MATH1030: Linear Algebra I

onenylxus/math-notes

August 31, 2020

Remarks

(1) Context of this document is based on university course *MATH1030: Linear Algebra I* from *Department of Mathematics, The Chinese University of Hong Kong (CUHK)*. The original source can be found at `https://www.math.cuhk.edu.hk/course`. The author does not own the source.

(2) This document is assumed unavailable for unauthorized parties that have not attended the university course. It is prohibited to share, including distributing or copying this document to unauthorized parties in any means for any non-academic purpose.

(3) Context of this doucment may not be completely accurate. The author assumes no responsibility or liability for any errors or omissions in the context of this document.

(4) This document is under license CC-BY-SA 4.0. It is allowed to make any editions on this document, as long as terms of the license is not violated.

(5) Starting from academic year 2020-2021, a series of Honours courses (with coursecode ending with 8 instead of 0) are introduced. However, Honours courses are assumed to have equivalent contents to their corresponding courses in the past. Hence, current contents will not be modified unless significant difference is found. For more details to Honours courses, please visit `https://www.math.cuhk.edu.hk/undergraduates/honours-courses/overview-honours-courses`.

Source

The latest version of this document can be found at `https://www.github.com/onenylxus/math-notes`.

Contents

1 Systems of Linear Equations **5**
1.1 Technique of Solving Linear Equations . . . 5
1.1.1 System of Linear Equations with Two Unknowns . . . 5
1.1.2 System of Linear Equations with Three Unknowns . . . 6
1.2 Geometric Interpretations . . . 8
1.2.1 Intersections of Lines . . . 8
1.2.2 Intersections of Planes . . . 9
1.2.3 Addition of Vectors . . . 10
1.3 Introduction to Systems of Linear Equations . . . 12
1.3.1 Definition of Systems of Linear Equations . . . 12
1.3.2 Possibilities of Solution Sets . . . 12
1.3.3 Equivalent Systems and Equation Operations . . . 13

2 Linear Algebra with Matrices **14**
2.1 Introduction to Matrices . . . 14
2.1.1 Matrix and Vector Notations . . . 14
2.1.2 Partitions of Matrices . . . 15
2.1.3 Matrix Representations of Linear Systems . . . 15
2.1.4 Row Operations . . . 16
2.2 Matrix Operations . . . 17
2.2.1 Matrix Equality, Addition and Scalar Multiplication . . . 17
2.2.2 Transposes and Symmetric Matrices . . . 18
2.3 Matrix Multiplication . . . 19
2.3.1 Matrix-Vector Product . . . 19
2.3.2 Introduction to Matrix Multiplication . . . 20
2.3.3 Properties of Matrix Multiplication . . . 20
2.3.4 Row Operations and Matrix Multiplication . . . 21
2.4 Reduced Row Echelon Form . . . 21
2.4.1 Introduction to Reduced Row Echelon Form . . . 21
2.4.2 Transforming Matrices to Reduced Row Echelon Form . . . 22
2.4.3 Uniqueness of Reduced Row Echelon Form . . . 24
2.5 Types of Solution Sets . . . 24
2.5.1 Consistent Systems . . . 24
2.5.2 Free and Dependent Variables . . . 24
2.5.3 Interpretations of Reduced Row Echelon Form . . . 25
2.6 Homogeneous Systems of Equations . . . 26
2.6.1 Definition of Homogeneous Systems . . . 26
2.6.2 Solutions of Homogeneous Systems . . . 27
2.6.3 Null Space of Matrices . . . 27
2.6.4 Augmented Matrices and Coefficient Matrices . . . 27
2.6.5 Nonsingular Matrices . . . 28
2.6.6 Null Space of Nonsingular Matrices . . . 28

3 Linear Algebra with Vectors **30**
3.1 Vector Space and Subspaces . . . 30
3.1.1 Definition of Vector Space . . . 30
3.1.2 Properties of Vector Space . . . 30
3.1.3 Abstract Definition of Vector Space . . . 31
3.1.4 Basic Properties of Vector Space . . . 32
3.1.5 Subspaces . . . 33
3.2 Linear Combinations . . . 33
3.2.1 Linear Combinations of Column Vectors . . . 33
3.2.2 Vector Form of Solution Sets . . . 34
3.3 Spanning Sets . . . 36
3.3.1 Spans of Sets of Vectors . . . 36
3.3.2 Spanning Sets of Null Spaces . . . 37
3.3.3 Reduction of Spans . . . 38
3.4 Linear Independence . . . 38
3.4.1 Linearly Independent Sets . . . 38
3.4.2 Linear Independence of Non-singular Matrices . . . 40
3.4.3 Linear Independence with Null Space and Spans . . . 40
3.5 Linear Dependence . . . 41
3.5.1 Dependency in Linearly Dependent Sets . . . 41
3.5.2 Elimination in Linearly Dependent Sets . . . 41
3.6 Column and Row Space . . . 43
3.6.1 Column Space . . . 43
3.6.2 Row Space . . . 45
3.7 Bases . . . 47
3.7.1 Definition of Bases . . . 47
3.7.2 Bases for Spans of Column Vectors . . . 48
3.7.3 Bases and Nonsingular Matrices . . . 49
3.7.4 Dimension . . . 49
3.7.5 Rank and Nullity . . . 50
3.8 Linear Transformation . . . 51
3.8.1 Matrices as Linear Transformation . . . 51
3.8.2 Geometric Interpretations . . . 52
3.8.3 Coordinate Vector . . . 53
3.8.4 Change of Basis . . . 54

4 Matrix Properties **56**
4.1 Inverses . . . 56
4.1.1 Solution Inverse . . . 56
4.1.2 Matrix Inverse . . . 56
4.1.3 Computation of Inverses . . . 56
4.1.4 Properties of Matrix Inverses . . . 58
4.1.5 Invertibility and Nonsingular Matrices . . . 59
4.2 Matrix Determinants . . . 60
4.2.1 Definition of Determinant . . . 60
4.2.2 Computation of Determinants . . . 62
4.2.3 Properties of Matrix Determinants . . . 63
4.3 Eigenvalues and Eigenvectors . . . 64

4.3.1 Eigenvalues and Eigenvectors of Matrices 64
4.3.2 Existence of Eigenvalues and Eigenvectors 64
4.3.3 Computation of Eigenvalues and Eigenvectors 65
4.3.4 Similar Matrices and Eigenvalues 67
4.3.5 Diagonalizability 68
4.3.6 Powers of Matrices 69
4.4 Inner Products and Orthogonaliztion 69
4.4.1 Inner Products 69
4.4.2 Norms 70
4.4.3 Orthogonal Sets 71
4.4.4 Orthogonal Bases 72
4.4.5 Gran-Schmidt Orthogonaliztion Process 73

References **74**

1 Systems of Linear Equations

1.1 Technique of Solving Linear Equations

1.1.1 System of Linear Equations with Two Unknowns

There are two common methods of solving a system of linear equations with two unknowns, which is **substitution** and **elimination**.

(a) **Substitution**

Example 1.1. The following is a system of linear equations with two unknowns:

$$\begin{cases} 3x + 4y = 2 \\ 4x + 5y = 3 \end{cases}$$

First, solve the second equation for y in terms of x:

$$y = \frac{3}{5} - \frac{4}{5}x$$

After that, substitute the above equation into the first equation, and the number of unknowns in the first equation will reduce to one:

$$\begin{aligned} 3x + 4\left(\frac{3}{5} - \frac{4}{5}x\right) &= 2 \\ \frac{12}{5} - \frac{x}{5} &= 2 \\ x &= 2 \end{aligned}$$

Hence the solution $x = 2$ is obtained, and finally solve for y:

$$y = \frac{3}{5} - \frac{4}{5} \times 2 = -1$$

Therefore the solution is $x = 2, y = 1$.

There are four ways of substitution to obtain the solution effectively:

(1) Solve x by the first equation in terms of y and substitute it into the second equation.

(2) Solve y by the first equation in terms of x and substitute it into the second equation.

(3) Solve x by the second equation in terms of y and substitute it into the first equation.

(4) Solve y by the second equation in terms of x and substitute it into the first equation.

However, the solution cannot be obtained if the substitution is performed at the same equation (for example, substituting the first equation back into the first equation).

(b) **Elimination**

Example 1.2. The following is the exact same system of linear equations with two unknowns as in the example above:

$$\begin{cases} 3x + 4y = 2 \\ 4x + 5y = 3 \end{cases}$$

This time the elimination is performed by combining both equations in order to eliminate the unknown x:

$$\begin{aligned} 3x + 4y &= 2 \\ 3x + \frac{15}{4}y &= \frac{9}{4} \end{aligned}$$

and by subtracting the first equation to the second equation,

$$\begin{aligned} \frac{1}{4}y &= -\frac{1}{4} \\ y &= -1 \end{aligned}$$

Hence the solution $y = -1$ is obtained, and finally solve for x:

$$\begin{aligned} 3x + 4(-1) &= 2 \\ x &= 2 \end{aligned}$$

Therefore the solution is $x = 2, y = 1$.

1.1.2 System of Linear Equations with Three Unknowns

Solving a system of linear equations with three unknowns may seem to be different to that with two unknowns, however the problem can be reduced to a previously solved problem by performing substitution or elimination. In this way, the problem will be familiar again.

Example 1.3. The following is a system of linear eqaution with three unknowns:

$$\begin{cases} x + 2y + 2z = 4 \\ x + 3y + 3z = 5 \\ 2x + 6y + 5z = 6 \end{cases}$$

The system is solved by the following sequence of equation operations:

(1) -1 times equation 1, add to equation 2:

$$\begin{aligned} 1x + 2y + 2z &= 4 \\ 0x + 1y + 1z &= 1 \\ 2x + 6y + 5z &= 6 \end{aligned}$$

(2) -2 times equation 1, add to equation 3:

$$\begin{aligned} 1x + 2y + 2z &= 4 \\ 0x + 1y + 1z &= 1 \\ 0x + 2y + 1z &= -2 \end{aligned}$$

(3) -2 times equation 2, add to eqaution 3:

$$\begin{aligned} 1x + 2y + 2z &= 4 \\ 0x + 1y + 1z &= 1 \\ 0x + 0y - 1z &= -4 \end{aligned}$$

(4) -1 times equation 3:

$$\begin{aligned} 1x + 2y + 2z &= 4 \\ 0x + 1y + 1z &= 1 \\ 0x + 0y + 1z &= 4 \end{aligned}$$

which can be rewritten as

$$\begin{cases} x + 2y + 2z = 4 \\ y + z = 1 \\ z = 4 \end{cases}$$

In fact, there are three types of system of linear equations, which is listed as follows:

(a) Some systems of linear equations have exactly one solution.

(b) Some systems of linear equations have no solution.

(c) Some systems of linear equations have infinitely many solutions which can be expressed in terms of variables.

1.2 Geometric Interpretations

1.2.1 Intersections of Lines

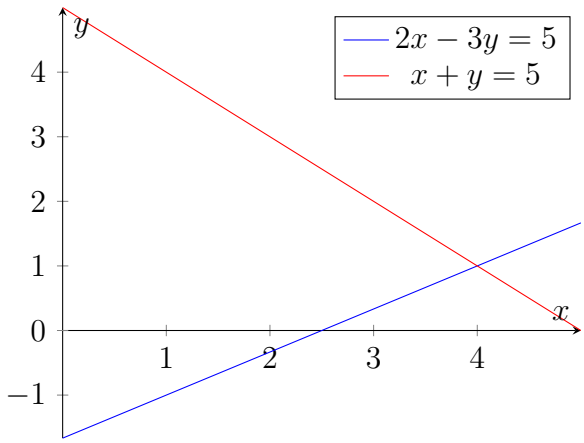
It is not the only way to solve systems of linear eqautions (where in this case, two unknowns and three unknowns) algebraically, but also by geometric interpretations.

Example 1.4. The following is a system of linear equations with two unknowns:

$$\begin{cases} 2x - 3y = -5 \\ x + y = 5 \end{cases}$$

By plotting the two equations as two different straight lines on the same graph (see *Figure 1.5*), where the green line is $2x - 3y = 5$ and the blue line is $x + y = 5$, a unique intersection $(2, 3)$ can be found, which is the solution.

Figure 1.5.



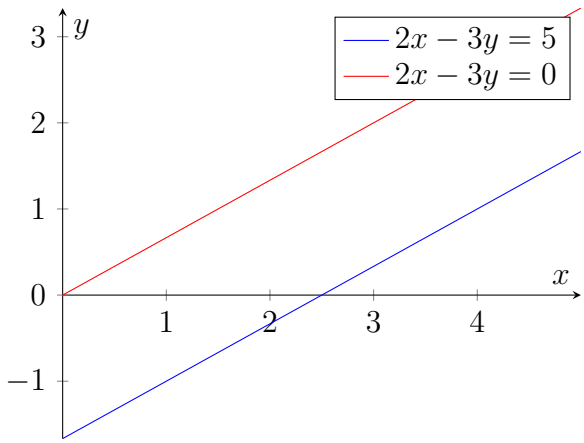
Below is another example where the equations have no solution:

Example 1.6. The following is a system of linear equations with two unknowns:

$$\begin{cases} 2x - 3y = -5 \\ 2x - 3y = 0 \end{cases}$$

By plotting the two equations as two different straight lines on the same graph (see *Figure 1.7*), where the green line is $2x - 3y = 5$ and the blue line is $2x - 3y = 0$, no solution can be found since they are parallel lines.

Figure 1.7.



Below is the third example where the equations have infinitely many solutions:

Example 1.8. The following is a system of linear equations with two unknowns:

$$\begin{cases} 2x - 3y = -5 \\ 2x - 3y = -5 \end{cases}$$

In this case, the system has infinite solutions: for any real number t, $(t, \frac{2t+5}{3})$ is a solution. In a geometrical perspective, since the two lines coincide, their intersection is the whole line $2x - 3y = -5$.

1.2.2 Intersections of Planes

Different to systems of linear variables of two unknowns which can be represented in 2-dimensional space, systems of linear variables of three unknowns have to be represented in 3-dimensional space, and each of the equations are represented by a plane.

Example 1.9. The following is a system of linear equations with three unknowns:

$$\begin{cases} 2x + 3y + 3z = 3 \\ 2x - y = 0 \end{cases}$$

The intersection of the two planes is a straight line, which can be expressed as

$$\left\{ \left(\frac{3 - 3t}{8}, \frac{3 - 3t}{4}, t \right) \mid t \in \mathbb{R} \right\}$$

Example 1.10. The following is a system of linear equations with three unknowns:

$$\begin{cases} 2x + 3y + 3z = 3 \\ 2x + 3y + 3z = 5 \end{cases}$$

Since the corresponding planes of the equations are parallel, the system has no solution.

Suppose there are three equations with three variables that corresponds to three different planes. There are five types of solutions to this situation:

(a) All three planes are parallel to each other. The system has no solution.

(b) Two of the three planes are parallel to each other and the other plane does not. The system has solution of two straight lines.

(c) Intersection of any two planes is a straight line. The three lines from the intersection of planes are parallel to each other. The system has solution of three straight lines.

(d) The intersection of three planes is a straight line. The system has solution of one straight line.

(e) The intersection of three planes is a point. The system has a unique solution.

1.2.3 Addition of Vectors

Beside from intersections of lines and planes, solving systems of equations with vectors is also a considerable choice.

Example 1.11. The following is a system of linear equations with two unknowns, written in a single vector form equation:

$$x\begin{bmatrix}2\\1\end{bmatrix}+y\begin{bmatrix}-3\\1\end{bmatrix}=\begin{bmatrix}-5\\5\end{bmatrix}=b$$

The problem now becomes finding the combination of the column vectors on the left side that produces the vector b on the right side.

Here are the system of linear equations used in examples in previous section:

Example 1.12. The following is a system of linear equations with two unknowns, written in a single vector form equation:

$$x\begin{bmatrix}2\\2\end{bmatrix}+y\begin{bmatrix}-3\\-3\end{bmatrix}=\begin{bmatrix}-5\\0\end{bmatrix}$$

$$x\begin{bmatrix}2\\2\end{bmatrix}+y\begin{bmatrix}-3\\-3\end{bmatrix}=\begin{bmatrix}-5\\-5\end{bmatrix}$$

In both of the examples, the two vectors on the left side are colinear. For the first example, since the target vector is not colinear with the two vectors, there are no

solutions for the system. For the second example, since the target vector is colinear with the two vectors, there are infinite solutions for the system.

In fact, this method can also be applied to 3-dimensional cases:

Example 1.13. The following is a system of linear equations with three unknowns, written in a single vector form equation:

$$x \begin{bmatrix} 2 \\ 4 \\ -2 \end{bmatrix} + y \begin{bmatrix} 1 \\ -6 \\ 7 \end{bmatrix} + z \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} = \begin{bmatrix} 5 \\ -2 \\ 9 \end{bmatrix}$$

By finding the combination of column vectors on the left side, the solution set for the equation above is

$$\begin{cases} x = 1 \\ y = 1 \\ z = 2 \end{cases}$$

Example 1.14. In the following example, the three vectors are coplanar, which means they all lies on the same plane:

$$x \begin{bmatrix} \frac{3}{2} \\ -1 \\ 0 \end{bmatrix} + y \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} + z \begin{bmatrix} \frac{3}{2} \\ 0 \\ -1 \end{bmatrix} = b$$

Note that the three vectors lies on the plane $2x + 3y + 3z = 0$. Similar to colinear cases in 2-dimensional space, the resulting vector b determines number of solutions. For example, when

$$b = b_0 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$$

the system has no solution, and when

$$b = b_\infty = \begin{bmatrix} 3 \\ -1 \\ -1 \end{bmatrix}$$

the system has infinite solutions.

1.3 Introduction to Systems of Linear Equations

1.3.1 Definition of Systems of Linear Equations

Definition 1.15. A **system of linear equations** is a collection of m equations in the variable quantities $x_1, x_2, x_3, \cdots, x_n$ of the form

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

where the values of a_{ij}, b_i and x_j, $1 < i < m, 1 < j < n$ are from the set of real numbers $\mathbb{R}$.

1.3.2 Possibilities of Solution Sets

Definition 1.16. A set $(s_1, s_2, \cdots, s_n)$ is said to be a **solution** of a system of linear equations in n variables if every equation in the system is valid after substitution. The **solution set** of a linear equation is the combination of all solutions of the system of linear equations.

Example 1.17. The following is a system of linear equations:

$$\begin{aligned} x_1 + 2x_2 + x_4 &= 7 \\ x_1 + x_2 + x_3 - x_4 &= 3 \\ 3x_1 + x_2 + 5x_3 - 7x_4 &= 1 \end{aligned}$$

and it can be rewritten as

$$\begin{aligned} 1x_1 + 2x_2 + 0x_3 + 1x_4 &= 7 \\ 1x_1 + 1x_2 + 1x_3 - 1x_4 &= 3 \\ 3x_1 + 1x_2 + 5x_3 - 7x_4 &= 1 \end{aligned}$$

Note that the system has infinitely many solutions, and the solution set is

$$\{(-1 - 2a + 3b, 4 + a - 2b, a, b) \mid a, b \text{ real numbers}\}$$

Here the theorem about the possibilities of solution sets is introduced:

Theorem 1.18. Any system of linear equations with real numbers as coefficients can have a unique solution, infinite solutions or no solution.

Therefore it is impossible for a system of linear equations to have exactly two solutions, for example. Also note that the above theorem is only true when the coefficients are real

numbers. In higher mathematics, coefficients can also be finite fields which will affect numbers of solutions.

1.3.3 Equivalent Systems and Equation Operations

Definition 1.19. Two systems of linear equations are said to be **equivalent** if their solution sets are equal, and the two systems are said to be **equivalent systems**.

Definition 1.20. Given a system of linear equations, the following three operations will transform a system into another one:

(a) Swap the position of any two equations in the list of equations.

(b) Perform a non-zero scalar multiplication to an equation in the list of equations.

(c) Perform a scalar multiplication to an equation, and then add to another equation on both sides of the equality.

The above operations are all called **equation operations**.

Note that the following theorem about equivalent systems and equation operations is true:

Proposition 1.21. If the transformed system is created by applying any one of three equation operations to an original system of linear equations, then the original system and the transformed system are equivalent.

2 Linear Algebra with Matrices

2.1 Introduction to Matrices

2.1.1 Matrix and Vector Notations

In order to solve systems of linear equations systematically, a useful and powerful tool called matrix is used.

Definition 2.1. A $m \times n$ **matrix** is a rectangular layout of numbers having m rows and n columns. The matrix is bracketed with either parentheses or brackets. Rows are counted from top to bottom and columns are counted from left to right, and for a matrix A, the notation $[A]_{ij}$ refers to the number in row i and column j of the matrix A.

Example 2.2. The following is a matrix

$$B = \begin{bmatrix} 1 & 0 & 2 & 5 \\ 3 & -2 & -7 & 4 \\ -1 & 4 & 6 & -3 \end{bmatrix}$$

with $m = 3$ rows and $n = 4$ columns. Note that $[B]_{2,3} = -7$ and $[B]_{3,4} = -3$.

In fact, during the solving process of systems of linear equations, the names of the variables do not matter at all. On the other hand, using matrix notations can describe linear systems, solve the systems and describe the solution sets with much ease.

Definition 2.3. A **column vector** of size m is an ordered list of m numbers in a vertical order from top to bottom.

At times, a column vector is also simply called a vector. Commonly, the notation of a vector is usually small letters in bold ($\mathbf{v}$) or with arrows on top of the symbol ($\vec{v}$). The entry of vector $\mathbf{v}$ in location i is in notation $[\mathbf{v}]_i$.

Definition 2.4. The **zero vector** of size m is the column vector of size m with each entry is zero:

$$\mathbf{0} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

or in entry notations, $[\mathbf{0}]_i = 0$ for $1 \leq i \leq m$.

2.1.2 Partitions of Matrices

Definition 2.5. A **partition** of a matrix is a horizontal or vertical line dividing the matrix into different areas. A matrix is the same with or without partitions, but partitions can provide better reading if required.

Example 2.6. Suppose A is a matrix and $\mathbf{u}$ is a column vector where

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \mathbf{u} = \begin{bmatrix} 5 \\ 6 \end{bmatrix}$$

Then

$$[A \,|\, \mathbf{u}] = \left[\begin{array}{cc|c} 1 & 2 & 5 \\ 3 & 4 & 6 \end{array}\right]$$

2.1.3 Matrix Representations of Linear Systems

Definition 2.7. For a system of linear eqautions

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

the **coefficient matrix** is the $m \times n$ matrix

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{bmatrix}$$

the **vector of constants** is the column vector of size m

$$\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

and the **solution vector** is the column vector of size n

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

Definition 2.8. If A is the coefficient matrix of a system of linear equations and b is the vector of constants, then $LS(A, \mathbf{b})$ is the shorthand expression for the system of linear equations, which is referred as the **matrix representation** of the linear system. Also, the **augmented matrix** of the linear system $[A\,|\,\mathbf{b}]$ is the $m \times (n+1)$ matrix with matrix A on the left and column vector $\mathbf{b}$ on the right.

2.1.4 Row Operations

An augmented matrix can represent the system of linear equations, which can also perform row operations as before. Recall the three equation operations:

(a) Swap the position of any two equations in the list of equations.

(b) Perform a non-zero scalar multiplication to an equation in the list of equations.

(c) Perform a scalar multiplication to an equation, and then add to another equation on both sides of the equality.

For augmented matrices, the following operations works similarly as the equation opertions above, but to transform matrices into another one with the same size:

(a) Swap any two rows in a matrix.

(b) Perform a non-zero scalar multiplication to every entry of a row in a matrix.

(c) Perform a scalar multiplication to every entry of a row in a matrix, and then add to another row in the same columns. The first row will remain unchanged while the second row is replaced by new values.

The three operations above are called **row operations**. There are symbolic shorthands to describe these row operations:

(a) $R_i \leftrightarrow R_j$: Swap row i and row j.

(b) αR_i : Multiply row i by the non-zero scalar α.

(c) $\alpha R_i + R_j$: Multiply row i by the scalar α and add to row j.

Suppose there are two matrices A and B of the same size. A and B are said to be **row-equivalent** if one can be obtained from the other by a sequence of row operations. Note that all row operations are reversible, so the distinction between A is row-equivalent to B and B is row-equivalent to A does not matter. With this definition, the following theorem applies:

Proposition 2.9. Suppose A and B are row-equivalent augmented matrices, then the systems of linear equations that the two matrices represent are equivalent systems.

2.2 Matrix Operations

2.2.1 Matrix Equality, Addition and Scalar Multiplication

Definition 2.10. Recall M_{mn} is the set of $m \times n$ matrices with real entries. Suppose there are two matrices A and B with size $m \times n$. The following definitions can be applied:

(a) The two matrices A and B are said to be **equal**, written as $A = B$ if

$$[A]_{ij} = [B]_{ij} \quad \text{for all} 1 \leq i \leq m, 1 \leq j \leq n$$

(b) The **sum** of two matrices A and B, written as $A + B$, is a matrix with size $m \times n$ where

$$[A + B]_{ij} = [A]_{ij} + [B]_{ij} \quad \text{for all} 1 \leq i \leq m, 1 \leq j \leq n$$

(c) The **scalar multiple** of the matrix A by α, where $\alpha \in \mathbb{R}$, is a matrix with size $m \times n$, written as αA where

$$[\alpha A]_{ij} = \alpha[A]_{ij} \quad \text{for all} 1 \leq i \leq m, 1 \leq j \leq n$$

(d) The $m \times n$ **zero matrix** is written as $O = O_{m\times n}$ and defined by $[O]_{ij} = 0$ for all $1 \leq i \leq m, 1 \leq j \leq n$.

(e) The **addictive inverse** of the matrix A, written as $-A$, is defined by $-A = (-1)A$.

Proposition 2.11. Below are some properties satisfied by addition and scalar multiplication:

(a) **Commutativity, Matrices**
If $A, B \in M_{mn}$, then $A + B = B + A$.

(b) **Addictive Associativity, Matrices**
If $A, B, C \in M_{mn}$, then $A + (B + C) = (A + B) + C$.

(c) **Zero Matrix, Matrices**
$A + O = A$ for all $A \in M_{mn}$.

(d) **Addictive Inverse, Matrices**
$A + (-A) = O$ for all $A \in M_{mn}$.

(e) **Scalar Multiplication Associativity, Matrices**
If $\alpha, \beta \in \mathbb{R}$ and $A \in M_{mn}$, then $\alpha(\beta A) = (\alpha\beta)A$.

(f) **Distributivity across Matrix Addition, Matrices**
If $\alpha \in \mathbb{R}$ and $A, B \in M_{mn}$, then $\alpha(A + B) = \alpha A + \alpha B$.

(g) **Distributivity across Scalar Addition, Matrices**
If $\alpha, \beta \in \mathbb{R}$ and $A \in M_{mn}$, then $(\alpha + \beta)A = \alpha A + \beta A$.

(h) **One, Matrices**
If $A \in M_{mn}$, then $1A = A$.

2.2.2 Transposes and Symmetric Matrices

Suppose there is a matrix A of size $m \times n$. The **transpose** of the matrix A is the matrix A^t of size $n \times m$ where

$$[A^t]_{ij} = [A]_{ji} \quad \text{for all } 1 \leq i \leq n, 1 \leq j \leq m$$

With the definition of transposes of matrices, a matrix A is said to be **symmetric** if $A = A^t$. Note that the following theorem is always applicable to symmetric matrices:

Proposition 2.12. Suppose A is a symmetric matrix, then A is also a square matrix.

Proof. Suppose A is a $n \times m$ matrix. Then A^t is a $m \times n$ matrix. In order for A and A^t to be equal, they must be in the same dimension. Hence $m = n$ and A is a square matrix.

The following are three theorems that are easy to prove matrix equalities with learnt notations and techniques:

Proposition 2.13. Suppose A and B are $m \times n$ matrices, then $(A + B)^t = A^t + B^t$.

Proof. For $1 \leq i \leq n, 1 \leq j \leq m$,

$$\begin{aligned}
[(A + B)^t]_{ij} &= [(A + B)]_{ji} \\
&= [A]_{ji} + [B]_{ji} \\
&= [A^t]_{ij} + [B^t]_{ij} \\
&= [A^t + B^t]_{ij}
\end{aligned}$$

Since every entry of the matrices $(A+B)^t$ and A^t+B^t agrees, the matrices are equal.

Proposition 2.14. Suppose $\alpha \in \mathbb{R}$ and A is an $m \times n$ matrix, then $(\alpha A)^t = \alpha A^t$.

Proof. For $1 \leq i \leq n, 1 \leq j \leq m$,

$$\begin{aligned}[(\alpha A)^t]_{ji} &= [\alpha A]_{ij} \\ &= \alpha[A]_{ij} \\ &= \alpha[A^t]_{ji} \\ &= [\alpha A^t]_{ji}\end{aligned}$$

Since every entry of the matrices $(\alpha A)^t$ and αA^t agrees, the matrices are equal.

Proposition 2.15. Suppose A is an $m \times n$ matrix, then $(A^t)^t = A$.

Proof. For $1 \leq i \leq m, 1 \leq j \leq n$,

$$\begin{aligned}[(A^t)^t]_{ij} &= [A^t]_{ji} \\ &= [A]_{ij}\end{aligned}$$

Since every entry of the matrices $(A^t)^t$ and A agrees, the matrices are equal.

2.3 Matrix Multiplication

2.3.1 Matrix-Vector Product

Definition 2.16. Suppose that A is an $m \times n$ matrix with columns $A_1, A_2, \cdots, A_n$ and $\mathbf{u}$ is a vector of size n, then the **matrix-vector product** of A with $\mathbf{u}$ is the linear combination

$$A\mathbf{u} = [\mathbf{u}]_1 A_1 + [\mathbf{u}]_2 A_2 + \cdots + [\mathbf{u}]_n A_n$$

As shown above, the matrix-vector product is yet another version of multiplication. Note that an $m \times n$ matrix multiplies with a vector of size n will produce a vector of size m.

With the definition of matrix-vector product, systems of linear equations can be represented compactly with a matrix-vector product and column vector equality. This yields a very popular alternative to the unconventional $LS(A, \mathbf{b})$ notation.

Proposition 2.17. The set of solutions to the linear system $LS(A, \mathbf{b})$ equals the set of solutions for x in the vector equation $A\mathbf{x} = \mathbf{b}$.

Proof.

$$\begin{aligned}&\mathbf{x} \text{ is a solution to } LS(A, \mathbf{b}) \\ &\Leftrightarrow [\mathbf{x}]_1 A_1 + [\mathbf{x}]_2 A_2 + [\mathbf{x}]_n A_n = b \\ &\Leftrightarrow \mathbf{x} \text{ is a solution to } A\mathbf{x} = \mathbf{b}\end{aligned}$$

Proposition 2.18. Suppose A and B are $m \times n$ matrices such that $A\mathbf{x} = B\mathbf{x}$ for any $\mathbf{x} \in \mathbb{R}^n$, then $A = B$.

2.3.2 Introduction to Matrix Multiplication

With the method of matrix-vector product, such method can also apply to multiplication of two matrices. Suppose A is an $m \times n$ matrix and B is an $n \times p$ matrix with columns $B_1, B_2, \cdots, B_p$, then the matrix product of A and B is the $m \times p$ matrix where

$$AB = A[B_1 \,|\, B_2 \,|\, \cdots \,|\, B_p] = [AB_1 \,|\, AB_2 \,|\, \cdots \,|\, AB_p]$$

Notice that the matrix product BA does not exist due to inappropriate size of the two matrices. Even if the size allows the matrix product to exist, the matrix product will not be the same as the other matrix product. This is because many of the ideas about multiplication will not apply to matrix multiplication. In this case, matrix multiplication is not commutative, which means that the order matters.

Proposition 2.19. Suppose A is an $m \times n$ matrix and B is an $n \times p$ matrix, then for $1 \leq i \leq m, 1 \leq j \leq p$, the individual entries of AB is given by

$$\begin{aligned}[AB]_{ij} &= [A]_{i1}[B]_{1j} + [A]_{i2}[B]_{2j} + \cdots + [A]_{in}[B]_{nj} \\ &= \sum_{k=1}^{n} [A]_{ik}[B]_{kj}\end{aligned}$$

In fact, this is often used as the definition of the matrix product AB.

2.3.3 Properties of Matrix Multiplication

Proposition 2.20. Suppose A is an $m \times n$ matrix, then the following equalities apply:

(a) $AO_{n\times p} = A_{m\times p}$

(b) $O_{p\times m}A = O_{p\times n}$

Proposition 2.21. Suppose A is an $m \times n$ matrix, then the following equalities apply:

(a) $AI_n = A$

(b) $I_m A = A$

Note that the matrix I is given the name of identity matrix since it behaves with matrix multiplication like the scalar 1 does in scalar multiplication, as shown in the

previous theorem. Multiplying by the identity matrix is to have no effect on the other matrix.

Proposition 2.22. Suppose A is an $m \times n$ matrix, B and C are $n \times p$ matrices, and D is a $p \times s$ matrix, then the following equalities apply:

(a) $A(B + C) = AB + AC$

(b) $(B + C)D = BD + CD$

Proposition 2.23. Suppose A is an $m \times n$ matrix, B is an $n \times p$ matrix and α is a scalar, then $\alpha(AB) = (\alpha A)B = A(\alpha B)$.

Proposition 2.24. Suppose A is an $m \times n$ matrix, B is an $n \times p$ matrix and D is a $p \times s$ matrix, then $(AB)D = A(BD)$.

Proposition 2.25. Suppose A is an $m \times n$ matrix and B is an $n \times p$ matrix, then $(AB)^t = B^t A^t$.

2.3.4 Row Operations and Matrix Multiplication

Recall the elementary row operations:

(a) Swap any two rows in a matrix.

(b) Perform a non-zero scalar multiplication to every entry of a row in a matrix.

(c) Perform a scalar multiplication to every entry of a row in a matrix, and then add to another row in the same columns. The first row will remain unchanged while the second row is replaced by new values.

Proposition 2.26. Let matrix $A \in M_{mn}$, B be a matrix obtained by applying one of the row operations on A, and J be a matrix obtained by applying the exact same row operation on I_m, then $JA = B$.

2.4 Reduced Row Echelon Form

2.4.1 Introduction to Reduced Row Echelon Form

Here are some terminology before introducing reduced row echelon form:

Definition 2.27.

(a) A **zero row** is a row consists of only 0.

(b) The **leftmost entry** of a row is the first nonzero entry of a row.

(c) The **index** of the leftmost entry of a row is the index of the first nonzero entry,

denoted by l_i for row i.

Example 2.28. The following is a matrix

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

Note that the index of leftmost entry of row 1 is $l_1 = 2$, that of row 2 is $l_2 = 4$ and that of row 3 is $l_3 = 3$. Row 4 is a zero row.

Definition 2.29. A matrix is said to be in **reduced row echelon form**, RREF in short, if it meets all of the conditions below:

(a) If there is a row where every entry is zero, then this row lies below any other row with nonzero entries.

(b) The leftmost nonzero entry of a row is equal to 1.

(c) The leftmost nonzero entry of a row is the only nonzero entry in its column.

(d) If $i < j$ and both row i and row j are not zero rows, then $l_i < l_j$.

With the definition of reduced row echelon form above, there are more definition based on reduced row echelon form:

Definition 2.30.

(a) A row of only entries of 0 is a zero row, and the leftmost nonzero entry of a nonzero row is a **leading 1**.

(b) A column that contains a leading 1 is called a **pivot column**, denoted by $D = \{d_1, d_2, \cdots, d_r\}$ where in ascending order.

(c) Columns that are not pivot columns are denoted by $F = \{f_1, f_2, \cdots, f_{n-r}\}$ where in ascending order.

Note that the number of nonzero rows is denoted by r as above, and it should always equal to the number of leading 1's and number of pivot columns.

2.4.2 Transforming Matrices to Reduced Row Echelon Form

Below is a theorem that states that any matrix can be transformed to reduced row echelon form:

Theorem 2.31. Suppose A is a matrix, then there exists a matrix B such that

(a) A and B are row-equivalent.

(b) B is in reduced row echelon form.

With the theorem above, suppose A has m rows and n columns. There is always a process to convert A into B via row operations. Such procedure is known as **Gaussian elimination** or **Gauss-Jordan elimination**. The following is an example of the procedure mentioned:

Example 2.32. The following is a matrix

$$A = \begin{bmatrix} 0 & 0 & 1 & 1 & 4 \\ 0 & 0 & 1 & 1 & 3 \\ 1 & 1 & 2 & 4 & 8 \\ 2 & 2 & 5 & 9 & 19 \end{bmatrix}$$

Consider the first row, note that the first entry of the row is 0, therefore consider the first column. If the first entry of the row is nonzero, this step can be skipped. The third entry of the column is a nonzero entry, hence perform $R_1 \leftrightarrow R_3$:

$$\xrightarrow{R_1 \leftrightarrow R_3} \begin{bmatrix} 1 & 1 & 2 & 4 & 8 \\ 0 & 0 & 1 & 1 & 3 \\ 0 & 0 & 1 & 1 & 4 \\ 2 & 2 & 5 & 9 & 19 \end{bmatrix}$$

Note that the fourth entry of the column is also a nonzero entry, but in this case choosing an entry that is equal to 1 or -1 will make future steps easier. Now check the leftmost entry of the first row, and peform $\frac{1}{a}R_1$ where a is the entry. This step is skipped in this example since the leftmost entry of the first row is already 1.

Now eliminate the remaining nonzero entries within the same column. In this example, the fourth entry of the column is nonzero, thus perform $-2R_1 + R_4$:

$$\xrightarrow{-2R_1 + R_4} \begin{bmatrix} 1 & 1 & 2 & 4 & 8 \\ 0 & 0 & 1 & 1 & 3 \\ 0 & 0 & 1 & 1 & 4 \\ 0 & 0 & 1 & 1 & 3 \end{bmatrix}$$

The first row and first column is completed and can be ignored:

$$\begin{bmatrix} * & * & * & * & * \\ * & 0 & 1 & 1 & 3 \\ * & 0 & 1 & 1 & 4 \\ * & 0 & 1 & 1 & 3 \end{bmatrix}$$

Note that the first leftmost entry of the row is at row 2 and column 3, so consider the third column and perform elimination, including the ignored row 1:

$$\xrightarrow{-2R_2+R_1,-R_2+R_3,-R_2+R_4} \begin{bmatrix} 1 & 1 & 0 & 2 & 2 \\ 0 & 0 & 1 & 1 & 3 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Repeat the process:

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

$$\xrightarrow{-2R_3+R_1,-3R_3+R_2} \begin{bmatrix} 1 & 1 & 0 & 2 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

until the matrix is in reduced row echelon form.

2.4.3 Uniqueness of Reduced Row Echelon Form

Theorem 2.33. Suppose A is an $m \times n$ matrix, and B and C are$m \times n$ matrix that are row-equivalent to A and in reduced row echelon form, then $B = C$.

With the above theorem, we can solve systems of equations in matrix form by reduced row echelon form.

2.5 Types of Solution Sets

2.5.1 Consistent Systems

With the reduced row echelon form, augmented matrices of systems of linear equations can be analyzed more systematically and carefully. In this section, the fact that every system of linear equations has either zero, one or infinitely many solutions will be proven in detail, thus routinely solve any linear system.

Definition 2.34. A system is said to be **consistent** if it has at least one solution. A system with no solution is **inconsistent**.

2.5.2 Free and Dependent Variables

Recall the following terms about reduced row echelon form:

Definition 2.35. Let A be a matrix in reduced row echelon form.

(a) The number of nonzero rows is called the **rank** of A and is denoted by r.

(b) The set of the column indexes for the pivot columns is denoted by

$$D = \{(d_1, d_2, \cdots, d_r) \mid d_1 < d_2 < \cdots < d_r\}$$

(c) The set of the column indexes that are not pivot columns is denoted by

$$D = \{(f_1, f_2, \cdots, f_{n-r}) \mid f_1 < f_2 < \cdots < f_{n-r}\}$$

With the definition above, free and dependent variables can also be defined as follows:

Definition 2.36. Suppose A is the augmented matrix of a consistent system of linear equations and B is a row-equivalent matrix in reduced row echelon form. If j is the index of a pivot column of B, then the variable x_j is **dependent**. A variable that is not dependent is **independent** or **free**.

Theorem 2.37. Suppose the matrix A is the augmented matrix of a system of linear equations with n variables, and B is a row-equivalent matrix in reduced row echelon form with r nonzero rows, then the following is true:

(a) The system of equations is inconsistent if and only if column $n+1$ (i.e. the last column) of B is a pivot column.

(b) The system of equations is also inconsistent if and only if the last nonzero row is $(0, 0, \cdots, 0, 1)$.

(c) Equivalently a system is consistent if and only if column $n+1$ is not a pivot column of B.

2.5.3 Interpretations of Reduced Row Echelon Form

Theorem 2.38. Suppose A is the augmented matrix of a consistent system of linear equations with n variables, and B is a row-equivalent matrix in reduced row echelon form with r pivot columns. Then $r \leq n$. If $r = n$, then the system has unique solution, and if $r < n$, then the system has infinitely many solutions.

Proof. Notice that B has $n+1$ columns, so there can be at most $n+1$ pivot columns, i.e. $r \leq n+1$. If $r = n+1$, then every column in B is pivot column, in particular, the last column is also a pivot column. However, the previous theorem tells us that the system is inconsistent, which is contrary to the hypothesis. Then $r \leq n$.

When $r = n$, then there are $n - r = 0$ free variables, and the solution is unique. On the other hand, when $r < n$, there are $n - r > 0$ free variables, and these free variables can be adjusted to create infinitely many solutions.

The next theorem simply states a conclusion from the previous proof, in order to state explicitly the number of free variables in a consistent system.

Theorem 2.39. Suppose A is the augmented matrix of a consistent system of linear equations with n variables, and B is a row-equivalent matrix in reduced row echelon form with r nonzero rows, then the solution set can be described with $n - r$ free variables.

With the number of free variables, possible solution sets for linear systems can be easily found, by the following theorem:

Theorem 2.40. A system of linear equations has no solutions, a unique solution or infinitely many solutions.

Proof. If the system is inconsistent, then the system has no solutions. Now suppose the system is consistent, if it has no free variables, then it has a unique solution. Otherwise, it has infinitely many solutions.

Theorem 2.41. Suppose a consistent system of linear equations has m equations in n variables, then if $n > m$, then the system has infinitely many solutions.

Proof. Suppose there is an augmented matrix of system of linear equations that is row-equivalent to a matrix in reduced row echelon form with r nonzero rows. Since the original matrix has m rows in total, the number of nonzero rows should be less than or equal to m, which is $r \leq m$. Note that the number of free variables

$$n - r \geq n - m > 0$$

based on the hypothesis $n > m$. A consistent system with free variables will then have infinitely many solutions, according to *Theorem 2.36*.

2.6 Homogeneous Systems of Equations

2.6.1 Definition of Homogeneous Systems

Definition 2.42. A system of linear equations $LS(A, \mathbf{b})$ is **homogeneous** if the vector of constants is the zero vector, or in other words, $\mathbf{b} = \mathbf{0}$. The **homogeneous system** of $LS(A, \mathbf{b})$ can be expressed as $LS(A, \mathbf{0})$.

Note that homogeneous systems have an identical property:

Proposition 2.43. Suppose a system of linear equation is homogeneous, then the system is consistent.

Proof. By setting every variable to 0, the left hand side of all equations will equal to 0, so does the right hand side.

2.6.2 Solutions of Homogeneous Systems

With the above theorem, the definition of trivial solution is introduced:

Definition 2.44. It is always possible to put the zero vector $\mathbf{0}$ into any homogeneous system as a solution. Such a solution is called a **trivial solution**.

Theorem 2.45. Suppose a homogeneous system of linear equations has m equations in n variables with $n > m$, then the system has infinitely many solutions.

Proof. By *Proposition 2.41*, the system is consistent, along with *Theorem 2.39*, the system has infinitely many solutions.

2.6.3 Null Space of Matrices

Definition 2.46. The **null space** of a matrix A, denoted by $N(A)$, is the set of all vectors that are solutions to the homogeneous system $LS(A, \mathbf{0})$.

Example 2.47. The following is a matrix

$$A = \begin{bmatrix} 1 & 4 & 0 & -1 & 0 & 7 & -9 \\ 2 & 8 & -1 & 3 & 9 & -13 & 7 \\ 0 & 0 & 2 & -3 & -4 & 12 & -8 \\ -1 & -4 & 2 & 4 & 8 & -31 & 37 \end{bmatrix}$$

Note that

$$\mathbf{x} = \begin{bmatrix} 3 \\ 0 \\ -5 \\ -6 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \ \mathbf{y} = \begin{bmatrix} -4 \\ 1 \\ -3 \\ -2 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

are both in $N(A)$ since $A\mathbf{x} = \mathbf{0}$ and $A\mathbf{y} = \mathbf{0}$.

2.6.4 Augmented Matrices and Coefficient Matrices

The augmented matrix for the homogeneous system of linear equaations $LS(A, \mathbf{0})$ is $[A \,|\, \mathbf{0}]$. Note that any row operations will not have any effect on the last zero column. If

$$A \xrightarrow{\text{row operations}} B$$

then

$$[A \,|\, \mathbf{0}] \xrightarrow{\text{same row operations}} [B \,|\, \mathbf{0}]$$

2.6.5 Nonsingular Matrices

In this section it is specialized that only matrices that have equal number of rows and columns, or in other words, systems with equal number of equations and variables are considered.

Definition 2.48. A matrix with m rows and n columns is a **square matrix** if $m = n$. In such a case, the matrix is said to have **size** n. In order to emphasize the situation where the matrix is not square, such a matrix can be said as a **rectangular matrix**.

Definition 2.49. Suppose matrix A is a square matrix, and the solution set to the homogeneous linear system of equations $LS(A, \mathbf{0})$ is $\{\mathbf{0}\}$, or in other words, it has only the trivial solution, then A is a **nonsingular matrix**. Otherwise, A is a **singular matrix**.

Also recall the identity matrix, with the definition of square matrices:

Definition 2.50. The **identity matrix** with size m, I_m, is defined by

$$[I_m]_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases} \qquad 1 \leq i, j \leq m$$

Note that every identity matrix is square, in reduced row echelon form, and every column is a unique pivot column.

Theorem 2.51. Suppose A is a square matrix and B is a row-equivalent matrix in reduced row echelon form, then A is nonsingular if and only if B is the identity matrix.

Proof. ($\Leftarrow$) Suppose B is the identity matrix. When the augmented matrix $[A \,|\, \mathbf{0}]$ is reduced, the result is $[B \mid \mathbf{0}] = [I_m \mid \mathbf{0}]$. The number of nonzero rows is equal to the number of eqautions in the linear system of eqautions $LS(A, \mathbf{0})$, so it has no free variables. Then the homogeneous system $LS(A, \mathbf{0})$ has only one unique solution, which is the trivial solution.

($\Rightarrow$) If A is nonsingular, then the homogeneous system $LS(A, \mathbf{0})$ has only unique solution, which means there are no free variables in the description of the solution set. By *Theorem 2.37*, the system has $n - r$ free variables, in which $n = r$ in this case. Thus B has n pivot columns among its n columns, forcing B to be the identity matrix.

2.6.6 Null Space of Nonsingular Matrices

Below are the theorems about null space and nonsingular matrices:

Theorem 2.52. Suppose A is a square matrix, then A is nonsingular if and only if the null space of A is the set containing only the zero vector, i.e. $N(A) = \{\mathbf{0}\}$.

Proof. Note that the null space of A is equal to the set of solutions to the homogeneous system $LS(A, \mathbf{0})$. Also a matrix is nonsingular if and only if the set of solution to the homogeneous system $LS(A, \mathbf{0})$ has only trivial solution. With these observations, both sides of the proof can be constructed easily.

Theorem 2.53. Suppose A is a square matrix, then A is nonsingular if and only if the system $LS(A, \mathbf{b})$ has a unique solution for any choice of the constant vector $\mathbf{b}$.

The following theorem is a simple conclusion of nonsingular matrix equivalences:

Proposition 2.54. Suppose A is a square matrix, then the following are equivalent to each other:

(a) A is nonsingular.

(b) A can be row-reduced to the identity matrix.

(c) The null space of A contains only the zero vector, i.e. $N(A) = \{\mathbf{0}\}$.

(d) The linear system $LS(A, \mathbf{b})$ has a unique solution for any choice of the constant vector $\mathbf{b}$.

Sometimes it is much more efficient to find just one solution and then find all the solutions to the corresponding homogeneous system, rather than finding all the solutions directly. This also explains part of our interest in the null sapce, the set of all solutions in the homogeneous system.

Theorem 2.55. Suppose that w is one solution to the linear system of equations $LS(A, \mathbf{b})$, then y is a solution to $LS(A, \mathbf{b})$ if and only if $\mathbf{y} = \mathbf{w} + \mathbf{z}$ for some vector $\mathbf{z} \in N(A)$, i.e.

(a) If y is a solution to $LS(A, \mathbf{b})$, then $\mathrm{y} - \mathrm{w} \in N(A)$.

(b) If $z \in N(A)$, then $\mathbf{w} + \mathbf{z}$ is a solution to $LS(A, \mathbf{b})$.

3 Linear Algebra with Vectors

3.1 Vector Space and Subspaces

3.1.1 Definition of Vector Space

Definition 3.1. The **vector space** $\mathbb{R}^m$ is the set of all column vectors with size m with entries from the set of real numbers $\mathbb{R}$. $\mathbb{R}^m$ is also called the **Euclidean m-space**.

Definition 3.2. Suppose that $\mathbf{u}, \mathbf{v} \in \mathbb{R}^m$, then $\mathbf{u}$ and $\mathbf{v}$ are **equal**, written as $\mathbf{u} = \mathbf{v}$ if

$$[\mathbf{u}]_i = [\mathbf{v}]_i \quad \text{for all } 1 \leq i \leq m$$

Definition 3.3. Suppose that $\mathbf{u}, \mathbf{v} \in \mathbb{R}^m$, then the **sum** of $\mathbf{u}$ and $\mathbf{v}$ is the vector $\mathbf{u} + \mathbf{v}$ defined by

$$[\mathbf{u} + \mathbf{v}]_i = [\mathbf{u}]_i + [\mathbf{v}]_i \quad \text{for all } 1 \leq i \leq m$$

Definition 3.4. Suppose that $\mathbf{u} \in \mathbb{R}^m$ and $\alpha \in \mathbb{R}$, then the **scalar multiple** of $\mathbf{u}$ by α is the vector $\alpha\mathbf{u}$ defined by

$$[\alpha\mathbf{u}]_i = \alpha[\mathbf{u}]_i \quad \text{for all } 1 \leq i \leq m$$

3.1.2 Properties of Vector Space

Comment. This section is for potential major in mathematics and it is optional.

With the definitions of vector addition and scalar multiplication, a list of properties, stated in the following proposition, can be stated:

Proposition 3.5. Suppose that $\mathbb{R}^m$ is the set of column vectors with size m with vector addition and scalar multiplication as defined in *Definition 3.3* and *Definition 3.4*, then

(a) **ACC: Addictive Closure, Column Vectors**
If $\mathbf{u}, \mathbf{v} \in \mathbb{R}^m$, then $\mathbf{u} + \mathbf{v} \in \mathbb{R}^m$.

(b) **SCC: Scalar Closure, Column Vectors**
If $\mathbf{u} \in \mathbb{R}^m$ and $\alpha \in \mathbb{R}$, then $\alpha\mathbf{u} \in \mathbb{R}^m$.

(c) **CC: Commutativity, Column Vectors**
If $\mathbf{u}, \mathbf{v} \in \mathbb{R}^m$, then $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$.

(d) **AAC: Addictive Associativity, Column Vectors**
If $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^m$, then $\mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w}$.

(e) **ZC: Zero Vector, Column Vectors**
There is a vector $\mathbf{0}$, called the **zero vector**, such that $\mathbf{u}+\mathbf{0} = \mathbf{u}$ for all $\mathbf{u} \in \mathbb{R}^m$.

(f) **AIC: Addictive Inverses, Column Vectors**
If $\mathbf{u} \in \mathbb{R}^m$, then there exists a vector $-\mathbf{u} \in \mathbb{R}^m$ so that $\mathbf{u} + -\mathbf{u} = 0$.

(g) **SMAC: Scalar Multiplication Associativity, Column Vectors**
If $\mathbf{u} \in \mathbb{R}^m$ and $\alpha, \beta \in \mathbb{R}$, then $\alpha(\beta\mathbf{u}) = (\alpha\beta)\mathbf{u}$.

(h) **DVAC: Distributivity across Vector Addition, Column Vectors**
If $\mathbf{u}, \mathbf{v} \in \mathbb{R}^m$ and $\alpha \in \mathbb{R}$, then $\alpha(\mathbf{u} + \mathbf{v}) = \alpha\mathbf{u} + \alpha\mathbf{v}$.

(i) **DSAC: Distributivity across Scalar Addition, Column Vectors**
If $\mathbf{u} \in \mathbb{R}^m$ and $\alpha, \beta \in \mathbb{R}$, then $(\alpha + \beta)\mathbf{u} = \alpha\mathbf{u} + \beta\mathbf{u}$.

(j) **OC: One, Column Vectors**
If $\mathbf{u} \in \mathbb{R}^m$, then $1\mathbf{u} = \mathbf{u}$.

3.1.3 Abstract Definition of Vector Space

Comment. This section is also for potential major in mathematics and it is optional.

The following is the abstract definition of vector space. In fact, a lot of different algebraic objects (for example polynomials, matrices, sequences and functions) share similar properties with the set of column vectors. Therefore, it may be found similar with *Proposition 3.4*.

Proposition 3.6. Suppose that V is a set with vector addition and scalar multiplication, then V, with the two opeations, is a **vector space** over $\mathbb{R}$ if the following properties hold:

(a) **AC: Addictive Closure**
If $\mathbf{u}, \mathbf{v} \in V$, then $\mathbf{u} + \mathbf{v} \in V$.

(b) **SC: Scalar Closure**
If $\mathbf{u} \in V$ and $\alpha \in \mathbb{R}$, then $\alpha\mathbf{u} \in V$.

(c) **C: Commutativity**
If $\mathbf{u}, \mathbf{v} \in V$, then $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$.

(d) **AA: Addictive Associativity**
If $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$, then $\mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w}$.

(e) **Z: Zero Vector**
There is a vector $\mathbf{0}$, called the **zero vector**, such that $\mathbf{u}+\mathbf{0}=\mathbf{u}$ for all $\mathbf{u}\in V$.

(f) **AI: Addictive Inverses**
If $\mathbf{u}\in V$, then there exists a vector $-\mathbf{u}\in V$ so that $\mathbf{u}+-\mathbf{u}=0$.

(g) **SMA: Scalar Multiplication Associativity**
If $\mathbf{u}\in V$ and $\alpha,\beta\in\mathbb{R}$, then $\alpha(\beta\mathbf{u})=(\alpha\beta)\mathbf{u}$.

(h) **DVA: Distributivity across Vector Addition**
If $\mathbf{u},\mathbf{v}\in V$ and $\alpha\in\mathbb{R}$, then $\alpha(\mathbf{u}+\mathbf{v})=\alpha\mathbf{u}+\alpha\mathbf{v}$.

(i) **DSA: Distributivity across Scalar Addition**
If $\mathbf{u}\in V$ and $\alpha,\beta\in\mathbb{R}$, then $(\alpha+\beta)\mathbf{u}=\alpha\mathbf{u}+\beta\mathbf{u}$.

(j) **O: One**
If $\mathbf{u}\in V$, then $1\mathbf{u}=\mathbf{u}$.

3.1.4 Basic Properties of Vector Space

Comment. This section is for potential major in mathematics and it is optional.

Theorem 3.7. If $\mathbf{u},\mathbf{v}$ and $\mathbf{w}$ are vectors in a vector space V such that $\mathbf{u}+\mathbf{w}=\mathbf{v}+\mathbf{w}$, then $\mathbf{u}=\mathbf{v}$.

Theorem 3.8. Let V be a vector space, then the vector $\mathbf{0}$ that satisfies property **Z** is unique.

Theorem 3.9. Let V be a vector space and $\mathbf{u},\mathbf{v},\mathbf{w}\in V$. If $\mathbf{u}+\mathbf{w}=\mathbf{0}$ and $\mathbf{v}+\mathbf{w}=\mathbf{0}$, then $\mathbf{u}=\mathbf{v}$. In other words, the addictive inverse of $\mathbf{w}$ is unique.

Theorem 3.10. Let V be a vector space, α be a real number, and $\mathbf{v}\in V$, then the following statements are true:

(a)
$$0\mathbf{v}=\mathbf{0}$$

(b)
$$\alpha\mathbf{0}=\mathbf{0}$$

(c)
$$(-\alpha)\mathbf{v}=-(\alpha\mathbf{v})=\alpha(-\mathbf{v})$$

3.1.5 Subspaces

Definition 3.11. Let V be a vector space. A subset W of V is said to be a **subspace** of V if the following conditions are satisfied:

(a) W is nonempty.

(b) For $\mathbf{v}, \mathbf{w} \in W$, then $\mathbf{v} + \mathbf{w} \in W$.

(c) For $\alpha \in \mathbb{R}$ and $\mathbf{v} \in W$, then $\alpha\mathbf{v} \in W$.

Proposition 3.12. Let V be a vector space and W be a subspace of V, then $\mathbf{0} \in W$.

With the definition and propostion above, there is a systematic way to check for subspaces:

Theorem 3.13. Let V be a vector space and W be a subset of V, then W is a subspace if and only if the following conditions are satisfied:

(a) W is nonempty.

(b) For any $\alpha \in \mathbb{R}$ and $\mathbf{v}, \mathbf{w} \in W$, $\alpha\mathbf{v} + \mathbf{w} \in W$.

There are some theorems on subspaces:

Theorem 3.14. Let $A \in M_{mn}$, then $W = N(A)$ is a subspace of $\mathbb{R}^n$.

The following theorem is about span, which will be further introduced and discussed later. This can be skipped for now.

Theorem 3.15. Let $S = \{\mathbf{u}_1, \cdots, \mathbf{u}_k\} \in V = \mathbb{R}^m$, then $\langle S \rangle$ is a subspace of V.

Corollary. Let $A \in M_{mn}$, then $C(A)$ is a subspace of $\mathbb{R}^m$.

3.2 Linear Combinations

3.2.1 Linear Combinations of Column Vectors

Definition 3.16. Let $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n$ be vectors in $\mathbb{R}^m$ and $\alpha_1, \alpha_2, \cdots, \alpha_n$ be scalars, then the **linear combination** of the vectors and the scalars is

$$\alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \cdots + \alpha_n\mathbf{u}_n$$

in $\mathbb{R}^m$.

With the definition above, systems of linear equations can be solved by rewriting it into a form of linear combination.

Example 3.17. Solve the system of linear equations

$$\begin{cases} x_1 - x_2 + 2x_3 = 1 \\ 2x_1 + x_2 + x_3 = 8 \\ x_1 + x_2 = 5 \end{cases}$$

Answer. The system of linear equations can be rewritten as a linear combination

$$x_1 \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} + x_2 \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} + x_3 \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 8 \\ 5 \end{bmatrix}$$

and by row-reducing the augmented matrix,

$$\begin{cases} x_1 = 2 \\ x_2 = 3 \\ x_3 = 1 \end{cases} \quad \text{or} \quad \begin{cases} x_1 = 3 \\ x_2 = 2 \\ x_3 = 0 \end{cases}$$

Theorem 3.18. Denote the columns of the $m \times n$ matrix be $A_1, A_2, \cdots, A_n$, then $\mathbf{x} \in \mathbb{R}^n$ is a solution to the linear system $A\mathbf{x} = \mathbf{b}$ if and only if $\mathbf{b}$ is equal to the linear combination of columns of A and entries of $\mathbf{x}$. In other words,

$$\mathbf{x}_1 A_1 + \mathbf{x}_2 A_2 + \cdots + \mathbf{x}_n A_n = \mathbf{b}$$

Proof. ($\Rightarrow$) If $\mathbf{x}$ is a solution of $A\mathbf{x} = \mathbf{b}$, then

$$\mathbf{b} = \mathbf{x}_1 A_1 + \mathbf{x}_2 A_2 + \cdots + \mathbf{x}_n A_n$$

where $\mathbf{b}$ is the linear combination of columns of A and entries of $\mathbf{x}$.

($\Leftarrow$) If $\mathbf{b}$ is the linear combination of columns of A and entries of $\mathbf{x}$, which is

$$\mathbf{b} = \mathbf{x}_1 A_1 + \mathbf{x}_2 A_2 + \cdots + \mathbf{x}_n A_n$$

then $\mathbf{b} = A\mathbf{x}$ and therefore $\mathbf{x}$ is a solution of $\mathbf{b} = A\mathbf{x}$.

3.2.2 Vector Form of Solution Sets

Consider the following example:

Example 3.19. Solve the system of linear equations

$$\begin{cases} 2x_1 + x_2 + 7x_3 - 7x_4 = 8 \\ -3x_1 + 4x_2 - 5x_3 - 6x_4 = -12 \\ x_1 + x_2 + 4x_3 - 5x_4 = 4 \end{cases}$$

Answer. By row-reducing the system of linear equations, the augmented matrix yields

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

Note that there are 2 pivot columns in the augmented matrix above. $D = \{1, 2\}$ and $F = \{3, 4, 5\}$ shows that the dependent variables are x_1, x_2 and the free variables are x_3, x_4. Hence, x_1, x_2 can be represented in the form of x_3, x_4:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 4 - 3x_3 + 2x_4 \\ -x_3 + 3x_4 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 4 \\ 0 \\ 0 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -3 \\ -1 \\ 1 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} 2 \\ 3 \\ 0 \\ 1 \end{bmatrix}$$

With the example above, the general procedure of finding the vector form of solution sets is as follows:

(1) Write the vector of variables as the sum of a constant vector and linear combination of vectors of free variables.

(2) Fill in 0 and 1 in the corresponding entries so that the free variables are independent to each other.

(3) Convert equations of the row-reduced augmented matrix to the corresponding row of each of the dependent variables.

Theorem 3.20. Suppose that $[A \mid \mathrm{b}]$ is the augmented matrix for a consistent linear system $A\mathbf{x} = \mathbf{b}$ of m equations in n variables. Let B be a row-equivalent matrix to A in reduced row-echelon form. Suppose B has r pivot columns, and let $D = \{d_1, d_2, \cdots, d_r\}$ and $F = \{f_1, f_2, \cdots, f_{n-r}\}$ be the set of dependent and free variables respectively, then define vectors $\mathbf{c}$ and $\mathbf{u}_j$ for $1 \leq j \leq n - r$ by

$$c_i = \begin{cases} 0 & \text{if } i \in F \\ B_{k,n+1} & \text{if } i \in D, i = d_k \end{cases}$$

$$(\mathbf{u}_j)_i = \begin{cases} 1 & \text{if } i \in F, i = f_j \\ 0 & \text{if } i \in F, i \neq f_j \\ -B_{k,f_j} & \text{if } i \in D, i = d_k \end{cases}$$

Then the set of solutions to the system of equations $A\mathbf{x} = \mathbf{b}$ is

$$S = \{\mathbf{c} + \alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \cdots + +\alpha_{n-r}\mathbf{u}_{n-r} \mid \alpha_1, \alpha_2, \cdots, \alpha_{n-r} \in \mathbb{R}\}$$

3.3 Spanning Sets

3.3.1 Spans of Sets of Vectors

Given that the solution set of a homogeneous system can be described as all possible linear combinations of sets of vectors, the definition of spans is as follows:

Definition 3.21. Let $S = \{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n\}$ be a set of vectors, then their **span**, denoted by $\langle S \rangle$, is the set of all linear combinations of the vectors within S. In other words,

$$\begin{aligned}\langle S \rangle &= \{\alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \cdots + \alpha_n\mathbf{u}_n \mid \alpha_1, \alpha_2, \cdots, \alpha_n \in \mathbb{R}\} \\ &= \left\{\sum_{i=1}^{n} \alpha_i\mathbf{u}_i \mid \alpha_i \in \mathbb{R}, 1 \leq i \leq n\right\}\end{aligned}$$

Theorem 3.22. Let $S = \{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n\} \subset V \in \mathbb{R}^m$, then $\langle S \rangle$ is a subspace of V.

Proof. Since S is not a null set, $\langle S \rangle$ is nonempty. Let $c \in \mathbb{R}$ and $\mathbf{v}, \mathbf{w} \in \langle S \rangle$ such that there exists α_i, β_i for $1 \leq i \leq n$ such that

$$\begin{cases}\mathbf{v} = \alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \cdots + \alpha_n\mathbf{u}_n \\ \mathbf{w} = \beta_1\mathbf{u}_1 + \beta_2\mathbf{u}_2 + \cdots + \beta_n\mathbf{u}_n\end{cases}$$

Combining both equations in the form $c\mathbf{v} + \mathbf{w}$ gets

$$c\mathbf{v} + \mathbf{w} = (c\alpha_1 + \beta_1)\mathbf{u}_1 + (c\alpha_2 + \beta_2)\mathbf{u}_2 + \cdots + (c\alpha_n + \beta_n)\mathbf{u}_n \in \langle S \rangle$$

Therefore $\langle S \rangle$ is a subspace of V.

Theorem 3.23. Suppose that $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n$ are in $\mathbb{R}^m$. Let A be the $m \times n$ matrix whose i-th column is $\mathbf{u}_i$, then $\mathbf{v} \in \langle S \rangle$ if and only if the equation $A\mathbf{x} = \mathbf{v}$ is consistent.

Theorem 3.24. Given a set of m vectors $S = \{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_m\}$, and let A be the $m \times m$ square matrix whose i-th column is $\mathbf{u}_i$, then A is nonsingular if and only if $\langle S \rangle = \mathbb{R}^m$.

3.3.2 Spanning Sets of Null Spaces

The following theorem describes how to express a null space as a spanning set:

Theorem 3.25. Let A be an $m \times n$ matrix and B be a row-equivalent matrix in reduced row-echelon form. Suppose B has r pivot columns with indices $D = \{d_1, d_2, \cdots, d_n\}$, and $n - r$ non-pivot columns with indices $F = \{f_1, d_2, \cdots, f_{n-r}\}$, then the null space of A is given by

$$N(A) = \langle \{\mathbf{z}_1, \mathbf{z}_2, \cdots, \mathbf{z}_{n-r}\} \rangle$$

where

$$[\mathbf{z}_j]_i = \begin{cases} 1 & \text{if } i \in F, i = f_j \\ 0 & \text{if } i \in F, i \neq f_j \\ -[B]_{k,f_j} & \text{if } i \in D, i = d_k \end{cases}$$

for $1 \leq j \leq n - r$.

Example 3.26. Find a set of vectors S such that the null space of the matrix

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

is the span of S, or $\langle S \rangle = N(A)$.

Answer. Note that the null space of A is the set of all solutions to the homogeneous system $A\mathbf{x} = \mathbf{0}$. By row-reducing the matrix A,

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

which means $D = \{1, 2, 4\}$ and $F = 3, 5$. In other words, x_1, x_2 and x_4 can be expressed in terms of x_3 and x_5. The solution vector can be written as

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = x_3 \begin{bmatrix} -6 \\ 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -4 \\ 2 \\ 0 \\ -3 \\ 1 \end{bmatrix}$$

Therefore

$$N(A) = \left\langle \left\{ \begin{bmatrix} -6 \\ 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -4 \\ 2 \\ 0 \\ -3 \\ 1 \end{bmatrix} \right\} \right\rangle$$

3.3.3 Reduction of Spans

Example 3.27. Let

$$T = \{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3, \mathbf{w}_4\} = \left\{ \begin{bmatrix} 2 \\ -3 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 4 \\ 1 \end{bmatrix}, \begin{bmatrix} 7 \\ -5 \\ 4 \end{bmatrix}, \begin{bmatrix} -7 \\ -6 \\ -5 \end{bmatrix} \right\}$$

and

$$D = \begin{bmatrix} 2 & 1 & 7 & -7 \\ -3 & 4 & -5 & -6 \\ 1 & 1 & 4 & -5 \end{bmatrix}$$

Row-reduce D gives

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

with

$$N(D) = \left\langle \left\{ \mathbf{z}_1 = \begin{bmatrix} -3 \\ -1 \\ 1 \\ 0 \end{bmatrix}, \mathbf{z}_2 = \begin{bmatrix} 2 \\ 3 \\ 0 \\ 1 \end{bmatrix} \right\} \right\rangle$$

Since $\mathbf{z}_1$ and $\mathbf{z}_2$ are solutions to the homogeneous equation $D\mathbf{x} = \mathbf{0}$, $\mathbf{w}_3$ and $\mathbf{w}_4$ can be expressed as $\mathbf{w}_1$ and $\mathbf{w}_2$. Therefore the infinite set $W = \langle\{\mathbf{w}_1, \mathbf{w}_2\}\rangle$.

3.4 Linear Independence

3.4.1 Linearly Independent Sets

Linear independence is one of the most fundamental conceptual ideas in linear algebra besides spans.

Definition 3.28. Let $S = \{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n\}$ be a set of vectors, then the equality

$$\alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \cdots + \alpha_n\mathbf{u}_n = 0$$

is a **relation of linear independence** on S. If the equality is satisfied only when $\alpha_i = 0$ for all $1 \leq i \leq n$, then such relation is a **trivial relation of linear independence** on S.

Definition 3.29. Let $S = \{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n\}$ be a set of vectors, then S is **linearly independent** if there is only one relation, namely the trivial relation of linear independence. If there exists any relations other than the trivial relation, S is **linearly dependent**.

Example 3.30. Find the linear independence of

$$S = \left\{ \begin{bmatrix} 2 \\ -1 \\ 3 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ -1 \\ 5 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ -3 \\ 6 \\ 1 \end{bmatrix}, \begin{bmatrix} -6 \\ 7 \\ -1 \\ 0 \\ 1 \end{bmatrix} \right\}$$

Answer. By row-reduce the associated matrix

$$A = \begin{bmatrix} 2 & 1 & 2 & -6 \\ -1 & 2 & 1 & 7 \\ 3 & -1 & -3 & -1 \\ 1 & 5 & 6 & 0 \\ 2 & 2 & 1 & 1 \end{bmatrix} \xrightarrow{\text{RREF}} \begin{bmatrix} 1 & 0 & 0 & -2 \\ 0 & 1 & 0 & 4 \\ 0 & 0 & 1 & -3 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

there exists a nontrivial solution

$$\mathbf{x} = \begin{bmatrix} 2 \\ -4 \\ 3 \\ 1 \end{bmatrix}$$

such that it will form a relation of linear independence. Therefore S is linearly dependent.

By the example above, linear independence can be figured out by solving a homogeneous system, which leads to the following theorem:

Theorem 3.31. Let $S = \{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n\}$ be a set of vectors in $\mathbb{R}^m$, and A be the $m \times n$ matrix whose columns are vectors in S, then S is linear independent if and only if the homogeneous system $LS(A, \mathbf{0})$ has a unique solution.

Corollary. S is linear independent if and only if number of vectors in S is equal to number of pivot columns in the matrix row-equivalent to A and in reduced row-echelon form.

Theorem 3.32. Let $S = \{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n\}$ be a set of vectors in $\mathbb{R}^m$, then S is linear dependent if $n > m$.

3.4.2 Linear Independence of Non-singular Matrices

Theorem 3.33. Let A be a square matrix, then A is nonsingular if and only if the columns of A form a linear independent set.

In general there are more ways to show a matrix is nonsingular:

Proposition 3.34. Let A be a square matrix, then the following are equivalent:

(a) A is nonsingular.

(b) A can be row-reduced to the identity matrix I.

(c) Null space of A contains only the zero vector, or $N(A) = \{\mathbf{0}\}$.

(d) The linear system $LS(A, \mathbf{b})$ has unique solution for all possible $\mathbf{b}$.

(e) Columns of A form a linear independent set.

Example 3.35. Do the columns of the matrix

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

form a linear independent set?

Answer. By *Theorem 3.33*, since A is singular, columns of the matrix form a linear dependent set.

3.4.3 Linear Independence with Null Space and Spans

Theorem 3.36. Let A be a $m \times n$ matrix, B be a row-equivalent matrix to A in reduced row-echelon form with r pivot columns. Further denote $D = \{d_1, d_2, \cdots, d_r\}$ be set of indices that are pivot columns in B, and $F = \{f_1, f_2, \cdots, f_{n-r}\}$ be set of indices that are not. If define

$$[\mathbf{z}_j]_i = \begin{cases} 1 & \text{if } i \in F, i = f_j \\ 0 & \text{if } i \in F, i \neq f_j \\ -[B]_{k,f_j} & \text{if } i \in D, i = d_k \end{cases}$$

for $1 \leq j \leq n-r$, and $S = \{\mathbf{z}_1, \mathbf{z}_2, \cdots, \mathbf{z}_{n-r}\}$, then $N(A) = \langle S \rangle$ and S is linear independent.

3.5 Linear Dependence

3.5.1 Dependency in Linearly Dependent Sets

Theorem 3.37. Let $S = \{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n\}$ be a set of vectors, then S is linearly dependent if and only if there is $\mathbf{u}_t$ such that $\mathbf{u}_t$ is a linear combination of the vectors in S excluding itself.

Example 3.38. Reduce

$$R = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\} = \left\{ \begin{bmatrix} 1 \\ 2 \\ -1 \\ 3 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 3 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ -7 \\ 6 \\ -11 \\ -2 \end{bmatrix}, \begin{bmatrix} 4 \\ 1 \\ 2 \\ 1 \\ 6 \end{bmatrix} \right\}$$

Answer. Let V be the span of R, and

$$D = \begin{bmatrix} 1 & 2 & 0 & 4 \\ 2 & 1 & -7 & 1 \\ -1 & 3 & 6 & 2 \\ 3 & 1 & -11 & 1 \\ 2 & 2 & -2 & 6 \end{bmatrix}$$

be a matrix whose columns are vectors in R. Row-reduce D gives

$$\begin{bmatrix} 1 & 0 & 0 & 4 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Note that $\mathbf{v}_4 = 4\mathbf{v}_1 + \mathbf{v}_3$, so $V = \langle R \rangle = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$.

Note that in the example above, either one of $\mathbf{v}_1$, $\mathbf{v}_3$ and $\mathbf{v}_4$ can be reduced since it can be expressed in terms of the other two vectors. However, $\mathbf{v}_2$ cannot be reduced because $\mathbf{v}_2$ has a zero coefficient.

3.5.2 Elimination in Linearly Dependent Sets

Since there are many ways to reduce a set into a span, the following theorem provides an effective way to perform elimination in linearly dependent sets:

Theorem 3.39. Let $S = \{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n\}$ be a set of column vectors and A be the matrix whose columns are vectors in S. Further let B be the reduced row-echelon form of A with $D = \{d_1, d_2, \cdots, d_r\}$ be the set of indices of pivot columns of B, then span of S, $\langle S \rangle = \langle \{\mathbf{v}_{d_1}, \mathbf{v}_{d_2}, \cdots, \mathbf{v}_{d_r}\} \rangle$.

Example 3.40. Reduce

$$R = \left\{ \begin{bmatrix} 1 \\ 1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 2 \\ 4 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 7 \\ 1 \\ -1 \\ 4 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \\ 5 \\ 1 \end{bmatrix} \right\}$$

Answer. By placing the column vectors in R into a matrix and row-reducing the matrix,

$$\begin{bmatrix} 1 & 2 & 2 & 7 & 0 \\ 1 & 2 & 0 & 1 & 2 \\ 2 & 4 & -1 & -1 & 5 \\ 1 & 2 & 1 & 4 & 1 \end{bmatrix} \xrightarrow{\text{RREF}} \begin{bmatrix} 1 & 2 & 0 & 1 & 2 \\ 0 & 0 & 1 & 3 & -1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

and along with *Theorem 3.39*, span of R

$$\langle R \rangle = \left\langle \left\{ \begin{bmatrix} 1 \\ 1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ -1 \\ 1 \end{bmatrix} \right\} \right\rangle$$

Example 3.41. Is the vector

$$\mathbf{y} = \begin{bmatrix} 9 \\ 2 \\ 1 \\ -3 \end{bmatrix}$$

within the span of

$$R = \left\{ \begin{bmatrix} 2 \\ 1 \\ 3 \\ 2 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} -8 \\ -1 \\ -9 \\ -4 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \\ -1 \\ -2 \end{bmatrix}, \begin{bmatrix} -10 \\ -1 \\ -1 \\ 4 \end{bmatrix} \right\}$$

Answer. Following the process in *Example 3.40*,

$$\langle R \rangle = \left\langle \left\{ \begin{bmatrix} 2 \\ 1 \\ 3 \\ 2 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \\ -1 \\ -2 \end{bmatrix} \right\} \right\rangle$$

By expressing $\langle R \rangle$ and $\mathbf{y}$ in an augmented matrix and row-reducing the matrix,

$$\begin{bmatrix} 2 & -1 & 3 & 9 \\ 1 & 1 & 1 & 2 \\ 3 & 0 & -1 & 1 \\ 2 & 1 & -2 & -3 \end{bmatrix} \xrightarrow{\text{RREF}} \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Therefore $\mathbf{y}$ can be expressed in terms of vectors in $\langle R \rangle$, which is

$$\mathbf{y} = \begin{bmatrix} 2 \\ 1 \\ 3 \\ 2 \end{bmatrix} - \begin{bmatrix} -1 \\ 1 \\ 0 \\ 1 \end{bmatrix} + 2 \begin{bmatrix} 3 \\ 1 \\ -1 \\ -2 \end{bmatrix}$$

Note that the example above implicitly shows that $\mathbf{y}$ in a linear combination of vectors in $\langle R \rangle$ is actually unique. This is a consequence of linear independence in the set of vectors in $\langle R \rangle$.

3.6 Column and Row Space

3.6.1 Column Space

Definition 3.42. Let A be an $m \times n$ matrix with $\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_n$ as columns, then the **column space** of A, denoted by $C(A)$, is the subset of $\mathbb{R}^m$ containing all linear combinations of columns of A. In other words,

$$C(A) = \langle \{ \mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_n \} \rangle$$

Theorem 3.43. Let A be an $m \times n$ matrix and $\mathbf{b}$ be a vector of size m, then $\mathbf{b} \in C(A)$ if and only if $LS(A, \mathbf{b})$ is consistent.

Since there is the theorem above, an alternative way to define column space of A is

$$C(A) = \{ A\mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n \} \subset \mathbb{R}^m$$

Example 3.44. Let

$$A = \begin{bmatrix} 1 & 2 & 7 & 1 & -1 \\ 1 & 1 & 3 & 1 & 0 \\ 3 & 2 & 5 & -1 & 9 \\ 1 & -1 & -5 & 2 & 0 \end{bmatrix}$$

Find the column space of A.

Answer. Let

$$\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{bmatrix}$$

By row-reducing the augmented matrix $[A\,|\,\mathbf{v}]$, the augmented matrix becomes

$$\begin{bmatrix} 1 & 0 & -1 & 0 & 3 & -3v_1 + 5v_2 - v_4 \\ 0 & 1 & 4 & 0 & -1 & v_1 - v_2 \\ 0 & 0 & 0 & 1 & -2 & 2v_1 - 3v_2 + v_4 \\ 0 & 0 & 0 & 0 & 0 & 9v_1 - 16v_2 + v_3 + 4v_4 \end{bmatrix}$$

and is in reduced row-echelon form if and only if $9v_1 - 16v_2 + v_3 + 4v_4 = 0$, since the last column is not a pivot column. If $9v_1 - 16v_2 + v_3 + 4v_4 \neq 0$, the corresponding system of linear equations is inconsistent. Therefore,

$$C(A) = N\left(\begin{bmatrix} 9 \\ -16 \\ 1 \\ 4 \end{bmatrix}\right)$$

On the other hand, the column space can be written in terms of set of pivot columns, which is

$$C(A) = \langle \{\mathbf{A}_1, \mathbf{A}_2, \mathbf{A}_4\} \rangle$$

With the example above, one can relate column space with span, and it is given by the following theorem:

Theorem 3.45. Let A be an $m \times n$ matrix with columns $\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_n$, and B be its row-equivalent matrix in reduced row-echelon form. Further let $D = \{d_1, d_2, \cdots, d_r\}$ be the set of indices for pivot columns of B, and $T = \{\mathbf{A}_{d_1}, \mathbf{A}_{d_2}, \cdots, \mathbf{A}_{d_r}\}$, then $C(A) = \langle T \rangle$.

Theorem 3.46. Let A be a square matrix of size n, then A is nonsingular if and only if $C(A) = \mathbb{R}^n$.

3.6.2 Row Space

Definition 3.47. Let A be an $m \times n$ matrix, then the **row space** of A, denoted by $R(A)$, is given by the column space of transpose matrix of A. In other words, $R(A) = C(A^t)$.

Proposition 3.48. Let A be a matrix, then $C(A) = R(A^t)$.

With the definition above, it is fairly easy to find row space with the technique of finding column space. However, $C(A)$ and $R(A)$ are rarely equal, at least A has to be a square matrix.

Example 3.49. Find the row space of

$$A = \begin{bmatrix} 1 & 4 & 0 & 1 & 0 & 7 & -9 \\ 2 & 8 & -1 & 3 & 9 & -13 & 7 \\ 0 & 0 & 2 & -3 & -4 & 12 & -8 \\ -1 & -4 & 2 & 4 & 8 & -31 & 37 \end{bmatrix}$$

Answer. By finding the transpose of A,

$$A^t = \begin{bmatrix} 1 & 2 & 0 & -1 \\ 4 & 8 & 0 & -4 \\ 0 & -1 & 2 & 2 \\ 1 & 3 & -3 & 4 \\ 0 & 9 & -4 & 8 \\ 7 & -13 & 12 & -31 \\ -9 & 7 & -8 & 37 \end{bmatrix}$$

Using the technique of finding the column space,

$$R(A) = C(A^t) = \left\langle \left\{ \begin{bmatrix} 1 \\ 4 \\ 0 \\ 1 \\ 0 \\ 7 \\ -9 \end{bmatrix}, \begin{bmatrix} 2 \\ 8 \\ -1 \\ 3 \\ 9 \\ -13 \\ 7 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 2 \\ -3 \\ -4 \\ 12 \\ -8 \end{bmatrix} \right\} \right\rangle$$

In fact, the definition of row space, even when there is already defintion of column space, has its own use. Consider the following theorem:

Theorem 3.50. If A and B are row-equivalent matrices, then their row spaces are equal, or $R(A) = R(B)$.

Theorem 3.51. Let A be a matrix and B be a row-equivalent matrix of A in reduced row-echelon form. Further let S be the set of nonzero columns of B^t, then $R(A) = \langle S \rangle$ and S is linearly independent.

One of the applications of row spaces is to improve a span, even the vectors given cannot be further reduced explicitly.

Example 3.52. Improve the span

$$X = \left\langle \left\{ \begin{bmatrix} 1 \\ 2 \\ 1 \\ 6 \\ 6 \end{bmatrix}, \begin{bmatrix} 3 \\ -1 \\ 2 \\ -1 \\ 6 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 0 \\ -1 \\ -2 \end{bmatrix}, \begin{bmatrix} -3 \\ 2 \\ -3 \\ 6 \\ -10 \end{bmatrix} \right\} \right\rangle$$

Answer. Let A be the matrix whose rows are vectors in the set of X,

$$A = \begin{bmatrix} 1 & 2 & 1 & 6 & 6 \\ 3 & -1 & 2 & -1 & 6 \\ 1 & -1 & 0 & -1 & -2 \\ -3 & 2 & -3 & 6 & -10 \end{bmatrix}$$

By row-reducing the matrix A,

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

By *Theorem 3.51*, X can be improved to nonzero columns of B^t, which is

$$X = R(A) = R(B) = \left\langle \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \\ 2 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 3 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ -2 \\ 5 \end{bmatrix} \right\} \right\rangle$$

Example 3.53. Let

$$A = \begin{bmatrix} 1 & 4 & 0 & 1 & 0 & 7 & -9 \\ 2 & 8 & -1 & 3 & 9 & -13 & 7 \\ 0 & 0 & 2 & -3 & -4 & 12 & -8 \\ -1 & -4 & 2 & 4 & 8 & -31 & 37 \end{bmatrix}$$

and

$$\mathbf{b} = \begin{bmatrix} 3 \\ 9 \\ 1 \\ 4 \end{bmatrix}$$

Find out whether $\mathbf{b}$ is in column space of A or not.

Answer. Using *Proposition 3.48*, by row-reducing the transpose matrix of A,

$$B^t = \begin{bmatrix} 1 & 0 & 0 & -\frac{31}{7} \\ 0 & 1 & 0 & \frac{12}{7} \\ 0 & 0 & 1 & \frac{13}{7} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

and hence the column space of A

$$C(A) = R(B^t) = \left\langle \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \\ -\frac{31}{7} \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ \frac{12}{7} \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ \frac{13}{7} \end{bmatrix} \right\} \right\rangle$$

by *Theorem 3.51*. Finally, in order to check whether $\mathbf{b}$ is in column space of A, simply check whether both sides of the substituted equation are equal or not. It turns out that

$$3\left(-\frac{31}{7}\right) + 9\left(\frac{12}{7}\right) + \left(\frac{13}{7}\right) = 4$$

is correct, so $\mathbf{b} \in C(A)$.

3.7 Bases

3.7.1 Definition of Bases

Definition 3.54. Let V be a vector space, then S as a subset of V is said to be a **basis** for V if S is linearly independent and $\langle S \rangle = V$. In other words, S can span (or generate) V.

Definition 3.55. Let $V = \mathbb{R}^m$, then $S = \{\mathbf{e}_1, \mathbf{e}_2, \cdots, \mathbf{e}_m\}$ (where $\mathbf{e}_i$ is a vector whose entries are all zero except the i-th entry is 1) is a **standard basis** for V.

Note that bases of a vector space is not unique. For example, if $V = \mathbb{R}^2$, then

$$\left\{\begin{bmatrix}1\\0\end{bmatrix}, \begin{bmatrix}0\\1\end{bmatrix}\right\} \text{ and } \left\{\begin{bmatrix}1\\0\end{bmatrix}, \begin{bmatrix}1\\1\end{bmatrix}\right\}$$

are some possible bases for V.

3.7.2 Bases for Spans of Column Vectors

Let $S = \{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n\}$ be a subset of R^m. In previous sections, there are two methods to find a subset T of S such that T is linearly independent and is a basis for $\langle S \rangle$:

(a) Let A be a matrix with columns of vectors in S. Find the set of vectors T by finding the set of indices of pivot columns of A in reduced row-echelon form, then T is a basis for $\langle S \rangle$.

(b) Let A be a matrix with columns of vectors in S. Find the set of vectors T by finding the transpose of nonzero columns of A^t in reduced row-echelon form, then T is a basis for $\langle S \rangle$.

Below is an example demonstrating the methods:

Example 3.56. Find a basis for

$$S = \left\langle \left\{ \begin{bmatrix}1\\2\\0\\-1\end{bmatrix}, \begin{bmatrix}4\\8\\0\\-4\end{bmatrix}, \begin{bmatrix}0\\-1\\2\\2\end{bmatrix}, \begin{bmatrix}-1\\3\\-3\\4\end{bmatrix}, \begin{bmatrix}0\\9\\-4\\8\end{bmatrix}, \begin{bmatrix}7\\-13\\12\\-31\end{bmatrix}, \begin{bmatrix}-9\\7\\-8\\37\end{bmatrix} \right\} \right\rangle$$

Answer. Part (a) Row-reducing the corresponding matrix gives

$$\begin{bmatrix} 1 & 4 & 0 & 0 & 2 & 1 & -3 \\ 0 & 0 & 1 & 0 & 1 & -3 & 5 \\ 0 & 0 & 0 & 1 & 2 & -6 & 6 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Therefore

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

is a basis for S.

Part (b) Row-reducing the corresponding transpose matrix gives

$$\begin{bmatrix} 1 & 0 & 0 & -\frac{31}{7} \\ 0 & 1 & 0 & \frac{12}{7} \\ 0 & 0 & 1 & \frac{13}{7} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Therefore

$$\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \\ -\frac{31}{7} \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ \frac{12}{7} \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ \frac{13}{7} \end{bmatrix} \right\}$$

is a basis for S.

Theorem 3.57. Let S be a finite subset of $\mathbb{R}^m$, then basis for $\langle S \rangle$ exists. Furthermore, there exists a subset T of S such that T is a basis for $\langle S \rangle$.

3.7.3 Bases and Nonsingular Matrices

Theorem 3.58. Let A be a square matrix of size n, then the set of columns of A is a basis for $\mathbb{R}^n$ if and only if A is nonsingular.

Example 3.59. Since

$$A = \begin{bmatrix} 2 & 2 & 3 \\ -1 & 6 & -4 \\ 8 & 0 & 1 \end{bmatrix}$$

is nonsingular, the set of columns of A is a basis for $\mathbb{R}^3$.

3.7.4 Dimension

Definition 3.60. Let V be a vector space and S is a basis for V, then V is a **finite dimensional vector space**, and the **dimension** of V, denoted by $\dim(V)$, is the size of the set S.

In order to make the definition of dimension well-defined, if A of size m and B of size n are both bases for a vector space, then $m = n$. Also note that $\dim(\mathbb{R}^n) = n$.

Theorem 3.61. A set S whose vectors are all from a vector space V is linearly dependent if its size is larger than the dimension of V.

Proposition 3.62. Let V be a vector space and $\mathbf{u}, \mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n \in V$. If $S = \{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n\}$ is linearly independent and $\mathbf{u} \notin \langle S \rangle$, then $S' = \{\mathbf{u}, \mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n\}$ is also linearly independent.

Theorem 3.63. If V is a subspace of $\mathbb{R}^n$, then there exists a basis for V.

Proposition 3.64. Let $S = \{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n\}$ be a subset of a finite dimensional vector space V, then $\dim \langle S \rangle \leq n$.

Proof. By *Theorem 3.57*, there exists a subset T of S such that T is a basis for $\langle S \rangle$, then dimension of $\langle S \rangle$ equals to the size of T, which is smaller than or equal to the size of S.

3.7.5 Rank and Nullity

Definition 3.65. Let $A \in M_{mn}$, then the **nullity** of A, denoted by $n(A)$, is the dimension of null space of A, and **rank** of A, denoted by $r(A)$, is the dimension of column space of A. In other words, $n(A) = \dim(N(A))$ and $r(A) = \dim(C(A))$.

Theorem 3.66. Let $A \in M_{mn}$ be a matrix and B be a row-equivalent matrix of A in reduced row-echelon form, then rank of A is the number of pivot columns in B and nullity of A is the number of free variables in B.

Corollary. Let $A \in M_{mn}$ be a matrix, then $r(A) + n(A) = n$.

Theorem 3.67. Let A be an $m \times n$ matrix, then $r(A) = r(A^t)$. Equivalently, $\dim(C(A)) = \dim(R(A))$.

Proof. Let B be a row-equivalent matrix of A in reduced row-echelon form with r pivot columns, then $r = r(A)$. By *Theorem 3.51*, the first r columns of B^t form a basis for row space of A, which is $r = r(A^t)$. Combining both equations, $r(A) = r(A^t)$.

The following proposition will conclude the section of rank and nullity:

Proposition 3.68. Let A be a square matrix of size n, then the following are equivalent:

(a) A is nonsingular.

(b) A can be row-reduced to the identity matrix I.

(c) Null space of A contains only the zero vector, or $N(A) = \{\mathbf{0}\}$.

(d) The linear system $LS(A, \mathbf{b})$ has unique solution for all possible $\mathbf{b}$.

(e) Columns of A form a linear independent set.

(f) A is invertible.

(g) Column space of A is $\mathbb{R}^n$, or $C(A) = \mathbb{R}^n$.

(h) Columns of A form a basis for $\mathbb{R}^n$.

(i) Rank of A is n, or $r(A) = n$.

(j) Nullity of A is zero, or $n(A) = 0$.

3.8 Linear Transformation

3.8.1 Matrices as Linear Transformation

There are a lot of types of linear transformations, so the left multiplication is first introduced:

Definition 3.69. Let A be an $m \times n$ matrix and $\mathbf{v}$ be a vector of size n, then the **left multiplying map** of A, denoted by L_A, is defined as

$$L_A(\mathbf{v}) = A\mathbf{v} \quad , L_A : \mathbb{R}^n \to \mathbb{R}^m$$

Example 3.70. Let

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

and

$$\mathbf{v} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

then the left multiplying map L_A is from $\mathbb{R}^3$ to $\mathbb{R}^4$, where

$$L_A(\mathbf{v}) = \begin{bmatrix} y - z \\ z - x \\ x - y \\ x + y + z \end{bmatrix}$$

Note that left multiplication has the following properties:

Proposition 3.71. Let A be an $m \times n$ matrix, $\mathbf{u}, \mathbf{v}$ be a vector of size n and $\alpha \in \mathbb{R}$, then the following properties are satisfied:

(a) $L_A(\mathbf{u} + \mathbf{v}) = L_A(\mathbf{u}) + L_A(\mathbf{v})$

(b) $L_A(\alpha\mathbf{v}) = \alpha L_A(\mathbf{v})$

With the proposition above, the definition of a linear transformation is given as follows:

Definition 3.72. Let V, W be vector spaces and T is a map from V to W, then T is said to be a **linear transformation** if the following properties are satisfied:

(a) $T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$

(b) $T(\alpha\mathbf{v}) = \alpha T(\mathbf{v})$

where $\mathbf{u}, \mathbf{v} \in V$ and $\alpha \in \mathbb{R}$.

3.8.2 Geometric Interpretations

In this section some linear transformations will be performed to a point in the xy-plane.

Example 3.73. Denote a point (x, y) as

$$\begin{bmatrix} x \\ y \end{bmatrix}$$

below are some types of linear transformations:

(a) **Rotation**
In order to rotate a point by θ in anticlockwise direction, the transformation matrix is

$$A = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}$$

(b) **Reflection**
In order to reflect a point through the x-axis, the transformation matrix is

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

Similarly, to reflect a point through the y-axis, the transformation matrix is

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

(c) **Stretching**
In order to stretch x-axis by a factor of p and y-axis by a factor of q, the transformation matrix is

$$A = \begin{bmatrix} p & 0 \\ 0 & q \end{bmatrix}$$

3.8.3 Coordinate Vector

Theorem 3.74. Let $\mathbf{v} \in V$ defined as a linear combination

$$\mathbf{v} = \alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \cdots + \alpha_n \mathbf{v}_n$$

where $\{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n\}$ form a basis for V, then every constant in the linear combination is unique.

With the theorem above, the following definition can be introduced:

Definition 3.75. Let V be a vector space, $B = \{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n\}$ be a basis for V and $\mathbf{v} \in V$, then the coordinate vector of $\mathbf{v}$ relative to the basis B is defined as

$$\rho_B(\mathbf{v}) = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix}$$

where

$$\mathbf{v} = \alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \cdots + \alpha_n \mathbf{v}_n$$

Example 3.76. Let

$$B = \left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \right\}$$

be a basis for $\mathbb{R}^3$, find

$$\rho_B \left(\begin{bmatrix} 3 \\ 4 \\ 3 \end{bmatrix} \right)$$

Answer. By solving a system of linear equations,

$$\begin{bmatrix}3\\4\\3\end{bmatrix} = 2\begin{bmatrix}1\\1\\0\end{bmatrix} + \begin{bmatrix}1\\2\\3\end{bmatrix}$$

therefore

$$\rho_B\left(\begin{bmatrix}3\\4\\3\end{bmatrix}\right) = \begin{bmatrix}2\\0\\1\end{bmatrix}$$

3.8.4 Change of Basis

Definition 3.77. Let V be a vector space and $B = \{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n\}, B' = \{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n\}$ be bases for V, then the matrix of **change of basis** from B to B' is defined as

$$C_{B,B'} = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} \end{bmatrix}$$

where

$$a_{i,j} = [\rho_{B'}(\mathbf{u}_j)]_i$$

Theorem 3.78. Let B, B' be bases for V, and $\mathbf{v} \in V$, then

$$\rho_{B'}(\mathbf{v}) = C_{B,B'}\rho_B(\mathbf{v})$$

Example 3.79. Using the assumption in *Example 3.76*, and further let

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

Find $\rho_{B'}\left(\begin{bmatrix}3\\4\\3\end{bmatrix}\right)$ by change of basis.

Answer. Note that

$$C_{B,B'} = \begin{bmatrix} 1 & 2 & 6 \\ 0 & 0 & -1 \\ 1 & 1 & 2 \end{bmatrix}$$

therefore

$$\rho_{B'}\left(\begin{bmatrix} 3 \\ 4 \\ 3 \end{bmatrix}\right) = \begin{bmatrix} 1 & 2 & 6 \\ 0 & 0 & -1 \\ 1 & 1 & 2 \end{bmatrix} \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 8 \\ -1 \\ 4 \end{bmatrix}$$

4 Matrix Properties

4.1 Inverses

4.1.1 Solution Inverse

Consider a simple case of solving $2x = 4$, if x can be made into subject, which is multiplying both sides with $\frac{1}{2}$ and giving $x = 2$, the solution is much easier to calculate. In the same sense, for $A\mathbf{x} = \mathbf{b}$, if $\mathbf{x}$ can be made into subject by left-multiplying both sides with a matrix B such that BA equals to the identity matrix I, then $\mathbf{x} = B\mathbf{b}$ is much easier to calculate.

In the above cases, $\frac{1}{2}$ is said to be the **multiplicative inverse** (or **solution inverse**, if considered $2x = 4$ as the case) of 2. In the following section, matrix inverse will be introduced to relate A with B.

4.1.2 Matrix Inverse

Definition 4.1. Let A and B be square matrices with size n, then B is the **matrix inverse** of A, denoted by A^{-1}, if the $AB = BA = I_n$. If A has a matrix inverse, A is said to be **invertible**.

Example 4.2. Let

$$A = \begin{bmatrix} 1 & -1 & 2 \\ 2 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix}$$

By row-reducing the augmented matrix,

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

shows that the inverse of A does not exist to solve the linear system $LS(A, \mathbf{b})$. Therefore, A is not invertible.

Example 4.3. Let

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & -1 \\ 0 & 1 & 1 \end{bmatrix}, B = \begin{bmatrix} 1 & 0 & -1 \\ -1 & 1 & 2 \\ 1 & -1 & -1 \end{bmatrix}$$

Since $AB = BA = I_3$, B is the inverse of A.

4.1.3 Computation of Inverses

For simplicity, matrix inverse for square matrix of size 2 will be considered first:

Theorem 4.4. Let

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

be a square matrix of size 2, then A is invertible if and only if $ad - bc \neq 0$. When A is invertible, its inverse

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

Example 4.5. Find the inverse of

$$A = \begin{bmatrix} 1 & 2 \\ -2 & 4 \end{bmatrix}$$

Answer. Since $ad - bc = 8 \neq 0$, A has an inverse. By *Theorem 4.4*,

$$A^{-1} = \frac{1}{8} \begin{bmatrix} 4 & -2 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & -\frac{1}{4} \\ \frac{1}{4} & \frac{1}{8} \end{bmatrix}$$

For square matrices of larger sizes, an equation cannot simply describe their inverses, so the following method is introduced:

Theorem 4.6. Let A be a nonsingular square matrix of size n, then the inverse of A is the last n columns of augmented matrix $[A \,|\, I_n]$ in reduced row-echelon form.

Example 4.7. Find the inverse of

$$A = \begin{bmatrix} -7 & -6 & -12 \\ 5 & 5 & 7 \\ 1 & 0 & 4 \end{bmatrix}$$

The augmented matrix for finding the inverse according to *Theorem 4.6* is

$$\begin{bmatrix} -7 & -6 & -12 & 1 & 0 & 0 \\ 5 & 5 & 7 & 0 & 1 & 0 \\ 1 & 0 & 4 & 0 & 0 & 1 \end{bmatrix}$$

Row-reduce the augmented matrix gives

$$\begin{bmatrix} 1 & 0 & 0 & -10 & -12 & -9 \\ 0 & 1 & 0 & \frac{13}{2} & 8 & \frac{11}{2} \\ 0 & 0 & 1 & \frac{5}{2} & 3 & \frac{5}{2} \end{bmatrix}$$

Therefore

$$A^{-1} = \begin{bmatrix} -10 & -12 & -9 \\ \frac{13}{2} & 8 & \frac{11}{2} \\ \frac{5}{2} & 3 & \frac{5}{2} \end{bmatrix}$$

Note that the left side of the augmented matrix in reduced row-echelon form should be an identity matrix. If it is impossible to make an identity matrix with any row operations, the matrix is not invertible.

4.1.4 Properties of Matrix Inverses

Proposition 4.8. If a square matrix A has an inverse, the inverse is unique.

Proof. Suppose B and C are inverses of A, then

$$B = BI = B(AC) = (BA)C = IC = C$$

which means inverse of A is unique.

Proposition 4.9. If A and B are invertible matrices of the same size, then $(AB)^{-1} = B^{-1}A^{-1}$.

Proof. Note that

$$(AB)(B^{-1}A^{-1}) = (B^{-1}A^{-1})(AB) = I$$

Proposition 4.10. If A has an inverse A^{-1}, then A^{-1} also have an inverse and the inverse is A, or $(A^{-1})^{-1} = A$.

Proof. Note that

$$AA^{-1} = (A^{-1})^{-1}A^{-1} = I$$

shows that $(A^{-1})^{-1} = A$.

Proposition 4.11. If A has an inverse A^{-1}, then A^t also have an inverse and $(A^t)^{-1} = (A^{-1})^t$.

Proof. Note that

$$(A^{-1})^t A^t = (A^{-1}A)^t = I^t = I$$

shows that $(A^{-1})^t$ is the inverse of A^t.

Proposition 4.12. If A has an inverse A^{-t}, then for any nonzero constant α, αA has an inverse and $(\alpha A)^{-1} = \frac{1}{\alpha}A^{-1}$.

Proof. Note that

$$(\frac{1}{\alpha}A^{-1})(\alpha A) = (\alpha A)(\frac{1}{\alpha}A^{-1}) = I$$

shows that $\frac{1}{\alpha}A^{-1}$ is the inverse of αA.

Although there are many properties of matrix inverses, this does not mean any kind of operation is applicable. For example, $(A + B)^{-1} \neq A^{-1} + B^{-1}$ for some matrices A and B.

4.1.5 Invertibility and Nonsingular Matrices

Consider α and β be real numbers, if α and β are nonzero, their product $\alpha\beta$ is also nonzero. However, if one of them are zero, their product will be equal to zero. Nonsingular matrices have similar results:

Theorem 4.13. Let A and B be square matrices of size n, then AB is nonsingular if and only if A and B are both nonsingular.

Theorem 4.14. Let A and B be square matrices of size n. If $AB = I$, then $BA = I$.

Proof. By *Theorem 4.13*, A and B are nonsingular. Let C be a square matrix of size n such that $BC = I$. For now, A is the left inverse of B and C is the right inverse of B. Note that

$$BA = BA(BC) = B(AB)C = BC = I$$

which means A is also the right inverse of B.

Theorem 4.15. Let A be a square matrix, then A is nonsingular if and only if A is invertible.

With the theorem above, if a matrix is found to be singular, it is a waste of time finding its inverse since it does not exist. This will be very helpful in later sections.

Theorem 4.16. Let A is an invertible matrix, then the unique solution of $LS(A, \mathbf{b})$ is $A^{-1}\mathbf{b}$.

Proof. Putting $A^{-1}\mathbf{b}$ into the solution, then

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

Example 4.17. Solve

$$\begin{cases} x_1 + x_2 - x_3 + 4x_4 = 1 \\ x_1 - x_2 + 2x_3 + 3x_4 = 2 \\ 2x_1 + x_2 + x_3 + x_4 = 0 \\ 2x_1 + 2x_2 + 2x_3 - 9x_4 = -1 \end{cases}$$

Answer. Let A be the coefficient matrix

$$\begin{bmatrix} 1 & 1 & -1 & 4 \\ 1 & -1 & 2 & 3 \\ 2 & 1 & 1 & 1 \\ 2 & 2 & 2 & -9 \end{bmatrix}$$

Its inverse is

$$A^{-1} = \begin{bmatrix} -33 & -22 & 45 & -17 \\ 35 & 23 & -47 & 18 \\ 25 & 17 & -34 & 13 \\ 6 & 4 & -8 & 3 \end{bmatrix}$$

and by *Theorem 4.16*, the solution is

$$A^{-1} \begin{bmatrix} 1 \\ 2 \\ 0 \\ -1 \end{bmatrix} = \begin{bmatrix} -60 \\ 63 \\ 46 \\ 11 \end{bmatrix}$$

4.2 Matrix Determinants

4.2.1 Definition of Determinant

Definition 4.18. Let A be an $m \times n$ matrix, then the submatrix $A(i \mid j)$ is the $(m-1) \times (n-1)$ matrix which equals to A without row i and column j.

Example 4.19. Let

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{bmatrix}$$

then

$$A(1\,|\,2) = \begin{bmatrix} 5 & 7 & 8 \\ 9 & 11 & 12 \end{bmatrix}, A(2\,|\,1) = \begin{bmatrix} 2 & 3 & 4 \\ 10 & 11 & 12 \end{bmatrix}$$

With the definition of submatrices, determinant can be defined in a recursive way:

Definition 4.20. Let A be a square matrix of size n, then the **determinant** of A, denoted by $\det(A)$ or $|A|$, is defined in the following methods:

(a) If $n = 1$, then $\det(A)$ is simply the value inside the matrix.

(b) If $n \geq 2$, then

$$\det(A) = \sum_{i=1}^{n} \left((-1)^{i+1}[A]_{1,i} \det(A(1\,|\,i))\right)$$

Theorem 4.21. Let

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

be a square matrix of size 2, then

$$det(A) = ad - bc$$

Proof. Note that

$$det(A) = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = a\,|d| - b\,|c| = ad - bc$$

Theorem 4.22. Let

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

be a square matrix of size 3, then

$$det(A) = aei + bfg + cdh - ceg - bdi - afh$$

4.2.2 Computation of Determinants

There are mainly two ways to compute determinants of large matrices: expansion and reduction.

Theorem 4.23. Let A be a square matrix of size n, then by **Laplace expansion**, determinant of A can be determined in one of the two methods:

(a) **Expansion about Rows**
By expanding about row i,

$$\begin{aligned}\det(A) =&(-1)^{i+1}[A]_{i,1}\det(A(i\,|\,1)) + (-1)^{i+2}[A]_{i,2}\det(A(i\,|\,2)) \\ &+ \cdots + (-1)^{i+n}[A]_{i,n}\det(A(i\,|\,n))\end{aligned}$$

(b) **Expansion about Columns**
By expanding about column j,

$$\begin{aligned}\det(A) =&(-1)^{1+j}[A]_{1,j}\det(A(1\,|\,j)) + (-1)^{2+j}[A]_{2,j}\det(A(2\,|\,j)) \\ &+ \cdots + (-1)^{n+j}[A]_{n,j}\det(A(n\,|\,j))\end{aligned}$$

Theorem 4.24. Let A be an upper triangular matrix, which is

$$[A]_{i,j} = 0 \quad \text{if } i > j$$

then the determinant of A

$$det(A) = \sum_{i=1}^{n}[A]_{i,i}$$

Corollary. Let A be a lower triangular matrix, which is

$$[A]_{i,j} = 0 \quad \text{if } i < j$$

then the determinant of A

$$det(A) = \sum_{i=1}^{n} [A]_{i,i}$$

4.2.3 Properties of Matrix Determinants

Theorem 4.25. Let A be a square matrix, then $\det(A) = \det(A^t)$.

In fact, the theorem above shows why the corollary of *Theorem 4.24* is valid.

Theorem 4.26. Let A be a square matrix with a row or a column where every entry is 0, then $\det(A) = 0$.

Proof. By performing Laplace expansion about the zero row or zero column, every term in the determinant is multiplied with 0, which means the determinant is 0.

Recall the row operations, and the following theorem gives the effect of the determinant:

Theorem 4.27. Let A be a square matrix, and B be the matrix after a row operation, then the following applies:

(a) $R_i \leftrightarrow R_j$: $\det(B) = -\det(A)$.

(b) αR_i : $\det(B) = \alpha \det(A)$.

(c) $\alpha R_i + R_j$: $\det(B) = \det(A)$.

Note that the theorem above also applies to column operations because of *Theorem 4.25*.

Theorem 4.28. Let A be a square matrix with two identical rows or columns, then $\det(A) = 0$.

Proof. This is a result from *Theorem 4.26* and *Theorem 4.27*. Subtract one of the identical rows (columns) will not affect the determinant but giving a zero row (column).

In fact, using the method or row reduction is efficient to find the determinant of a matrix, since a row-reduced matrix is already an upper triangular matrix.

Example 4.29. Compute the determinant of

$$\begin{bmatrix} 2 & 2a & 2a^2 \\ 2 & 2b & 2b^2 \\ 2 & 2c & 2c^2 \end{bmatrix}$$

Answer. Note that

$$\begin{aligned}\begin{vmatrix}2 & 2a & 2a^2\\ 2 & 2b & 2b^2\\ 2 & 2c & 2c^2\end{vmatrix} &= 2^3\begin{vmatrix}1 & a & a^2\\ 1 & b & b^2\\ 1 & c & c^2\end{vmatrix} = 8\begin{vmatrix}1 & 0 & 0\\ 1 & b-a & b^2-ab\\ 1 & c-a & c^2-ac\end{vmatrix}\\ &= 8\begin{vmatrix}b-a & b(b-a)\\ c-a & c(c-a)\end{vmatrix} = 8(b-a)(c-a)\begin{vmatrix}1 & b\\ 1 & c\end{vmatrix}\\ &= 8(b-a)(c-a)(c-b)\end{aligned}$$

4.3 Eigenvalues and Eigenvectors

4.3.1 Eigenvalues and Eigenvectors of Matrices

Definition 4.30. Let A be a square matrix of size n, $\mathbf{x}$ be a nonzero vector of size n and λ be a constant, then $\mathbf{x}$ is said to be an **eigenvector** with λ as **eigenvalue** if

$$A\mathbf{x} = \lambda\mathbf{x}$$

Example 4.31. Find the eigenvectors with their eigenvalues of

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

Answer. The eigenvectors are

$$\begin{bmatrix}1\\ 0\end{bmatrix}, \begin{bmatrix}1\\ -1\end{bmatrix}$$

with eigenvalues 3 and 2 respectively, which satisfies

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

The following section will show how to find the eigenvectors.

4.3.2 Existence of Eigenvalues and Eigenvectors

Observe that

$$A\mathbf{x} = \lambda\mathbf{x} \Rightarrow (A - \lambda I)\mathbf{x} = \mathbf{0}$$

In particular, the matrix $A - \lambda I$ is necessarily singular, hence $\det(A - \lambda I) = 0$. The definition below is an extent from this observation:

Definition 4.32. Let A be a square matrix, then the **characteristic polynomial** of A, denoted by p_A, is the polynomial

$$p_A(\lambda) = \det(A - \lambda I)$$

Theorem 4.33. Let A be a square matrix, then λ is an eigenvalue if and only if $p_A(\lambda) = 0$.

Theorem 4.34. Let A be a square matrix of size n, then the characteristic polynomial p_A has degree n.

Example 4.35. Find the eigenvalues of

$$A = \begin{bmatrix} -13 & -8 & -4 \\ 12 & 7 & 4 \\ 24 & 16 & 7 \end{bmatrix}$$

Answer. The characteristic polynomial

$$p_A(x) = -(x-3)(x+1)^2$$

shows that the eigenvalues are the roots of the polynomial, which is $x = 3$ and $x = -1$. Moreover, there are no eigenvalues other than the roots mentioned.

4.3.3 Computation of Eigenvalues and Eigenvectors

Definition 4.36. Let A be a square matrix and λ be an eigenvalue of A, then the **eigenspace** of A for λ, denoted by $\varepsilon_A(\lambda)$, is the set of all eigenvectors of A with eigenvalue λ, together with the zero vector.

In other words, if $\mathbf{x}$ is found to be an eigenvector of A with eigenvalue λ, then

$$\varepsilon_A(\lambda) = \{k\mathbf{x} \mid k \text{ constant}\}$$

Theorem 4.37. Let A be a square matrix and λ be an eigenvalue of A, then

$$\varepsilon_A(\lambda) = N(A - \lambda I)$$

Corollary. Every eigenspace is a subspace.

Proof. Every eigenspace is a null space, and every null space is a subspace.

However, for simplicity, usually only one eigenvector is picked (and counted) from its corresponding eigenspace.

Example 4.38. Find the eigenvectors of

$$A = \begin{bmatrix} -13 & -8 & -4 \\ 12 & 7 & 4 \\ 24 & 16 & 7 \end{bmatrix}$$

Answer. From *Example 4.35*, the eigenvalues are $\lambda = 3$ and $\lambda = -1$. When $\lambda = 3$,

$$A - 3I = \begin{bmatrix} -16 & -8 & -4 \\ 12 & 4 & 4 \\ 24 & 16 & 4 \end{bmatrix} \xrightarrow{\text{RREF}} \begin{bmatrix} 1 & 0 & \frac{1}{2} \\ 0 & 1 & -\frac{1}{2} \\ 0 & 0 & 0 \end{bmatrix}$$

Therefore

$$\varepsilon_A(3) = \left\langle \begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix} \right\rangle$$

Similarly, when $\lambda = -1$,

$$A + I = \begin{bmatrix} -12 & -8 & -4 \\ 12 & 8 & 4 \\ 24 & 16 & 8 \end{bmatrix} \xrightarrow{\text{RREF}} \begin{bmatrix} 1 & \frac{2}{3} & \frac{1}{3} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Therefore

$$\varepsilon_A(-1) = \left\langle \begin{bmatrix} -2 \\ 3 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 3 \end{bmatrix} \right\rangle$$

Example 4.39. Consider the identity matrix of size 2

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

its characteristic polynomial $p_I(x) = 1$. Notice that for any $\mathbf{x} \in \mathbb{R}^2$,

$$I\mathbf{x} = \mathbf{x}$$

4.3.4 Similar Matrices and Eigenvalues

Definition 4.40. Let A and B be square matrices of size n, then A and B are **similar** if there exists a nonsingular matrix S such that $A = S^{-1}BS$. The process of transforming B to A is called **similarity transformation**.

Theorem 4.41. Similarity transformation is an equivalence relation. For instance, let A, B and C be square matrices of size n, then the following applies:

(a) **Reflexive**
A is similar to A.

(b) **Symmetric**
If A is similar to B, B is similar to A.

(c) **Transitive**
If A is similar to B and B is similar to C, A is similar to C.

Theorem 4.42. Let A and B are similar matrices, then their characteristic polynomial are the same.

Corollary. Similar matrices have same eigenvalues.

Note that A and B may not be similar if their characteristic polynomial are the same. For example,

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

shares the same characteristic polynomial, but

$$A \neq S^{-1}BS = I$$

shows that A and B are not similar.

4.3.5 Diagonalizability

Definition 4.43. Let A be a square matrix of size n, then A is a **diagonal matrix** if $[A]_{i,j} = 0$ if $i \neq j$. In other words,

$$A = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

is a diagonal matrix, denoted by $\operatorname{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)$.

Definition 4.44. Let A be a square matrix, then A is **diagonalizable** if there is a diagonal matrix D such that A is similar to D. In other words, there exists a nonsingular matrix S such that

$$D = S^{-1}AS$$

Theorem 4.45. Let A be a square matrix of size n, then A is diagonalizable if and only if there exists a linearly independent set T that contains n eigenvectors of A.

Theorem 4.46. Let A be a square matrix of size n with eigenvalues $\lambda_1, \lambda_2, \cdots, \lambda_n$, then A is diagonalizable if and only if

$$\sum_{i=1}^{k} \dim(\varepsilon_A(\lambda_i)) = n$$

Example 4.47. Diagonalize

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

Answer. Note that the eigenvalues are $\lambda = -1$, $\lambda = 1$ and $\lambda = 3$. Their bases are

$$\left\{ \begin{bmatrix} -1 \\ -1 \\ 1 \\ 1 \end{bmatrix} \right\}, \left\{ \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ -1 \\ 1 \end{bmatrix} \right\}, \left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \right\}$$

Therefore

$$S = \begin{bmatrix} -1 & -1 & 0 & 1 \\ -1 & 1 & 0 & 1 \\ 1 & 0 & -1 & 1 \\ 1 & 0 & 1 & 1 \end{bmatrix}, D = \operatorname{diag}(-1, 1, 1, 3)$$

4.3.6 Powers of Matrices

Proposition 4.48. Let A be a diagonalizable matrix with transformation matrix S and diagonal matrix D, then for any positive integer n,

$$A^n = SD^nS^{-1}$$

Proof. Since $D = S^{-1}AS$,

$$\begin{aligned} D^n &= \underbrace{(S^{-1}AS)(S^{-1}AS)\cdots(S^{-1}AS)}_{\text{n times}} \\ SD^nS^{-1} &= A^n \end{aligned}$$

4.4 Inner Products and Orthogonaliztion

4.4.1 Inner Products

Definition 4.49. Let $\mathbf{u}$ and $\mathbf{v}$ are vectors of size n, then their **inner product**, denoted by $\langle \mathbf{u}, \mathbf{v} \rangle$, is defined as

$$\langle \mathbf{u}, \mathbf{v} \rangle = \sum_{i=1}^{n} [\mathbf{u}]_i [\mathbf{v}]_i$$

Example 4.50. Let

$$\mathbf{u} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$$

then their inner product is

$$\langle \mathbf{u}, \mathbf{v} \rangle = 1(3) + 2(4) = 11$$

Proposition 4.51. Let $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{w}$ are vectors of size n, $\alpha \in \mathbb{R}$, then the following properties apply:

(a)
$$\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle$$

(b)
$$\langle \alpha \mathbf{u}, \mathbf{v} \rangle = \alpha \langle \mathbf{u}, \mathbf{v} \rangle$$

(c)
$$\langle \mathbf{u} + \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{w} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle$$

(d)
$$\langle \mathbf{v}, \mathbf{v} \rangle \geq 0$$

Proposition 4.52. Let $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{w}$ are vectors of size n, $\alpha, \beta \in \mathbb{R}$, then the following properties apply:

(a)
$$\langle \alpha \mathbf{u} + \beta \mathbf{v}, \mathbf{w} \rangle = \alpha \langle \mathbf{u}, \mathbf{w} \rangle + \beta \langle \mathbf{v}, \mathbf{w} \rangle$$

(b)
$$\langle \mathbf{0}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{0} \rangle = 0$$

(c)
$$\langle \mathbf{u}, \mathbf{v} \rangle = 0 \quad \text{for all } \mathbf{v} \Rightarrow \mathbf{u} = \mathbf{0}$$

(d)
$$\langle \mathbf{u}, \mathbf{w} \rangle = \langle \mathbf{v}, \mathbf{w} \rangle \quad \text{for all } \mathbf{w} \Rightarrow \mathbf{u} = \mathbf{v}$$

4.4.2 Norms

Definition 4.53. Let $\mathbf{v}$ be a vector, then the **norm** (or **length**) of $\mathbf{v}$, denoted by $\|\mathbf{v}\|$, is defined as

$$\|\mathbf{v}\| = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle}$$

Proposition 4.54. Let $\mathbf{v}$ be a vector and $\alpha \in \mathbb{R}$, then the following properties apply:

(a) $\|\mathbf{v}\| = 0$ if and only if $\mathbf{v} = \mathbf{0}$.

(b) $\|\alpha \mathbf{v}\| = |\alpha| \, \|\mathbf{v}\|$.

(c) Suppose $\mathbf{v}$ is nonzero and $\alpha = \frac{1}{\|\mathbf{v}\|}$, then $\|\alpha \mathbf{v}\| = 1$.

Definition 4.55. Let $\mathbf{v}$ be a vector, then $\mathbf{v}$ is an **unit vector** if $\|\mathbf{v}\| = 1$.

Note that any nonzero vector can be transformed to a unit vector by scaling the vector by the reciprocal of its norm. Such process is called **normalization**.

Example 4.56. Normalize

$$\mathbf{v} = \begin{bmatrix} -2 \\ 1 \\ 2 \end{bmatrix}$$

Answer. Note that $\|\mathbf{v}\| = 3$, therefore the unit vector is

$$\frac{\mathbf{v}}{\|\mathbf{v}\|} = \begin{bmatrix} -\frac{2}{3} \\ \frac{1}{3} \\ \frac{2}{3} \end{bmatrix}$$

4.4.3 Orthogonal Sets

Definition 4.57. Let $\mathbf{u}$ and $\mathbf{v}$ are vectors of size n, then $\mathbf{u}$ and $\mathbf{v}$ are said to be **orthogonal** (or **perpendicular**), denoted by $\mathbf{u} \perp \mathbf{v}$, if their inner product $\langle \mathbf{u}, \mathbf{v} \rangle = 0$.

Example 4.58. Let

$$\mathbf{u} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}$$

Since $\langle \mathbf{u}, \mathbf{v} \rangle = 1(-1) + 2(-1) + 3(1) = 0$, $\mathbf{u}$ and $\mathbf{v}$ are perpendicular.

Definition 4.59. A subset $S = \{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k\}$ of nonzero vectors in $\mathbb{R}^n$ is said to be **orthogonal** if $\mathbf{v}_i \perp \mathbf{v}_j$ for all distinct i, j.

Theorem 4.60. Let S be an orthogonal subset of $\mathbb{R}^n$, then S is linearly independent.

Theorem 4.61. Let $S = \{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k\}$ be an orthogonal subset of $\mathbb{R}^n$ and $\mathbf{u}$ be a vector. If $\mathbf{u}$ is within the span of S, $\mathbf{u}$ can be expressed as the linear combination

$$\mathbf{u} = \sum_{i=1}^{k} \frac{\langle \mathbf{u}, \mathbf{v}_i \rangle}{\|\mathbf{v}_i\|^2} \mathbf{v}_i$$

The technique in the theorem above is called **projection** of matrix, and will be useful across the sections about orthogonaliztion.

Example 4.62. Let

$$S = \left\{ \mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix} \right\}$$

be the orthogonal subset of $\mathbb{R}^3$, and

$$\mathbf{u} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

Express $\mathbf{u}$ in terms of vectors in S.

Answer. By *Theorem 4.61*,

$$\begin{aligned} \mathbf{u} &= \frac{\langle \mathbf{u}, \mathbf{v}_1 \rangle}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 + \frac{\langle \mathbf{u}, \mathbf{v}_2 \rangle}{\|\mathbf{v}_2\|^2}\mathbf{v}_2 + \frac{\langle \mathbf{u}, \mathbf{v}_3 \rangle}{\|\mathbf{v}_3\|^2}\mathbf{v}_3 \\ &= 2\mathbf{v}_1 - \frac{1}{2}\mathbf{v}_2 - \frac{1}{2}\mathbf{v}_3 \end{aligned}$$

4.4.4 Orthogonal Bases

Definition 4.63. Let V be a subspace of $\mathbb{R}^n$ and S be a subset of V, then S is said to be an **orthogonal basis** for V if S is a basis for V and S is orthogonal.

Theorem 4.64. Let V be a subspace of $\mathbb{R}^n$ and S be an orthogonal subset of V, then S is an orthogonal basis if and only if $\langle S \rangle = V$.

Corollary. If S is an orthogonal basis for V, S is a basis of $\langle S \rangle$, and for any vector $\mathbf{v} \in V$, there is always a linear combination in terms of S for the vector.

Definition 4.65. Let S be a subset of $\mathbb{R}^n$, then S is said to be **orthonormal** if S is orthogonal and every vector in S is a unit vector. Furthermore, S is said to be an **orthonormal basis** for V if S is a basis for V and S is orthonormal.

Theorem 4.66. Let $S = \{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k\}$ be an orthonormal subset of $\mathbb{R}^n$ and $\mathbf{u}$ be a vector within the span of S, then

$$\mathbf{u} = \sum_{i=1}^{k} \langle \mathbf{u}, \mathbf{v}_i \rangle \mathbf{v}_i$$

Proof. This is a result from *Theorem 4.61* where the norms are all 1.

4.4.5 Gran-Schmidt Orthogonaliztion Process

Consider the case where a vector $\mathbf{u}$ is not within an orthogonal subset, then the theorem in previous section cannot be used. Luckily, there is a way to fix such problem.

Theorem 4.67. Let $S = \{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k\}$ be an orthogonal subset of $\mathbb{R}^n$ and $\mathbf{u}$ be a vector in $\mathbb{R}^n$, then the vector

$$\mathbf{u}' = \mathbf{u} - \sum_{i=1}^{k} \frac{\langle \mathbf{u}, \mathbf{v}_i \rangle}{\|\mathbf{v}_i\|^2} \mathbf{v}_i$$

is perpendicular to all vectors in S.

With the theorem above, the **Gran-Schmidt orthogonaliztion process** is introduced:

Theorem 4.68. Let $S = \{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k\}$ be an linearly independent subset of V. Further let

$$\mathbf{u}_i = \begin{cases} \mathbf{v}_1 & \text{if } i = 1 \\ \mathbf{u}_i - \sum_{j=1}^{i-1} (\langle \mathbf{u_i}, \mathbf{v}_j \rangle / \|\mathbf{v}_j\|^2) \mathbf{v}_j & \text{if } 1 < i \leq k \end{cases}$$

then $S' = \{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_k\}$ is an orthogonal set. Moreover, $\langle S' \rangle = \langle S \rangle$.

Corollary. Let V be a subspace of $\mathbb{R}^n$, then there exists an orthogonal (or orthonormal) basis for V.

References

The following are the references of the context of this document:

(a) Professor(s) associated to *MATH1030: Linear Algebra I*

(b) Robert A. Beezer, *A First Course in Linear Algebra*, Congruent Press (Version 3.50), 2015

(c) David C. Lay, Steven R. Lay, Judi J. McDonald, *Linear Algebra and Its Applications*, Pearson (5th Edition), 2018