Matrix Algebra and Applications

MATH 2111 Lecture Notes

LI, Hongrui

May 6, 2025

Contents

1 Systems of Linear Equations

1.1 Systems of Linear Equations

At the start of this course, let's take a look at the word "Linear". Among all the types of single-variable functions and equations we have learnt before, the most simple and basic one is exactly the linear function, which is a function that can be written in the form $f(x) = mx + b$. The graph of a linear function is a straight line, and the slope of the line is m. In general, when we have more than one variable, a linear equation is an equation that can be written in the form

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b$$

, where $a_1, a_2, \cdots, a_n$ and b are constants (which we call coefficients), and $x_1, x_2, \cdots, x_n$ are variables.

A system of linear equations is a set of two or more linear equations with the same variables. The solution to a system of linear equations is the set of all values of the variables that satisfy all the equations in the system. And a system of linear equations can have

1. No solution, or

2. Exactly one solution, or

3. Infinitely many solutions.

If there is no solution, we say the system is **inconsistent**; otherwise, we say the system is **consistent**. If there is exactly one solution, we say the system is **consistent and independent**. If there are infinitely many solutions, we say the system is **consistent and dependent**.

Previously, we have learnt how to solve a system of linear equations by using the method of substitution or elimination, which can be quite tedious and time-consuming when the number of equations and variables is large. Hence, we need to find a systematic and efficient way to solve a system of linear equations

In this course, we will use the language of matrices and vectors to represent and solve systems of linear equations. In particular, we will use the following notation:

- A system of linear equations will be represented by a matrix equation $A\mathbf{x} = \mathbf{b}$, where A is a matrix, $\mathbf{x}$ is a vector of variables, and $\mathbf{b}$ is a vector of constants.

- The solution to a system of linear equations will be represented by a set of vectors.

- The solution set of a system of linear equations will be represented by a subspace of $\mathbb{R}^n$, which can be denoted by a linear combination of vectors, a parametric vector form, or a set of vectors.

We will cover all the topics above in this chapter, and let's start with the matrix notation of a system of linear equations, coefficient matrix and augmented matrix.

> **Example 1.1. Matrix Notation of a System of Linear Equations**
> Consider the following system of linear equations:
>
> $$2x_1 + 3x_2 = 5$$
> $$4x_1 + 5x_2 = 6$$
>
> We can write this system of linear equations in matrix notation as
>
> $$\begin{pmatrix} 2 & 3 \\ 4 & 5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 5 \\ 6 \end{pmatrix}$$
>
> where the matrix on the left side is called the **coefficient matrix**, the matrix on the right side is called the **constant matrix**. Furthermore, we can write the **augmented matrix** of this system of linear equations as
>
> $$\begin{pmatrix} 2 & 3 & 5 \\ 4 & 5 & 6 \end{pmatrix}$$

Recall how we solve a system of linear equations by using the method of substitution or elimination before. Representing the rules we follow previously in the language of matrices, we have the **Elementary Row Operations** as follows:

1. **Interchange** Interchange the i-th row and the j-th row.

$$\begin{pmatrix} \vdots & \vdots & \vdots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \vdots & \vdots \\ a_{j1} & a_{j2} & \cdots & a_{jn} \\ \vdots & \vdots & \vdots & \vdots \end{pmatrix} \rightarrow \begin{pmatrix} \vdots & \vdots & \vdots & \vdots \\ a_{j1} & a_{j2} & \cdots & a_{jn} \\ \vdots & \vdots & \vdots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \vdots & \vdots \end{pmatrix}$$

2. **Scaling** Multiply the i-th row by a nonzero constant c.

$$\begin{pmatrix} \vdots & \vdots & \vdots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \vdots & \vdots \end{pmatrix} \rightarrow \begin{pmatrix} \vdots & \vdots & \vdots & \vdots \\ ca_{i1} & ca_{i2} & \cdots & ca_{in} \\ \vdots & \vdots & \vdots & \vdots \end{pmatrix}$$

3. **Replacement** Replace the i-th row by the sum of the i-th row and a multiple of the j-th row.

$$\begin{pmatrix} \vdots & \vdots & \vdots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \vdots & \vdots \\ a_{j1} & a_{j2} & \cdots & a_{jn} \\ \vdots & \vdots & \vdots & \vdots \end{pmatrix} \rightarrow \begin{pmatrix} \vdots & \vdots & \vdots & \vdots \\ a_{i1}+ca_{j1} & a_{i2}+ca_{j2} & \cdots & a_{in}+ca_{jn} \\ \vdots & \vdots & \vdots & \vdots \\ a_{j1} & a_{j2} & \cdots & a_{jn} \\ \vdots & \vdots & \vdots & \vdots \end{pmatrix}$$

1.2   Row Reduction and Echelon Forms

Previously, after solving a system of linear equations we may represent the solution set by

$$x = m$$
$$y = n$$

, a simple form of linear equations where m and n are constants. The ultimate goal of all your efforts spent in solving the problem is to write x, y like something above.

Similarly, when solving by using the matrix notation, we want to get a matrix in a form that is as simple as possible.

Definition 1.2. Row Echelon Form
A rectangular matrix is in **echelon form** or **row echelon form** if it has the following three properties:

1. All zero rows are at the bottom of the matrix.

2. Each leading entry of a row is in a column to the right of the leading entry of the previous row.

3. All entries in a column below a leading entry are zeros.

If a matrix in echelon form satisfies the following additional properties, it is said to be in **reduced row echelon form**:

1. The leading entry in each nonzero row is 1.

2. Each leading 1 is the only nonzero entry in its column.

Note that **leading entry** is the first nonzero entry in a row.

Definition 1.3. Pivot Position
A **pivot position** in a matrix is a location in the matrix that corresponds to a leading 1 in the reduced row echelon form of the matrix. A **pivot column** is a column that contains a pivot position.

We have a theorem that each matrix is row equivalent to one and only one reduced echelon matrix. If an echelon matrix U can be achieved by row operations from a matrix A, we say U is called an echelon form of A.

So far, we have all the theoretical fundations needed for introducing the **Row Reduction Algorithm**, the key of solving a system of linear equations by using the matrix notation.

Example 1.2. Apply elementary row operations to transform the following matrix first into echelon form and then into reduced echelon form:

$$\begin{pmatrix} 0 & 3 & -6 & 6 & 4 & -5 \\ 3 & -7 & 8 & -5 & 8 & 9 \\ 3 & -9 & 12 & -9 & 6 & 15 \end{pmatrix}$$

Solution. Step 1: Begin with the leftmost nonzero column. This is a pivot column. The pivot position is at the top.

Step 2: Select a nonzero entry in the pivot column as a pivot. If necessary, interchange rows to move this entry into the pivot position.

$$\begin{pmatrix} 3 & -9 & 12 & -9 & 6 & 15 \\ 3 & -7 & 8 & -5 & 8 & 9 \\ 0 & 3 & -6 & 6 & 4 & -5 \end{pmatrix}$$

Step 3: Use row replacement operations to create zeros in all positions below the pivot.

$$\begin{pmatrix} 3 & -9 & 12 & -9 & 6 & 15 \\ 0 & 2 & -4 & 4 & 2 & -6 \\ 0 & 3 & -6 & 6 & 4 & -5 \end{pmatrix}$$

Step 4: Cover (or ignore) the row containing the pivot position and cover all rows, if any, above it. Apply steps 1-3 to the submatrix that remains. Repeat the process until there are no more nonzero rows to modify.

$$\begin{pmatrix} 3 & -9 & 12 & -9 & 6 & 15 \\ 0 & 2 & -4 & 4 & 2 & -6 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{pmatrix}$$

Step 5: Begin with the rightmost pivot and create zeros above it.

$$\begin{pmatrix} 3 & -9 & 12 & -9 & 0 & -9 \\ 0 & 2 & -4 & 4 & 0 & -14 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{pmatrix} \rightarrow \cdots \rightarrow \begin{pmatrix} 1 & 0 & -2 & 3 & 0 & -24 \\ 0 & 1 & -2 & 2 & 0 & -7 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{pmatrix}$$

Remark. Step 1-4 are the steps to transform the matrix into echelon form, which is called the forward phase of the row reduction algorithm, and step 5 is the step to transform the matrix into reduced echelon form, which is called the backward phase.

1.3 Vector Equations

As we know, Vectors in $\mathbb{R}^n$ is an ordered list of n numbers. The handwriten notes spend a lot of time on the arithmetic of vectors, which is quite trivial. This note will skip this part and start from the Linear combinations.

Definition 1.4. In general, given vectors $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_p$ in $\mathbb{R}^n$ and given scalars $c_1, c_2, \cdots, c_p$, the vector $\mathbf{w}$

$$\mathbf{w} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_p\mathbf{v}_p$$

is called a **linear combination** of $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_p$ with weights coefficients $c_1, c_2, \cdots, c_p$.

Example 1.3. Determin whether $\mathbf{b}$ is a linear combination of $\mathbf{a}_1$ and $\mathbf{a}_2$.

$$\mathbf{a}_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \mathbf{a}_2 = \begin{pmatrix} 3 \\ 4 \end{pmatrix}, \mathbf{b} = \begin{pmatrix} 5 \\ 6 \end{pmatrix}$$

Solution. We need to find c_1 and c_2 such that $c_1\mathbf{a}_1 + c_2\mathbf{a}_2 = \mathbf{b}$. This is equivalent to solving the following system of linear equations:

$$c_1 + 3c_2 = 5$$
$$2c_1 + 4c_2 = 6$$

The augmented matrix of this system is

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

We can row reduce this matrix to

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

So $c_1 = -1$ and $c_2 = 2$. Therefore, $\mathbf{b}$ is a linear combination of $\mathbf{a}_1$ and $\mathbf{a}_2$.

In general, a vector equation

$$\mathbf{b} = x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_p\mathbf{a}_p \quad \text{or} \quad \mathbf{b} = \sum_{i=1}^{p} x_i\mathbf{a}_i$$

has the same solution set as the linear system whose augmented matrix is

$$\begin{pmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_p & \mathbf{b} \end{pmatrix}$$

In particular, $\mathbf{b}$ can be generated by a linear combination of $\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_p$ if and only if there exists a solution to the linear system corresponding to the augmented matrix above.

Definition 1.5. Span
If $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_p$ are in $\mathbb{R}^n$ and $c_1, c_2, \cdots, c_p$ are scalars, then the set of all possible linear combinations of $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_p$ is called the **span** of $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_p$ and is denoted by Span $\{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_p\}$.

That is, Span $\{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_p\}$ is the set of all vectors that can be written in the form $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_p\mathbf{v}_p$ for some scalars $c_1, c_2, \cdots, c_p$.

Example 1.4.

$$\text{Span} \left\{ \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 \\ 4 \\ 0 \end{pmatrix} \right\} = \left\{ \begin{pmatrix} x \\ y \\ 0 \end{pmatrix} \mid x, y \in \mathbb{R} \right\}$$

Next let's discuss a geometric interpretation of the span of vectors.

- Let $\mathbf{v}$ be a nonzero vector in $\mathbb{R}^n$. The span of $\mathbf{v}$ is the line through the origin in the direction of $\mathbf{v}$.

- Let $\mathbf{v}_1$ and $\mathbf{v}_2$ be two nonzero vectors in $\mathbb{R}^n$ that are not scalar multiples of each other. The span of $\mathbf{v}_1$ and $\mathbf{v}_2$ is the plane through the origin that contains $\mathbf{v}_1$ and $\mathbf{v}_2$.

Example 1.5. For what values of h will $\mathbf{y}$ be in the Span $\{\mathbf{v_1}, \mathbf{v_2}, \mathbf{v_3}\}$, if

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}, \mathbf{v}_2 = \begin{pmatrix} 5 \\ -4 \\ -7 \end{pmatrix}, \mathbf{v}_3 = \begin{pmatrix} -3 \\ 1 \\ 0 \end{pmatrix}, \text{ and } \mathbf{y} = \begin{pmatrix} -4 \\ 3 \\ h \end{pmatrix}$$

Solution. The vector $\mathbf{y}$ belongs to Span $\{\mathbf{v_1}, \mathbf{v_2}, \mathbf{v_3}\}$ if and only if there exist scalars x_1, x_2, x_3 such that

$$x_1 \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} + x_2 \begin{pmatrix} 5 \\ -4 \\ -7 \end{pmatrix} + x_3 \begin{pmatrix} -3 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} -4 \\ 3 \\ h \end{pmatrix},$$

which indicates that the following system of linear equations must have a solution:

$$x_1 + 5x_2 - 3x_3 = -4$$
$$-x_1 - 4x_2 + x_3 = 3$$
$$2x_1 - 7x_2 = h$$

The augmented matrix of this system is

$$\begin{pmatrix} 1 & 5 & -3 & -4 \\ -1 & -4 & 1 & 3 \\ 2 & -7 & 0 & h \end{pmatrix}$$

We can row reduce this matrix to

$$\begin{pmatrix} 1 & 5 & -3 & -4 \\ 0 & 1 & -2 & -1 \\ 0 & 0 & 0 & h-5 \end{pmatrix}$$

So the system has a solution if and only if $h - 5 = 0$, i.e. $h = 5$.

Example 1.6. The handwritten notes also provide examples about the application in "reality scenarios", please refer to file1.1

1.4 The Matrix Equation

Study of the Matrix Equation

Definition 1.6. Matrix-Vector Product
A is a $m \times n$ matrix with column $\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_n \in \mathbb{R}^m$, i.e. $A = [\mathbf{a}_1, \cdots, \mathbf{a}_n]$, and $\mathbf{x}$ is a vector in $\mathbb{R}^n$. Then we have

$$A \cdot \mathbf{x} = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n.$$

i.e. the product $A \cdot \mathbf{x}$ is a linear combination of the columns of A with weights given by the entries of $\mathbf{x}$.
The form $A\mathbf{x} = \mathbf{b}$ is called a **matrix equation**.

Example 1.7. Let $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix}$ and $\mathbf{x} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$. Then

$$A \cdot \mathbf{x} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \end{pmatrix} = 2 \begin{pmatrix} 1 \\ 3 \\ 5 \end{pmatrix} + 1 \begin{pmatrix} 2 \\ 4 \\ 6 \end{pmatrix} = \begin{pmatrix} 4 \\ 10 \\ 16 \end{pmatrix}$$

In the example above, we calculated the matrix-vector Multiplication by strictly following the definination, which is a good way to avoid making mistakes in calculations or losing marks due to the steps issue.

From the defination, the equivalance of the matrix-vector multiplication and the linear combination of the columns of the matrix is quite trivial, and it's quite clear that the equation $A \cdot \mathbf{x} = \mathbf{b}$ has the same solution set as the vector equation

$$\mathbf{b} = x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n,$$

which, in turn, has the same solution set as the linear system whose augmented matrix is

$$\begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n & \mathbf{b} \end{bmatrix}.$$

Simply following this defination, we have the following theorem.

Proposition 1.2. Let A be an $m \times n$ matrix, $\mathbf{b}$ be a vector in $\mathbb{R}^m$, and $\mathbf{x}$ be a vector in $\mathbb{R}^n$. Then the following statements are equivalent:

1. For each $\mathbf{b}$ in $\mathbb{R}^m$, the equation $A \cdot \mathbf{x} = \mathbf{b}$ has a solution.

2. Each $\mathbf{b}$ in $\mathbb{R}^m$ is a linear combination of the columns of A.

3. The columns of A span $\mathbb{R}^m$.

4. A has a pivot position in every row.

Proof. The proof of 1. $\iff$ 2. $\iff$ 3. is quite trivial, and the proof of 2. $\iff$ 4. has already been included in previous chapters. The handwritten lecture notes present a proof of (1.) $\iff$ (4.), but it seems to have assumed that (1.) $\iff$ (2.) and (2.) $\iff$ (4.) are true, and is not very clear in the perspective of rigorous reasoning.

Computation of Matrix-Vector Multiplication

So far, we have been regarding $A\mathbf{x}$ as a linear combination of the columns of A. Following this rule, we can write down the matrix-vector multiplication concisely where the i-th row, j-th column number in A is denoted by a_{ij}:

$$A\mathbf{x} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{pmatrix}.$$

Take a closer look at each element in the result vector, we can see that the i-th element in the result vector is the dot product of the i-th row of A and the vector $\mathbf{x}$. Therefore, we can also write the matrix-vector multiplication as

Proposition 1.3. Row vector Rule for Matrix-Vector Product

$$A\mathbf{x} = \begin{pmatrix} \mathbf{a}_1 \cdot \mathbf{x} \\ \mathbf{a}_2 \cdot \mathbf{x} \\ \vdots \\ \mathbf{a}_m \cdot \mathbf{x} \end{pmatrix}, \text{ where } \mathbf{a}_i \text{ is the i-th row of } A.$$

As all other operational rules, in the end of this subchapter, we present the properties of matrix-vector product as follows, which can be easily proved by examining each element in the result vector.

Proposition 1.4. Properties of Matrix-Vector Product

Let A and B be matrices, and let $\mathbf{x}$ and $\mathbf{y}$ be vectors in $\mathbb{R}^n$. Then the following properties hold:

1. $A(\mathbf{x} + \mathbf{y}) = A\mathbf{x} + A\mathbf{y}$.

2. $A(c\mathbf{x}) = c(A\mathbf{x})$.

3. $(A + B)\mathbf{x} = A\mathbf{x} + B\mathbf{x}$.

4. $A(B\mathbf{x}) = (AB)\mathbf{x}$.

5. $I\mathbf{x} = \mathbf{x}$, where I is the identity matrix.

Remark: Property 3-5 are to be explained after defining the matrix-matrix operational rules in the next chapter.

1.5   Solution Sets of Linear Systems

Homogeneous Linear Systems

Definition 1.7. Homogeneous Linear System

A linear system $A\mathbf{x} = \mathbf{0}$ is called **homogeneous** if the vector of constants on the right side of the equation is the zero vector.

The solution set of a homogeneous linear system is a subspace of $\mathbb{R}^n$. It's obvious that the Homogeneous Linear System always has a solution $\mathbf{x} = \mathbf{0}$, which is called the **trivial solution**.

The homogeneous linear system $A\mathbf{x} = \mathbf{0}$ has a nontrivial solution *if and only if* the system has at least one free variable.

Example 1.8. Determine whether the following homogeneous linear system has a nontrivial solution:

$$3x_1 + 5x_2 - 4x_3 = 0$$
$$-3x_1 - 2x_2 + 4x_3 = 0$$
$$6x_1 + x_2 - 8x_3 = 0$$

Solution. The augmented matrix of this system is

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

We can row reduce this matrix to

$$\begin{pmatrix} 1 & 0 & -\frac{4}{3} & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Hence, we get

$$x_1 = \frac{4}{3}x_3$$
$$x_2 = 0$$
$$x_3 \text{ is a free variable}$$

Therefore, the system has a nontrivial solution, and the solution set is

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} \frac{4}{3}x_3 \\ 0 \\ x_3 \end{pmatrix} = x_3 \begin{pmatrix} \frac{4}{3} \\ 0 \\ 1 \end{pmatrix}.$$

Note. **Parametric Vector Form** of the solution set of a homogeneous linear system is the form of the solution set that is expressed as a linear combination of one or more vectors, like

$$\mathbf{x} = s\mathbf{v}_1 + t\mathbf{v}_2 + \cdots \quad (s, t, \cdots \in \mathbb{R})$$

Solutions of Nonhomogeneous Linear Systems

Example 1.9. Change the right side of the previous example to $\begin{pmatrix} 7 \\ -1 \\ -4 \end{pmatrix}$, and determine whether the system has a solution.

Solution. Skip the row reduction part, we get

$$x_1 - \frac{4}{3}x_3 = -1$$
$$x_2 = 2$$
$$x_3 \text{ is a free variable}$$

Therefore, the system has a solution, and the solution set is

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -1 + \frac{4}{3}x_3 \\ 2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} \frac{4}{3} \\ 0 \\ 1 \end{pmatrix}.$$

In this example, we wrote in the form of

$$\mathbf{x} = \mathbf{p} + t\mathbf{v} \ \text{ where } \ t = x_3,$$

in which:

- $\mathbf{p}$ is a particular solution of the non-homogeneous equation $A\mathbf{x} = \mathbf{0}$, and

- $\mathbf{v}$ is the general solution of the corresponding homogeneous equation $A\mathbf{x} = \mathbf{0}$.

Now we can generalize this to the following theorem.

Proposition 1.5. General Solution of a Nonhomogeneous Linear System
Suppose $A\mathbf{x} = \mathbf{b}$ is consistent for some given vector $\mathbf{b}$, and let $\mathbf{p}$ be a particular solution. Then the general solution set of $A\mathbf{x} = \mathbf{b}$ is the set of all vectors of the form

$$\mathbf{p} + t\mathbf{v_n}$$

, where $\mathbf{v_n}$ is any solution of the corresponding homogeneous equation $A\mathbf{x} = \mathbf{0}$.

Hence, we end this topic by showing the following summary given in the handwritten notes.

- Row reduce the augmented matrix of the system to reduced row-echelon form.

- Express each basic variable in terms of any free variables.

- Write a typical solution $\mathbf{x}$ as a vector whose entries depend on the free variables, if any.

- Decompose $\mathbf{x}$ into a linear combination of vectors (with numeric entries) using the free variables as parameters.

P.S. For situations with more than 1 free variables, the handwritten notes present the solution in this way

$$\cdots = x_2 \begin{pmatrix} 4 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + x_6 \begin{pmatrix} -5 \\ 0 \\ 1 \\ 0 \\ 4 \\ 1 \end{pmatrix}$$

1.6 Linear Independence

Example 1.10.

$$\mathbf{v_1} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \mathbf{v_2} = \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix}, \mathbf{v_3} = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}$$

If $x_1\mathbf{v_1} + x_2\mathbf{v_2} + x_3\mathbf{v_3} = \mathbf{0}$, are the solutions unique?

Solution. The vector equation $x_1\mathbf{v_1} + x_2\mathbf{v_2} + x_3\mathbf{v_3} = \mathbf{0}$ is equivalent to the homogeneous linear system whose augmented matrix is

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

We can row reduce this matrix to

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

Hence, we have

$$x_1 = 2x_3$$
$$x_2 = -x_3$$

In particular, we can write $\mathbf{v_3} = -2\mathbf{v_1} + \mathbf{v_2}$, which means $\text{Span}\{\mathbf{v_1}, \mathbf{v_2}\} = \text{Span}\{\mathbf{v_1}, \mathbf{v_2}, \mathbf{v_3}\}$ is a plane. Therefore, $\mathbf{v_1}, \mathbf{v_2}, \mathbf{v_3}$ are linearly dependent.

From the example above, we can see that the vectors $\mathbf{v_1}, \mathbf{v_2}, \mathbf{v_3}$ are linearly dependent, which means that at least one of the vectors can be written as a linear combination of the others. From this observation, we can generalize the following definition.

Definition 1.8. Linear Independence
A set of vectors $\mathbf{v_1}, \mathbf{v_2}, \cdots, \mathbf{v_p}$ in $\mathbb{R}^n$ is said to be **linearly independent** if the vector equation

$$x_1\mathbf{v_1} + x_2\mathbf{v_2} + \cdots + x_p\mathbf{v_p} = \mathbf{0}$$

has only the trivial solution $x_1 = x_2 = \cdots = x_p = 0$.
A set of vectors that is said to be **linearly dependent** if there exist weights $c_1, c_2, \cdots, c_p$, not all zero, such that
$$c_1\mathbf{v_1} + c_2\mathbf{v_2} + \cdots + c_p\mathbf{v_p} = \mathbf{0}.$$

For particular situations of the general defination above, we'll discuss the sets of one or two vectors as follows, the *proof* of which just follows the defination.

- A set of one vector $\mathbf{v}$ is linearly independent if and only if $\mathbf{v} \neq \mathbf{0}$.

- A set of two vectors $\mathbf{v_1}$ and $\mathbf{v_2}$ is linearly independent if and only if neither vector is a scalar multiple of the other. i.e. $\{\mathbf{v_1}, \mathbf{v_2}\}$ are linearly dependent if at least one of them is a scalar multiple of the other.

Moreover, for a set of three or more vectors, we can use the following proposition to determine whether they are linearly independent.

Proposition 1.6. A set of three or more vectors $\mathbf{v_1}, \mathbf{v_2}, \cdots, \mathbf{v_p}$ is linearly dependent if and only if at least one of the vectors can be written as a linear combination of the others.

Given the deductions above, if we want to prove a set of vectors are linearly dependent, one simple way is to find one vector that can be written as a multiple of another (for system with 2 vectors) or a linear combination of the others (for system with 3 or more vectors).

After getting familiar with the basic concepts, lets move on to some specific situations.

Proposition 1.7. Theorems for determining linear dependence

- For $\{\mathbf{v_1}, \cdots, \mathbf{v_p}\}_{p \geq 2}$, where $\mathbf{v_1}, \mathbf{v_2}, \cdots, \mathbf{v_p}$ are vectors in $\mathbb{R}^n$, if $p > n$, then $\{\mathbf{v_1}, \cdots, \mathbf{v_p}\}$ must be linearly dependent.

- If a set of vectors $\{\mathbf{v_1}, \cdots, \mathbf{v_p}\}$ in $\mathbb{R}^n$ contains the zero vector, then $\{\mathbf{v_1}, \cdots, \mathbf{v_p}\}$ is linearly dependent.

1.7   Introduction to Linear Transformations

Definition 1.9. Transformation
A transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$, $T : \mathbb{R}^n \mapsto \mathbb{R}^m$, is a rule that assigns to each vector $\mathbf{x}$ in $\mathbb{R}^n$ to a vector $T(\mathbf{x})$ in $\mathbb{R}^m$. We call

$$\mathbb{R}^n \text{ the } \mathbf{domain} \text{ of } T,$$
$$\mathbb{R}^m \text{ the } \mathbf{codomain} \text{ of } T.$$

For $\mathbf{x} \in \mathbb{R}^n$, $T(\mathbf{x})$ is called the **image** of $\mathbf{x}$ under T, and the set of all images $T(\mathbf{x})$ is called the **range** of T.
For example,
$$A : \text{matrix } m \times n, \quad \mathbf{x} \in \mathbb{R}^n, \quad T : \mathbb{R}^n \mapsto \mathbb{R}^m, \quad T(\mathbf{x}) \triangleq A\mathbf{x}$$

Note. The handwritten notes provide an example of $T : \mathbb{R}^2 \mapsto \mathbb{R}^2$ regarding the geometric interpretation of the transformation, please refer to file1.2. The key ideas are:

- T deforms the unit square into a parallelogram.

- T maps a line segment to a line segment.

Definition 1.10. Linear Transformation
A transformation $T : \mathbb{R}^n \mapsto \mathbb{R}^m$ is called a **linear transformation** if

$$T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v}) \quad \text{for all } \mathbf{u}, \mathbf{v} \in \mathbb{R}^n,$$
$$T(c\mathbf{u}) = cT(\mathbf{u}) \quad \text{for all } \mathbf{u} \in \mathbb{R}^n \text{ and all scalars } c.$$

i.e. a transformation T is linear if it preserves vector addition and scalar multiplication.

Deduction. If T is a linear transformation, then

$$T(\mathbf{0}) = \mathbf{0},$$
$$T(c_1\mathbf{v}_1 + c_2\mathbf{v}_2) = c_1 T(\mathbf{v}_1) + c_2 T(\mathbf{v}_2),$$
$$T(c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_p\mathbf{v}_p) = c_1 T(\mathbf{v}_1) + c_2 T(\mathbf{v}_2) + \cdots + c_p T(\mathbf{v}_p).$$

Example 1.11. Geometric meaning of linear transformation

- Given a scalar, define $T : \mathbb{R}^2 \mapsto \mathbb{R}^2$ by $T(\mathbf{x}) = r\mathbf{x}$. T is called contraction if $0 < r < 1$, and dilation if $r > 1$.

- Define $T : \mathbb{R}^2 \mapsto \mathbb{R}^2$ by $T(\mathbf{x}) = A\mathbf{x}$, where $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. Here T is a rotation.

1.8 The Matrix of a Linear Transformation

> **Proposition 1.8. Matrix of a Linear Transformation**
>
> Let $T : \mathbb{R}^n \mapsto \mathbb{R}^m$ be a linear transformation. Then there exists a unique $m \times n$ matrix A such that $T(\mathbf{x}) = A\mathbf{x}$ for all $\mathbf{x}$ in $\mathbb{R}^n$.
>
> In fact, $A = [T(\mathbf{e}_1), T(\mathbf{e}_2), \cdots, T(\mathbf{e}_n)]$, where $\mathbf{e}_1, \mathbf{e}_2, \cdots, \mathbf{e}_n$ are the standard unit vectors in $\mathbb{R}^n$, which means
>
> $$\mathbf{e_1} = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \mathbf{e_2} = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \cdots, \mathbf{e_n} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}.$$

Proof. The key of this proof is to make use of the linearity of T stated in the Defination of Linear Transformation. Denote $\mathbf{x} = (x_1, x_2, \cdots, x_n)^T$, then we have

$$
\begin{aligned}
T(\mathbf{x}) &= T(x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \cdots + x_n\mathbf{e}_n) \\
&= x_1 T(\mathbf{e}_1) + x_2 T(\mathbf{e}_2) + \cdots + x_n T(\mathbf{e}_n) \\
&= \begin{pmatrix} T(\mathbf{e}_1) & T(\mathbf{e}_2) & \cdots & T(\mathbf{e}_n) \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \\
&= A\mathbf{x}.
\end{aligned}
$$