

Linear Algebra and Analytic Geometry

Lorenzo Arcioni

November 2, 2024

Preface

In the contemporary information age, where the digital landscape is saturated with an incessant stream of data, extracting and connecting key concepts together becomes an almost impossible challenge. Discerning crucial information from background noise becomes increasingly challenging as the volume of accessible information grows. So, the primary objective of this book is to synthesize the key concepts of linear algebra and analytic geometry, presenting them in a cohesive and meaningful framework. In a society inundated with vast amounts of data, often presented in a haphazard manner, the book serves as a guiding beacon to distill order from the chaos.

The overarching mission is to create a structured pathway through the fundamental principles of linear algebra and analytic geometry. Recognizing the challenges posed by the disorganized nature of information in the modern era, this book endeavors to provide a clear and comprehensible narrative. The intention is not merely to impart theoretical knowledge but to interconnect these essential mathematical concepts, allowing learners to derive a coherent understanding from the seemingly disparate pieces of information. It is important to note that the purpose of this book is not to replace a comprehensive university textbook but rather to serve as a collection of notes, handouts, and concepts acquired throughout the study journey, logically interconnected.

Linear Algebra and Analytic Geometry offer a structured approach to understanding and modeling complex patterns inherent in our data-driven world. As we confront an era marked by the ubiquity of information, these mathematical disciplines become more than theoretical framework, they become indispensable tools for organizing, interpreting, and extracting meaningful insights from the vast sea of data.

The integration of the Python programming language, with its emphasis on clarity and versatility, serves as a catalyst in this process. Through Jupyter notebooks and practical exercises, learners are not only equipped with theoretical knowledge but also empowered to apply these concepts to real-world scenarios. This dynamic approach not only aligns with the demands of our information-rich

environment but also fosters a deeper understanding by bridging theory with hands-on application.

In essence, the reflection on this theme underscores the importance of equipping individuals with the skills to navigate through the complexities of the modern world. The ability to distill key concepts from the informational noise becomes a crucial element in academic pursuits, professional endeavors, and everyday decision-making. As we embark on the journey through linear algebra and analytic geometry, we recognize that these mathematical tools not only enhance our capacity to comprehend but also refine our capability to filter and prioritize, allowing us to extract meaningful patterns and make informed decisions in a world saturated with data.

Best wishes on your mathematical exploration.

Lorenzo Arcioni

Prerequisites

To fully engage with the content presented in "Linear Algebra and Analytic Geometry", readers are encouraged to possess a solid understanding of the following prerequisites:

- **High School Mathematics:** A foundational grasp of high school mathematics is crucial. Readers should be comfortable with algebraic concepts, geometry, trigonometry, and basic calculus. Familiarity with concepts such as functions, equations, and coordinate geometry will provide a strong basis for delving into the more advanced topics covered in this book.
- **Set Theory:** A basic understanding of set theory is recommended. Concepts such as sets, subsets, intersections, unions, and the fundamental operations involving sets will be encountered throughout the book. This knowledge forms the basis for many abstract mathematical structures, providing a solid foundation for the discussions on vector spaces and transformations.
- **Python Programming:** Proficiency in the Python programming language is essential for extracting maximum value from the practical applications and exercises presented in this book. While not mandatory, a basic understanding of Python syntax, data structures, and fundamental programming concepts will greatly enhance the reader's ability to implement and experiment with the mathematical principles covered.

Readers who meet these prerequisites will find themselves well-prepared to navigate the material in "Linear Algebra and Analytic Geometry". For those seeking a refresher on any of these topics, additional resources are recommended to ensure a smooth and enriching learning experience. As we embark on this mathematical journey, a solid foundation in these fundamental areas will contribute to a deeper appreciation and mastery of the concepts explored in the subsequent chapters.

Contents

Chapter 0

Introduction

Welcome to the world of "Linear Algebra and Analytic Geometry", a journey that unfolds the elegance and applicability of two foundational branches of mathematics. This book is designed to be your companion as you explore the intricacies of linear algebra and analytic geometry, essential disciplines that underpin a myriad of scientific and engineering applications.

In this introduction, we set the stage for a comprehensive exploration of mathematical concepts that form the backbone of various fields. Linear algebra, with its rich tapestry of vector spaces, matrices and vector spaces, provides a versatile toolkit for problem-solving and modeling. Analytic geometry, the seamless integration of algebra and geometry, empowers us to visualize and analyze mathematical structures in a way that transcends abstract representation.

As we embark on this journey, you will find a distinctive feature of this book: the integration of the powerful Python programming language. Python serves as our tool to bridge theory and practice, enabling you to implement and experiment with the concepts covered in a hands-on, interactive manner. Whether you are a student seeking a solid foundation or a practitioner aiming to refresh and deepen your knowledge, this book caters to a diverse audience with varying levels of expertise. The chapters ahead will guide you through the fundamental principles, practical applications, and real-world implications of linear algebra and analytic geometry. We will explore the interconnected nature of these subjects, providing not only theoretical insights but also practical skills that can be directly applied in scientific research, data science, computer engineering, and many other domains. The intersection of linear algebra and analytic geometry forms a rich tapestry where abstract mathematical structures meet real-world geometric representations. In Figure 0.1, a conceptual map visually encapsulates this symbiotic relationship, illustrating how vectors, matrices, and systems of linear equations seamlessly connect with geometric entities like lines, planes, and conics. Linear algebra provides the theoretical framework, offering tools to understand transformations and relationships among vectors. Simultaneously, analytic geometry provides a tangible canvas

where these abstract principles manifest visually. This convergence enhances our ability to analyze spatial relationships, solve complex problems, and bridge the gap between mathematical abstraction and geometric reality.

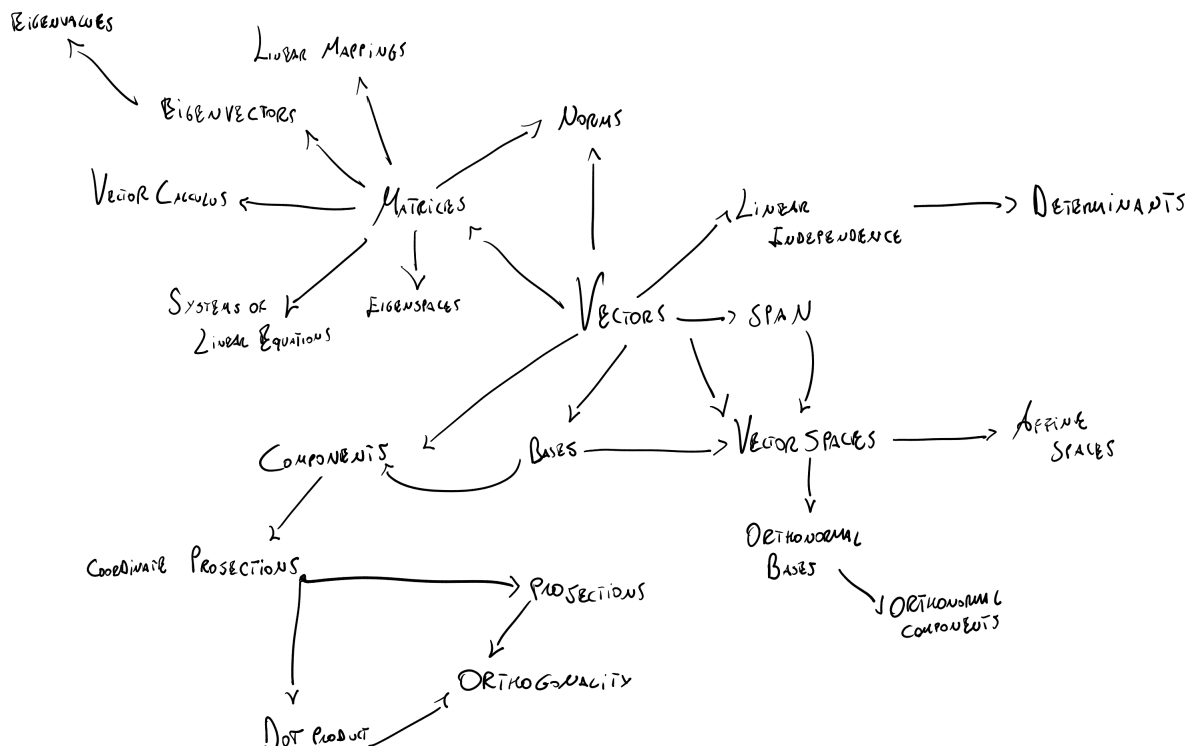


Figure 0.1: Conceptual map of intersection of Linear Algebra and Analytic Geometry

Join me on this mathematical odyssey, where clarity, intuition, and computational prowess converge to make the abstract tangible. Let's unravel the beauty of linear algebra and analytic geometry, armed with the simplicity of Python programming, as we navigate through the pages of discovery and learning.

0.1 Objects

In the realm of mathematics, fundamental objects like scalars, vectors, matrices, and tensors serve as the building blocks for expressing a wide range of mathematical and scientific concepts. These entities encapsulate different levels of complexity, from singular values to multi-dimensional structures, each offering unique properties and applications.

0.1.1 Scalars

Scalars are elements of a field, like \mathbb{R} , that represent quantities that only have magnitude, such as temperature, mass, distance, or time. Scalars have zero dimensions. They are considered point values in a mathematical space.

Since a scalar has no dimensions, its shape is represented as an empty set.

Scalars are typically denoted by lowercase letters, e.g., a, b, c and they are treated as constants; they can be real numbers, complex numbers, or elements from other mathematical fields.

For example

$$a \in \mathbb{F}$$

where a is a scalar in a field \mathbb{F} .

0.1.2 Vectors

A vector is an ordered collection of elements of a field that represents a quantity with both magnitude and direction. Vectors are used to describe various physical quantities, such as displacement, velocity, and force. In a geometric sense, they can be visualized as directed line segments with a specific length and direction in space.

Vectors have one dimension, so the shape of a vector is a 1-dimensional tuple (n), where n is the number of elements in the vector, also known as vector size.

Here are a few key points to understand about vectors:

- **Magnitude:** This is the length of the vector.
- **Direction:** This indicates the way the vector points.
- **Components:** The individual parts that make up a vector.

They are typically denoted by a letter with an arrow above. For example

$$\vec{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \in \mathbb{F}^n,$$

where \vec{v} is a vector in \mathbb{F}^n , (n) is the vector shape, n is the vector size and $v_1, v_2, \dots, v_n \in \mathbb{F}$ are vector components.

If we wish to access the component within the vectors we can employ the following mathematical notation:

$$\vec{v}^{(1)} = v_1$$

And to access all elements of a dimension we can use ":" instead of numerical indexes:

$$\vec{v}^{(:)} = \vec{v}$$

0.1.3 Matrices

A matrix is a two-dimensional array of elements of a field, organized in rows and columns. It is used to represent and manipulate data in various applications, including linear transformations and systems of equations.

The shape of a matrix is a 2-dimensional tuple (m, n) , where m is the number of rows and n is the number of columns.

Matrices are typically denoted by uppercase letters, e.g., A, B, C . For example

$$A_{m,n} = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix} \in \mathbb{F}^{m \times n}$$

where $A_{m,n}$ is a matrix in $\mathbb{F}^{m \times n}$, (m, n) is the matrix shape and $a_{i,j} \in \mathbb{F}$.

If we wish to access the component within the matrix we can employ the following mathematical notation:

$$A_{m,n}^{(i,j)} = a_{i,j}$$

And to access all elements of a dimension we can use ":" instead of numerical indexes:

$$A_{m,n}^{(i,:)} = \begin{bmatrix} a_{i,1} \\ a_{i,2} \\ \vdots \\ a_{i,n} \end{bmatrix}, \quad A_{m,n}^{(:,j)} = \begin{bmatrix} a_{1,j} \\ a_{2,j} \\ \vdots \\ a_{m,j} \end{bmatrix}, \quad A_{m,n}^{(:, :) } = A_{m,n}$$

Square Matrices

A square matrix $A_{m,n}$ is a matrix with the same number of rows and columns, i.e., $m = n$.

Triangular Matrices

A lower triangular matrix A is a square matrix where all entries above the main diagonal are zero, i.e., $a_{i,j} = 0$ for $i < j$.

An upper triangular matrix A is a square matrix where all entries below the main diagonal are zero, i.e., $a_{i,j} = 0$ for $i > j$.

Diagonal Matrices

A diagonal matrix D is a square matrix where all entries outside the main diagonal are zero, i.e., $a_{i,j} = 0$ for $i \neq j$.

Symmetric Matrices

A symmetric matrix $A_{n,n}$ is a square matrix where the elements are symmetric with respect to the main diagonal. In other words, if $a_{i,j}$ is an entry in the matrix at the i -th row and j -th column, then $a_{i,j} = a_{j,i}$ for all i and j in the matrix. This symmetry implies that the values on one side of the main diagonal mirror the values on the other side.

0.2 Basic Operations

In the world of linear algebra, basic operations form the cornerstone of mathematical manipulations applied to our objects (scalar, vectors and matrices). These operations serve as fundamental tools for analyzing and transforming these mathematical entities, unlocking insights across various disciplines.

0.2.1 Transposition

In linear algebra, transposition refers to the operation of switching the rows and columns of a mathematical object. This operation is commonly applied to matrices, and geometrically can be interpreted as reflecting the matrix over its main diagonal.

Matrices Transposition

For a given matrix $A_{m,n} \in \mathbb{F}^{m \times n}$, the transpose, denoted as $A_{m,n}^T$, is obtained by swapping its rows with columns. Mathematically, each $a_{i,j}$ in $A_{m,n}$, would be mapped into a_{ji} . In other words, the element in the i -th row and j -th column of $A_{m,n}^T$ is the same as the element in the j -th row and i -th column of $A_{m,n}$.

So $A_{m,n}^T = A'_{n,m}$.

Example:

Consider the following matrix:

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

The transpose of matrix A is:

$$A^T = \begin{bmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{bmatrix}$$

Properties:

- $(A^T)^T = A$: Transposing a matrix twice results in the original matrix.
- $(cA)^T = cA^T$: Transposing a scaled matrix is equivalent to scaling the transpose.
- $(A + B)^T = A^T + B^T$: Transposing the sum of two matrices is equivalent to the sum of their transposes.

These properties are fundamental in various mathematical applications, including linear algebra and optimization.

0.2.2 Addition

In linear algebra, the sum of two objects is a fundamental operation that combines their respective components. We are very familiar with the common sum of scalars:

$$a + b = c,$$

but what happens when we need to apply the sum to objects of dimension greater than 0?

Vector sum, for example is defined as follow.

Let

$$\vec{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}, \quad \vec{w} = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix}$$

be two vectors in \mathbb{F}^n . The sum of two vectors \vec{v} and \vec{w} , denoted as $\vec{v} + \vec{w}$, is defined component-wise:

$$\vec{v} + \vec{w} = \begin{bmatrix} v_1 + w_1 \\ v_2 + w_2 \\ \vdots \\ v_n + w_n \end{bmatrix}$$

Geometrically, vector addition corresponds to placing the initial point of the second vector at the terminal point of the first vector and connecting the initial point of the first vector to the terminal point of the second vector (as shown in Figure 0.2).

Now, let's consider an example of vector addition using two real vectors, \vec{a} and \vec{b} . The sum of \vec{a} and \vec{b} will be:

$$\vec{a} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \quad \vec{b} = \begin{bmatrix} 2 \\ -3 \end{bmatrix}$$

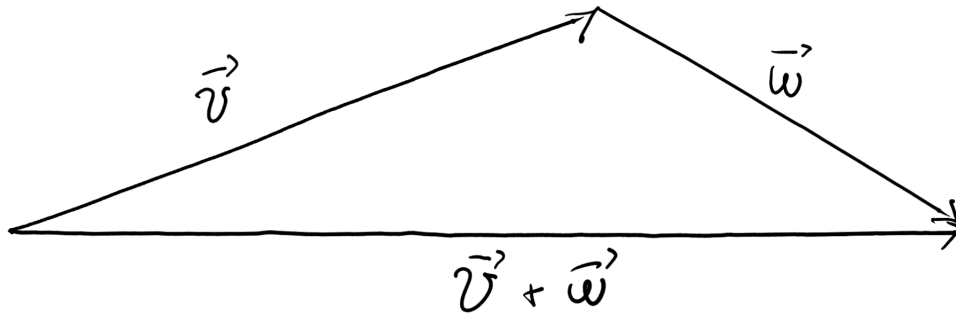


Figure 0.2: Triangle Law of Vector Addition

$$\vec{a} + \vec{b} = \begin{bmatrix} a_0 + b_0 \\ a_1 + b_1 \end{bmatrix} = \begin{bmatrix} 1 + 2 \\ 2 + (-3) \end{bmatrix} = \begin{bmatrix} 3 \\ -1 \end{bmatrix}$$

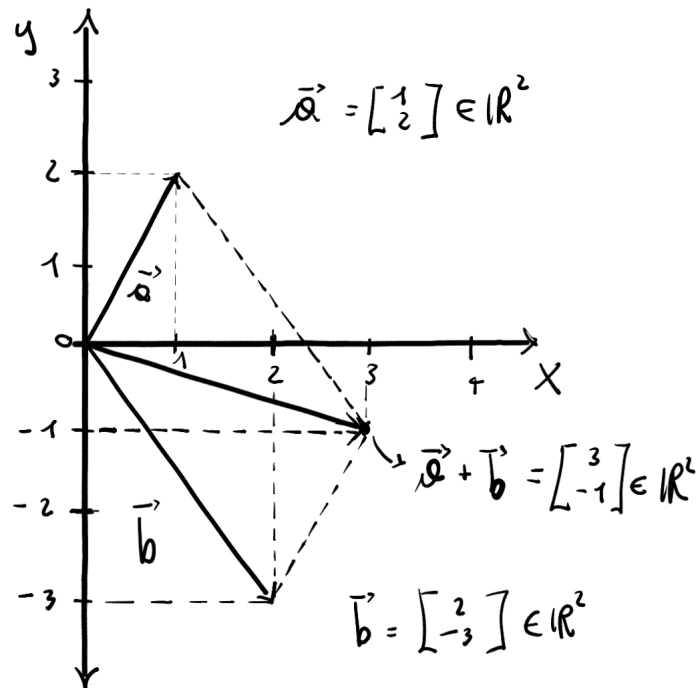


Figure 0.3: Parallelogram Law of Addition of Vectors

The Figure 0.3 shows the vector sum graphically by using the Parallelogram law. Note that we can sum together only vectors that have the same shape.

The same concept can be applied to matrices, in fact given two matrices

$$A_{m,n} = \begin{bmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \dots & a_{m,n} \end{bmatrix}, \quad B_{m,n} = \begin{bmatrix} b_{1,1} & b_{1,2} & \dots & b_{1,n} \\ b_{2,1} & b_{2,2} & \dots & b_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m,1} & b_{m,2} & \dots & b_{m,n} \end{bmatrix}$$

their sum is defined as

$$A_{m,n} + B_{m,n} = \begin{bmatrix} a_{1,1} + b_{1,1} & \dots & a_{1,n} + b_{1,n} \\ a_{2,1} + b_{2,1} & \dots & a_{2,n} + b_{2,n} \\ \vdots & \ddots & \vdots \\ a_{m,1} + b_{m,1} & \dots & a_{m,n} + b_{m,n} \end{bmatrix} = \begin{bmatrix} c_{1,1} & \dots & c_{1,n} \\ c_{2,1} & \dots & c_{2,n} \\ \vdots & \ddots & \vdots \\ c_{m,1} & \dots & c_{m,n} \end{bmatrix} = C_{m,n}$$

or, more mathematically, the generic $c_{i,j}$ of $C_{m,n}$ of new matrices is equal to $a_{i,j} + b_{i,j}$. Note that we can sum together only matrices that have the same shape.

0.2.3 Sum Over Axes

Now, let's introduce the concept of *sum over axes*.

The *sum over axes* operation plays a crucial role, particularly when dealing with high dimensionality objects. This operation allows us to efficiently contract algebraic objects along specified axes, resulting in a new object with reduced dimension.

Consider a very simple example, imagining having a vector $\vec{v} \in \mathbb{F}^n$ and wanting to perform a *sum over axes* on its (only) axis, i.e., axis 0.

$$\vec{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \in \mathbb{F}^n \xrightarrow{\text{Sum over 0 axis}} = \sum_{i=0}^n v_i$$

After the sum over the 0 axis of the vector \vec{v} , we obtain a scalar belonging to the same field as the vector components.

So, for example:

$$\vec{v} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \in \mathbb{F}^3 \xrightarrow{\text{Sum over 0 axis}} [1 + 2 + 3] = 6 \in \mathbb{F}$$

Consider another very simple example, imagining having a matrix $A_{m,n}$ and wanting to perform a *sum over axes* on axis 0.

$$A_{m,n} = \begin{bmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \dots & a_{m,n} \end{bmatrix} \in \mathbb{F}^{m \times n} \xrightarrow{\text{Sum over 0 axis}} A'_{1,n} = \begin{bmatrix} \sum_{i=0}^m a_{i,1} & \sum_{i=0}^m a_{i,2} & \dots & \sum_{i=0}^m a_{i,n} \end{bmatrix}$$

For example:

$$A_{3,4} = \begin{bmatrix} 1 & 2 & 3 & -8 \\ 2 & 5 & -1 & 9 \\ 3 & 3 & -7 & -2 \end{bmatrix} \in \mathbb{F}^{3 \times 4} \xrightarrow{\text{Sum over 0 axis}} A'_{1,4} = \begin{bmatrix} 1 & 2 & 3 & -8 \\ + & + & + & + \\ 2 & 5 & -1 & 9 \\ + & + & + & + \\ 3 & 3 & -7 & -2 \end{bmatrix} = \begin{bmatrix} 6 & 10 & -5 & -1 \end{bmatrix}$$

Thus, we have obtained the *sum over axes* on axis 0 of the matrix $A_{5,4}$, obtaining a row vector A' of one dimension and size 4.

Similarly, if we wanted to perform a *sum over axes* on axis 1, we would get

$$A_{mn} = \begin{bmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \dots & a_{m,n} \end{bmatrix} \in \mathbb{F}^{m \times n} \xrightarrow{\text{Sum over 1 axis}} A'_{m,1} = \begin{bmatrix} \sum_{i=0}^n a_{1,i} \\ \sum_{i=0}^n a_{2,i} \\ \vdots \\ \sum_{i=0}^n a_{m,i} \end{bmatrix}$$

For example:

$$A_{3,4} = \begin{bmatrix} 1 & 2 & 3 & -8 \\ 2 & 5 & -1 & 9 \\ 3 & 3 & -7 & -2 \end{bmatrix} \in \mathbb{F}^{3 \times 4} \xrightarrow{\text{Sum over 1 axis}} A'_{3,1} = \begin{bmatrix} 1 + 2 + 3 - 8 \\ 2 + 5 - 1 + 9 \\ 3 + 3 - 7 - 2 \end{bmatrix} = \begin{bmatrix} -2 \\ 15 \\ -3 \end{bmatrix}$$

0.2.4 Multiplication

When we talk about product, we are, of course, referring to the multiplication between two algebraic objects. When these two objects are scalars, we don't need to consider the issue of dimensionality, as scalars are objects with no dimension, and therefore, their product will inevitably have no dimension. However, as the complexity (and hence, dimension) of the objects we want to multiply increases and differs between the two objects, we need to take into account not only the product itself, but also the dimensional differences involved. In fact, as the shape of the objects increases (and differs), the possible ways of performing their product also increase.

In the following sections, it will become clearer in what ways it is possible to perform the standard product between two objects (scalar, vectors or matrices) and what is the difference with standard product, **Hadamard** product and **Kronecker** product.

Esistono molti modi di calcolare il prodotto tra due oggetti algebrici, ma la caratteristica che li accomuna è che tutti i prodotti si basano su due operazioni fondamentali:

- Sum over axes.
- Moltiplicazione tra scalari.

Dove la *sum over axes*, precedentemente introdotta nella sottosezione 0.2.3, ci permette di eseguire le operazioni di contrazione degli assi necessarie al calcolo del prodotto di due oggetti algebrici.

Mentre il prodotto standard tra scalari è il normale prodotto a cui siamo abituati; definito per due generici elementi a, b del campo \mathbb{F} come:

$$a \cdot b = c \in \mathbb{F}$$

Tramite queste due operazioni siamo in grado di eseguire il prodotto di due elementi algebrici di dimensionalità arbitraria.

Compatibilità tra Assi

Prima di addentrarci nel calcolo vero e proprio, dobbiamo chiarire anche un altro punto, senza il quale non è possibile eseguire tutte le tipologie di prodotto. Introduciamo quindi il concetto di *compatibilità tra assi*: due assi (dimensioni) si dicono compatibili se e solo se hanno la stessa size (contengono lo stesso numero di elementi). Per asse si intende la dimensione, ad esempio l'asse 0 (l'unico) del vettore $\vec{v} \in \mathbb{F}^n$ è la sua dimensione, l'asse del vettore $\vec{w} \in \mathbb{F}^m$ sono compatibili se e solo se $n = m$.

Facciamo un altro esempio, siano $A_{3,2} \in \mathbb{F}^{3 \times 2}$ e $B_{2,3} \in \mathbb{F}^{2 \times 3}$ due matrici, in questo caso l'asse 0 della matrice $A_{3,2}$ e l'asse 1 della matrice $B_{2,3}$ sono compatibili, ma anche l'asse 1 della matrice $A_{3,2}$ e l'asse 0 della matrice $B_{2,3}$ lo sono.

Mentre eseguiamo il prodotto di due oggetti algebrici, può essere necessario tenere conto di quali assi sono compatibili, perchè alcuni tipi di prodotto necessitano tale requisito.

Prodotto

In generale il prodotto tra due generici oggetti algebrici viene calcolato attraverso una prima fase di scelta degli assi compatibili sui quali effettuare la contrazione, e una seconda fase nella quale si effettua il broadcast sulle dimensioni che non sono state scelte per la contrazione. In generale il risultato è dato da un oggetto che ha per dimensioni tutte le dimensioni "non scelte" in ordine di moltiplicazione.

In questa sezione ci limiteremo a descrivere le principali metodologie di calcolo del prodotto.

L'esempio più semplice (escluso il prodotto tra scalari) è il prodotto di uno scalare per un vettore. In questo caso non si scelgono assi compatibili (perché non ve ne sono) quindi non si effettua la contrazione, ma si passa direttamente alla fase di broadcast dello scalare per ogni componente del vettore. Tale prodotto è definito tra uno scalare $a \in \mathbb{F}$ e un vettore $\vec{v} \in \mathbb{F}^n$ e si scrive

$$a \cdot \vec{v} = a \cdot \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} a \cdot v_1 \\ a \cdot v_2 \\ \vdots \\ a \cdot v_n \end{bmatrix} \in \mathbb{F}^n.$$

Il risultato è un vettore con la stessa size del vettore moltiplicato. Infatti se prendiamo la shape dello scalare () e la shape del vettore (n) e le mettiamo insieme otteniamo la shape del vettore risultato, che, dato che lo scalare non ha dimensioni, è proprio (n).

Allo stesso modo, il prodotto tra uno scalare $a \in \mathbb{F}$ e una matrice $B_{m,n} \in \mathbb{F}^{m \times n}$ è definito come

$$a \cdot B_{m,n} = a \cdot \begin{bmatrix} b_{1,1} & \dots & b_{1,n} \\ b_{2,1} & \dots & b_{2,n} \\ \vdots & \ddots & \vdots \\ b_{m,1} & \dots & b_{m,n} \end{bmatrix} = \begin{bmatrix} a \cdot b_{1,1} & \dots & a \cdot b_{1,n} \\ a \cdot b_{2,1} & \dots & a \cdot b_{2,n} \\ \vdots & \ddots & \vdots \\ a \cdot b_{m,1} & \dots & a \cdot b_{m,n} \end{bmatrix} \in \mathbb{F}^{m \times n}.$$

Da notare che, come nel caso precedente, lo scalare e la matrice non hanno dimensioni (assi) compatibili, quindi lo scalare viene moltiplicato per tutti i valori della matrice (broadcast). Come risultato finale abbiamo un oggetto che ha per dimensioni le dimensioni dello scalare, che non ne ha, e le dimensioni della matrice, e quindi (m, n) .

Un altro importante prodotto è il prodotto tra vettori. Il **Dot Product** tra vettori è definito come segue: dati due vettori $\vec{v} \in \mathbb{F}^n$ e $\vec{w} \in \mathbb{F}^n$, il loro prodotto è definito come:

$$\vec{v} \cdot \vec{w} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} \cdot \begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix} = \sum_{i=0}^n v_i \cdot w_i$$

Come possiamo osservare, scegliendo come assi compatibili: l'asse 0 del primo vettore e l'asse 0 del secondo; possiamo calcolare il prodotto element-wise delle componenti dei due vettori e successivamente, tramite la sum over axes, effettuiamo una contrazione sugli assi compatibili. Come notiamo, non ci sono dimensioni oltre a quelle che abbiamo selezionato, quindi il risultato avrà zero dimensioni e dunque sarà uno scalare.

È possibile anche la moltiplicazione tra una matrice ed un vettore. Siano $A_{m,n} \in \mathbb{F}^{m \times n}$ una matrice e $\vec{v} \in \mathbb{F}^n$ un vettore. Il loro prodotto è definito come segue:

$$A_{m,n} \cdot \vec{v} = \begin{bmatrix} a_{1,1} & \dots & a_{1,n} \\ a_{2,1} & \dots & a_{2,n} \\ \vdots & \ddots & \vdots \\ a_{m,1} & \dots & a_{m,n} \end{bmatrix} \cdot \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} \sum_{i=0}^n a_{1,i} \cdot v_i \\ \sum_{i=0}^n a_{2,i} \cdot v_i \\ \vdots \\ \sum_{i=0}^n a_{m,i} \cdot v_i \end{bmatrix} \in \mathbb{F}^m$$

Notiamo che come assi compatibili abbiamo scelto l'unico asse (0) del vettore \vec{v} e l'asse 1 della matrice $A_{m,n}$, e quindi la contrazione avviene su questa coppia di assi. Il risultato, come vediamo, ha per dimensioni gli assi non scelti come compatibili nel prodotto, quindi solo l'asse 0 della matrice, che ha come size m .

Vediamo ora un esempio con due matrici: siano $A_{m,n} \in \mathbb{F}^{m \times n}$ e $B_{r,s} \in \mathbb{F}^{r \times s}$ due matrici con un asse compatibile (quindi ad esempio $n = r$), il loro prodotto è definito come

$$A_{m,n} \cdot B_{r,s} = \begin{bmatrix} \sum_{i=0}^n a_{1,i} \cdot b_{i,1} & \dots & \sum_{i=0}^n a_{1,i} \cdot b_{i,s} \\ \vdots & \ddots & \vdots \\ \sum_{i=0}^n a_{m,i} \cdot b_{i,1} & \dots & \sum_{i=0}^n a_{m,i} \cdot b_{i,s} \end{bmatrix} = C_{m,s}.$$

Abbiamo ottenuto una matrice $C_{m,s}$ che ha lo stesso numero di righe di $A_{m,n}$ e lo stesso numero di colonne di $B_{r,s}$.

Questo è il prodotto matriciale standard più comunemente utilizzato; ve ne sono molti altri che si ottengono facendo una scelta diversa degli assi compatibili su cui fare la contrazione.

C'è poi un altro particolare prodotto element-wise, che si effettua sugli assi compatibili, ma senza alcuna contrazione. Questo tipo di prodotto si chiama **Hadamard product** ed è definito solamente per oggetti algebrici con la stessa shape (stesso numero di assi e stessa size per ogni asse).

L'**Hadamard product** per la generica coppia di vettori $\vec{v}, \vec{w} \in \mathbb{F}^n$ è definito come

$$\vec{v} \odot \vec{w} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \odot \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix} = \begin{bmatrix} v_1 \cdot w_1 \\ v_2 \cdot w_2 \\ \vdots \\ v_n \cdot w_n \end{bmatrix} \in \mathbb{F}^n.$$

Analogamente nel caso di due matrici $A_{m,n}, B_{m,n}$ il loro **Hadamard product** è definito come

$$A_{m,n} \odot B_{m,n} = \begin{bmatrix} a_{1,1} \cdot b_{1,1} & \dots & a_{1,n} \cdot b_{1,n} \\ \vdots & \ddots & \vdots \\ a_{m,1} \cdot b_{m,1} & \dots & a_{m,n} \cdot b_{m,n} \end{bmatrix} \in \mathbb{F}^{m \times n}.$$

In questa sezione sono state illustrate solamente alcune delle tipologie di prodotto che è possibile effettuare nell'algebra lineare. Questo perché si tratta di un testo introduttivo, dove vengono trattati solamente gli aspetti basici dell'algebra lineare che è una materia molto profonda.

0.3 Gaussian Elimination

Gaussian elimination, also known as row reduction, is an algorithm primarily used for solving systems of linear equations. It involves a series of row-wise operations performed on the corresponding matrix of coefficients. This method can be extended to compute the rank of a matrix,

Where the first nonzero element of each row is equal to 1 and has $j_k \geq i$; and all a_{r,j_k} with $r \neq i$ are equal to zero.

0.4 Inverse of a Matrix

The Gauss-Jordan elimination method is a systematic way to find the inverse of a matrix by transforming it into reduced row-echelon form. Let's denote the given matrix as A :

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

To find the inverse of A using Gauss-Jordan elimination, we will form an augmented matrix by appending the identity matrix of the same size next to A . Then, we will perform row operations until A becomes the identity matrix, and the inverse of A will be on the other side of the augmented matrix.

Example: Let's find the inverse of the matrix $A = \begin{bmatrix} 2 & 3 \\ 1 & 2 \end{bmatrix}$ using Gauss-Jordan elimination.

We start with the augmented matrix:

$$\left[\begin{array}{cc|cc} 2 & 3 & 1 & 0 \\ 1 & 2 & 0 & 1 \end{array} \right]$$

We perform row operations to transform the left side of the matrix into the identity matrix:

1. Subtract $\frac{1}{2}$ times the first row from the second row:

$$\left[\begin{array}{cc|cc} 2 & 3 & 1 & 0 \\ 0 & 1 & -\frac{1}{2} & 1 \end{array} \right]$$

2. Subtract 3 times the second row from the first row:

$$\left[\begin{array}{cc|cc} 2 & 0 & \frac{7}{2} & -3 \\ 0 & 1 & -\frac{1}{2} & 1 \end{array} \right]$$

3. Divide the first row by 2:

$$\left[\begin{array}{cc|cc} 1 & 0 & \frac{7}{4} & -\frac{3}{2} \\ 0 & 1 & -\frac{1}{2} & 1 \end{array} \right]$$

Therefore, the inverse of matrix A is $\begin{bmatrix} \frac{7}{4} & -\frac{3}{2} \\ -\frac{1}{2} & 1 \end{bmatrix}$.

Chapter 1

Vector spaces

1.1 Introduction

A vector space over a field \mathbb{F} is a non-empty set V together with two binary operations. It is indicated as a triplet: $(V, +, \cdot)$.

A vector space must satisfy the ten axioms listed below.

$$\forall \vec{v}, \vec{w}, \vec{u} \in V \wedge \forall a, b \in \mathbb{F}$$

- **Closure under ”+” operation:** $\vec{v} + \vec{w} \in V$
- **Closure under ”.” operation:** $a \cdot \vec{v} \in V$
- **Associativity of vector addition:** $\vec{u} + (\vec{v} + \vec{w}) = (\vec{u} + \vec{v}) + \vec{w}$
- **Commutativity of vector addition:** $\vec{u} + \vec{v} = \vec{v} + \vec{u}$
- **Identity element of vector addition:** There exists an element $\vec{0} \in V$, called the zero vector, such that $\vec{v} + \vec{0} = \vec{v}$.
- **Inverse elements of vector addition:** $\exists -\vec{v} \in V$ called the additive inverse of \vec{v} , such that $\vec{v} + (-\vec{v}) = \vec{0} \in V$.
- **Compatibility of scalar multiplication with field multiplication:** $a(b\vec{v}) = (ab)\vec{v}$.
- **Identity element of scalar multiplication:** $1\vec{v} = \vec{v}$ where 1 denotes the multiplicative identity in \mathbb{F} .

- **Distributivity of scalar multiplication with respect to vector addition:** $a(\vec{u} + \vec{v}) = a\vec{u} + a\vec{v}$.
- **Distributivity of scalar multiplication with respect to field addition:** $(a + b)\vec{v} = a\vec{v} + b\vec{v}$.

In this context, the elements of V are commonly called vectors, and the elements of \mathbb{F} are called scalars. A vector space with one element ($\{\vec{0}\}$) is the *trivial* vector space.

Remark. We will indicate a vector space $(V, +, \cdot)$ by V when $+$ and \cdot are the standard vector addition and scalar multiplication. And we will write $\vec{x} \in V$ for vectors in V to simplify notation.

Example of Vector Space

Consider a vector space V defined as:

$$V = \{\vec{v} \in \mathbb{R}^3 \mid v_3 = v_1\}$$

The vector space V is a subset of \mathbb{R}^3 that consists of all vectors \vec{v} of size 3, where the v_3 component is equal to v_1 component. The resulting vector space is a two-dimensional vector space over the \mathbb{R} field.

Now let's check if it is a vector space by proving all vector space's axioms!

Consider two generic vectors v and w in V and a scalar in \mathbb{R} . We need to verify closure under vector addition and scalar multiplication for the vector space V (other axiom verifications are left as exercises).

$$\vec{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_1 \end{bmatrix} \in V, \quad \vec{w} = \begin{bmatrix} w_1 \\ w_2 \\ w_1 \end{bmatrix} \in V, \quad k \in \mathbb{R}$$

$$\vec{v} + \vec{w} = \begin{bmatrix} v_1 + w_1 \\ v_2 + w_2 \\ v_1 + w_1 \end{bmatrix} = \vec{z}$$

As we can see, the z_3 component of \vec{z} is equal to z_1 component of \vec{z} , so $\vec{z} \in V$.

$$k \cdot \vec{v} = \begin{bmatrix} k \cdot v_1 \\ k \cdot v_2 \\ k \cdot v_1 \end{bmatrix} \in V$$

Once again, the z_3 component of \vec{z} is equal to z_1 component of \vec{z} , so $\vec{z} \in V$.

The graphical representation of vector space V , shown in Figure 1.1, is a plane.

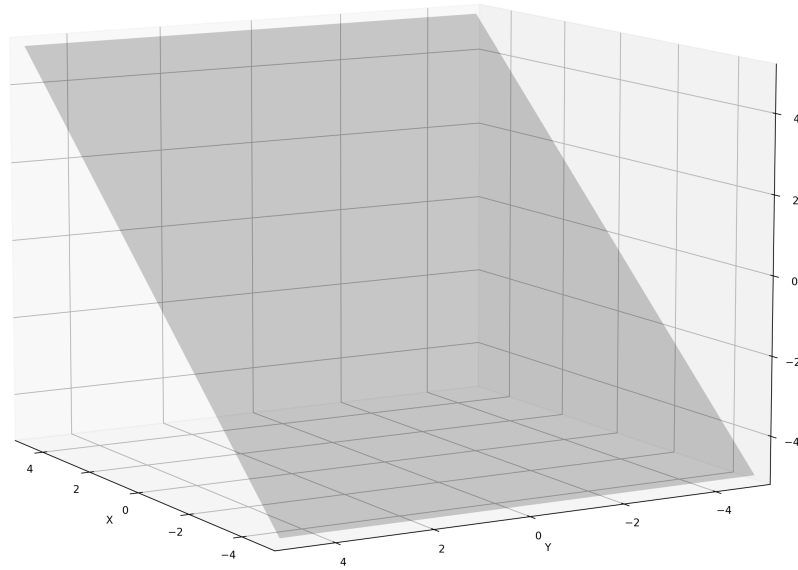


Figure 1.1: Graphical representation of V vector space.

1.1.1 Vector Subspaces

Let V be a vector space over a field \mathbb{F} , and let W be a non-empty subset of V . W is called a *vector subspace* of V if it satisfies the following properties:

1. **Closure under Vector Addition:** For any $\vec{u}, \vec{v} \in W$, their sum $\vec{u} + \vec{v}$ is also in W .

2. **Closure under Scalar Multiplication:** For any $\vec{v} \in W$ and any scalar $c \in \mathbb{F}$, the scalar product $c\vec{v}$ is in W .

3. **Contains the Zero Vector:** The zero vector $\vec{0}$ of V is in W .

If these conditions are met, W is a vector subspace of V and is denoted as $W \subseteq V$.

Example of Vector Subspace

Consider W as a vector subspace of previously defined space V :

$$W \subseteq V = \{\vec{v} \in \mathbb{R}^3 \mid v_3 = v_1\}$$

$$W = \{\vec{w} \in \mathbb{R}^3 \mid w_3 = w_1 \wedge w_1 + w_2 = 0\}$$

In other words, W consists of all vectors in \mathbb{R}^3 whose component w_3 is equal to w_1 and w_1 and w_2 components sum to zero.

Let's check the three conditions for W to be a vector subspace:

1. Closure under Addition:

- Take $\vec{v} = \begin{bmatrix} v_1 \\ -v_1 \\ v_1 \end{bmatrix}$ and $\vec{w} = \begin{bmatrix} w_1 \\ -w_1 \\ w_1 \end{bmatrix}$ where $\vec{v}, \vec{w} \in W$.
- Their sum $\vec{v} + \vec{w} = \begin{bmatrix} v_1 + w_1 \\ -(v_1 + w_1) \\ v_1 + w_1 \end{bmatrix} = \vec{z}$ is also in W because $z_2 = -z_1 \wedge z_3 = z_1$.

2. Closure under Scalar Multiplication:

- Take $\vec{w} = \begin{bmatrix} w_1 \\ -w_1 \\ w_1 \end{bmatrix} \in W$ and $k \in \mathbb{R}$.
- The scalar multiple $k \cdot \vec{w} = \begin{bmatrix} kw_1 \\ -(kw_1) \\ kw_1 \end{bmatrix} = \vec{z}$ is also in W since $z_1 = -z_2 \wedge z_3 = z_1$.

3. Contains the Zero Vector:

- The zero vector $\vec{0} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ is in W because $z_1 = -z_2 \wedge z_3 = z_1$.

Therefore, W satisfies all conditions and is indeed a vector subspace of V .

The graphical representation of the vector subspace W is a one-dimensional line as shown in Figure 1.2.

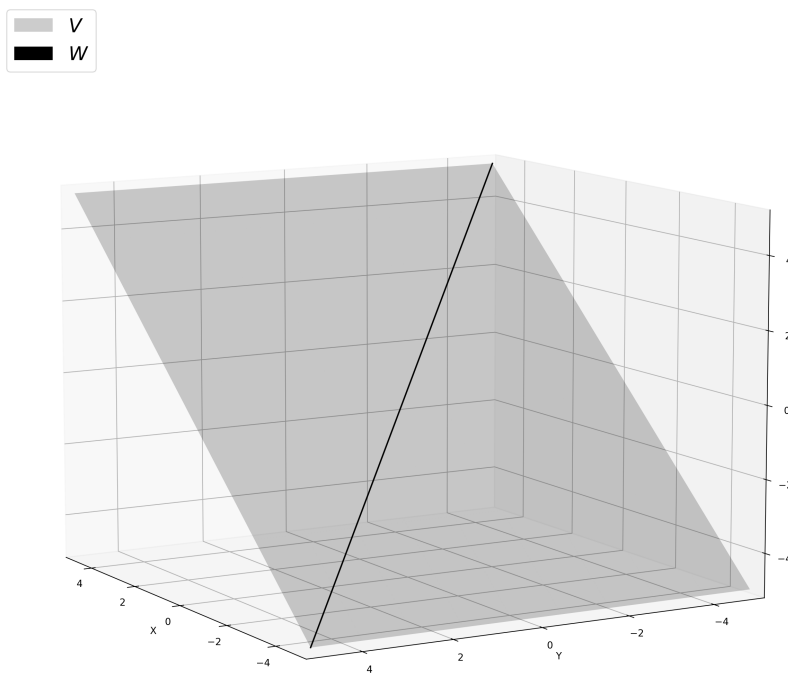


Figure 1.2: Graphical representation of V vector space (gray) and W vector subspace (black).

Sum of Subspaces

Let U_1, \dots, U_m be subspaces of V . The sum $U_1 + \dots + U_m$ is a vector subspace of V , defined as the set of all possible sums of elements from U_1, \dots, U_m , denoted by:

$$U_1 + \dots + U_m = \{u_1 + \dots + u_m : u_1 \in U_1, \dots, u_m \in U_m\}$$

This sum represents the space formed by combining elements from each subspace.

The union of basis vectors (of each base) of each subspace is a spanning set for the new sum vector subspace $U_1 + \cdots + U_m$.

Note that if

1. $U_1 + \cdots + U_m = V$

2. $U_1 \cap \cdots \cap U_m = \{\vec{0}\}$

The sum $U_1 + \cdots + U_m$ is called **direct sum** and it is denoted as

$$U_1 \oplus \cdots \oplus U_m = \{u_1 + \cdots + u_m : u_1 \in U_1, \dots, u_m \in U_m\}$$

.

In a direct sum, every element of $U_1 \oplus \cdots \oplus U_m$ can be uniquely expressed as a sum $u_1 + \cdots + u_m$, where each u_j belongs to U_j . And the only way to write 0 as a sum $u_1 + \cdots + u_m$ is by taking each u_j equal to 0. So the union of the basis vectors (of each basis) of each vector space, form a basis for the new sum vector space.

This represents a decomposition of the vector space V into mutually exclusive subspaces, where each element has a unique representation as a sum of elements from the respective subspaces.

Intersection of Subspaces

Consider subspaces V and W of a vector space U . The intersection $V \cap W$ is also a subspace of U . Since both V and W are subspaces, they contain the zero vector, making $V \cap W$ nonempty. Let $\vec{v}_1, \vec{v}_2 \in V \cap W$. As V and W are closed under addition, $\vec{v}_1 + \vec{v}_2 \in V$ and $\vec{v}_1 + \vec{v}_2 \in W$, implying $\vec{v}_1 + \vec{v}_2 \in V \cap W$. Thus, $V \cap W$ is closed under addition.

Consider $c \in \mathbb{R}$ and $v \in V \cap W$. Since V and W are closed under scalar multiplication, $c\vec{v} \in V$ and $c\vec{v} \in W$, indicating $c\vec{v} \in V \cap W$. Therefore, $V \cap W$ is closed under scalar multiplication.

These closures demonstrate that the intersection of subspaces preserves the subspace property within the overarching vector space U .

Union of Subspaces

In contrast, the union of subspaces V and W may not form a subspace of U . For instance, consider two distinct lines passing through the origin in \mathbb{R}^2 . While each line individually forms a subspace, their union lacks closure under addition, failing to be a subspace.

1.2 Linear Combinations

In the realm of vector spaces, the concept of a *linear combination* holds significant importance. A *linear combination* involves combining vectors in a space by scaling each vector by a scalar and then summing them up.

Let V be a vector space over the field \mathbb{F} . If $\vec{v}_1, \dots, \vec{v}_s \in V$ are vectors and $a_1, \dots, a_s \in \mathbb{F}$ are scalars, then the **linear combination** of those vectors with those scalars as coefficients is

$$a_1\vec{v}_1 + \dots + a_s\vec{v}_s \in V, \quad a_1, \dots, a_s \in \mathbb{F}.$$

Understanding linear combinations is foundational in linear algebra, providing insights into the structure and properties of vector spaces. Linear combinations play a crucial role in defining concepts such as span, linear independence, and solutions to systems of linear equations.

1.2.1 Span

The **span** of a nonempty subset S of a vector space V over a field \mathbb{F} , is the set of all possible linear combinations of those vectors.

$$\text{span}(\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k) = \left\{ \sum_{i=1}^k \lambda_i \vec{v}_i \mid k \in \mathbb{N}, \vec{v}_i \in V, \lambda_i \in \mathbb{F} \right\}$$

The span of the empty subset of a vector space is its trivial subspace.

In simpler terms, the span is the set of all possible vectors that can be formed by linear combinations of the given vectors.

Let $S \subseteq V$. The set S is a **spanning set** of V if *every* $\vec{v} \in V$ can be written as linear combination of vectors in S .

The Span of a Subset of a Vector Space is a Vector Subspace

Let S be a set of vectors of vector space V . The $\text{span}(S)$ is a vector subspace.

It is easy to show that the span of S contain the $\vec{0}$ and is closed under addition and scalar multiplication.

All the other axioms are inherited from V .

The span plays a crucial role in determining the extent or reach of a set of vectors. If the span of a set of vectors is the entire vector space V , then the vectors are said to *span* the space.

1.3 Linear (in)dependence

In a vector space V , a subset of vectors is deemed *linearly independent* if no vector within the subset can be expressed as a linear combination of others in the same subset. Conversely, if such a linear combination exists, the subset is *linearly dependent*.

Now, let's explore some implications of these definitions!

Consider a subset S of vectors in a vector space V . If S is linearly independent, then the only way to obtain the zero vector through a linear combination of its elements is by using all coefficients as zero:

$$c_1\vec{s}_1 + \cdots + c_n\vec{s}_n = \vec{0}, \quad \text{with } c_1 = \cdots = c_n = 0.$$

This condition serves as a hallmark of linear independence.

Conversely, if a subset S is not linearly independent, it implies the existence of at least one vector that can be expressed as a linear combination of the others. For example the vector

$$\vec{s}_i = c_1\vec{s}_1 + \cdots + c_{i-1}\vec{s}_{i-1} + c_{i+1}\vec{s}_{i+1} + \cdots + c_n\vec{s}_n,$$

is linearly dependent from vectors in $\{\vec{s}_1, \dots, \vec{s}_{i-1}, \vec{s}_{i+1}, \dots, \vec{s}_n\}$.

In this case, the subset $\{\vec{s}_1, \dots, \vec{s}_n\}$ is termed linearly dependent. This dependence is evident when one of the vectors can be obtained by combining the rest.

Furthermore, the span of a subset of vectors in a vector space, denoted as $\text{span}(S)$, encompasses all possible linear combinations of the vectors in S . Interestingly, the span itself forms a vector subspace of V , regardless of the choice of subset S . This property underscores the structural coherence of vector spaces under linear combinations.

Moreover, adding a vector \vec{v} to a set S expands its span only if \vec{v} is not already in the span of S . Conversely, removing a vector from a set S shrinks its span if and only if the removed vector was not a linear combination of the others. In both this cases, the vector \vec{v} is linearly independent from vectors in S . This observation leads to a significant corollary: in any vector space, every finite set of vectors possesses a linearly independent subset with the same span.

Additionally, the removal of a vector from a linearly independent set preserves its linear independence. Conversely, any superset of a linearly dependent set inherits its linear dependence. These observations highlight the robustness of linear independence and dependence under subset and superset operations.

1.4 Basis and Dimension

1.4.1 Basis

A maximal linearly independent subset of vectors of V that spans V is called a **basis** of the vector space V .

This means that a subset B of V is a **basis** if it satisfies the two following conditions:

- **Linear independence:** For every finite subset $\{\vec{\beta}_1, \dots, \vec{\beta}_n\}$ of B , if $c_1\vec{\beta}_1 + \dots + c_n\vec{\beta}_n = \vec{0}$ for some c_1, \dots, c_n in \mathbb{F} , then $c_1 = \dots = c_n = 0$.
- **Spanning property:** For every vector \vec{v} in V , one can choose a_1, \dots, a_n in \mathbb{F} and $\vec{\beta}_1, \dots, \vec{\beta}_n$ in B such that $\vec{v} = a_1\vec{\beta}_1 + \dots + a_n\vec{\beta}_n$.

The scalars a_i are called the coordinates of the vector \vec{v} with respect to the **basis** B , and by the first property they are uniquely determined.

Note that it is more useful to consider a basis as a sequence of vectors instead of a set. This is because when considering the coordinates of a vector, we have an implicit need to know the order of the

basis vectors for applying coordinates to the right basis vectors. For this reason, we will denote a basis with angle brackets $\langle \beta_1, \beta_2, \dots, \beta_n \rangle$.

In a vector space V with a basis $B = \langle \vec{\beta}_1, \dots, \vec{\beta}_n \rangle$, the coordinates of a generic vector $\vec{v} \in V$ are the coefficients used to express \vec{v} as a linear combinations of the basis vectors. So the coordinates of the vector $\vec{v} \in V$ with respect to basis B are represented by the column vector

$$[\vec{v}]_B = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$$

In a vector space, the coordinates of a vector with respect to a given *basis* are unique. Specifically, let V be a vector space and $B = \langle \vec{\beta}_1, \vec{\beta}_2, \dots, \vec{\beta}_n \rangle$ be a *basis* for V . For any vector $\vec{v} \in V$, there exists a unique set of scalars c_1, c_2, \dots, c_n such that:

$$\vec{v} = c_1 \vec{\beta}_1 + c_2 \vec{\beta}_2 + \dots + c_n \vec{\beta}_n$$

This uniqueness property ensures that the coordinates of \vec{v} in terms of the *basis* B are well-defined and distinct.

We can view a basis of a vector space as a minimal subset of vectors with which we can control all the dimensions of a vector space. In particular, each vector in the basis is responsible for "controlling" one dimension of the vector space. This way, through coefficients, we can express each dimension of the vector space. From the perspective of an individual component of a vector: the basis vector responsible for the i -th component of vector \vec{v} will be a vector from the basis (let's denote it for simplicity as $\vec{\beta}_x \in B$). In this case, the i -th component of the basis vector $\vec{\beta}_x$ and the coefficient α_x will be responsible for generating the i -th component of vector \vec{v} in a similar way to this: $v_i = \alpha_x b_{x_i}$.

However, since there are other combinations that indirectly affect the same dimension managed by the basis vector $\vec{\beta}_x$, we also need to include these coefficients in the equation. Thus, we obtain $v_i = \alpha_1 \vec{\beta}_1^i + \dots + \alpha_x \vec{\beta}_x^i + \dots + \alpha_n \vec{\beta}_n^i$. Therefore, to find all the coefficients of the basis for vector \vec{v} (across all components), one needs to solve the linear system $v_i = \alpha_1 \vec{\beta}_1^i + \dots + \alpha_x \vec{\beta}_x^i + \dots + \alpha_n \vec{\beta}_n^i$ for each i from 1 to n .

$$\begin{cases} \alpha_1 \vec{\beta}_1^1 + \dots + \alpha_x \vec{\beta}_x^1 + \dots + \alpha_n \vec{\beta}_n^1 = v_1 \\ \vdots \\ \alpha_1 \vec{\beta}_1^n + \dots + \alpha_x \vec{\beta}_x^n + \dots + \alpha_n \vec{\beta}_n^n = v_n \end{cases}$$

And since it is a linear system of n equations and n variables (where all equations are linearly independent, as we mentioned that B is a minimal set that can control all dimensions of V), we obtain a unique solution for each vector in V .

For each \mathbb{R}^n vector space, the vectors

$$\xi_n = \left\langle \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \dots, \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix} \right\rangle$$

are a **basis** (known as *canonical basis*) also known as ξ_n . We denote these vectors $\vec{e}_1, \dots, \vec{e}_n$.

Every non-trivial vector space posses a basis. The trivial one, instead, doesn't have a basis (its basis is \emptyset). Existence of a basis in a vector space is guaranteed by the properties of spanning and linear independence.

For a given basis B of V with n elements, a linear combination of vectors $a_1 \vec{v}_1 + \dots + a_k \vec{v}_k = \vec{0}$ if and only if $a_1 [\vec{v}_1]_B + \dots + a_k [\vec{v}_k]_B = \vec{0}$.

In fact, let $\langle \vec{\beta}_1, \dots, \vec{\beta}_n \rangle$ be a basis of V vector space. Each vector $\vec{v}_i \in V$ can be expressed as a linear combination of basis vector, so:

$$\begin{aligned} a_1 \vec{v}_1 + \dots + a_k \vec{v}_k &= a_1 (\alpha_{1,1} \vec{\beta}_1 + \dots + \alpha_{1,n} \vec{\beta}_n) + \dots + a_k (\alpha_{k,1} \vec{\beta}_1 + \dots + \alpha_{k,n} \vec{\beta}_n) = \\ &= (a_1 \alpha_{1,1} + \dots + a_k \alpha_{k,1}) \vec{\beta}_1 + \dots + (a_1 \alpha_{1,n} + \dots + a_k \alpha_{k,n}) \vec{\beta}_n = \vec{0}. \end{aligned}$$

This implies that

$$\begin{cases} a_1\alpha_{1,1} + \cdots + a_k\alpha_{k,1} = 0 \\ \vdots \\ a_1\alpha_{1,n} + \cdots + a_k\alpha_{k,n} = 0 \end{cases}$$

that can be easily written as follow:

$$a_1 \begin{bmatrix} \alpha_{1,1} \\ \vdots \\ \alpha_{1,n} \end{bmatrix} + \cdots + a_k \begin{bmatrix} \alpha_{k,1} \\ \vdots \\ \alpha_{k,n} \end{bmatrix} = a_1[\vec{v}_1]_B + \cdots + a_k[\vec{v}_k]_B = \vec{0}.$$

Geometrically, the whole matter can be interpreted in this way: each coordinate vector of every vector in the initial linear combination is "telling" how much "force" to apply to each base vector that "controls" a certain coordinate. Let's take an example: consider a generic component of every vector in the initial combination. We know that the combination $a_1v_i + \cdots + a_kv_i$ equals zero, which tells us that the forces acting on the i -th component of the combination cancel out. However, this component (and all other components) stem from a sum of a linear combination of the i -th components of the base vectors. This implies that the force applied to any generic component is equally expressed by the same base vectors. So, for example $6 = 3 \cdot 2$ and $3 = 3 \cdot 1$, the combination $2 \cdot 6 + (-4) \cdot 3 = 0$, so also $2 \cdot (3 \cdot 2) + (-4) \cdot (3 \cdot 1) = 0$; now you can note that the number 3 is our basis number, it can express 3 and 6. The coordinate of 6 is 2 and the coordinate of 3 is 1. Substitute now 6 with 3 and 3 with 1 in the first combination and tah dah: $2 \cdot 2 + (-4) \cdot 1 = 0$. That is obvious because we have scaled the **same** basis number by two factors that cancel each other out.

Steinitz Lemma

This lemma describes a relationship between the sizes of linearly independent sets and spanning sets within a vector space. It states that if you have a linearly independent set U and a spanning set W in a vector space V , then the size of U cannot exceed the size of W . Furthermore, it asserts that by removing a specific number of vectors from W , you can create a new set that, when combined with U , still spans the entire space V . This lemma highlights the balance between linear independence and spanning capabilities in constructing bases for vector spaces.

This lemma can be understood in terms of spanning and linear independence in vector spaces. Imagine a vector space as a geometric plane or higher-dimensional space. A set of vectors is linearly independent if none of them can be expressed as a linear combination of the others. This implies that each vector in the set contributes a unique direction to the space. On the other hand,

a spanning set encompasses vectors that collectively cover the entire space, allowing you to reach any point within it through linear combinations of those vectors.

The Steinitz Lemma highlights a balance between these two aspects: linear independence and spanning capability. It states that if you have a set of vectors that are linearly independent and another set that spans the space, you cannot have more linearly independent vectors than those that span the space. In other words, you can't introduce more unique directions than necessary to cover the entire space. Moreover, the lemma suggests that you can fine-tune the spanning set by removing redundant vectors while still ensuring that it covers the space adequately.

Exchange Lemma

This lemma addresses the process of exchanging vectors within a basis while maintaining the basis' fundamental properties. It states that given a vector space V and a basis B , if there exists a vector \vec{v} that can replace one of the basis vectors while preserving linear independence, then the resulting set after the exchange still forms a basis for V . This lemma offers an intuitive understanding of how bases can be modified without compromising their essential properties. Think of a basis as a set of vectors that serves as the "building blocks" for constructing any vector within the space through linear combinations. The Exchange Lemma states that if you have a basis and you encounter a vector that can replace one of the basis vectors while preserving linear independence, you can make that exchange without losing the ability to represent all vectors in the space.

Geometrically, this means that if you have a set of vectors forming a basis, each vector contributes a unique direction to the space. If you find another vector that can take the place of one of the basis vectors without introducing any redundant direction, you can seamlessly integrate it into the basis. This process allows for flexibility in choosing basis vectors while ensuring that the resulting set still forms a complete "framework" for the vector space.

Theorem on Basis Size Consistency

This theorem establishes a fundamental property of bases in finite-dimensional vector spaces. It asserts that all bases of a finite-dimensional vector space have the same number of elements. In other words, the size of a basis is an intrinsic property of the vector space itself and does not depend on the specific choice of basis vectors. This theorem underscores the uniformity and consistency of bases across a given vector space.

Geometrically, this theorem reflects the uniformity of bases across a given vector space. Imagine a basis as a set of vectors that forms the "standard directions" or "axes" along which you can measure and describe vectors within the space. The theorem asserts that, regardless of the specific choice of basis vectors, the number of these "standard directions" remains constant for a finite-dimensional

vector space.

In essence, this theorem emphasizes the intrinsic properties of vector spaces that dictate the number of dimensions needed to describe them fully. It implies that the dimensionality of a space is a well-defined characteristic that transcends the individual choice of basis vectors, highlighting the consistency and universality of bases within the same vector space.

1.4.2 Dimension

The *dimension* of a vector space is a fundamental concept that quantifies the "size" or "degree of freedom" of the space. For a vector space V , the dimension is defined as the cardinality of a basis of V . The *dimension* of the trivial vector space is equal to 0, since its basis is \emptyset .

A vector space V with a basis B is *finite-dimensional* if the basis B has finitely many vectors

Dimensionality Corollary: Limits of Linear Independence

The corollary regarding the dimension of linearly independent sets states that in a vector space, the maximum number of linearly independent vectors within a set is bounded by the dimensionality of the space. This implies that adding more vectors beyond this limit leads to linear dependence, compromising the independence of the set.

Geometrically, it signifies that in a space of a given dimension, there is a finite capacity for representing distinct directions or degrees of freedom with linearly independent vectors. Algebraically, it underscores the role of dimensionality as an upper bound for the size of independent sets, reflecting the inherent structure of the space. This corollary guides the analysis of systems of equations, subspaces, and transformations by providing insight into the relationship between linear independence and space dimension.

Ultimately, it illuminates the fundamental connection between the structure of vector spaces and the properties of their subsets, offering a foundational principle for mathematical reasoning and analysis.

Dimensionality Corollary: Minimum Spanning Sets

The corollary states that any spanning set in a vector space must contain at least as many vectors as the dimension of the space. In a vector space of dimension n , a spanning set cannot have fewer than n vectors. This is because a spanning set must be able to generate all vectors within the space. If it had fewer than n vectors, there would not be enough basis vectors to span the entire

space. Thus, the minimum size of a spanning set is the dimension of the vector space.

This corollary highlights a fundamental property of spanning sets: they must have sufficient vectors to cover the entire space, ensuring that any vector in the space can be expressed as a linear combination of those spanning vectors. It underscores the relationship between the dimension of a vector space and the minimum number of vectors required to span that space effectively.

Interconnection of Linear Independence and Spanning Sets in Vector Spaces

Combining the two previous results yields a fundamental understanding of linear independence and spanning sets in a vector space. The first corollary establishes that the maximum number of linearly independent vectors within a set is bounded by the dimensionality of the vector space. It implies that exceeding this limit results in linear dependence, compromising the set's independence. Conversely, the second result, the theorem, states that any spanning set must contain at least n vectors in a vector space of dimension n . If it had fewer than n vectors, there would not be enough vectors to span the entire space.

Therefore, the duality between linear independence and spanning sets is clear: a set of n linearly independent vectors is guaranteed to span a space of dimension n , while a spanning set with n vectors ensures linear independence and sufficiency for spanning the space. This interconnection underscores the importance of these concepts in understanding the structure and properties of vector spaces.

Dimension Inequality

If V is finite-dimensional and U is a subspace of V , then $\dim U \leq \dim V$. This inequality highlights that the dimension of a subspace cannot exceed the dimension of its parent vector space.

Geometrically, this statement reflects the idea that a subspace cannot "stick out" beyond the dimensions of its parent space. If you imagine a three-dimensional room (parent space) and a two-dimensional wall within it (subspace), the wall cannot extend beyond the boundaries of the room. Therefore, the dimension of the subspace must be less than or equal to the dimension of the parent space.

Equality of Dimensions

Suppose V is finite-dimensional and U is a subspace of V such that $\dim U = \dim V$. Then $U = V$. This equality asserts that when the dimensions of a subspace and its parent space are equal, the subspace spans the entire parent space.

This concept can be visualized as a smaller room (subspace) filling up the entire space of a larger room (parent space). If both rooms have the same dimensions, it means the smaller room perfectly fits within the larger one, occupying all available space. Geometrically, this represents a complete overlap between the subspace and the parent space.

Dimension of Sum of Subspaces

For subspaces V_1 and V_2 of a finite-dimensional vector space, $\dim(V_1 + V_2) = \dim V_1 + \dim V_2 - \dim(V_1 \cap V_2)$. This equation determines the dimension of the sum of two subspaces based on their individual dimensions and their intersection.

Geometrically, this theorem describes how the dimensions of subspaces interact when they are combined. It's like stacking two rooms (subspaces) on top of each other within a building (parent space). The combined space's dimension is determined by adding the individual dimensions of each room and then adjusting for any shared areas between them. This reflects the total usable space created by the combination of both subspaces.

1.5 Rows and Columns Space

In this section, we will investigate the vector space spanned by the row vectors (or column vectors) of a matrix. And, in the next chapter, we will see how such spaces relate to solutions of systems of linear equations.

To begin, you need to know some terminology. For an $m \times n$ matrix

$$A = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{bmatrix},$$

the n -tuples corresponding to the rows of A are called the row vectors of A .

Row Vectors of A :

$$\begin{aligned}\vec{r}_1 &= \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \end{pmatrix} \\ \vec{r}_2 &= \begin{pmatrix} a_{2,1} & a_{2,2} & \dots & a_{2,n} \end{pmatrix} \\ &\vdots \\ \vec{r}_m &= \begin{pmatrix} a_{m,1} & a_{m,2} & \dots & a_{m,n} \end{pmatrix}\end{aligned}$$

Similarly, the columns of A are called the column vectors of A . You will find it useful to preserve the column notation for these column vectors.

Column Vectors of A :

$$\vec{c}_1 = \begin{pmatrix} a_{1,1} \\ a_{2,1} \\ \vdots \\ a_{m,1} \end{pmatrix} \quad \begin{pmatrix} a_{1,2} \\ a_{2,2} \\ \vdots \\ a_{m,2} \end{pmatrix} \quad \dots \quad \begin{pmatrix} a_{1,n} \\ a_{2,n} \\ \vdots \\ a_{m,n} \end{pmatrix}$$

Two fundamental concepts associated with matrices are the row space and column space, which offer deep insights into their structure and behavior.

Consider a matrix A with dimensions $m \times n$. The row space of A , denoted as $\text{Row}(A)$, encompasses all possible linear combinations of its row vectors. Mathematically, it can be represented as the subspace spanned by the row vectors of A in \mathbb{R}^n .

Formally:

$$\text{Row}(A) = \text{span}\{\vec{r}_1, \vec{r}_2, \dots, \vec{r}_m\}$$

where \vec{r}_i represents the i -th row vector of A .

Similarly, the column space of A , denoted as $\text{Col}(A)$, captures all possible linear combinations of its column vectors. It forms a subspace spanned by the column vectors of A in \mathbb{R}^m .

Formally:

$$\text{Col}(A) = \text{span}\{\vec{c}_1, \vec{c}_2, \dots, \vec{c}_n\}$$

where \vec{c}_i represents the i -th column vector of A .

1.5.1 Row-Equivalent Matrices and Row Space

The row space of a matrix refers to the space formed by all possible linear combinations of its row vectors. Two matrices are considered row-equivalent if one can be obtained from the other through elementary row operations, such as multiplying a row by a number or adding one row to another.

If matrix A can be turned into matrix B through these operations, then the row space of A is exactly the same as the row space of B .

This is because the rows of B can be represented as combinations of the rows of A , and conversely, the rows of A can be expressed as combinations of the rows of B . Hence, the row spaces of A and B are identical.

Remark. The concept previously highlighted is significant: it indicates that elementary row operations don't change the row space of a matrix. However, it's important to note that these operations might impact the column space.

1.5.2 Basis for Row Space

When a matrix is in row-echelon form, its nonzero row vectors form a set that is both linearly independent and spans the row space. This set serves as a basis for the row space.

So, if a matrix A is row-equivalent to a matrix B in row-echelon form, then the nonzero row vectors of B form a basis for the row space of A .

1.5.3 Basis for Column Space

Just calculate the transpose of the matrix and find a basis for row space.

1.5.4 Row and Column Spaces Dimension

The row space and column space of a matrix A are fundamental concepts representing the span of vectors formed by its rows and columns, respectively. Let A be an $m \times n$ matrix, the row space and column space of A share a key property: they have the same dimension.

Let $\vec{r}_1, \vec{r}_2, \dots, \vec{r}_m$ denote the row vectors and $\vec{c}_1, \vec{c}_2, \dots, \vec{c}_n$ denote the column vectors of A .

Suppose the row space of A has dimension s , spanned by linearly independent vectors $\vec{b}_1, \vec{b}_2, \dots, \vec{b}_s$. Each row vector of A can then be expressed as a linear combination of these basis vectors.

By representing each row vector as a linear combination of the basis vectors,

$$\vec{r}_i = \alpha_{i,1}\vec{b}_1 + \cdots + \alpha_{i,s}\vec{b}_s$$

we can analyze the relationships between the entries of A and the basis vectors.

Isolating the entries corresponding to each column of A , we obtain systems of scalar equations relating the entries of A to the basis vectors.

So for each column j in the matrix A , we obtain the following system of linear equations:

$$\begin{cases} \alpha_{1,1}\vec{b}_1^j + \alpha_{1,2}\vec{b}_2^j + \cdots + \alpha_{1,s}\vec{b}_s^j = \vec{r}_1^j \\ \alpha_{2,1}\vec{b}_1^j + \alpha_{2,2}\vec{b}_2^j + \cdots + \alpha_{2,s}\vec{b}_s^j = \vec{r}_2^j \\ \vdots \\ \alpha_{m,1}\vec{b}_1^j + \alpha_{m,2}\vec{b}_2^j + \cdots + \alpha_{m,s}\vec{b}_s^j = \vec{r}_m^j \end{cases}$$

where $\vec{c}_j = [\vec{r}_1^j, \vec{r}_2^j, \dots, \vec{r}_m^j]^T$. With this in mind, we can now write each column vector c_j of A as

$$\vec{c}_j = \begin{bmatrix} \alpha_{1,1} \\ \alpha_{2,1} \\ \vdots \\ \alpha_{m,1} \end{bmatrix} \vec{b}_1^j + \begin{bmatrix} \alpha_{1,2} \\ \alpha_{2,2} \\ \vdots \\ \alpha_{m,2} \end{bmatrix} \vec{b}_2^j + \cdots + \begin{bmatrix} \alpha_{1,s} \\ \alpha_{2,s} \\ \vdots \\ \alpha_{m,s} \end{bmatrix} \vec{b}_s^j$$

So the column space is spanned by s vectors; the dimension of the column space of A cannot exceed the dimension of the row space of A , i.e., $\dim(\text{column space of } A) \leq \dim(\text{row space of } A)$. Considering the transpose of A , denoted as A^T , and repeating the analysis, we conclude that $\dim(\text{column space of } A^T) \leq \dim(\text{row space of } A^T)$.

This mutual limitation implies that the dimensions of the row space and column space of A are equal, highlighting the profound connection between the rows and columns of a matrix.

1.5.5 The Rank of a Matrix

Since the column-spanned vector space has the same dimension as the row-spanned vector space, for each matrix we have a new feature to describe and analyse.

Definition

The rank of a matrix A is the dimension of the vector space spanned by its column/row vectors. In simpler terms, it is the maximum number of linearly independent columns (or rows) in the matrix A .

Computing the Rank

To compute the rank of a matrix, we typically use row reduction techniques to transform the matrix into row-echelon form or reduced row-echelon form. The number of non-zero rows (or pivot columns) in the reduced form is equal to the rank of the matrix.

1.6 Intuitive Approach to Basis Change

Let $B = \{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n\}$ and $B' = \{\vec{u}_1, \vec{u}_2, \dots, \vec{u}_n\}$ be two bases for a vector space V .

Compute the coordinates of a vector with respect to a basis is a way to represent a vector as a linear combination of the basis vectors, finding the coefficient to apply to each basis vector.

We can find a way to "translate" the coordinates of a vector \vec{x} from a basis B to a basis B' (holding the same vector). Suppose \vec{x} has coordinates $[\vec{x}]_B$ with respect to B and coordinates $[\vec{x}]_{B'}$ with respect to B' . Then, there exists a matrix P , such that:

$$[\vec{x}]_{B'} = P[\vec{x}]_B$$

where P is the transition matrix from B to B' .

To find out how the vector coordinates change from a basis to an other, we must know the coordinates of the basis B vectors with respect to B' basis.

Let $\vec{x} = d_1\vec{v}_1 + d_2\vec{v}_2 + \dots + d_n\vec{v}_n$ be an arbitrary vector in V . The coordinate vector of \vec{x} with respect to the basis B is

$$[\vec{x}]_B = \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_n \end{bmatrix}.$$

If

$$\begin{aligned}\vec{v}_1 &= c_{1,1}\vec{u}_1 + c_{2,1}\vec{u}_2 + \dots + c_{n,1}\vec{u}_n \\ \vec{v}_2 &= c_{1,2}\vec{u}_1 + c_{2,2}\vec{u}_2 + \dots + c_{n,2}\vec{u}_n \\ &\vdots \\ \vec{v}_n &= c_{1,n}\vec{u}_1 + c_{2,n}\vec{u}_2 + \dots + c_{n,n}\vec{u}_n,\end{aligned}$$

then you can write

$$\begin{aligned}\vec{x} &= d_1\vec{v}_1 + d_2\vec{v}_2 + \dots + d_n\vec{v}_n \\ &= d_1(c_{1,1}\vec{u}_1 + \dots + c_{n,1}\vec{u}_n) + \dots + d_n(c_{1,n}\vec{u}_1 + \dots + c_{n,n}\vec{u}_n) \\ &= (d_1c_{1,1} + \dots + d_nc_{1,n})\vec{u}_1 + \dots + (d_1c_{n,1} + \dots + d_nc_{n,n})\vec{u}_n,\end{aligned}$$

which implies

$$[\vec{x}]_{B'} = \begin{bmatrix} c_{1,1}d_1 + c_{1,2}d_2 + \dots + c_{1,n}d_n \\ c_{2,1}d_1 + c_{2,2}d_2 + \dots + c_{2,n}d_n \\ \vdots \\ c_{n,1}d_1 + c_{n,2}d_2 + \dots + c_{n,n}d_n \end{bmatrix} = \begin{bmatrix} c_{1,1} & c_{1,2} & \dots & c_{1,n} \\ c_{2,1} & c_{2,2} & \dots & c_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{n,1} & c_{n,2} & \dots & c_{n,n} \end{bmatrix} \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_n \end{bmatrix} = P[\vec{x}]_B.$$

So, the transition matrix from B to B' is

$$P = \begin{bmatrix} c_{1,1} & c_{1,2} & \dots & c_{1,n} \\ c_{2,1} & c_{2,2} & \dots & c_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{n,1} & c_{n,2} & \dots & c_{n,n} \end{bmatrix}.$$

In the section 2.3 of the next chapter, we introduce a general method to compute the transition matrices between two coordinate systems of a vector space.

Chapter 2

Linear Systems

2.1 Definition

A system of linear equations consists of a set of linear equations involving the same variables. Consider a set of m linear equations in n variables, a linear system can be defined as follows:

$$\begin{cases} a_{1,1}x_1 + a_{1,2}x_2 + \dots + a_{1,n}x_n = b_1 \\ a_{2,1}x_1 + a_{2,2}x_2 + \dots + a_{2,n}x_n = b_2 \\ \vdots \\ a_{m,1}x_1 + a_{m,2}x_2 + \dots + a_{m,n}x_n = b_m \end{cases}$$

where x_1, x_2, \dots, x_n are the variables, $a_{ij} \in \mathbb{F}$ are the coefficients, and $b_1, b_2, \dots, b_m \in \mathbb{F}$ are the constants. Each equation in the system is a linear combination of the variables with coefficients a_{ij} , and the objective is to find values of the variables x_1, x_2, \dots, x_n that satisfy all equations simultaneously. These values are called the solutions to the linear system.

A linear system can be represented in matrix form as $\mathbf{A}\vec{x} = \vec{b}$, where \mathbf{A} is a $m \times n$ coefficient matrix, \vec{x} is the column vector of variables, and \vec{b} is the column vector of constants.

So it can be expressed in this form

$$\begin{bmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \dots & a_{m,n} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}.$$

Our goal is to find values for each x_i that satisfy the equality.

When the constants vector, often referred to as the right-hand side vector or the vector of constants, is the zero vector ($\vec{0}$), the system is indeed called a homogeneous system of linear equations. In such a system, all equations are equal to zero.

2.1.1 Geometric Interpretation

In a geometric sense, each equation in the linear system represents a hyperplane in the n -dimensional space, where n is the number of variables. For example, in two dimensions ($n = 2$), each equation represents a line, while in three dimensions ($n = 3$), each equation represents a plane.

The solution to the linear system corresponds to the intersection of these hyper-planes.

So there are three main cases:

- If all hyper-planes intersect at a single point, then the system has a unique solution.
- If the hyper-planes are coincident in some way, then the system has infinitely many solutions. Geometrically, this means that the hyper-planes overlap, forming a line, plane, or higher-dimensional subspace of solutions.
- If the hyper-planes do not intersect, then the system has no solution. This means that there is no common solution to all equations.

Given a linear system of equations, just one of these cases can be satisfied.

2.1.2 Associated Homogeneous System

The associated homogeneous system to a system of linear equations is obtained from the original system by substituting all the known terms with zeros.

For example, for the system:

$$\begin{bmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \dots & a_{m,n} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

the associated homogeneous system is:

$$\begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

The solutions to a homogeneous system of linear equations can fall into one of two categories:

- **The trivial solution:** In which all variables are equal to zero. This solution is always present in a homogeneous system.
- **Non-trivial solutions:** In which at least one variable is not zero. The existence of non-trivial solutions depends on the properties of the coefficient matrix and is determined through techniques such as Gaussian elimination or matrix inversion.

If A is an $m \times n$ matrix, then the set of all solutions of the homogeneous system of linear equations $\mathbf{A}\vec{x} = \vec{0}$ is a subspace of \mathbb{R}^n called the *nullspace* of A and is denoted by $N(A)$. So,

$$N(A) = \{\vec{x} \in \mathbb{R}^n : \mathbf{A}\vec{x} = \vec{0}\}.$$

The dimension of the *nullspace* of A is called the **nullity** of A .

Because A is an $m \times n$ matrix, you know that \vec{x} has size $n \times 1$. So, the set of all solutions of the system is a subset of \mathbb{R}^n . This set is clearly nonempty, because $A\vec{0} = \vec{0}$. You can verify that it is a subspace of \mathbb{R}^n by showing that it is closed under the operations of addition and scalar multiplication.

If the rank of matrix A is equal to r , then the dimension of the solution space of $Ax = 0$ is $n - r$. That is,

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

Because A has rank r , you know it is row-equivalent to a reduced row-echelon matrix B with r nonzero rows. No generality is lost by assuming that the upper left corner of B has the form of

the $r \times r$ identity matrix I_r (sorting the columns if it's need). Moreover, because the zero rows of B contribute nothing to the solution, you can discard them to form the $r \times n$ matrix B' , where $B' = [I_r \mid C]$. The matrix C has $n - r$ columns corresponding to the variables $x_{r+1}, x_{r+2}, \dots, x_n$.

So, the solution space of $A\vec{x} = \vec{0}$ can be represented by the system:

$$\begin{aligned} x_1 + & c_{11}x_{r+1} + c_{12}x_{r+2} + \cdots + c_{1,n-r}x_n = 0 \\ x_2 + & c_{21}x_{r+1} + c_{22}x_{r+2} + \cdots + c_{2,n-r}x_n = 0 \\ & \vdots \\ x_r + & c_{r1}x_{r+1} + c_{r2}x_{r+2} + \cdots + c_{r,n-r}x_n = 0. \end{aligned}$$

Solving for the first r variables in terms of the last $n - r$ variables produces $n - r$ vectors in the basis of the solution space. Consequently, the solution space - the nullspace - has dimension $n - r$.

Ok, but, intuitively, why is the nullity equals to $n - r$?

When we write the pivot variables in terms of free variables:

$$\begin{aligned} x_1 &= a_{1,1}x_{r+1} + a_{1,2}x_{r+2} + \cdots + a_{1,n-r}x_n \\ x_2 &= a_{2,1}x_{r+1} + a_{2,2}x_{r+2} + \cdots + a_{2,n-r}x_n \\ &\vdots \\ x_r &= a_{r,1}x_{r+1} + a_{r,2}x_{r+2} + \cdots + a_{r,n-r}x_n \end{aligned}$$

So we can write the solutions in this way

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_r \\ x_{r+1} \\ \vdots \\ x_n \end{bmatrix} = x_{r+1} \begin{bmatrix} a_{1,1} \\ a_{2,1} \\ \vdots \\ a_{r,1} \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + x_{r+2} \begin{bmatrix} a_{1,2} \\ a_{2,2} \\ \vdots \\ a_{r,2} \\ 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix} + \cdots + x_n \begin{bmatrix} a_{1,n} \\ a_{2,n} \\ \vdots \\ a_{r,n} \\ 0 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$$

As you can see, the solution/null space is spanned by $n - r$ linearly independent vectors. Thus, this is a basis for nullspace, this means that the dimension/nullity of the solution space is $n - r$.

2.2 Solving Linear Systems

In this section, we will understand how the Gauss-Jordan algorithm can be a powerful method for solving linear systems by transforming the augmented matrix into reduced row-echelon form. We will walk through each step of the algorithm, providing detailed explanations and examples to illustrate its implementation and effectiveness in finding solutions to linear systems.

Here's how the process works:

Step 1: Augmented Matrix

Start with the augmented matrix of the linear system, combining the coefficient matrix and the constants vector:

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

Step 2: Gaussian-Jordan Elimination

Perform row operations to transform the augmented matrix into reduced row-echelon form, similar to the Gaussian elimination method. The objective is to create leading ones (pivots) in each row and eliminate entries up and below the pivots:

1. Start with the leftmost column.
2. Make the first nonzero entry in the column (the pivot) equal to 1 by scaling the row.
3. Use row operations to eliminate all entries up and below the pivot, making them zero.
4. Move to the next column and repeat the process until the matrix is in reduced row-echelon form.

Step 3: Determine the solution set

Once the matrix is in reduced row-echelon form, we have we have a clear understanding of the system's solution space and can easily determine the presence of unique solutions, infinite solutions, or inconsistency.

In fact, at the end of the Gauss-Jordan elimination algorithm, there are a few possible outcomes that determine the nature of the solution set for the system of linear equations:

- **Unique Solution:** If every column contains a leading 1 (pivot) and every other entry in the column is 0, then the system has a unique solution. In this case, there are no free variables, and the solution set consists of a single point in n -dimensional space. In this case, the rank of the coefficient matrix equals the number of unknowns (i.e., the number of columns), so the system has a unique solution. This occurs when the columns of the matrix are linearly independent, allowing us to uniquely determine the values of the unknowns.
- **Inconsistent System:** If during the elimination process, a row of the form $[0 \cdots 0 \mid d]$, where $d \neq 0$, is encountered, then the system is inconsistent and has no solution. This indicates that the system of equations represents parallel planes in space, which do not intersect. In this case, the rank of the coefficient matrix is less than the number of unknowns and also less than the number of non-zero rows in the augmented matrix, so the system is inconsistent and has no solution. This occurs when the equations represented by the rows of the matrix are linearly dependent, leading to contradictory constraints.
- **Infinite Solutions:** If there are one or more rows with a leading 1 followed by columns with non-leading entries (not all zeros), then the system has infinitely many solutions. These non-leading entries correspond to free variables. The solution set is expressed parametrically in terms of these free variables. In this case, the rank of the coefficient matrix is less than the number of unknowns but greater than the number of non-zero rows in the augmented matrix (indicating the presence of free variables after row reduction), so the system has infinitely many solutions. In this case, the rank deficiency corresponds to the dimension of the solution space's non-trivial part.

In the first case, the solution set contain a single column vector called particular solution.

Generally, the solution set can be expressed in terms of both a particular solution and the parameters corresponding to the free variables. This involves finding a particular solution by setting the free variables to arbitrary values and then expressing the general solution as a linear combination of the particular solution and the solutions corresponding to the free variables in a way that

General = Particular + Homogeneous.

Remark. The general solution is given by the sum of a particular solution and the homogeneous solution set. When the system has just one solution, the homogeneous solution set contain only the trivial solution.

Let's denote the particular solution as \vec{p} and the solutions corresponding to the free variables as $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k$. Then the general solution set can be expressed as:

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

where c_1, c_2, \dots, c_k are arbitrary constants.

So, the general solution set of a system of linear equation equals

$$\{\vec{p} + \vec{h} \mid \vec{h} \text{ satisfies the associated homogeneous system}\}.$$

We can interpret solving linear systems as a way to find the coordinates of the constants vector with respect to a given column vectors span. When the column vectors are linearly independent, we have just one way to represent the constants vector, because we have a basis for the column-spanned vector space (due to the uniqueness of coordinates of a vector with respect to a basis). In the case of an inconsistent system, we have a column-spanned vector space that has a smaller dimension than vector dimension, so (except for the 0 case), we can't control all dimensions of the constants vector with the given vectors (matrix of coefficients). When we have infinitely many solutions, we have infinitely many ways to represent the constants vector, because the column vectors are linearly dependent. So, for example, if we have one free parameter in our solution set, it means that we can shrink the set of column vectors to obtain a basis. With respect to this basis, we can represent all vectors in the vector space spanned by it. With our solution set, because of the free parameter vector, we can achieve infinitely many vectors; so to obtain infinitely many vectors, we must have infinitely many coordinates/solutions.

2.3 In depth Basis Change

In the previous chapter, we have introduced the main idea behind basis change and explained what transition matrix is. In this section, we will delve into the details of how the transformation matrix between one basis and another of a vector space is computed, and we will provide an intuitive interpretation of why this process works.

Let $B = \{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n\}$ and $B' = \{\vec{u}_1, \vec{u}_2, \dots, \vec{u}_n\}$ be two bases for \mathbb{R}^n .

How can we find the transformation matrix from B' to B without knowing the coordinates of vectors of B with respect to B' ?

We have to find a way to express the vectors of B' as a linear combination of vectors of B in some way. For each vector component we can write a system of n linear equations with n unknown.

$$\left\{ \begin{array}{l} x_{1,1}\vec{v}_1 + \dots + x_{n,1}\vec{v}_n = \vec{u}_1 \\ \vdots \\ x_{1,n}\vec{v}_1 + \dots + x_{n,n}\vec{v}_n = \vec{u}_n \end{array} \right. \longrightarrow \left\{ \begin{array}{l} x_{1,1}\vec{v}_1^1 + \dots + x_{n,n}\vec{v}_n^1 = \vec{u}_1^1 \\ \vdots \\ x_{1,1}\vec{v}_1^n + \dots + x_{n,n}\vec{v}_n^n = \vec{u}_1^n \\ \vdots \\ x_{n,1}\vec{v}_1^1 + \dots + x_{n,n}\vec{v}_n^1 = \vec{u}_n^1 \\ \vdots \\ x_{n,1}\vec{v}_1^n + \dots + x_{n,n}\vec{v}_n^n = \vec{u}_n^n \end{array} \right.$$

So, since we have common coefficients, we can put all known terms (the vectors \vec{v}_i in this case) in a matrix and augment it with all known constants (the vectors \vec{u}_i in this case):

$$\left[\begin{array}{cccc|cccc} \vec{v}_1^1 & \vec{v}_2^1 & \dots & \vec{v}_n^1 & \vec{u}_1^1 & \vec{u}_2^1 & \dots & \vec{u}_n^1 \\ \vec{v}_1^2 & \vec{v}_2^2 & \dots & \vec{v}_n^2 & \vec{u}_1^2 & \vec{u}_2^2 & \dots & \vec{u}_n^2 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \vec{v}_1^n & \vec{v}_2^n & \dots & \vec{v}_n^n & \vec{u}_1^n & \vec{u}_2^n & \dots & \vec{u}_n^n \end{array} \right].$$

$\underbrace{\hspace{10em}}_B$
 $\underbrace{\hspace{10em}}_{B'}$

After the Gauss-Jordan reduction we have this situation

$$\left[\begin{array}{cccc|cccc} 1 & 0 & \cdots & 0 & x_{1,1} & x_{1,2} & \cdots & x_{1,n} \\ 0 & 1 & \cdots & 0 & x_{2,1} & x_{2,2} & \cdots & x_{2,n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 & x_{n,1} & x_{n,2} & \cdots & x_{n,n} \end{array} \right]$$

As we can see, we have obtained a bijective correspondence between the vectors of B and the vectors of B' , because we can write the vectors of B' as a unique combinations of vectors of B in this way:

$$\begin{aligned} \vec{u}_1 &= x_{1,1}\vec{v}_1 + x_{1,2}\vec{v}_2 + \cdots + x_{1,n}\vec{v}_n \\ \vec{u}_2 &= x_{2,1}\vec{v}_1 + x_{2,2}\vec{v}_2 + \cdots + x_{2,n}\vec{v}_n \\ &\vdots \\ \vec{u}_n &= x_{n,1}\vec{v}_1 + x_{n,2}\vec{v}_2 + \cdots + x_{n,n}\vec{v}_n \end{aligned}.$$

As we have previously ascertained, our transition matrix from B' to B is

$$X = \begin{bmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,n} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n,1} & x_{n,2} & \cdots & x_{n,n} \end{bmatrix}.$$

So we can transform the coordinates of a generic vector \vec{w} , expressed through the B' basis, in the coordinates with respect to B basis in this way:

$$[\vec{w}]_B = X[\vec{w}]_{B'}.$$

The same process can be applied to find the transition matrix from B to B' , just swap the two bases in the augmented matrix and solve the linear systems.

Chapter 3

Linear Maps

3.1 Introduction to Linear Maps

A linear map T between vector spaces V and W is a function that preserves vector addition and scalar multiplication. Intuitively, this means that if you apply the linear map to a combination of vectors (either by adding them or scaling them), the result is the same as applying the linear map to each individual vector and then combining the results.

Every linear transformation can be represented by a matrix. Let

$$T : \mathbb{R}^n \rightarrow \mathbb{R}^m$$

be a linear transformation. Then there exists an $m \times n$ matrix A such that for any vector \vec{x} in \mathbb{R}^n , $T(\vec{x}) = A\vec{x}$.

So, for any linear transformation T that maps from an n -dimensional space to an m -dimensional space, there exists a specific $m \times n$ matrix A such that when you multiply this matrix by any vector \vec{x} from the original space, you get the transformed vector $T(\vec{x})$ in the new space. This matrix A essentially encodes all the information about how the linear transformation T operates on the vectors in the original space.

The reason why this is true intuitively can be understood by considering how linear transformations operate on vectors and how matrices represent these transformations.

A matrix essentially encodes the action of the transformation on the basis vectors of the original space. When you multiply this matrix by a vector, you're essentially applying the transformation to each component of the vector and then combining the results. This process is consistent with the properties of linear transformations, as it preserves vector addition and scalar multiplication.

Therefore, the result in the new space is a combination of the transformed components, which accurately reflects the action of the linear transformation on the original vector.

3.2 Properties of Linear Maps

Consider a linear map

$$T : V \rightarrow W,$$

where V and W are two vector spaces over the same field; and two vectors: $\vec{v} \in V$ and $\vec{w} \in W$.

The linear map T exhibit several important properties:

- **Linearity:** A linear map preserves vector addition and scalar multiplication. In other words, for any vectors \vec{w} and \vec{v} and any scalar c , we have:

$$T(\vec{w} + \vec{v}) = T(\vec{w}) + T(\vec{v})$$

$$T(c \cdot \vec{w}) = c \cdot T(\vec{w})$$

This property ensures that the structure of the vector space is preserved under the linear map.

- **Kernel:** The kernel (or null space) of a linear map T consists of all vectors \vec{v} such that $T(\vec{v}) = \vec{0}$. Intuitively, the kernel represents the set of vectors that are mapped to the zero vector by T . The kernel is a subspace of the domain vector space.
- **Image:** The image (or range) of a linear map T is the set of all vectors \vec{w} in W that can be expressed as $T(\vec{v})$ for some \vec{v} in V . In other words, the image represents the set of all possible outputs of T . It is a subspace of the codomain vector space.

3.2.1 Find the Kernel of a Linear Map

To find the kernel of a linear map, we use homogeneous systems of linear equations. Let $T : V \rightarrow W$ be a linear map between vector spaces V and W . The kernel of T , denoted $\ker(T)$, is defined as the set of all vectors in V that map to the zero vector in W .

Consider the homogeneous system of linear equations associated with the linear map T :

$$T(\mathbf{x}) = \mathbf{0}$$

where \mathbf{x} is a vector in V and $\mathbf{0}$ is the zero vector in W .

To find the kernel of T , we solve this system of equations. This can be done by expressing T in terms of its matrix representation A with respect to some basis of V and W . Let $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ be a basis for V , and $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_m$ be a basis for W . Then the matrix representation of T is given by:

$$A = \begin{bmatrix} | & | & & | \\ [T(\mathbf{v}_1)]_B & [T(\mathbf{v}_2)]_B & \cdots & [T(\mathbf{v}_n)]_B \\ | & | & & | \end{bmatrix}$$

where $[T(\mathbf{v}_i)]_B$ represents the coordinate vector of $T(\mathbf{v}_i)$ with respect to the basis B of W .

Once we have the matrix A , we solve the system of linear equations $A\mathbf{x} = \mathbf{0}$ to find a basis for the kernel of T . The solutions to this system represent the vectors \mathbf{x} in V that map to the zero vector in W , i.e., elements of $\ker(T)$.

3.2.2 Find the Image of a Linear Map

To find the image of a linear map, we can utilize the transpose of the matrix representation of the linear transformation and row reduction techniques. Let $T : V \rightarrow W$ be a linear map between vector spaces V and W . The image of T , denoted $\text{Im}(T)$, is defined as the set of all vectors in W that are obtained by applying T to vectors in V .

To find the image of T , we first express T in terms of its matrix representation A with respect to some basis of V and W . Let $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ be a basis for V , and $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_m$ be a basis for W . Then the matrix representation of T is given by:

$$A = \begin{bmatrix} | & | & & | \\ [T(\mathbf{v}_1)]_B & [T(\mathbf{v}_2)]_B & \cdots & [T(\mathbf{v}_n)]_B \\ | & | & & | \end{bmatrix}$$

where $[T(\mathbf{v}_i)]_B$ represents the coordinate vector of $T(\mathbf{v}_i)$ with respect to the basis B of W .

Next, we perform row reduction on A to obtain its row echelon form (or reduced row echelon form). This process helps us to identify the linearly independent columns of A .

The image of T is then given by the span of the pivot columns of A . These columns form a basis for the image of T , and thus, $\text{Im}(T)$ can be represented as:

$$\text{Im}(T) = \text{span}\{\text{pivot columns of } A\}.$$

To find a basis for $\text{Im}(A)$ you can also find A^T first, and then perform the row-reduction. The first nonzero rows of $\mathbf{RREF}(A^T)$ form a basis for $\text{Im}(A)$. This because with transposition we exchange the columns with rows, and then the row-reduction, that don't change the row space of a matrix, allow us to find linearly independent rows.

Chapter 4

Eigenvalues and Eigenvectors

Chapter 5

Eigenvalues and Eigenvectors

Chapter 6

Inner Product Spaces

Chapter 7

Eigenvalues and Eigenvectors

Chapter 8

Euclidean vector space

Chapter 9

Analytic Geometry in the plane