helpful for decision making.

For more reference, can visit this site.

17 Week 11

17.1 First Order Differential Systems

A first order linear system of differential equations (with variable t) in its standard form is given by

$$\begin{cases} y_1'(t) = a_{11}(t)y_1(t) + \cdots + a_{1n}(t)y_n(t) + g_1(t) \\ y_2'(t) = a_{21}(t)y_1(t) + \cdots + a_{2n}(t)y_n(t) + g_2(t) \\ \qquad \vdots \\ y_n'(t) = a_{n1}(t)y_1(t) + \cdots + a_{nn}(t)y_n(t) + g_n(t) \end{cases}$$

This system can be rewritten in matrix form as

$$\begin{pmatrix} y_1'(t) \\ y_2'(t) \\ \vdots \\ y_n'(t) \end{pmatrix} = \begin{pmatrix} a_{11}(t) & a_{12}(t) & \cdots & a_{1n}(t) \\ a_{21}(t) & a_{22}(t) & \cdots & a_{2n}(t) \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1}(t) & a_{n2}(t) & \cdots & a_{nn}(t) \end{pmatrix} \begin{pmatrix} y_1(t) \\ y_2(t) \\ \vdots \\ y_n(t) \end{pmatrix} + \begin{pmatrix} g_1(t) \\ g_2(t) \\ \vdots \\ g_n(t) \end{pmatrix}$$

In compact vector notation, this can also be written as

$$\mathbf{y}'(t) = \mathbf{A}(t)\mathbf{y}(t) + \mathbf{g}(t)$$

where $\mathbf{A}(t)$ is the coefficient matrix, which is generally a matrix-valued function of t; $\mathbf{y}(t)$ is the vector of unknown functions; $\mathbf{A}(t)$ governs the system dynamics; and, $\mathbf{g}(t)$ is a given forcing term.

17.2 Classification: Homogeneous Differential Systems

A first order linear system of differential equations is **homogeneous** if there is no forcing term, $\mathbf{g}(t) = 0$. The equation in compact vector notation can be written as $\mathbf{y}'(t) = \mathbf{A}(t)\mathbf{y}(t)$.

17.3 Classification: Differential Systems with Constant Coefficient

A first order linear system of differential equations has constant coefficients if the entries of $\mathbf{A}(t)$ are constants. Simply speaking, $\mathbf{A}$ is a real-valued constant matrix.

17.4 Homogeneous First Order Differential Systems with Constant Coefficient

17.4.1 The Definition

A first order linear system of differential equations (with variable t) in its standard form is given by

$$\begin{cases} y_1'(t) = a_{11}y_1(t) + \cdots + a_{1n}y_n(t) \\ y_2'(t) = a_{21}y_1(t) + \cdots + a_{2n}y_n(t) \\ \qquad \vdots \\ y_n'(t) = a_{n1}y_1(t) + \cdots + a_{nn}y_n(t) \end{cases}$$

This system can be rewritten in matrix form as

$$\begin{pmatrix} y_1'(t) \\ y_2'(t) \\ \vdots \\ y_n'(t) \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} y_1(t) \\ y_2(t) \\ \vdots \\ y_n(t) \end{pmatrix}$$

In compact vector notation, this can also be written as

$$\mathbf{y}'(t) = \mathbf{A}\mathbf{y}(t)$$

17.4.2 What is a solution?

A function valued vector $\mathbf{x}(t)$ is called a solution to the first order homogeneous linear differential systems with constant coefficient:

$$\mathbf{y}'(t) = \mathbf{A}\mathbf{y}(t)$$

if it satisfies

$$\mathbf{x}'(t) = \mathbf{A}\mathbf{x}(t)$$

If $\mathbf{y}(t_0) = \mathbf{a} \in \mathbb{R}^n$ is an initial condition for the system, then $\mathbf{x}(t)$ is a solution if

$$\mathbf{x}(t_0) = \mathbf{a}$$

17.4.3 A Theorem

Suppose $\mathbf{v}$ is an eigenvector associated with the eigenvalue λ of a matrix $\mathbf{A}$. Then, the function

$$\mathbf{x}(t) = \mathbf{v}e^{\lambda t}$$

is a solution to the first-order homogeneous linear system of differential equations with constant coefficients

$$\mathbf{y}' = \mathbf{A}\mathbf{y}$$

Proof:

From Chapter 17.4.2, we know that $\mathbf{x}(t)$ is a solution to $\mathbf{y}' = \mathbf{A}\mathbf{y}$ if and only if $\mathbf{x}'(t) = \mathbf{A}\mathbf{x}(t)$. Therefore, let's find $\mathbf{x}'(t)$, that is $\frac{d}{dt}[\mathbf{v}e^{\lambda t}]$!

First of all, because $\mathbf{v}$ is just a constant with respect to t, we can pull it out.

$$\mathbf{x}'(t) = \frac{d}{dt}[\mathbf{v}e^{\lambda t}] = \mathbf{v}\frac{d}{dt}[e^{\lambda t}] = \mathbf{v}\lambda e^{\lambda t} = \lambda \mathbf{v}e^{\lambda t}$$

By the definition of eigenvalues and eigenvectors, $\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$. Therefore,

$$\mathbf{x}'(t) = \lambda\mathbf{v}e^{\lambda t} = \mathbf{A}\mathbf{v}e^{\lambda t} = \mathbf{A}\mathbf{x}(t)$$

17.4.4 The Superposition Principle

Let $\mathbf{y}' = \mathbf{A}\mathbf{y}$ be a first order homogeneous linear differential system with constant coefficient. Suppose $\mathbf{x}_1(t)$ and $\mathbf{x}_2(t)$ are solutions to $\mathbf{y}' = \mathbf{A}\mathbf{y}$. For any $\alpha, \beta \in \mathbb{R}$,

$$\alpha\mathbf{x}_1(t) + \beta\mathbf{x}_2(t)$$

is also a solution to the system of differential equations.

Proof:

We know that $\mathbf{x}_1(t)$ and $\mathbf{x}_2(t)$ are solutions to $\mathbf{y}' = \mathbf{A}\mathbf{y}$. By the definition of solution, we have $\mathbf{x}_1'(t) = \mathbf{A}\mathbf{x}_1(t)$ and $\mathbf{x}_2'(t) = \mathbf{A}\mathbf{x}_2(t)$. By this information, we will prove that $\mathbf{x}(t) = \alpha\mathbf{x}_1(t) + \beta\mathbf{x}_2(t)$ is also a solution to $\mathbf{y}' = \mathbf{A}\mathbf{y}$. We will prove this by showing $\mathbf{x}'(t) = \mathbf{A}\mathbf{x}(t)$.

$$\begin{aligned}
\mathbf{x}'(t) &= \frac{d}{dt}[\alpha\mathbf{x}_1(t) + \beta\mathbf{x}_2(t)] \\
&= \alpha\mathbf{x}_1'(t) + \beta\mathbf{x}_2'(t) \\
&= \alpha\mathbf{A}\mathbf{x}_1(t) + \beta\mathbf{A}\mathbf{x}_2(t) \\
&= \mathbf{A}(\alpha\mathbf{x}_1(t) + \beta\mathbf{x}_2(t)) \\
&= \mathbf{A}\mathbf{x}(t)
\end{aligned}$$

17.5 Fundamental Set of Solutions

17.5.1 Zero Vector in Function-valued Vector

A function-valued vector $\mathbf{v}(t)$ is zero if for every t, $\mathbf{v}(t) = \mathbf{0} \in \mathbb{R}^n$.

17.5.2 Linear Independence in Function-valued Vector

A set $\{\mathbf{x}_1(t), \mathbf{x}_2(t), ..., \mathbf{x}_k(t)\}$ of function-valued vector is linearly independent if whenever $c_1, c_2, ..., c_k \in \mathbb{R}$ are real numbers such that

$$c_1\mathbf{x}_1(t) + c_2\mathbf{x}_2(t) + ... + c_k\mathbf{x}_k(t) = \mathbf{0}$$

for all t in the domain, necessarily $c_1 = c_2 = c_3 = \cdots = c_k = 0$. Otherwise, we say that it is linearly independent.

Note: This is actually similar to the definition of linear independence in normal vector. The only difference is that now we generalize it so that it applies for every value of t in every function-valued vector.

17.5.3 A Corollary

For a function-valued vector $\mathbf{v}(t)$ to be zero, $\mathbf{v}(t) = 0$ for all $t \in \mathbb{R}$. Therefore, for $\{\mathbf{x}_1(t), \mathbf{x}_2(t), ..., \mathbf{x}_k(t)\}$ to be linearly dependent, we can find some $c_1, c_2, ..., c_k$ not all zero, and there exists one value of $t \in \mathbb{R}$ such that

$$c_1\mathbf{x}_1(t) + c_2\mathbf{x}_2(t) + ... + c_k\mathbf{x}_k(t) = \mathbf{0}$$

17.5.4 A Problem

Observe what happen when we represent the equation

$$c_1\mathbf{x}_1(t) + c_2\mathbf{x}_2(t) + ... + c_k\mathbf{x}_k(t) = 0$$

as an augmented matrix.

Each entry in the LHS will be represented in terms of t as $\mathbf{x}_1(t), \mathbf{x}_2(t), ..., \mathbf{x}_k(t)$ depend on t. Then, because we don't know what the value of t is, we cannot perform elementary row operations. Therefore, we cannot find the RREF for checking linear independence.

17.5.5 The Solution for Special Cases

Let's consider a special case where $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n \in \mathbb{R}^n$.

Construct $\mathbf{A} = \begin{pmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n \end{pmatrix}$. $\mathbf{A}$ is a square matrix of order n with $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n$ be the rows of $\mathbf{A}$.

By the equivalent statements of invertibility, $\{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n\}$, the rows of $\mathbf{A}$, is linearly independent if and only if $\det(\mathbf{A})$ is nonzero.

17.5.6 The Definition of Wronskian

Inspired by what discussed in Chapter 17.5.5, let $S = \{\mathbf{x}_1(t), \mathbf{x}_2(t), ..., \mathbf{x}_n(t)\}$ be a set containing n functioned-valued vectors with n coordinates. Define the Wronskian of S to be

$$W(\mathbf{x}_1(t), \mathbf{x}_2(t), ..., \mathbf{x}_n(t)) = \det(\begin{pmatrix} \mathbf{x}_1(t) & \mathbf{x}_2(t) & ... & \mathbf{x}_n(t) \end{pmatrix})$$

Then, $S = \{\mathbf{x}_1(t), \mathbf{x}_2(t), ..., \mathbf{x}_n(t)\}$ is linearly independent if the Wronskian is not the constant zero function,

$$W(\mathbf{x}_1(t), \mathbf{x}_2(t), ..., \mathbf{x}_n(t)) \neq 0$$

17.5.7 A Remark

To show that the Wronskian to not be the constant zero function, show that there exists one t in its domain such that $W(\mathbf{x}_1(t), \mathbf{x}_2(t), ..., \mathbf{x}_n(t)) \neq 0$.

17.5.8 A Strong Alert

$$\begin{aligned} \text{The Wronskian is not constant zero} &\Rightarrow S \text{ is linearly independent} \\ S \text{ is linearly independent} &\not\Rightarrow \text{the Wronskian is not constant zero} \end{aligned}$$

There might be a case where the Wronskian is a constant zero function, but the set is still linearly independent.

17.5.9 The Definition of Fundamental Set of Solutions

A set $S = \{\mathbf{x}_1(t), \mathbf{x}_2(t), ..., \mathbf{x}_n(t)\}$ of solutions to the differential system $\mathbf{y}' = \mathbf{A}\mathbf{y}$ is called a fundamental set of solutions if its Wronskian is nonzero.

17.5.10 General Solution from the Fundamental Set of Solutions

Let $S = \{\mathbf{x}_1(t), \mathbf{x}_2(t), ..., \mathbf{x}_n(t)\}$ be a fundamental set of solutions to a differential system $\mathbf{y}' = \mathbf{A}\mathbf{y}$. Then, the general solution to the differential system is

$$\mathbf{x}(t) = c_1\mathbf{x}_1(t) + c_2\mathbf{x}_2(t) + ... + c_n\mathbf{x}_n(t), \quad c_1, c_2, ..., c_n \in \mathbb{R}$$

In other words, any solution to the differential system $\mathbf{y}' = \mathbf{A}\mathbf{y}$ must be a linear combination of the function-valued vectors in the fundamental set of solutions.

17.5.11 A Theorem

Suppose $\mathbf{A}$ is diagonalizable. Let $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n$ be n linearly independent eigenvectors associated to real eigenvalues $\lambda_1, \lambda_2, ..., \lambda_n$ (not necessarily distinct). Then,

$$\{\mathbf{v}_1 e^{\lambda_1 t}, \mathbf{v}_2 e^{\lambda_2 t}, ..., \mathbf{v}_n e^{\lambda_n t}\}$$

is a fundamental set of solutions to the first order homogeneous system of differential equations $\mathbf{y}' = \mathbf{A}\mathbf{y}$ with constant coefficients, and

$$\mathbf{x}(t) = c_1\mathbf{v}_1 e^{\lambda_1(t)} + c_2\mathbf{v}_2 e^{\lambda_2(t)} + ... + c_n\mathbf{v}_n e^{\lambda_n(t)}, \quad c_1, c_2, ..., c_n \in \mathbb{R}$$

is the general solution.

Proof:

To prove that

$$\{\mathbf{v}_1 e^{\lambda_1 t}, \mathbf{v}_2 e^{\lambda_2 t}, ..., \mathbf{v}_n e^{\lambda_n t}\}$$

is a fundamental set of solutions, we need to prove that $W(\mathbf{v}_1 e^{\lambda_1 t}, \mathbf{v}_2 e^{\lambda_2 t}, ..., \mathbf{v}_n e^{\lambda_n t})$ is nonzero.

$$\begin{aligned}
W(\mathbf{v}_1 e^{\lambda_1 t}, \mathbf{v}_2 e^{\lambda_2 t}, ..., \mathbf{v}_n e^{\lambda_n t}) &= \det\big(\begin{pmatrix}\mathbf{v}_1 e^{\lambda_1 t} & \mathbf{v}_2 e^{\lambda_2 t} & \cdots & \mathbf{v}_n e^{\lambda_n t}\end{pmatrix}\big) \\
&= e^{\lambda_1 t} e^{\lambda_2 t} ... e^{\lambda_n t} \det\big(\begin{pmatrix}\mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n\end{pmatrix}\big) \\
&= e^{(\lambda_1+\lambda_2+\lambda_3+...+\lambda_n)t} \det\big(\begin{pmatrix}\mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n\end{pmatrix}\big)
\end{aligned}$$

Then, because $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v_n} \in \mathbb{R}^n$ is linearly independent, by the equivalent statements of invertibility, $\det\big(\begin{pmatrix}\mathbf{v}_1 & \mathbf{v}_2 & ... & \mathbf{v_n}\end{pmatrix}\big) \neq 0$. Therefore, because of this, $W(\mathbf{v}_1 e^{\lambda_1 t}, \mathbf{v}_2 e^{\lambda_2 t}, ..., \mathbf{v}_n e^{\lambda_n t}) \neq 0$.

17.6 Complex Eigenvalues and Eigenvectors

17.6.1 Complex Vectors

An n-complex vector is a collection of n ordered complex numbers,

$$\mathbf{v} = \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{pmatrix}, z_j \in \mathbb{C} \text{ for all } j = 1, 2, ..., n$$

The collection of all n-complex vectors is denoted as $\mathbb{C}^n$. Given any complex vector $\mathbf{v} \in \mathbb{C}^n$, we can split it into its real and imaginary parts,

$$\mathbf{v} = \begin{pmatrix} x_1 + iy_1 \\ x_2 + iy_2 \\ \vdots \\ x_n + iy_n \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} + i \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = Re(\mathbf{v}) + i\ Im(\mathbf{v})$$

where $Re(\mathbf{v}), Im(\mathbf{v}) \in \mathbb{R}^n$.

17.6.2 Complex Conjugate

Let

$$\mathbf{v} = \begin{pmatrix} x_1 + iy_1 \\ x_2 + iy_2 \\ \vdots \\ x_n + iy_n \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} + i \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = Re(\mathbf{v}) + i\ Im(\mathbf{v})$$

be a complex n-vector. It's complex conjugate is

$$\overline{\mathbf{v}} = \begin{pmatrix} \overline{x_1 + iy_1} \\ \overline{x_2 + iy_2} \\ \vdots \\ \overline{x_n + iy_n} \end{pmatrix} = \begin{pmatrix} x_1 - iy_1 \\ x_2 - iy_2 \\ \vdots \\ x_n - iy_n \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} - i \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = Re(\mathbf{v}) - i\ Im(\mathbf{v})$$

17.6.3 Complex eigenvector and eigenvalue (Definition)

Let $\mathbf{A}$ be an order n real square matrix. $\lambda \in \mathbb{C}$ is a complex eigenvalue of $\mathbf{A}$ if there is a nonzero vector $\mathbf{v} \in \mathbb{C}^n$, $\mathbf{v} \neq 0$ such that

$$\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$$

In this case, $\mathbf{v}$ is called a complex eigenvector of $\mathbf{A}$ associated to λ.

17.6.4 A Theorem

Let $\mathbf{A}$ be an order n square matrix with real entries.

1. Complex eigenvalues of $\mathbf{A}$ comes in conjugate pairs. If $\lambda \in \mathbb{C}$ is an eigenvalue of $\mathbf{A}$, then $\overline{\lambda}$ is also an eigenvalue of $\mathbf{A}$.

2. If $\mathbf{v} \in \mathbb{C}^n$ is an eigenvector associated to eigenvalue λ, then $\overline{\mathbf{v}}$ is an eigenvector associated to eigenvalue $\overline{\lambda}$.

Proof:

Let $\mathbf{v} \in \mathbb{C}^n$ be an eigenvector of $\mathbf{A}$ associated to eigenvalue λ. By the definition of eigenvalue and eigenvector, we have

$$\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$$

Observe this claim: because $\mathbf{A}$ is a real matrix, $\overline{\mathbf{A}} = \mathbf{A}$. Let's analyze the conjugation of the equation above and use this claim!

$$\begin{aligned}
\overline{\mathbf{A}\mathbf{v}} &= \overline{\lambda\mathbf{v}} \\
\overline{\mathbf{A}}\overline{\mathbf{v}} &= \overline{\lambda}\overline{\mathbf{v}} \\
\mathbf{A}\overline{\mathbf{v}} &= \overline{\lambda}\overline{\mathbf{v}}
\end{aligned}$$

By this, observe that $\overline{\mathbf{v}}$ is an eigenvector of $\mathbf{A}$ associated to eigenvalue $\overline{\lambda}$.

17.6.5 A Theorem

Suppose $\lambda \in \mathbb{C}$ is a complex eigenvalue of $\mathbf{A}$, a square matrix of order n, with eigenvector $\mathbf{v} \in \mathbb{C}^n$. Then, $\mathbf{v}e^{\lambda t}$ is a complex solution to the differential system $\mathbf{y}' = \mathbf{A}\mathbf{y}$.

Proof:

$$\begin{aligned}
\mathbf{y}' &= \frac{d}{dt}(\mathbf{v}e^{\lambda t}) \\
&= \mathbf{v}\frac{d}{dt}(e^{\lambda t}) \\
&= \mathbf{v}\lambda e^{\lambda t} \\
&= \lambda\mathbf{v}e^{\lambda t} \\
&= \mathbf{A}\mathbf{v}e^{\lambda t} && \text{by the definition of eigenvalue and eigenvector} \\
&= \mathbf{A}(\mathbf{v}e^{\lambda t}) \\
&= \mathbf{A}\mathbf{y} && \text{by the definition of } \mathbf{y}
\end{aligned}$$

17.6.6 Deriving Real Solutions from Complex Solution

Suppose $\lambda \in \mathbb{C}$ is a complex eigenvalue of $\mathbf{A}$ associated to complex eigenvector $\mathbf{v} \in \mathbb{C}^n$. Decompose λ and $\mathbf{v}$ into their real and imaginary parts,

$$\lambda = \lambda_r + i\lambda_i, \mathbf{v} = \mathbf{v}_r + i\mathbf{v}_i, \quad \lambda_r, \lambda_i \in \mathbb{R}, \quad \mathbf{v}_r, \mathbf{v}_i \in \mathbb{R}^n$$

Then,

$$\begin{aligned}
e^{\lambda t}\mathbf{v} &= e^{(\lambda_r + i\lambda_i)t}(\mathbf{v}_r + i\mathbf{v}_i) \\
&= e^{\lambda_r t}e^{i\lambda_i t}(\mathbf{v}_r + i\mathbf{v}_i)
\end{aligned}$$

Recall the euler identity, i.e., $e^{ix} = \cos(x) + i\sin(x)$. We will use this identity to break down $e^{i\lambda_i t}$.

$$\begin{aligned}
e^{\lambda t}\mathbf{v} &= e^{\lambda_r t}(\cos(\lambda_i t) + i\sin(\lambda_i t))(\mathbf{v}_r + i\mathbf{v}_i) \\
&= e^{\lambda_r t}(\mathbf{v}_r\cos(\lambda_i t) + i\mathbf{v}_r\sin(\lambda_i t) + i\mathbf{v}_i\cos(\lambda_i t) + i^2\mathbf{v}_i\sin(\lambda_i t)) \\
&= e^{\lambda_r t}(\mathbf{v}_r\cos(\lambda_i t) - \mathbf{v}_i\sin(\lambda_i t)) + ie^{\lambda_r t}(\mathbf{v}_r\sin(\lambda_i t) + \mathbf{v}_i\cos(\lambda_i t))
\end{aligned}$$

Define

$$\begin{aligned}
Re(e^{\lambda t}\mathbf{v}) = \mathbf{x}_r(t) = e^{\lambda_r t}(\mathbf{v}_r\cos(\lambda_i t) - \mathbf{v}_i\sin(\lambda_i t)) \\
Im(e^{\lambda t}\mathbf{v}) = \mathbf{x}_i(t) = e^{\lambda_r t}(\mathbf{v}_r\sin(\lambda_i t) + \mathbf{v}_i\cos(\lambda_i t))
\end{aligned}$$

where both of them are real function-valued vectors and $e^{\lambda t}\mathbf{v} = Re(e^{\lambda t}\mathbf{v}) + i\,Im(e^{\lambda t}\mathbf{v})$.

17.6.7 A Theorem

Both $\mathbf{x}_r(t)$ and $\mathbf{x}_i(t)$ from Chapter 17.6.6 are real solutions to the differential equations $\mathbf{y}' = \mathbf{A}\mathbf{y}$, and $\{\mathbf{x}_r(t), \mathbf{x}_i(t)\}$ is linearly independent.

17.6.8 A Corollary from Chapter 17.6.6 and 17.6.7

Suppose $\mathbf{A}$ is an order 2 square matrix and $\lambda \in \mathbb{C}$ is a nonreal complex eigenvalue $\lambda \notin \mathbb{R}$. Let $\mathbf{v} \in \mathbb{C}^2$ be an associated eigenvector. Write $e^{\lambda t}\mathbf{v} = \mathbf{x}_r(t) + i\mathbf{x}_i(t)$. Then, $\{\mathbf{x}_r(t), \mathbf{x}_i(t)\}$ is a fundamental set of solutions for $\mathbf{y}' = \mathbf{A}\mathbf{y}$, and

$$\mathbf{x} = c_1\mathbf{x}_1(t) + c_2\mathbf{x}_2(t)$$

is the general solution. Here,

$$\begin{aligned}
\mathbf{x}_r(t) = e^{\lambda_r t}(\cos(\lambda_i t)\mathbf{v}_r - \sin(\lambda_i t)\mathbf{v}_i) \\
\mathbf{x}_i(t) = e^{\lambda_r t}(\sin(\lambda_i t)\mathbf{v}_r + \cos(\lambda_i t)\mathbf{v}_i)
\end{aligned}$$

where $\lambda = \lambda_r + i\lambda_i, \quad \mathbf{v} = \mathbf{v}_r + i\mathbf{v}_i, \quad \lambda_r, \lambda_i \in \mathbb{R}, \quad \mathbf{v}_r, \mathbf{v}_i \in \mathbb{R}^n$

17.7 Repeated Eigenvalue and Generalized Eigenvector

17.7.1 The Definition of Repeated Eigenvalue

Let $\mathbf{A}$ be an order n square matrix. An eigenvalue λ of $\mathbf{A}$ is a repeated eigenvalue if the algebraic multiplicity of λ is more than 1, $r_\lambda > 1$.

17.7.2 The Definition of Generalized Eigenvector

Suppose now λ is a repeated eigenvalue such that the geometric multiplicity is strictly smaller than the algebraic multiplicity, $\dim(E_\lambda) < r_\lambda$.

In this case, we don't have enough linearly independent eigenvectors associated to λ to diagonalize $\mathbf{A}$ and as a consequence to form a fundamental set of solutions. This is where generalized eigenvector comes in to address this issue.

Let $\mathbf{v}_1$ be an eigenvector associated to λ. A vector $\mathbf{v}_2$ is a generalized eigenvector associated to λ if it is a solution to

$$(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{v}_1$$

17.7.3 A Theorem

Let λ be a repeated eigenvalue of an order n matrix $\mathbf{A}$, with eigenvector $\mathbf{v}_1$ and a generalized vector $\mathbf{v}_2$ associated to λ. Let

$$\begin{aligned}\mathbf{x}_1(t) &= e^{\lambda t}\mathbf{v}_1\\ \mathbf{x}_2(t) &= e^{\lambda t}(t\mathbf{v}_1 + \mathbf{v}_2)\end{aligned}$$

Then, $\{\mathbf{x}_1(t), \mathbf{x}_2(t)\}$ is a linearly independent set of real solutions to the differential equations $\mathbf{y}' = \mathbf{A}\mathbf{y}$.

17.7.4 A Corollary

Let $\mathbf{A}$ be an order 2 square matrix with a repeated eigenvalue λ. Suppose $\mathbf{v}_1$ is an eigenvector, and $\mathbf{v}_2$ is a generalized eigenvector associated to λ. Then, the set

$$\{\mathbf{x}_1(t), \mathbf{x}_2(t)\}$$

forms a fundamental set of solutions for $\mathbf{y}' = \mathbf{A}\mathbf{y}$.

18 Tutorial 5

18.1 Question 1a(i)

To check whether $span\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\} \subseteq span\{\mathbf{v}_1, \mathbf{v}_2\}$, we will check the consistency of this augmented matrix.

$$\left(\begin{array}{ccc|cc} 2 & -1 & 0 & 1 & 0 \\ -2 & 1 & 0 & -1 & 1 \\ 0 & -1 & 9 & -5 & 1 \end{array}\right) \xrightarrow{RREF} \left(\begin{array}{ccc|cc} 1 & 0 & -\frac{9}{2} & 3 & 0 \\ 0 & 1 & -9 & 5 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{array}\right)$$

This augmented matrix is indeed inconsistent; therefore, $span\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\} \not\subseteq span\{\mathbf{v}_1, \mathbf{v}_2\}$.

Next, to check whether $span\{\mathbf{v}_1, \mathbf{v}_2\} \subseteq span\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$, we will check the consistency of this augmented matrix.

$$\left(\begin{array}{cc|ccc} 1 & 0 & 2 & -1 & 0 \\ -1 & 1 & -2 & 1 & 0 \\ -5 & 1 & 0 & -1 & 9 \end{array}\right) \xrightarrow{RREF} \left(\begin{array}{cc|ccc} 1 & 0 & 0 & \frac{1}{5} & -\frac{9}{5} \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & -\frac{3}{5} & \frac{9}{10} \end{array}\right)$$

This augmented matrix is indeed inconsistent; therefore, $span\{\mathbf{v}_1, \mathbf{v}_2\} \not\subseteq span\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$.

18.2 Question 1b(i)

To describe the span geometrically, let's observe the RREF of this matrix ($span\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$).

$$\begin{pmatrix} 2 & -1 & 0 \\ -2 & 1 & 0 \\ 0 & -1 & 9 \end{pmatrix} \xrightarrow{RREF} \begin{pmatrix} 1 & 0 & -\frac{9}{2} \\ 0 & 1 & -9 \\ 0 & 0 & 0 \end{pmatrix}$$

From the RREF, it can be observed that $\mathbf{u}_3$ can be obtained from $\mathbf{u}_1$ and $\mathbf{u}_2$, meaning that $span\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\} = span\{\mathbf{u}_1, \mathbf{u}_2\}$. As the span of $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$ do not span the entire $\mathbb{R}^3$, it is a plane.

To find the equation of the plane, treat $\mathbf{u}_1, \mathbf{u}_2$ (we ignore $\mathbf{u}_3$ as it is dependent on $\mathbf{u}_1$ and $\mathbf{u}_2$) as a point where you will substitute to the general equation of the plane. Observe that the entire space is $\mathbb{R}^3$ and we are describing a subspace. Therefore, the general equation of the plane is $ax+by+cz = 0$. We do not add $+d$ into the general equation because span is a subspace, not an affine space.

Let's plug in these two points: $\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2 \\ -2 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \\ -1 \end{pmatrix}$.

We will obtain this linear system:

$$\begin{cases} 2a - 2b = 0 \\ -a + b - c = 0 \end{cases}$$

Let's represent this system as an augmented matrix.

$$\left(\begin{array}{ccc|c} 2 & -2 & 0 & 0 \\ -1 & 1 & -1 & 0 \end{array}\right) \xrightarrow{RREF} \left(\begin{array}{ccc|c} 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right)$$

Assume $b = p$. By observing the augmented matrix, we can obtain $a = p$ and $c = 0$. Therefore, we can formulate the equation of the plane to be $px + py = 0 \rightarrow x + y = 0$.

Next, to describe $span\{\mathbf{v}_1, \mathbf{v}_2\}$, let's observe the RREF of this matrix.

$$\begin{pmatrix} 1 & 0 \\ -1 & 1 \\ -5 & 1 \end{pmatrix} \xrightarrow{RREF} \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}$$

As the span of $\mathbf{v}_1$ and $\mathbf{v}_2$ do not span the entire $\mathbb{R}^3$, it is also a plane. Let's plug in these two points: $\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ -5 \end{pmatrix}$ and $\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$. We will obtain this linear system:

$$\begin{cases} a - b - 5c = 0 \\ b + c = 0 \end{cases}$$

Let's represent this system as an augmented matrix.

$$\left(\begin{array}{ccc|c} 1 & -1 & -5 & 0 \\ 0 & 1 & 1 & 0 \end{array}\right) \xrightarrow{RREF} \left(\begin{array}{ccc|c} 1 & 0 & -4 & 0 \\ 0 & 1 & 1 & 0 \end{array}\right)$$

Assume $c = p$. By observing the augmented matrix, we can obtain $a = 4p$ and $b = -p$. Therefore, we can formulate the equation of the plane to be

$$4px - py + pz = 0 \rightarrow 4x - y + z = 0$$

18.3 Question 2a

$$S = \left\{ \begin{pmatrix} p \\ q \\ p \\ q \end{pmatrix} \middle| \; p, q \in \mathbb{R} \right\} = \left\{ p \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} + q \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix} \middle| \; p, q \in \mathbb{R} \right\}$$

S is indeed a subspace with a spanning set (basis) of $\left\{ \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix} \right\}$.

18.4 Question 3a

To check for linear independence, let's observe the RREF of this matrix.

$$\begin{pmatrix} 2 & 0 & 2 & 3 \\ -1 & 3 & 4 & 6 \\ 0 & 2 & 3 & 6 \end{pmatrix} \xrightarrow{RREF} \begin{pmatrix} 1 & 0 & 0 & \frac{9}{2} \\ 0 & 1 & 0 & \frac{15}{2} \\ 0 & 0 & 1 & -3 \end{pmatrix}$$

Observe that the fourth column is a non-pivot column. Therefore, this set is not linearly independent. We can represent the fourth column as the linear combination of the first third columns as follows:

$$\begin{pmatrix} 3 \\ 6 \\ 6 \end{pmatrix} = \frac{9}{2} \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix} + \frac{15}{2} \begin{pmatrix} 0 \\ 3 \\ 2 \end{pmatrix} - 3 \begin{pmatrix} 2 \\ 4 \\ 3 \end{pmatrix}$$

with each entry in the fourth column as the coefficient for the linear combination.

18.5 Question 4a

$$V = \left\{ \begin{pmatrix} a+b \\ a+c \\ c+d \\ b+d \end{pmatrix} \middle| \; a, b, c, d \in \mathbb{R} \right\} = \left\{ a \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + b \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} + c \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} + d \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} \middle| \; a, b, c, d \in \mathbb{R} \right\}$$

In other words, $V = span\left\{\begin{pmatrix}1\\1\\0\\0\end{pmatrix}, \begin{pmatrix}1\\0\\0\\1\end{pmatrix}, \begin{pmatrix}0\\1\\1\\0\end{pmatrix}, \begin{pmatrix}0\\0\\1\\1\end{pmatrix}\right\}$.

Next, let's check the linear independence of the spanning set, and remove the redundant vectors.

$$\begin{pmatrix}1 & 1 & 0 & 0\\1 & 0 & 1 & 0\\0 & 0 & 1 & 1\\0 & 1 & 0 & 1\end{pmatrix} \overset{RREF}{\longrightarrow} \begin{pmatrix}1 & 0 & 0 & -1\\0 & 1 & 0 & 1\\0 & 0 & 1 & 1\\0 & 0 & 0 & 0\end{pmatrix}$$

The fourth vector is redundant. Throwing the fourth vector away, we can get the basis of V to be

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

18.6 Question 4b

Let's check the linear independence of the spanning set, and remove the redundant vectors. We can check this by observing this augmented matrix.

$$\begin{pmatrix}1 & -1 & 0 & 1\\0 & 2 & 3 & -1\\-1 & 3 & 0 & 1\end{pmatrix} \overset{RREF}{\longrightarrow} \begin{pmatrix}1 & 0 & 0 & 2\\0 & 1 & 0 & 1\\0 & 0 & 1 & -1\end{pmatrix}$$

The fourth vector is dependent on the first three; we can remove this. Therefore, the basis for this set is

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

18.7 Question 4c

Let's represent this linear system as an augmented matrix.

$$\left(\begin{array}{ccccc|c}1 & 0 & 1 & 1 & -1 & 0\\0 & 1 & 1 & 2 & 1 & 0\\1 & 1 & 2 & 1 & -2 & 0\end{array}\right) \overset{RREF}{\longrightarrow} \left(\begin{array}{ccccc|c}1 & 0 & 1 & 0 & -2 & 0\\0 & 1 & 1 & 0 & -1 & 0\\0 & 0 & 0 & 1 & 1 & 0\end{array}\right)$$

Assume $a_3 = p$ and $a_5 = q$. From the augmented matrix, we can obtain $a_1 = 2q - p$, $a_2 = q - p$, and $a_4 = -q$. By this, we can construct:

$$\begin{pmatrix}a_1\\a_2\\a_3\\a_4\\a_5\end{pmatrix} = \begin{pmatrix}2q-p\\q-p\\p\\-q\\q\end{pmatrix} = p\begin{pmatrix}-1\\-1\\1\\0\\0\end{pmatrix} + q\begin{pmatrix}2\\1\\0\\-1\\1\end{pmatrix}$$

By this, we will have

$$V = span\left\{\begin{pmatrix}-1\\-1\\1\\0\\0\end{pmatrix}, \begin{pmatrix}2\\1\\0\\-1\\1\end{pmatrix}\right\}$$

Because these two vectors are not a multiple of each other, they are linearly independent. Therefore, the basis of V is

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

18.8 Question 5a

To express $\mathbf{u}$ as a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4$, let's find c_1, c_2, c_3, c_4 such that

$$c_1\begin{pmatrix}3\\1\\1\\1\end{pmatrix} + c_2\begin{pmatrix}1\\2\\1\\1\end{pmatrix} + c_3\begin{pmatrix}1\\1\\0\\1\end{pmatrix} + c_4\begin{pmatrix}4\\3\\2\\2\end{pmatrix} = \begin{pmatrix}4\\4\\3\\2\end{pmatrix}$$

Express this as an augmented matrix.

$$\left(\begin{array}{cccc|c} 3 & 1 & 1 & 4 & 4\\ 1 & 2 & 1 & 3 & 4\\ 1 & 1 & 0 & 2 & 3\\ 1 & 1 & 1 & 2 & 2 \end{array}\right) \xrightarrow{RREF} \left(\begin{array}{cccc|c} 1 & 0 & 0 & 1 & 1\\ 0 & 1 & 0 & 1 & 2\\ 0 & 0 & 1 & 0 & -1\\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

Assume $c_4 = p$. From the augmented matrix, we can obtain $c_1 = 1-p$, $c_2 = 2-p$, $c_3 = -1$.

One way to express $\mathbf{u}$ as the linear combination of $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4$ is to choose $p = 1$. Then,

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

The other way is to choose $p = 0$. We can then obtain

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

18.9 Question 5b

Assume that $\mathbf{u}$ can be expressed as a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ in these two ways:

$$\mathbf{u} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3 \quad \cdots(1)$$
$$\mathbf{u} = d_1\mathbf{v}_1 + d_2\mathbf{v}_2 + d_3\mathbf{v}_3 \quad \cdots(2)$$

Let's observe $\cdots(1) - \cdots(2)$!

$$(c_1 - d_1)\mathbf{v}_1 + (c_2 - d_2)\mathbf{v}_2 + (c_3 - d_3)\mathbf{v}_3 = 0$$

Observe that we can express $\mathbf{u}$ as a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ in two distinct ways if there exists a non trivial solution for the equation above, i.e., $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ are linearly dependent.

By this observation, let's check the linear independence of $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ by observing at the RREF of this matrix.

$$\begin{pmatrix} 3 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix} \xrightarrow{RREF} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

$\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ is indeed linearly independent as there is no non pivot column in the RREF. Because they are linearly independent, we cannot express $\mathbf{u}$ as a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ in two distinct ways.

19 Tutorial 6

19.1 Question 1a

Recall that $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$ will form a basis if and only if the matrix below

$$\begin{pmatrix} a & -1 & 1 \\ 1 & a & -1 \\ -1 & 1 & a \end{pmatrix}$$

has no nonzero rows.

```
syms a

A = [a -1 1; 1 a -1; -1 1 a;];
A([1,2], :) = A([2,1], :);
A(2,:) = A(2,:) - a * A(1,:);
A(3,:) = A(3,:) + A(1,:);
A(3,:) = A(3,:) + (a+1)/(a^2+1) * A(2,:);
disp(simplify(A));
```

Observe that in the elementary row operations above, we did not require any constraint of a. That is, for any value of a we substitute in, the elementary row operations are valid. This is crucial!

Perform the Matlab code above, and we can simplify the matrix to be

$$\begin{pmatrix} 1 & a & -1 \\ 0 & -a^2-1 & a+1 \\ 0 & 0 & \frac{a(a^2+3)}{a^2+1} \end{pmatrix}$$

Now, so that the matrix does not have any zero row, $-a^2-1 \neq 0$, which is always the case, and $\frac{a(a^2+3)}{a^2+1} \neq 0 \rightarrow a \neq 0$.

19.2 Question 1b

$$\begin{aligned}
\begin{vmatrix} a & -1 & 1 \\ 1 & a & -1 \\ -1 & 1 & a \end{vmatrix} &= a\begin{vmatrix} a & -1 \\ 1 & a \end{vmatrix} - \begin{vmatrix} -1 & 1 \\ 1 & a \end{vmatrix} - \begin{vmatrix} -1 & 1 \\ a & -1 \end{vmatrix} \\
&= a(a^2+1) - (-a-1) - (1-a) \\
&= a^3 + a + a + 1 - 1 + a \\
&= a^3 + 3a \\
&= a(a^2+3)
\end{aligned}$$

The determinant will be nonzero for any value $a \in \mathbb{R}$ where $a \neq 0$.

19.3 Question 2a

To find the relative basis, let's solve this equation:

$$c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + c_3\mathbf{u}_3 = \begin{pmatrix} 1 \\ -2 \\ 6 \end{pmatrix}$$

Plugging in the value of $\mathbf{u}_1, \mathbf{u}_2$, and $\mathbf{u}_3$, the equation will become:

$$c_1\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + c_2\begin{pmatrix} 0 \\ 2 \\ 2 \end{pmatrix} + c_3\begin{pmatrix} 0 \\ 0 \\ 3 \end{pmatrix} = \begin{pmatrix} 1 \\ -2 \\ 6 \end{pmatrix}$$

Then, express this as an augmented matrix!

$$\left(\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 1 & 2 & 0 & -2 \\ 1 & 2 & 3 & 6 \end{array}\right) \xrightarrow{RREF} \left(\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & -\frac{3}{2} \\ 0 & 0 & 1 & \frac{8}{3} \end{array}\right)$$

By this, we will have

$$[\mathbf{v}]_S = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} 1 \\ -\frac{3}{2} \\ \frac{8}{3} \end{pmatrix}$$

19.4 Question 4a

19.4.1 Claim

$U \cup V$ is a subspace if and only if $U \subseteq V$ or $V \subseteq U$. This is because if $U \not\subseteq V$ and $V \not\subseteq U$, we can find $\mathbf{u} \in U$ where $\mathbf{u} \notin V$ and $\mathbf{v} \in V$ where $\mathbf{v} \notin U$ such that $\mathbf{u} + \mathbf{v} \notin U \cup V$.

19.4.2 Solution

By the claim above, we will first check whether $U \subseteq V$ by observing the consistency of the augmented matrix below.

$$\left(\begin{array}{cc|cc} 1 & 1 & 1 & 1 \\ 1 & 2 & 0 & 0 \\ 1 & 2 & 1 & 2 \\ 1 & 1 & 0 & -1 \end{array}\right) \xrightarrow{RREF} \left(\begin{array}{cc|cc} 1 & 0 & 0 & -2 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

The augmented matrix is indeed inconsistent. Therefore, $U \not\subseteq V$.

Next, let's check whether $V \subseteq U$ or not. Let's observe the augmented matrix below.

$$\left(\begin{array}{cc|cc} 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 2 \\ 1 & 2 & 1 & 2 \\ 0 & -1 & 1 & 1 \end{array}\right) \xrightarrow{RREF} \left(\begin{array}{cc|cc} 1 & 0 & 0 & -2 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

This augmented matrix is also inconsistent. Therefore, $U \not\subseteq V$.

Because $U \not\subseteq V$ and $U \not\subseteq V$, $U \cup V$ is not a subspace of $\mathbb{R}^4$.

19.5 Question 4b

We know that by definition

$$U + V = \{\mathbf{u} + \mathbf{v} \mid \mathbf{u} \in U \text{ and } \mathbf{v} \in V\}$$

Then, because

$$U = span \left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 2 \\ 1 \end{pmatrix} \right\}$$

we can represent any vector $\mathbf{u} \in U$ as

$$\mathbf{u} = a \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} + b \begin{pmatrix} 1 \\ 2 \\ 2 \\ 1 \end{pmatrix}$$

Also, because

$$V = span \left\{ \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 2 \\ -1 \end{pmatrix} \right\}$$

we can represent any vector $\mathbf{v} \in V$ as

$$\mathbf{v} = p \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} + q \begin{pmatrix} 1 \\ 0 \\ 2 \\ -1 \end{pmatrix}$$

Using this information, we can construct

$$\mathbf{u} + \mathbf{v} = a \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} + b \begin{pmatrix} 1 \\ 2 \\ 2 \\ 1 \end{pmatrix} + p \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} + q \begin{pmatrix} 1 \\ 0 \\ 2 \\ -1 \end{pmatrix}$$

By this, we can rewrite the solution set of $U + V$ as

$$\left\{ a \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} + b \begin{pmatrix} 1 \\ 2 \\ 2 \\ 1 \end{pmatrix} + p \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} + q \begin{pmatrix} 1 \\ 0 \\ 2 \\ -1 \end{pmatrix} \middle| \, a, b, p, q \in \mathbb{R} \right\}$$

Therefore, $U + V$ is indeed a subspace as it can be written as a span of set.

Next, to check the dimension, we will need to find the basis of this subspace first. Observe the RREF of this matrix!

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 0 & 0 \\ 1 & 2 & 1 & 2 \\ 1 & 1 & 0 & -1 \end{pmatrix} \xrightarrow{RREF} \begin{pmatrix} 1 & 0 & 0 & -2 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Notice that the fourth vector is redundant. The basis of the subspace is

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

The dimension is indeed 3.

19.6 Question 4c

From Question 4b, we know that the solution set of $U + V$ as

$$\left\{ a \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} + b \begin{pmatrix} 1 \\ 2 \\ 2 \\ 1 \end{pmatrix} + p \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} + q \begin{pmatrix} 1 \\ 0 \\ 2 \\ -1 \end{pmatrix} \middle| \, a, b, p, q \in \mathbb{R} \right\}$$

In other words, $U+V = span\left\{\begin{pmatrix}1\\1\\1\\1\end{pmatrix}, \begin{pmatrix}1\\2\\2\\1\end{pmatrix}, \begin{pmatrix}1\\0\\1\\0\end{pmatrix}, \begin{pmatrix}1\\0\\2\\-1\end{pmatrix}\right\}$.

Observe that

$$U = span\left\{\begin{pmatrix}1\\1\\1\\1\end{pmatrix}, \begin{pmatrix}1\\2\\2\\1\end{pmatrix}\right\} \subseteq span\left\{\begin{pmatrix}1\\1\\1\\1\end{pmatrix}, \begin{pmatrix}1\\2\\2\\1\end{pmatrix}, \begin{pmatrix}1\\0\\1\\0\end{pmatrix}, \begin{pmatrix}1\\0\\2\\-1\end{pmatrix}\right\} = U+V$$

$$V = span\left\{\begin{pmatrix}1\\0\\1\\0\end{pmatrix}, \begin{pmatrix}1\\0\\2\\-1\end{pmatrix}\right\} \subseteq span\left\{\begin{pmatrix}1\\1\\1\\1\end{pmatrix}, \begin{pmatrix}1\\2\\2\\1\end{pmatrix}, \begin{pmatrix}1\\0\\1\\0\end{pmatrix}, \begin{pmatrix}1\\0\\2\\-1\end{pmatrix}\right\} = U+V$$

Therefore, $U+V$ indeed contains U and V.

19.7 Question 4d

The two vectors in each U and V are linearly independent because they are not a multiple of each other. Therefore, $\dim(U) = \dim(V) = 2$.

19.8 Question 4e

Let $\mathbf{x}$ be a vector such that $\mathbf{x} \in U \cap V$. This means $\mathbf{x} \in U$ and $\mathbf{x} \in V$. Therefore we can express $\mathbf{x}$ as follows.

$$\mathbf{x} = a\begin{pmatrix}1\\1\\1\\1\end{pmatrix} + b\begin{pmatrix}1\\2\\2\\1\end{pmatrix} = p\begin{pmatrix}1\\0\\1\\0\end{pmatrix} + q\begin{pmatrix}1\\0\\2\\-1\end{pmatrix}$$

This means we are solving the equation below

$$a\begin{pmatrix}1\\1\\1\\1\end{pmatrix} + b\begin{pmatrix}1\\2\\2\\1\end{pmatrix} - p\begin{pmatrix}1\\0\\1\\0\end{pmatrix} - q\begin{pmatrix}1\\0\\2\\-1\end{pmatrix} = 0$$

Let's express this as an augmented matrix.

$$\left(\begin{array}{cccc|c} 1 & 1 & -1 & -1 & 0\\ 1 & 2 & 0 & 0 & 0\\ 1 & 2 & -1 & -2 & 0\\ 1 & 1 & 0 & 1 & 0 \end{array}\right) \xrightarrow{RREF} \left(\begin{array}{cccc|c} 1 & 0 & 0 & 2 & 0\\ 0 & 1 & 0 & -1 & 0\\ 0 & 0 & 1 & 2 & 0\\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

Assume $q = s$. By this augmented matrix, we will have $a = -2s$, $b = s$, $c = -2s$.

Now, pick the equation:

$$\mathbf{x} = a\begin{pmatrix}1\\1\\1\\1\end{pmatrix} + b\begin{pmatrix}1\\2\\2\\1\end{pmatrix} = -2s\begin{pmatrix}1\\1\\1\\1\end{pmatrix} + s\begin{pmatrix}1\\2\\2\\1\end{pmatrix} = \begin{pmatrix}-2s\\-2s\\-2s\\-2s\end{pmatrix} + \begin{pmatrix}s\\2s\\2s\\s\end{pmatrix} = \begin{pmatrix}-s\\0\\0\\-s\end{pmatrix} = s\begin{pmatrix}-1\\0\\0\\-1\end{pmatrix}$$

By this, we can have the solution set of $U \cap V$ as

$$\left\{ s\begin{pmatrix}-1\\0\\0\\-1\end{pmatrix} \;\middle|\; s \in \mathbb{R} \right\}$$

Therefore, $U \cap V$ is indeed a subspace as it can be written as a span of set. The dimension is indeed 1.

19.9 Question 6a

Let $[\mathbf{u}]_S = \begin{pmatrix}c_1\\c_2\\\vdots\\c_n\end{pmatrix}$ and $[\mathbf{v}]_S = \begin{pmatrix}d_1\\d_2\\\vdots\\d_n\end{pmatrix}$.

By the definition of relative basis, we will have

$$\begin{aligned}\mathbf{u} &= c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n \quad \cdots(1)\\ \mathbf{v} &= d_1\mathbf{v}_1 + d_2\mathbf{v}_2 + \cdots + d_n\mathbf{v}_n \quad \cdots(2)\end{aligned}$$

Adding these two equations up, i.e., $\mathbf{u} + \mathbf{v}$, we will have

$$\begin{aligned}\mathbf{u} + \mathbf{v} &= c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n + d_1\mathbf{v}_1 + d_2\mathbf{v}_2 + \cdots + d_n\mathbf{v}_n\\ \mathbf{u} + \mathbf{v} &= (c_1 + d_1)\mathbf{v}_1 + (c_2 + d_2)\mathbf{v}_2 + \cdots + (c_n + d_n)\mathbf{v}_n\end{aligned}$$

By the definition of relative basis, we can express $[\mathbf{u} + \mathbf{v}]_S$ as

$$\begin{pmatrix}c_1 + d_1\\c_2 + d_2\\\vdots\\c_n + d_n\end{pmatrix}, \text{which is equivalent to} \begin{pmatrix}c_1\\c_2\\\vdots\\c_n\end{pmatrix} + \begin{pmatrix}d_1\\d_2\\\vdots\\d_n\end{pmatrix}$$

19.10 Question 6b

Let $[\mathbf{u}]_S = \begin{pmatrix}c_1\\c_2\\\vdots\\c_n\end{pmatrix}$.

By the definition of relative basis, we will have

$$\mathbf{u} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n$$

Let's multiply both sides with c! We will get

$$c\mathbf{u} = cc_1\mathbf{v}_1 + cc_2\mathbf{v}_2 + \cdots + cc_n\mathbf{v}_n$$

By the definition of relative basis, we can express

$$[c\mathbf{u}]_S = \begin{pmatrix} cc_1 \\ cc_2 \\ \vdots \\ cc_n \end{pmatrix} = c\begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix}$$

20 Tutorial 8

20.1 Question 3

We will use Matlab to help us in our calculation. Therefore, in top of the program, declare these:

```
u1 = sym([1; 2; 2; -1]);
u2 = sym([1; 1; -1; 1]);
u3 = sym([-1; 1; -1; -1]);
u4 = sym([-2; 1; 1; 2]);
```

20.1.1 Part a

Then, for part a, we will use the dot function in the Matlab to check for every pair.

```
% Part a
disp('Part a')
disp(dot(u1, u2));
disp(dot(u1, u3));
disp(dot(u1, u4));
disp(dot(u2, u3));
disp(dot(u2, u4));
disp(dot(u3, u4));
```

20.1.2 Part b

Then, to check whether S is basis for $\mathbb{R}^4$ or not, we will check for the existence of non pivot columns in the RREF of $\begin{pmatrix}\mathbf{u}_1 & \mathbf{u}_2 & \mathbf{u}_3 & \mathbf{u}_4\end{pmatrix}$. Find the Matlab code below.

```
% Part b
disp('Part b')
A = [u1 u2 u3 u4];
disp(rref(A))
```

Using the Matlab code below, we can get that the RREF of the matrix is the identity matrix. Therefore, S indeed is the basis for $\mathbb{R}^4$.

20.1.3 Part c

Because S is a basis, S is linearly independent to $\mathbb{R}^4$. When we add $\mathbf{w}$ to S, now the set contains 5 vectors. 5 vectors cannot be linearly independent to $\mathbb{R}^4$, so $S \cup \{\mathbf{w}\}$ is not linearly independent.

Recall Theorem 14.6.1, i.e., if S is an orthogonal set, S is linearly independent. Consider the inverse of this statement, if S is linearly dependent, S is not an orthogonal set. Because $S \cup \{\mathbf{w}\}$ is not linearly independent, it is not an orthogonal set.

20.1.4 Part d

Run the Matlab code below.

```
% Part d
disp('Part d')
u1n = u1 / norm(u1);
u2n = u2 / norm(u2);
u3n = u3 / norm(u3);
u4n = u4 / norm(u4);
disp('Normalized u1');
disp(u1n);
disp('Normalized u2');
disp(u2n);
disp('Normalized u3');
disp(u3n);
disp('Normalized u4');
disp(u4n);
```

We will get

$$\mathbf{u}_1' = \begin{pmatrix} \frac{\sqrt{10}}{10} \\ \frac{\sqrt{10}}{5} \\ \frac{\sqrt{10}}{5} \\ -\frac{\sqrt{10}}{10} \end{pmatrix}, \mathbf{u}_2' = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ -\frac{1}{2} \\ \frac{1}{2} \end{pmatrix}, \mathbf{u}_3' = \begin{pmatrix} -\frac{1}{2} \\ \frac{1}{2} \\ -\frac{1}{2} \\ -\frac{1}{2} \end{pmatrix}, \mathbf{u}_4' = \begin{pmatrix} -\frac{\sqrt{10}}{5} \\ \frac{\sqrt{10}}{10} \\ \frac{\sqrt{10}}{10} \\ \frac{\sqrt{10}}{5} \end{pmatrix}$$

20.1.5 Part e

Use this code to find $[\mathbf{v}]_S$ and $[\mathbf{v}]_T$!

```
% Part e
disp('Part e')
v = [0; 1; 2; 3];
c1 = dot(u1, v) / norm(u1)^2;
c2 = dot(u2, v) / norm(u2)^2;
c3 = dot(u3, v) / norm(u3)^2;
c4 = dot(u4, v) / norm(u4)^2;
vs = [c1; c2; c3; c4];
disp('Vs');
disp(vs);

d1 = dot(u1n, v) / norm(u1n)^2;
d2 = dot(u2n, v) / norm(u2n)^2;
d3 = dot(u3n, v) / norm(u3n)^2;
d4 = dot(u4n, v) / norm(u4n)^2;
vt = [d1; d2; d3; d4];
disp('Vt');
disp(vt);
```

20.1.6 Part f

Recall that $S = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4\}$ and $T = \{\mathbf{u}'_1, \mathbf{u}'_2, \mathbf{u}'_3, \mathbf{u}'_4\}$.

Based on the formula in Chapter 14.7, we can obtain the coordinate relative to the orthogonal and orthonormal basis as follows:

$$[\mathbf{v}]_S = \begin{pmatrix} \frac{\mathbf{u}_1 \cdot \mathbf{v}}{||\mathbf{u}_1||^2} \\ \frac{\mathbf{u}_2 \cdot \mathbf{v}}{||\mathbf{u}_2||^2} \\ \frac{\mathbf{u}_3 \cdot \mathbf{v}}{||\mathbf{u}_3||^2} \\ \frac{\mathbf{u}_4 \cdot \mathbf{v}}{||\mathbf{u}_4||^2} \end{pmatrix}, [\mathbf{v}]_T = \begin{pmatrix} \mathbf{u}'_1 \cdot \mathbf{v} \\ \mathbf{u}'_2 \cdot \mathbf{v} \\ \mathbf{u}'_3 \cdot \mathbf{v} \\ \mathbf{u}'_4 \cdot \mathbf{v} \end{pmatrix}$$

Notice that $\mathbf{u}'_i = \frac{\mathbf{u}_i}{||\mathbf{u}_i||}$ for $1 \leq i \leq 4$. Therefore,

$$[\mathbf{v}]_T = \begin{pmatrix} \frac{\mathbf{u}_1 \cdot \mathbf{v}}{||\mathbf{u}_1||} \\ \frac{\mathbf{u}_2 \cdot \mathbf{v}}{||\mathbf{u}_2||} \\ \frac{\mathbf{u}_3 \cdot \mathbf{v}}{||\mathbf{u}_3||} \\ \frac{\mathbf{u}_4 \cdot \mathbf{v}}{||\mathbf{u}_4||} \end{pmatrix}$$

Observe the relationship of each entry in $[\mathbf{v}]_S$ and $[\mathbf{v}]_T$. To obtain $[\mathbf{v}]_T$, you can multiply each entry of $[\mathbf{v}]_S$ with $||\mathbf{u}_i||$ where i is the row index of the entry.

Motivated by this idea, here is the Matlab code to solve for this question!

```
% Part f
disp('Part f');
ws = [1; 2; 1; 1];
```

```
norms = [norm(u1); norm(u2); norm(u3); norm(u4)];
wt = ws.*norms;
disp(wt);
```

20.2 Question 4a

To perform the Gram-Schmidt Process faster, run the Matlab code below.

```
u1 = sym([1; 1; 1; 1]);
u2 = sym([1; -1; 1; 0]);
u3 = sym([1; 1; -1; -1]);
u4 = sym([1; 2; 0; 1]);

v1 = u1;
v2 = u2 - (dot(v1, u2) / norm(v1)^2) * v1;
v3 = u3 - (dot(v1, u3) / norm(v1)^2) * v1 - (dot(v2, u3) / norm(v2)^2) *
    v2;
v4 = u4 - (dot(v1, u4) / norm(v1)^2) * v1 - (dot(v2, u4) / norm(v2)^2) *
    v2 - (dot(v3, u4) / norm(v3)^2) * v3;

disp("Orthogonal set");
disp(v1);
disp(v2);
disp(v3);
disp(v4);

disp("Orthonormal set");
disp(v1 / norm(v1));
disp(v2 / norm(v2));
disp(v3 / norm(v3));
disp(v4 / norm(v4));
```

20.3 Question 4b

Perform the Matlab code below.

```
u1 = sym([1; 2; 2; 1]);
u2 = sym([1; 2; 1; 0]);
u3 = sym([1; 0; 1; 0]);
u4 = sym([1; 0; 2; 1]);

v1 = u1;
v2 = u2 - (dot(v1, u2) / norm(v1)^2) * v1;
v3 = u3 - (dot(v1, u3) / norm(v1)^2) * v1 - (dot(v2, u3) / norm(v2)^2) *
    v2;
v4 = u4 - (dot(v1, u4) / norm(v1)^2) * v1 - (dot(v2, u4) / norm(v2)^2) *
    v2 - (dot(v3, u4) / norm(v3)^2) * v3;
```

```
disp("Orthogonal set");
disp(v1);
disp(v2);
disp(v3);
disp(v4);

disp("Orthonormal set");
disp(v1 / norm(v1));
disp(v2 / norm(v2));
disp(v3 / norm(v3));
disp(v4 / norm(v4));
```

Observe that the fourth vector in the orthogonal set is the zero vector, meaning that it is contained on the span of the first three vectors already. Therefore, this set is not the orthonormal basis.

20.4 Question 6a

An orthogonal matrix $\mathbf{A}$ of order n is always invertible. By the equivalent statements of invertibility, the columns of $\mathbf{A}$ always is a basis for $\mathbb{R}^n$. Now, we are left to prove that the columns of $\mathbf{A}$ is orthonormal.

By the definition of orthogonal matrix, $\mathbf{A}$ is an orthogonal matrix of order n if and only if $\mathbf{A}^T = \mathbf{A}^{-1}$. Post-multiply and pre-multiply both sides with $\mathbf{A}$, we will get that $\mathbf{A}\mathbf{A}^T = \mathbf{I}$ and $\mathbf{A}^T\mathbf{A} = \mathbf{I}$.

By block multiplication, let $\mathbf{A} = \begin{pmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{pmatrix}$, so $\mathbf{A}^T = \begin{pmatrix} -\mathbf{a}_1- \\ -\mathbf{a}_2- \\ \vdots \\ -\mathbf{a}_n- \end{pmatrix}$. Let's use the information that $\mathbf{A}^T\mathbf{A} = \mathbf{I}$.

$$\mathbf{A}^T\mathbf{A} = \mathbf{I}$$

$$\begin{pmatrix} -\mathbf{a}_1- \\ -\mathbf{a}_2- \\ \vdots \\ -\mathbf{a}_n- \end{pmatrix} \begin{pmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{pmatrix} = \mathbf{I}$$

$$\begin{pmatrix} \mathbf{a}_1 \cdot \mathbf{a}_1 & \mathbf{a}_1 \cdot \mathbf{a}_2 & \cdots & \mathbf{a}_1 \cdot \mathbf{a}_n \\ \mathbf{a}_2 \cdot \mathbf{a}_1 & \mathbf{a}_2 \cdot \mathbf{a}_2 & \cdots & \mathbf{a}_2 \cdot \mathbf{a}_n \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{a}_n \cdot \mathbf{a}_1 & \mathbf{a}_n \cdot \mathbf{a}_2 & \cdots & \mathbf{a}_n \cdot \mathbf{a}_n \end{pmatrix} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

Let's observe these matrices row by row. For row i where $1 \leq i \leq n$,

$$\begin{cases} \mathbf{a}_i \cdot \mathbf{a}_i = 1 \\ \mathbf{a}_i \cdot \mathbf{a}_j = 0, \text{ for } j \neq i \end{cases}$$

Construct the set $S = \{\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_n\}$! S is the set of columns of $\mathbf{A}$, and S is an orthonormal set.

21 Tutorial 9

21.1 Question 1a

To find the basis for the nullspace of $\mathbf{A}$, let's solve for the augmented matrix below.

$$\left(\begin{array}{ccccc|c} 1 & 1 & 2 & 1 & 0 & 0 \\ 1 & 1 & 1 & 2 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 0 \end{array}\right) \xrightarrow{RREF} \left(\begin{array}{ccccc|c} 1 & 0 & 0 & 1 & -1 & 0 \\ 0 & 1 & 0 & 2 & 1 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 \end{array}\right)$$

Assume the fourth column is a and the fifth column is b. Then, the solution for the augmented matrix will be:

$$\begin{pmatrix} -a+b \\ -2a-b \\ a \\ a \\ b \end{pmatrix} = a \begin{pmatrix} -1 \\ -2 \\ 1 \\ 1 \\ 0 \end{pmatrix} + b \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

By this, we will have the basis for the nullspace to be:

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

21.2 Question 1b

Recall Chapter 15.4.6! From Question 1a, we have already obtained S, the basis for W. Construct $\mathbf{N}$ whose columns are vectors in S. With this, the projection of $\mathbf{e}_i$ onto W, call it $\mathbf{p}_i$ can be obtained with this formula:

$$\mathbf{p}_i = \mathbf{N}(\mathbf{N}^T\mathbf{N})^{-1}\mathbf{N}^T\mathbf{e}_i$$

By block multiplication, construct the matrix $\mathbf{B} = \begin{pmatrix} \mathbf{p}_1 & \mathbf{p}_2 & \mathbf{p}_3 & \mathbf{p}_4 & \mathbf{p}_5 \end{pmatrix}$. By the formula above, we can obtain

$$\mathbf{B} = \mathbf{N}(\mathbf{N}^T\mathbf{N})^{-1}\mathbf{N}^T$$

Next, let's run this Matlab code below.

```
% Part b
disp('Part b')
N = sym([-1 1; -2 -1; 1 0; 1 0; 0 1]);
disp(N*inv(N'*N)*N')
```

By this, we will obtain

$$\mathbf{B} = \begin{pmatrix} \frac{3}{5} & 0 & -\frac{1}{5} & -\frac{1}{5} & \frac{2}{5} \\ 0 & \frac{3}{4} & -\frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} \\ -\frac{1}{5} & -\frac{1}{4} & \frac{3}{20} & \frac{3}{20} & -\frac{1}{20} \\ -\frac{1}{5} & -\frac{1}{4} & \frac{3}{20} & \frac{3}{20} & -\frac{1}{20} \\ \frac{2}{5} & -\frac{1}{4} & -\frac{1}{20} & -\frac{1}{20} & \frac{7}{20} \end{pmatrix}$$

Each of the column vector $\mathbf{b}_i$ represent the projection of $\mathbf{e}_i$ to W.

21.3 Question 1c

By using the same Theorem as Question 1b, we can obtain

$$\mathbf{x}_p = \mathbf{N}(\mathbf{N}^T\mathbf{N})^{-1}\mathbf{N}^T\mathbf{x}$$

Use this Matlab code to obtain the answer:

```
% Part c
disp('Part c')
syms x1 x2 x3 x4 x5
A = N*inv(N'*N)*N';
x = [x1; x2; x3; x4; x5];
disp(simplify(A * x));
```

21.4 Question 2a

Observe that we can reconstruct $\mathbf{y} = a_1\mathbf{x} + a_0$ as a matrix equation below:

$$\begin{pmatrix} 1 & \mathbf{x}_1 \\ 1 & \mathbf{x}_2 \\ \vdots & \vdots \\ 1 & \mathbf{x}_m \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \end{pmatrix} = \begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_m \end{pmatrix}$$

Let's plug in the equation with the given data!

$$\begin{pmatrix} 1 & 0.01 \\ 1 & 0.012 \\ 1 & 0.015 \\ 1 & 0.02 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \end{pmatrix} = \begin{pmatrix} 2.75 \times 10^{-4} \\ 3.31 \times 10^{-4} \\ 3.92 \times 10^{-4} \\ 4.95 \times 10^{-4} \end{pmatrix}$$

For question a, we are asked to find the value of a_0 and a_1 that satisfy the equation. So, let's represent this equation as augmented matrix.

$$\left(\begin{array}{cc|c} 1 & 0.01 & 2.75 \times 10^{-4} \\ 1 & 0.012 & 3.31 \times 10^{-4} \\ 1 & 0.015 & 3.92 \times 10^{-4} \\ 1 & 0.02 & 4.95 \times 10^{-4} \end{array}\right)$$

Next, let's find the RREF of this augmented matrix using this Matlab code.

```
% Part a
disp("Part a");
C = sym([1 0.01; 1 0.012; 1 0.015; 1 0.02]);
V = sym([2.75 * 10^(-4); 3.31 * 10^(-4); 3.92 * 10^(-4); 4.95 * 10^(-4)])
    ;
A = [C V];
disp(rref(A));
```

The matrix is inconsistent, we cannot find a_0 and a_1 that satisfy the points.

21.5 Question 2b

Let $\mathbf{A}$ be the coefficient matrix $\begin{pmatrix} 1 & 0.01 \\ 1 & 0.012 \\ 1 & 0.015 \\ 1 & 0.02 \end{pmatrix}$ and $\mathbf{b}$ be the vector that we want to project, which is $\begin{pmatrix} 2.75 \times 10^{-4} \\ 3.31 \times 10^{-4} \\ 3.92 \times 10^{-4} \\ 4.95 \times 10^{-4} \end{pmatrix}$.

Using this information, let's find the least square solution to $\mathbf{Ax} = \mathbf{b}$ using the formula in Theorem 15.4.4! We will use this Matlab code below to help us compute!

```
% Part b
disp("Part b")
AUG = [A'*A A'*b];
disp(round(rref(AUG), 8))
```

By this, we can obtain the RREF of the matrix to be

$$\left(\begin{array}{cc|c} 1 & 0 & 0.00007 \\ 0 & 1 & 0.21617 \end{array}\right)$$

Therefore, the least square solution for a_0 and a_1 is 0.00007 and 0.21617, respectively.

21.6 Question 4

Run the Matlab code below!

```
A1 = [1 -3 3;
      3 -5 3;
      6 -6 4];

A2 = [9 8 6 3;
      0 -1 3 -4;
      0 0 3 0;
```

```
     0 0 0 2];

% Calculate eigenvalues
eig_A1 = eig(A1);
eig_A2 = eig(A2);

% Display eigenvalues
disp('Eigenvalues of matrix (a):');
disp(eig_A1);

disp('Eigenvalues of matrix (b):');
disp(eig_A2);
```