WARNING!

READ
CAREFULLY!

The following document has not been checked for mistakes, and although every effort possible has been made to ensure that the information in this document is correct, it might still contain some issues, either with the content or with the layout of the document. Therefore, it is given "as is", without any warranty whatsoever.

Neither the Professor nor the TAs have anything to do with such document. If you have issues, visit `https://github.com/pennal/LinAlgNotes`

# Linear Algebra
# Class Notes

*Based on Professor Pivkin's Material*

SPRING 2015
UNIVERSITÀ DELLA SVIZZERA ITALIANA

Version 0.1.24
Generated on July 2, 2015
For the latest version, visit `https://github.com/pennal/LinAlgNotes`

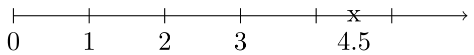
# Contents

# Chapter 1

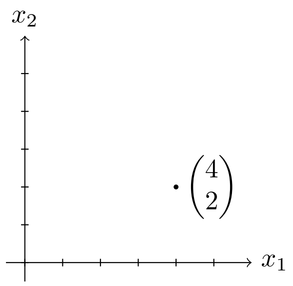
# Vectors

## 1.1 Introduction

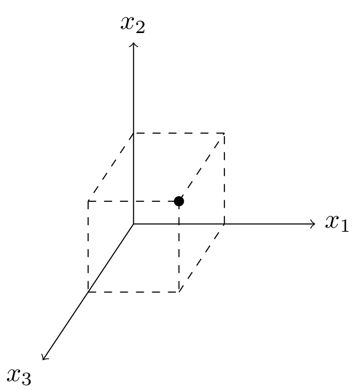
A real number can be represented by a point on a line, which is a 2-dimensional space, $\mathbb{R}$




a pair of real numbers can be represented by a point on a plane, which is a 2-dimensional space, $\mathbb{R}^2$




a triplet of real numbers can be represented by a point in 3D space, $\mathbb{R}^3$



**Definition**

A vector is an ordered collection of $n$ numbers

**Notation**

Usually vectors are given by letters, such as $u, v, w$. In textbooks vectors are written with bold font. In handwriting vectors are often written with a right arrow on top, such as $\overrightarrow{u}$. We will underline vectors, like so: $\underline{u}$.

□

**Definition**

Let us consider vector $\underline{u} \in \mathbb{R}^n$. The $i-$th component of vector

$$\underline{u} = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}$$

is $u_i$

**Example**

$$\underline{u} = \begin{pmatrix} 3 \\ 7 \\ 11 \end{pmatrix} \in \mathbb{R}^3 \Rightarrow u_1 = 3, u_2 = 7, u_3 = 11$$

**Definition**

Let us consider vectors $\underline{u} \in \mathbb{R}^n$ and $\underline{v} \in \mathbb{R}^n$. Vector $\underline{w} \in \mathbb{R}^n$ is a sum of $\underline{u}$ and $\underline{v}$, $\underline{w} = \underline{u} + \underline{v}$, if $w_i = u_i + v_i$ for all $i = 1, \ldots, n$

**Example**

$$\underline{u} = \begin{pmatrix} 3 \\ 5 \\ 1 \end{pmatrix}, \underline{v} = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}, \underline{w} = \underline{u} + \underline{v} = \begin{pmatrix} 3 + (-1) \\ 5 + 0 \\ 1 + 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 5 \\ 2 \end{pmatrix}$$

**Example**

$$\underline{u} = \begin{pmatrix} 3 \\ 9 \\ -2 \end{pmatrix}, \quad \underline{v} = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 0 \end{pmatrix}$$

$\underline{u} + \underline{v}$ is **not** defined: both vectors should have the same number of components.

**Definition**

1. Vectors $\underline{u} \in \mathbb{R}^n$ and $\underline{v} \in \mathbb{R}^n$ are equal, if $u_i = v_i$ for all $i = 1, \ldots, n$
2. A scalar is just another name for real number
3. Let us consider a scalar $\alpha \in \mathbb{R}$ and vector $\underline{u} \in \mathbb{R}^n$. A product of $\alpha$ and $\underline{u}$ is defined as:

$$\alpha \underline{u} = \alpha \cdot \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} = \begin{pmatrix} \alpha \cdot u_1 \\ \vdots \\ \alpha \cdot u_n \end{pmatrix}$$

**Example**

$$\alpha = 3, \underline{u} = \begin{pmatrix} -1 \\ 2 \\ 5 \\ 7 \end{pmatrix} \Rightarrow \alpha \cdot \underline{u} \begin{pmatrix} 3 \cdot -1 \\ 3 \cdot 2 \\ 3 \cdot 5 \\ 3 \cdot 7 \end{pmatrix} = \begin{pmatrix} -3 \\ 6 \\ 15 \\ 21 \end{pmatrix}$$

**Definition**

Let us consider scalars $\alpha$ and $\beta$, and vectors $\underline{u} \in \mathbb{R}^n$ and $\underline{v} \in \mathbb{R}^n$. A sum of $\alpha \cdot \underline{u} + \beta \cdot \underline{v}$ is called a linear combination of vectors $\underline{u}$ and $\underline{v}$.

**Example**

$$2 \cdot \begin{pmatrix} -1 \\ 3 \\ 5 \end{pmatrix} + 3 \cdot \begin{pmatrix} 7 \\ 2 \\ 1 \end{pmatrix} + 5 \cdot \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} = \begin{pmatrix} 24 \\ 12 \\ 8 \end{pmatrix}$$

**Example**

$$\underline{u} - \underline{v} = 1 \cdot \underline{u} + (-1) \cdot \underline{v} = \begin{pmatrix} u_1 - v_1 \\ \vdots \\ u_i - v_i \end{pmatrix}$$

**Note**

$$\underline{u} - \underline{u} = \begin{pmatrix} u_1 - u_1 \\ \vdots \\ u_i - u_i \end{pmatrix} = \underline{0}$$

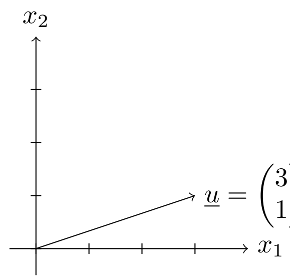
**Definition**

Vector $\underline{u} \in \mathbb{R}^n$ is called a zero vector if all $u_i = 0$, $i = 1, \ldots, n$. The zero vector is often written as $\underline{0} \in \mathbb{R}^n$

## 1.2 Vector Representations and Operations

A vector can be represented in the following ways

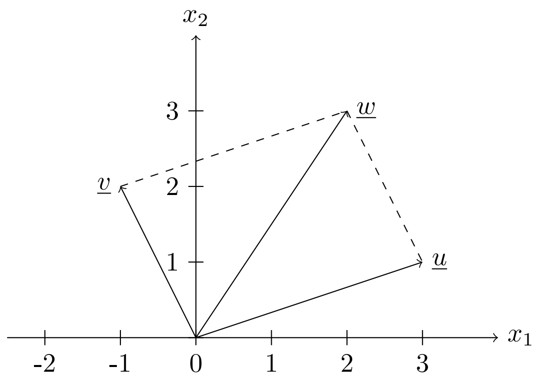
1. As ordered collection of numbers, for example $\underline{u} = \begin{pmatrix} 3 \\ 5 \end{pmatrix}$
2. As an arrow in space




3. As a point in space, the endpoint of a vector from the origin.

### 1.2.1 Combination of Vectors

Let us consider vectors $\underline{u} = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$, $\underline{v} = \begin{pmatrix} -1 \\ 2 \end{pmatrix}$ and $\underline{w} = \underline{u} + \underline{v} = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$. The following picture would represent the vectors



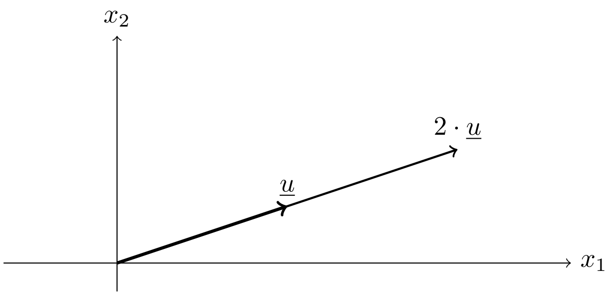

**Example**

Let us consider vector $\underline{u} = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$.

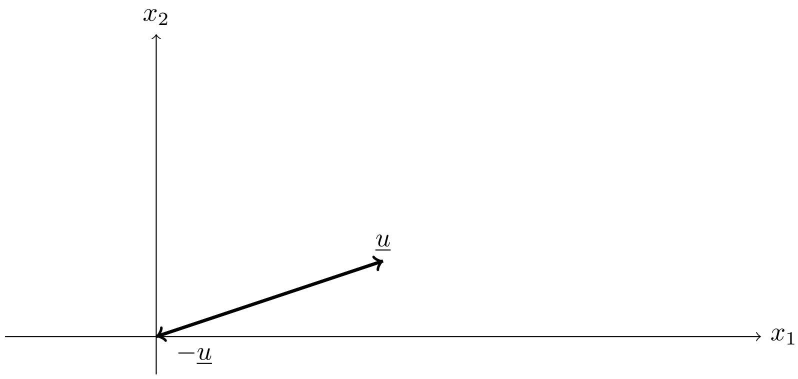
1. What is $2 \cdot \underline{u}$?
   We can calculate it as follows:

$$2 \cdot \underline{u} = 2 \cdot \begin{pmatrix} 3 \\ 1 \end{pmatrix} = \begin{pmatrix} 6 \\ 2 \end{pmatrix}$$

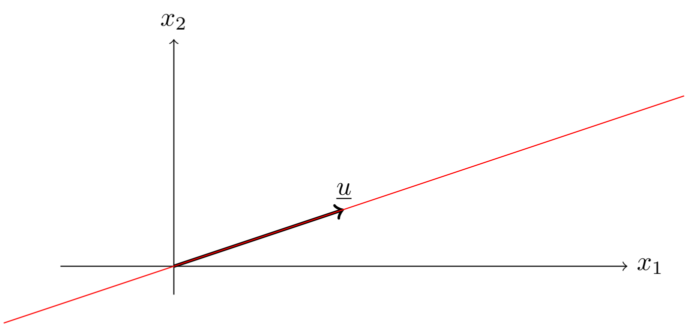
   We stretch vector $\underline{u}$ two times along the line defined by vector $\underline{u}$ as follows:




2. What is $-\underline{u}$?
   Simply reverse the direction of $\underline{u}$ as follows:



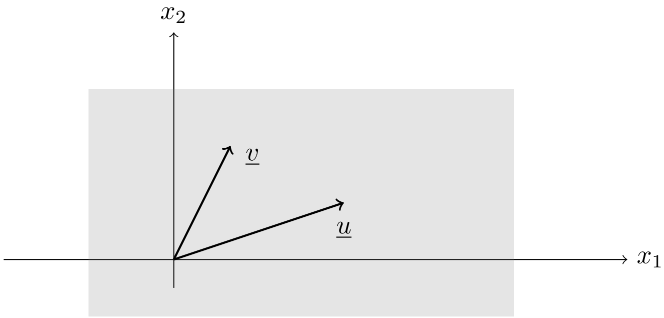

3. What will be the representation of $\alpha\underline{u}$, for all possible values of $\alpha$?
   An endless line as follows



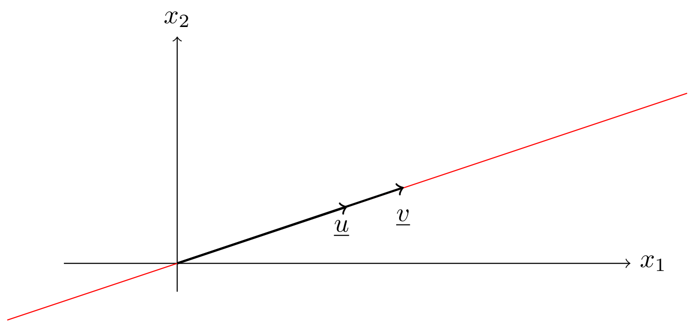
**Example**

Let us consider two vectors $\underline{u} \in \mathbb{R}^2$ and $\underline{v} \in \mathbb{R}^2$. What will be the representation of all linear combinations of $\underline{u}$ and $\underline{v}$, that is what will be $\alpha\underline{u} + \beta\underline{v}$, for all $\alpha$ and $\beta$?

1. A **plane**, if $\underline{u} \neq \underline{0}$ and $\underline{v} \neq \underline{0}$, and $\underline{u}$ and $\underline{v}$ are not on the same line.




2. A **line**, if $\underline{u}$ and $\underline{v}$ are on the same line.




**Note**

Consider $\underline{u}, \underline{v} \in \mathbb{R}^n$. $\underline{u}$ and $\underline{v}$ are on the same line if there exists scalars $\alpha$ and $\beta$ such that $\alpha\underline{u} + \beta\underline{v} = \underline{0}$, when $\alpha$ and $\beta \neq 0$

3. A **point**, if $\underline{u} = \underline{0}$ and $\underline{v} = \underline{0} \Rightarrow \alpha\underline{u} + \beta\underline{v} = \underline{0}$

## 1.3 Dot Product (Scalar product)

**Definition**

Let us consider two vectors $\underline{u} \in \mathbb{R}^n$ and $\underline{v} \in \mathbb{R}^n$. The dot (or scalar) product of vectors $\underline{u}$ and $\underline{v}$ is defined as

$$\langle \underline{u}, \underline{v} \rangle = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n = \sum_{i=1}^{n} u_i v_i$$

**Notation**

We will use $\langle \underline{u}, \underline{v} \rangle$ to denote the dot product, but sometimes $\underline{u} \cdot \underline{v}$ is used.

$\square$

**Example**

$$\underline{u} = \begin{pmatrix} 1 \\ -1 \\ 3 \end{pmatrix}, \underline{v} = \begin{pmatrix} 0 \\ \frac{1}{2} \\ -1 \end{pmatrix}$$

$$\langle \underline{u}, \underline{v} \rangle = 1 \cdot 0 + (-1) \cdot \frac{1}{2} + 3 \cdot (-1) = -3.5$$

**Example**

$$\underline{u} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \underline{v} = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$$

$$\langle \underline{u}, \underline{v} \rangle = 0$$

### 1.3.1 Properties of Dot Product

1. $\langle \alpha \cdot \underline{u}, \underline{v} \rangle = \alpha \cdot \langle \underline{u}, \underline{v} \rangle$ for any $\alpha \in \mathbb{R}, \underline{u} \in \mathbb{R}^n, \underline{v} \in \mathbb{R}^n$.

   **Proof**

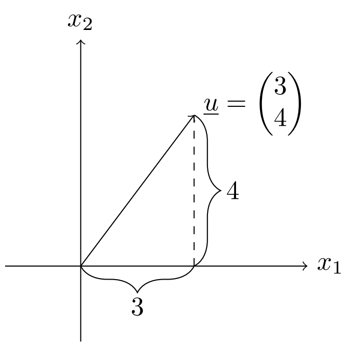
   $$\langle \alpha \cdot \underline{u}, \underline{v} \rangle = (\alpha u_1) \cdot v_1 + \cdots + (\alpha u_n) \cdot v_n = \alpha \cdot (u_1 \cdot v_1 + \cdots + u_n \cdot v_n) = \alpha \cdot \langle \underline{u}, \underline{v} \rangle$$

   $\square$

2. $\langle \underline{u}, \alpha \underline{v} \rangle = \alpha \langle \underline{u}, \underline{v} \rangle$ for any $\alpha \in \mathbb{R}, \underline{u}, \underline{v} \in \mathbb{R}^n$
3. $\langle \alpha \underline{u} + \beta \underline{v}, \underline{w} \rangle = \alpha \cdot \langle \underline{u}, \underline{w} \rangle + \beta \langle \underline{v}, \underline{w} \rangle, \forall \alpha \in \mathbb{R}, \forall \underline{u}, \underline{v}, \underline{w} \in \mathbb{R}^n$

**Example**

Let us consider $\underline{u} = \begin{pmatrix} 3 \\ 4 \end{pmatrix}$. $\langle \underline{u}, \underline{u} \rangle = 3 \cdot 3 + 4 \cdot 4 = 9 + 16 = 25 = 5^2$



## 1.4 Length of a Vector

**Definition**

The length of vector $\underline{u} \in \mathbb{R}^n$, $\|\underline{u}\|$, is defined as $\|\underline{u}\| = \sqrt{\langle \underline{u}, \underline{u} \rangle}$. Sometimes it is also called the Euclidian norm of $\underline{u}$.

## 1.5 Unit Vectors

**Definition**

A vector with length equal to 1 is called a unit vector

### 1.5.1 How to Normalize a Vector?

Sometimes we just want to deal with unit vectors, because some calculations might be simpler.
If we take the vector $\underline{u} \neq \underline{0}$, how do we make it a *unit vector*, or, in other words, how do we *normalise* it?
To normalise a vector $\underline{u}$, we simply have to multiply $\underline{u}$ by the inverse of its length, or $\frac{1}{\|\underline{u}\|}$, and we obtain

$$\hat{\underline{u}} = \frac{1}{\|\underline{u}\|} \cdot \underline{u} = \frac{\underline{u}}{\|\underline{u}\|}$$

**Note**

$\hat{\underline{u}}$ is **not** $\underline{u}$.

**Notation**

Usually unit vectors are denoted with an *hat* over them, like $\hat{\underline{v}}$.

□

**Example**

Consider the following vector

$$\underline{u} = \begin{pmatrix} 3 \\ 4 \end{pmatrix}$$

The *unit vector* $\hat{\underline{v}}$ with the same direction of $\underline{u}$ is then

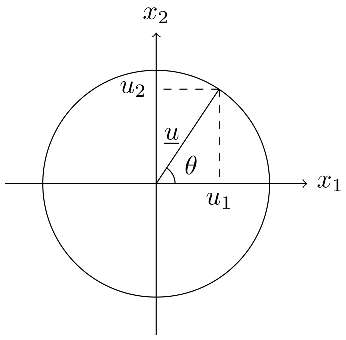
$$\hat{\underline{v}} = \frac{\underline{u}}{\|\underline{u}\|} = \frac{1}{5} \cdot \begin{pmatrix} 3 \\ 4 \end{pmatrix} = \begin{pmatrix} \frac{3}{5} \\ \frac{4}{5} \end{pmatrix} = \begin{pmatrix} 0.6 \\ 0.8 \end{pmatrix}$$

**Note**

You can check by yourself if $\hat{\underline{v}}$ is really a unit vector by calculating its length.

## 1.6 Angle between Vectors

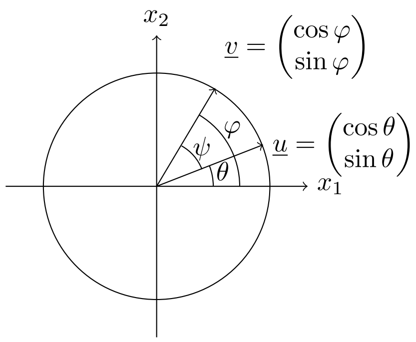
Let us consider $\mathbb{R}^2$. What is the set of all possible endpoints of unit vectors (vectors of length 1, usually denoted with an hat above them $\hat{u}$) in $\mathbb{R}^2$, originating from the origin? A unit circle or, in other words, a circle with radius 1.




$$\underline{u} = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$$
$$\cos(\theta) = \frac{u_1}{\|\underline{u}\|} = u_1$$
$$\sin(\theta) = \frac{u_2}{\|\underline{u}\|} = u_2$$
$$\underline{u} = \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix}$$

### 1.6.1 Angle between Unit Vectors

Let us consider two *unit vectors*




$$\begin{aligned} \langle \underline{u}, \underline{v} \rangle &= \cos(\theta)\cos(\varphi) + \sin(\theta)\sin(\varphi) \\ &= \cos(\theta - \varphi) = \cos(\psi) \\ &= \cos(\angle(\underline{u}, \underline{v})) \end{aligned}$$

### 1.6.2 Angle between Non-Unit Vectors

If $\underline{u} \neq \underline{0}$ or $\underline{v} \neq \underline{0}$ are **not** unit vectors, we can find the angle between them in the following way

$$\langle \underline{u}, \underline{v} \rangle = \left\langle \|\underline{u}\| \cdot \frac{1}{\|\underline{u}\|} \cdot \underline{u}, \|\underline{v}\| \cdot \frac{1}{\|\underline{v}\|} \cdot \underline{v} \right\rangle = \|\underline{u}\|\|\underline{v}\| \left\langle \underbrace{\frac{1}{\|\underline{u}\|} \cdot \underline{u}, \frac{1}{\|\underline{v}\|} \cdot \underline{v}}_{\text{Unit Vectors}} \right\rangle$$

$$\langle \underline{u}, \underline{v} \rangle = \|\underline{u}\|\|\underline{v}\| \cos\left(\angle(\underline{u}, \underline{v})\right)$$

**Lemma**

If $\underline{u} \neq \underline{0}, \underline{v} \neq \underline{0}, \underline{u} \in \mathbb{R}^n, \underline{v} \in \mathbb{R}^n$, then

$$\cos\left(\angle(\underline{u}, \underline{v})\right) = \frac{\langle \underline{u}, \underline{v} \rangle}{\|\underline{u}\|\|\underline{v}\|}$$

## 1.7 Cauchy Schwarz inequality

**Lemma**

The *Cauchy-Schwarz inequality* is a useful inequality encountered in many different settings, such as *linear algebra*, and it states that for any $\underline{u} \in \mathbb{R}^n$ and $\underline{v} \in \mathbb{R}^n$

$$|\langle \underline{u}, \underline{v} \rangle| \leq \|\underline{u}\| \|\underline{v}\|$$

**Remark:**

It is easy to see that *Cauchy-Schwarz inequality* is also correct for zero vectors.

**Note**

In the previous section, we have seen that

$$\langle \underline{u}, \underline{v} \rangle = \|\underline{u}\| \|\underline{v}\| \cdot \cos\left(\angle(\underline{u}, \underline{v})\right)$$

Now, let us take the absolute value of both sides

$$|\langle \underline{u}, \underline{v} \rangle| = |\|\underline{u}\| \|\underline{v}\| \cdot |\cos\left(\angle(\underline{u}, \underline{v})\right)|| = \|\underline{u}\| \|\underline{v}\| \cdot |\cos\left(\angle(\underline{u}, \underline{v})\right)|$$

Notice that $|\cos\left(\angle(\underline{u}, \underline{v})\right)| \leq 1$.

# Chapter 2

# Matrices

Let us consider a linear combination of vectors

$$x_1 \cdot \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} + x_2 \cdot \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} + x_3 \cdot \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$$

This can be written using matrices in the following way:

$$\begin{pmatrix} u_1 & v_1 & w_1 \\ \vdots & \vdots & \vdots \\ u_n & v_n & w_n \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$$

In matrix-vector multiplication, we take dot products of rows of matrices times the vector.

**Example**

1.
$$\begin{pmatrix} 1 & 0 & -1 \\ 3 & 1 & 2 \\ 1 & -1 & 5 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \cdot 1 + 0 \cdot 0 + (-1) \cdot 1 \\ 3 \cdot 1 + 1 \cdot 0 + 2 \cdot 1 \\ 1 \cdot 1 + (-1) \cdot 0 + 5 \cdot 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 5 \\ 6 \end{pmatrix}$$

2.
$$A = \begin{pmatrix} -1 & 2 & 3 \\ 0 & 1 & 0 \end{pmatrix}, \underline{x} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

$$\begin{aligned} A \cdot \underline{x} &= \begin{pmatrix} -1 & 2 & 3 \\ 0 & 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} (-1) \cdot 1 + 2 \cdot 1 + 3 \cdot 1 \\ 0 \cdot 1 + 1 \cdot 1 + 0 \cdot 1 \end{pmatrix} = \begin{pmatrix} 4 \\ 1 \end{pmatrix} \end{aligned}$$

For the product of matrix $A$ with vector $\underline{x}$ to exist, matrix $A$ should have the same number of columns as vector $\underline{x}$ has components.

**Notation**

- Matrices are usually written with capital letters, i.e. $A, B, C, \ldots$
- $A$ is an $n$ by $m$ matrix, $A \in \mathbb{R}^{n,m}$ if it has $n$ rows and $m$ columns.
- The element of matrix $A$ located in row $i$ and column $j$ is written as $a_{ij}$ or $(A)_{ij}$.

□

## 2.1 Matrix Operations

**Definition**

Let us consider matrices $A \in \mathbb{R}^{n,m}$ and $B \in \mathbb{R}^{n,m}$ where $n =$ rows, $m =$ columns. Matrix $C \in \mathbb{R}^{n,m}$ is a sum of $A$ and $B, C = A + B$, if $C_{ij} = A_{ij} + B_{ij}$ for all $i = 1, \ldots, n$, $j = 1, \ldots, m$

**Example**

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 5 \end{pmatrix}, B = \begin{pmatrix} -1 & 0 \\ 0 & -1 \\ -1 & 0 \end{pmatrix}, C = A + B = \begin{pmatrix} 0 & 2 \\ 3 & 3 \\ 4 & 6 \end{pmatrix}$$

**Definition**

A product of a scalar $\alpha$ and a matrix $A \in \mathbb{R}^{n,m}$ is defined as $(\alpha A)_{ij} = \alpha \cdot A_{ij}$, $\forall i = 1, \ldots, n; j = 1, \ldots, m$.

**Example**

$$\alpha = 3, A = \begin{pmatrix} 0 & 0 & 1 \\ 2 & 3 & 5 \end{pmatrix} \Rightarrow \alpha \cdot A = \begin{pmatrix} 0 & 0 & 3 \\ 6 & 9 & 15 \end{pmatrix}$$

**Properties**

1. $A \in \mathbb{R}^{n,m}$ and $B \in \mathbb{R}^{n,m}$: $A + B = B + A$
2. $A, B, C \in \mathbb{R}^{n,m}$: $(A + B) + C = A + (B + C)$
3. $\alpha \cdot (A + B) = \alpha A + \alpha B$ for $\forall \alpha \in \mathbb{R}, A, B \in \mathbb{R}^{n,m}$

**Proof**

1.
$$\begin{cases} (A + B)_{ij} = A_{ij} + B_{ij} \\ (B + A)_{ij} = B_{ij} + A_{ij} \end{cases}$$

□

## 2.2 Matrix-Matrix Multiplication

**Definition**

Let us consider matrix $A \in \mathbb{R}^{n,m}$ and $A \in \mathbb{R}^{m,l}$. Then $C = A \cdot B$ is an $n$ by $l$ matrix, $C \in \mathbb{R}^{n,l}$ such that

$$C_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$$

**Example**

1.

$$A = \begin{pmatrix} 1 & 2 \\ 0 & 1 \\ 3 & 4 \end{pmatrix} \in \mathbb{R}^{3,2}, B = \begin{pmatrix} 1 & 2 & 0 & 1 \\ -1 & 1 & 1 & 0 \end{pmatrix} \in \mathbb{R}^{2,4}$$

$$C = A \cdot B \in \mathbb{R}^{3,4} = \begin{pmatrix} 1 & 2 \\ 0 & 1 \\ 3 & 4 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 & 0 & 1 \\ -1 & 1 & 1 & 0 \end{pmatrix} = \begin{pmatrix} -1 & 4 & 2 & 1 \\ -1 & 1 & 1 & 0 \\ -1 & 10 & 4 & 3 \end{pmatrix}$$

2.

$$A = \begin{pmatrix} 1 & 2 \\ 0 & 1 \\ 3 & 4 \end{pmatrix}, B = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 2 & 2 & 2 \end{pmatrix}; AB = \text{ Not defined}$$

**Properties**

1. $AB$ is not always equal to $BA$. (most often, is the case).

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

$$AB = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, BA = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

2. $C(A + B) = CA + CB$
3. $(A + B)C = AC + BC$
4. $\alpha(AB) = A(\alpha B)$, $A \in \mathbb{R}^{n,m}$, $B \in \mathbb{R}^{m,l}$. Proof:

$$(\alpha(AB))_{ij} = \alpha \sum_{k=1}^{m} a_{ik} b_{kj} = \sum_{k=1}^{m} a_{ik} (\alpha b_{kj}) = A(\alpha B)$$

5. $(AB)C = A(BC)$

**Theorem**

Let us consider matrices $A \in \mathbb{R}^{n,n}$ and $B \in \mathbb{R}^{n,n}$, such that $A^{-1}$ and $B^{-1}$ exist. Then,

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

**Proof**

$$\left.\begin{aligned} (AB)(B^{-1}A^{-1}) = I \\ (B^{-1}A^{-1})(AB) = I \end{aligned}\right\} \text{Prove this}$$

$$\begin{aligned} (AB)(B^{-1}A^{-1}) &= A\underbrace{BB^{-1}}_{I}A^{-1} = A \cdot I \cdot A^{-1} = I \\ (B^{-1}A^{-1})(AB) &= B^{-1}\underbrace{A^{-1}A}_{I}B = \cdot B^{-1} \cdot I \cdot B = I \end{aligned}$$

$\Rightarrow$ According to the definition $B^{-1}A^{-1}$ is the inverse of $AB$

$\square$

**Lemma**

$A, B, C \in \mathbb{R}^{n,n}, \exists A^{-1}, \exists B^{-1}, \exists C^{-1}$

$$(ABC)^{-1} = C^{-1}B^{-1}A^{-1}$$

**Theorem**

Let us consider $A \in \mathbb{R}^{n,n}$. Let us consider that $B \in \mathbb{R}^{n,n}$ and $C \in \mathbb{R}^{n,n}$ are both inverses of $A$.Then $B = C$. (The inverse is unique)

**Proof**

$$AB = BA = I \qquad\qquad AC = CA = I$$

$$\underbrace{BA \times C}_{} = I \times C \qquad\qquad \underbrace{B \times AC}_{} = B \times I$$

$$\longrightarrow \underline{\underline{C = B}} \longleftarrow$$

$\square$

## 2.3 Linear System of Equations

Let us consider the following system of equations

$$\begin{cases} 2x_1 + 2x_2 + 4x_3 = 2 \\ \qquad\quad x_2 + 2x_3 = 3 \\ \qquad\qquad\quad 4x_3 = -1 \end{cases}$$

Find $x_1, x_2, x_3$. We can write this system in matrix form.

$$A = \begin{pmatrix} 2 & 2 & 4 \\ 0 & 1 & 2 \\ 0 & 0 & 4 \end{pmatrix} \in \mathbb{R}^{3,3}, \underline{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \underline{b} = \begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix} \Rightarrow A\underline{x} = \underline{b}$$

$A$ is an upper triangular matrix. We can use backward substitution to find the solution:

1. $x_3 = -\frac{1}{4} = \frac{b_3}{a_{33}}$
2. $x_2 = \frac{3-2x_3}{1} = \frac{3-2\cdot\left(-\frac{1}{4}\right)}{1} = 3.5 = \frac{b_2 - a_{33}\cdot x_3}{a_{22}}$
3. $x_1 = \frac{2-4x_3-2x_2}{2} = -2 = \frac{b_1 - a_{13}x_3 - a_{12}x_2}{a_{11}}$

In general, if $A \in \mathbb{R}^{n,n}$ is an upper triangular with $a_{ii} \neq 0, i = 1, \ldots, n$ then the backward substitution works as:

1. $x_n = \frac{b_n}{a_{nn}}$
2. $x_{n-1} = \frac{b_{n-1} - a_{n-1\ n}x_n}{a_{n-1\ n-1}} \quad \ldots \quad x_i = \frac{b_i - a_{in}x_n - \cdots - a_{i\ i+1} + x_{i+1}}{a_{ii}} \qquad i = 1, \ldots, n$

## 2.4 Inverse of a Matrix

**Definition**

Let us consider a matrix $A \in \mathbb{R}^{n,n}$ (square matrix). Matrix $B \in \mathbb{R}^{n,n}$ is called an inverse of $A$, if

$$A \cdot B = I \quad \text{AND} \quad B \cdot A = I$$

(Both conditions are vital)

**Notation**

Usually, the inverse of $A$ is written as $A^{-1}$

□

**Note**

Not all matrices have an inverse! In most cases, it is quite difficult to find an inverse matrix. But in some cases, the inverse is easy to find.

**Example**

$$A = \begin{pmatrix} a_{11} & 0 & \ldots & 0 \\ 0 & a_{22} & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \ldots & 0 & a_{nn} \end{pmatrix}, a_{ii} \neq 0, \forall i = 1, \ldots, n$$

Then

$$
\begin{aligned}
A &= \begin{pmatrix} a_{11}^{-1} & 0 & \dots & 0 \\ 0 & a_{22}^{-1} & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & a_{nn}^{-1} \end{pmatrix} \\
A \cdot A^{-1} &= \begin{pmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22} & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & a_{nn} \end{pmatrix} \cdot \begin{pmatrix} a_{11}^{-1} & 0 & \dots & 0 \\ 0 & a_{22}^{-1} & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & a_{nn}^{-1} \end{pmatrix} \\
&= \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & 1 \end{pmatrix} \\
A \cdot A^{-1} &= I
\end{aligned}
$$

## 2.5 Special Matrices

- Let us consider $A \in \mathbb{R}^{n,m}$ matrix. $A$ is called the zero matrix if all $a_{ij} = 0$, $\forall i = 1, \dots, n; j = 1, \dots, n$

- $D \in \mathbb{R}^{n,n}$ - square matrix is called diagonal matrix, if $d_{ij} = 0$ and if $i \neq j$

- Identity matrix:

$$I \in \mathbb{R}^{n,n}, I = \begin{pmatrix} 1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 1 \end{pmatrix}$$

- $L \in \mathbb{R}^{n,n}$ - lower triangular matrix, if

$$l_{ij} = 0, \forall i < j, L = \begin{pmatrix} * & \dots & 0 \\ \vdots & \ddots & \vdots \\ * & \dots & * \end{pmatrix}$$

- $U \in \mathbb{R}^{n,n}$ - upper triangular matrix, if

$$u_{ij} = 0, \forall i > j, L = \begin{pmatrix} * & \dots & * \\ \vdots & \ddots & \vdots \\ 0 & \dots & * \end{pmatrix}$$

**Remark:**

If $A, B \in \mathbb{R}^{n,n}$ are both upper (lower) triangular matrices, then $C = A \cdot B$ is an upper triangular (lower).

If $A$ is lower triangular, $A \in \mathbb{R}^{n,n}, a_{ii} \neq 0, i = 1, \ldots, n$ then we can use forward substitution, i.e.:

$$\begin{aligned} x_1 &= \frac{b_1}{a_{11}} \\ &\vdots \\ x_i &= \frac{b_i - a_{i1}x_1 - \cdots - a_{ii-1}x_{i-1}}{a_{ii}} \qquad \forall i = 2, \ldots, n \end{aligned}$$

## 2.6 Elementary Transition Matrices

Let us consider matrix

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

and matrix

$$I_{21} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Then

$$I_{21} \cdot A = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & -1 & 2 \\ 3 & 4 & 5 & 7 \\ 2 & -1 & 0 & 0 \\ -1 & 3 & 5 & -1 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & -1 & 2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

also

$$A \cdot I_{21} = \begin{pmatrix} 1 & 0 & -1 & 2 \\ 3 & 4 & 5 & 7 \\ 2 & -1 & 0 & 0 \\ -1 & 3 & 5 & -1 \end{pmatrix} \cdot \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 4 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 3 & 0 & 0 & 0 \end{pmatrix}$$

**Definition**

We can define the elementary transition matrix $I_{pq} \in \mathbb{R}^{n,n}$

$$(I_{pq}) = \begin{cases} 1 & i = p, q = j \\ 0 & \text{otherwise} \end{cases}$$

If we take a matrix $A \in \mathbb{R}^{n,n}$ then when calculating $I_{pq}$ we take row $q$ of $A$, put it into row $p$, replace everything else with 0.

We can also define:

$$E_{pq}(l) = I + l \cdot I_{pq}, l \in \mathbb{R} \text{ - scalar}$$
$$E_{pq}(l) \cdot A = (I + lI_{pq}) \cdot A = A + l \cdot I_{pq}A$$

We take row $q$ of $A$, multiply it by $l$, add it to row $p$ of $A$

$$E_{pq}^{-1}(l) = E_{pq}(-l)$$

If we have vector $\underline{b} \in \mathbb{R}^n$, then $I_{pq}\underline{b}$ - we take component $q$ of $\underline{b}$, put it into component $p$, replace everything else with zeros.

$$E_{pq}(l)\underline{b} \text{ - same as for matrices}$$

# Chapter 3

# Gaussian Elimination

**Example**

$$A = \begin{pmatrix} 2 & 4 & -2 \\ 4 & 9 & -3 \\ -2 & -3 & 7 \end{pmatrix} .\underline{b} = \begin{pmatrix} 2 \\ 8 \\ 10 \end{pmatrix}, A\underline{x} = b$$

We can write this as a system of equations:

$$\begin{cases} 2x_1 + 4x_2 - 2x_3 = 2 \\ 4x_1 + 9x_2 - 3x_3 = 8 \\ -2x_1 - 3x_2 7x_3 = 10 \end{cases}$$

We can multiply equation 1 by $-\frac{a_{21}}{a_{11}} = -\frac{4}{2} = -2$, and add to equation 2. This is equivalent to multiplying $A\underline{x} = \underline{b}$ by $E_{21}\left(-\frac{a_{21}}{a_{11}}\right)$ on the left.

$$\begin{cases} 2x_1 + 4x_2 - 2x_3 = 2 \\ 4x_1 + 9x_2 - 3x_3 = 8 \\ -2x_1 - 3x_2 7x_3 = 10 \end{cases} \Leftrightarrow E_{21}\left(-\frac{a_{21}}{a_{11}}\right) \times A\underline{x} = E_{21}\left(-\frac{a_{21}}{a_{11}}\right)\underline{b}$$

$$E_{21}\left(-\frac{a_{21}}{a_{11}}\right) = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\begin{cases} 2x_1 + 4x_2 - 2x_3 = 2 \\ x_2 + x_3 = 4 \\ x_2 + 5x_3 = 12 \end{cases}$$

$$\Leftrightarrow E_{31}\left(-\frac{a_{31}}{a_{11}}\right) E_{21}\left(-\frac{a_{21}}{a_{11}}\right) \times A\underline{x} = E_{31}\left(-\frac{a_{31}}{a_{11}}\right) E_{21}\left(-\frac{a_{21}}{a_{11}}\right)\underline{b}$$

$$E_{31}\left(-\frac{a_{31}}{a_{11}}\right) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

We are done with the first column. Let us denote the resulting matrix by $A^{(1)}$

$$\begin{cases} 2x_1 + 4x_2 - 2x_3 = 2 \\ \qquad\quad x_2 + x_3 = 4 \\ \qquad\qquad 4x_3 = 8 \end{cases}$$

$$\Leftrightarrow E_{32}\left(-\frac{a_{32}^{(1)}}{a_{22}^{(1)}}\right) E_{31}\left(-\frac{a_{31}}{a_{11}}\right) E_{21}\left(-\frac{a_{21}}{a_{11}}\right) \times \underbrace{A\underline{x}}_{\underline{b}}$$

$$E_{32}\left(-\frac{a_{32}^{(1)}}{a_{22}^{(1)}}\right) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix}$$

We are done with the second column, so we can denote the resulting matrix by $A^{(2)}$.

In fact, we got an upper triangular matrix. We can solve it using backward compatibility. Let us denote

$$E_{32}\left(-\frac{a_{32}^{(1)}}{a_{22}^{(1)}}\right) E_{31}\left(-\frac{a_{31}}{a_{11}}\right) E_{21}\left(-\frac{a_{21}}{a_{11}}\right) = U$$

where $U$ is the upper triangular matrix. Then the inverse of it is

$$\left[E_{32}\left(-\frac{a_{32}^{(1)}}{a_{22}^{(1)}}\right) E_{31}\left(-\frac{a_{31}}{a_{11}}\right) E_{21}\left(-\frac{a_{21}}{a_{11}}\right)\right]^{-1}$$
$$= E_{21}\left(-\frac{a_{21}}{a_{11}}\right) E_{31}\left(-\frac{a_{31}}{a_{11}}\right) E_{32}\left(-\frac{a_{32}^{(1)}}{a_{22}^{(1)}}\right)$$
$$A = \underbrace{E_{21}\left(-\frac{a_{21}}{a_{11}}\right) E_{31}\left(-\frac{a_{31}}{a_{11}}\right) E_{32}\left(-\frac{a_{32}^{(1)}}{a_{22}^{(1)}}\right)}_{L} \cdot U$$

All matrices $E_{xx}(x)$ are lower triangular $\rightarrow$ the product is also lower triangular $(A = l \cdot U)$. So using Gaussian elimination, we represented $A$ as a product of lower and upper triangular matrices

$$A\underline{x} = \underline{b} \Rightarrow LU\underline{x} = \underline{b}$$

Let us denote $U\underline{x}$ by $\underline{y}$, then we get

$$\begin{cases} l\underline{y} = \underline{b} & \text{Solve by forward substitution, find } \underline{y} \\ U\underline{x} = \underline{y} & \text{Solve by backward substitution} \end{cases}$$

**<u>Remark:</u>**

Gaussian elimination works if all elements $a_{11}, a_{22}^{(1)}, a_{33}^{(2)}, \ldots, a_{ii}^{(i-1)}$ are non-zero! These elements are called <u>PIVOT</u> elements.

**Example**

$$\begin{cases} 2x_1 + 4x_2 - 2x_3 = 2 \\ 4x_1 + 8x_2 - 3x_3 = 6 \\ -2x_1 - 3x_2 + 7x_3 = 10 \end{cases} \Leftrightarrow A\underline{x} = \underline{b}$$

$$\begin{cases} 2x_1 + 4x_2 - 2x_3 = 2 \\ x_3 = 2 \\ -2x_1 - 3x_2 + 7x_3 = 10 \end{cases} \Leftrightarrow E_{21}\left(-\frac{a_{21}}{a_{11}}\right)A\underline{x} = E_{21}\left(-\frac{a_{21}}{a_{11}}\right)$$

$$\begin{cases} 2x_1 + 4x_2 - 2x_3 = 2 \\ x_3 = 2 \\ x_2 + 5x_3 = 12 \end{cases} \Leftrightarrow E_{31}\left(-\frac{a_{31}}{a_{11}}\right)A\underline{x} = E_{31}\left(-\frac{a_{31}}{a_{11}}\right)E_{21}\left(-\frac{a_{21}}{a_{11}}\right)\underline{b}$$

We denote the resulting matrix by $A^{(1)}$. In order to proceed we need $a_{22}^{(1)} \neq 0$. Let us consider matrix $P_{pq}$-matrix, which you get from identity matrix by exchanging rows $p$ and $q$. It is easy to show that $P_{pq} \cdot A$ is equal to matrix $A$ with rows $p$ and $q$ exchanged.

**Definition**

Permutation matrix $P$ is an identity matrix with rows in any order.

**Remark:**

$P^{-1} = P$. The product of permutation on matrices is a permutation matrix.

We want to exchange rows 2 and 3. We need to multiply by the permutation matrix $P_{23}$

$$\begin{cases} 2x_1 + 4x_2 - 2x_3 = 2 \\ x_2 + 5x_3 = 12 \\ x_3 = 2 \end{cases}$$

$$\begin{aligned} \Leftrightarrow & P_{23} \cdot E_{31}\left(-\frac{a_{31}}{a_{11}}\right)E_{21}\left(-\frac{a_{21}}{a_{11}}\right)A\underline{x} \\ = & P_{23} \cdot E_{31}\left(-\frac{a_{31}}{a_{11}}\right)E_{21}\left(-\frac{a_{21}}{a_{11}}\right)\underline{b} \end{aligned}$$

In general, the Gaussian elimination proceeds like this:

$$E_{xx} \dots E_{xx} P_{xx} E_{xx} \dots E_{xx} A\underline{x} = E_{xx} \dots E_{xx} P_{xx} E_{xx} \dots E_{xx}\underline{b}$$

Turns out, that we can exchange the rows, or in other words multiply $A$ by

$(P_{xx} \dots P_{xx}$ before doing the Gaussian elimination

$$\underbrace{\underbrace{(E_{xx} \dots E_{xx})}_{E} \underbrace{(P_{xx} \dots P_{xx})}_{P} A}_{U} \underline{x} = (E_{xx} \dots E_{xx})(P_{xx} \dots P_{xx})\underline{b}$$

$$EPA = U$$
$$PA = E^{-1}U = LU \leftarrow \text{Lower triangular}$$

**Theorem**

There exists permutation matrix $P$, such that $PA = LU$. The only necessary condition for that is that $A^{-1}$ exists.

## 3.1 Matrix Transposition

**Definition**

Let us consider matrix $A \in \mathbb{R}^{m,n}$. Matrix $B \in \mathbb{R}^{n,m}$ is called the transpose of $A$ if $(B)_{ij} = (A)_{ji}, i = 1 \dots n, j = 1 \dots n$

**Notation**

Usually the transpose of $A$ is written as $A^T$

□

**Example**

$$A = \begin{pmatrix} 2 & 3 \\ 4 & 5 \\ 6 & 7 \\ 9 & 10 \end{pmatrix} \in \mathbb{R}^{4,2} \Rightarrow A = \begin{pmatrix} 2 & 4 & 6 & 9 \\ 3 & 5 & 7 & 10 \end{pmatrix} \in \mathbb{R}^{2,4}$$

**Properties**

1. $\left(A^T\right)^T = A$
2. $(A + B)^T = A^T + B^T$
3. $(AB)^T = B^T \cdot A^T$
4. $\left(A^T\right)^{-1} = \left(A^{-1}\right)^T$

**Proof**

3.
$$
\begin{aligned}
A \in \mathbb{R}^{m,n} &= \begin{pmatrix} -\text{row } 1 \rightarrow \\ \vdots \\ -\text{row } n \rightarrow \end{pmatrix}, B \in \mathbb{R}^{n,l} = \begin{pmatrix} \text{-col 1}\downarrow & \cdots & \text{-col n}\downarrow \end{pmatrix} \\
(AB)_{ij} &= \langle \text{row } i \text{ of } A, \text{ column } j \text{ of } B \rangle \\
\left((AB)^T\right)_{pq} &= (AB)_{qp} = \langle \text{row } q \text{ of } A, \text{ column } p \text{ of } B \rangle \\
B^T &= \begin{pmatrix} -\text{col } 1 \rightarrow \\ \vdots \\ -\text{col } n \rightarrow \end{pmatrix}, A^T = \begin{pmatrix} \text{-col 1}\downarrow & \cdots & \text{-col n}\downarrow \end{pmatrix} \\
\left(B^T A^T\right)_{pq} &= \langle \text{column } p \text{ of } B, \text{ row } q \text{ of } A \rangle \\
\Rightarrow \left((AB)^T\right)_{pq} &= \left(B^T A^T\right)_{pq}; p = 1, \ldots, l; q = 1, \ldots, m. \\
\Rightarrow (AB)^T &= B^T A^T
\end{aligned}
$$

4. Assume that $A \in \mathbb{R}^{n,n}, \exists A^{-1}$
$$
\begin{aligned}
AA^{-1} = I &\rightarrow (AA^{-1})^T = \left(A^{-1}\right)^T \cdot A^T = I^T = I \\
A^{-1}A = I &\rightarrow (A^{-1}A)^T = A^T \cdot \left(A^{-1}\right)^T = I^T = I \\
\left(A^T\right)^{-1} &= \left(A^{-1}\right)^T
\end{aligned}
$$

□

Let us consider vector $\underline{u} = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} \in \mathbb{R}^{n,1}$ - column vector. Then $\underline{u}^T \in \mathbb{R}^{1,n} =$ $(u_1 \ldots u_n)$ - row vector. Let us also consider $\underline{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \in \mathbb{R}^{n,1}$. Then

$$
\underline{u}^T \cdot \underline{v} = (u_1 \ldots u_n) \cdot \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n = \langle \underline{u}, \underline{v} \rangle
$$
$$
\underline{v} \cdot \underline{u}^T = n \times n \text{ matrix}
$$

**Definition**

Matrix $A$ is called symmetric if $A^t = A$. Matrix $A$ should be a square matrix, $A \in \mathbb{R}^{n,n}$

$$
\begin{aligned}
&\text{e.g. } A = \begin{pmatrix} 0 & 3 \\ 3 & 4 \end{pmatrix} \rightarrow A^T = \begin{pmatrix} 0 & 3 \\ 3 & 4 \end{pmatrix} \Rightarrow A^T = A \\
&\text{e.g. } A = I \in \mathbb{R}^{n,n} \rightarrow I^T = I
\end{aligned}
$$

# Chapter 4

# Vector Spaces

**Definition**

A vector space $V$ is a set of objects, such that any two objects can be added together, any object can be multiplied by a scalar.

If two objects belong to the vector space, then their sum also belongs to the vector space.

If an object belongs to $V$, then the product of any scalar with this object belongs to V and the following properties are satisfied:

1. $\forall \underline{u}, \underline{v}, \underline{w} \in V$; $(\underline{u} + \underline{v}) + \underline{w} = \underline{u} + (\underline{v} + \underline{w})$
2. $\forall \underline{u}, \underline{v} \in V$; $\underline{u} + \underline{v} = \underline{v} + \underline{u}$
3. There exists unique elements $\underline{0} \in V$, such that $\forall \underline{u} \in V$; $\underline{u} + \underline{0} = \underline{0} + \underline{u} = \underline{u}$
4. For any $\underline{u} \in V, \exists!(-\underline{u}) \in V$, such that $\underline{u} + (-\underline{u}) = \underline{0}$
5. $\forall \underline{u}, \underline{v} \in V$; $\forall \alpha \in \mathbb{R}$; $\alpha(\underline{u} + \underline{v}) = \alpha\underline{u} + \alpha\underline{v}$
6. $\forall \underline{u} \in V$; $\forall \alpha, \beta \in \mathbb{R}$; $(\alpha + \beta)\underline{u} = \alpha\underline{u} + \beta\underline{u}$
7. $\forall \underline{u} \in V$; $\forall \alpha, \beta \in \mathbb{R}$; $(\alpha\beta)\underline{u} = \alpha(\beta\underline{u})$
8. $\forall \underline{u} \in V$; $1 \cdot \underline{u} = \underline{u}$ (1 is a scalar here)

**Remark:**

The "vectors" in the vector space, are not necessarily vectors ($\in \mathbb{R}^n$) , but can be other objects, as long as the definition is satisfied.

**Example**

Let us consider a set of all $2 \times 2$ matrices. It is a vector space. Proof:

$$\text{If } A, B \in \mathbb{R}^{2,2} \Rightarrow (A+B) \in \mathbb{R}^{2,2}$$
$$\text{If } \alpha \in \mathbb{R}, A \in \mathbb{R}^{2,2} \Rightarrow \alpha A \in \mathbb{R}^{2,2}$$

1. $A, B, C \in \mathbb{R}^{2,2}$; $(A+B)+C = A+(B+C)$
2. ...
3. $$\underline{0} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \in \mathbb{R}^{2,2}, \forall A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \Rightarrow A + \underline{0} = A$$
4. $$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \Rightarrow (-A) = \begin{pmatrix} -a_{11} & -a_{12} \\ -a_{21} & -a_{22} \end{pmatrix}$$

The layout of this example is not clear

**Example**

Let us consider a set consisting of a single object, $\underline{0}$. It is a vector space.

**Note**

There is no vector space, which does not contain $\underline{0}$

## 4.1 Subspace of the Vector Space

**Definition**

A subspace $W$ of the vector space $V$, is a set of vectors in $V$, such that:

1. If $\underline{u}, \underline{v} \in W$ then $\underline{u} + \underline{v} \in W$
2. If $\alpha \in \mathbb{R}$, $\underline{u} \in W$ then $\alpha\underline{u} \in W$

**Definition**

Let us consider a set of vectors $\{\underline{u}_1, \ldots, \underline{u}_n\}$. The span of vectors $\{\underline{u}_1, \ldots, \underline{u}_n\}$ is defined as

$$\mathcal{S} = \text{span}\{\underline{u}_1, \ldots, \underline{u}_n\} = \{\alpha_1\underline{u}_1 + \cdots + \alpha_n\underline{u}_n \mid \forall\alpha_1 \ldots \alpha_n \in \mathbb{R}\}$$

**Example**

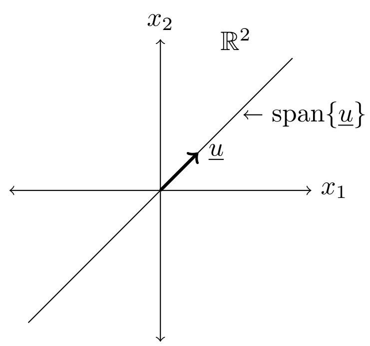
Is span$\{\underline{u}\}$ a subspace in $\mathbb{R}^2$? Proof:

$$\underline{v} = \alpha\underline{u} \in \text{span}\{\underline{u}\}$$
$$\underline{w} = \beta\underline{u} \in \text{span}\{\underline{u}\}$$

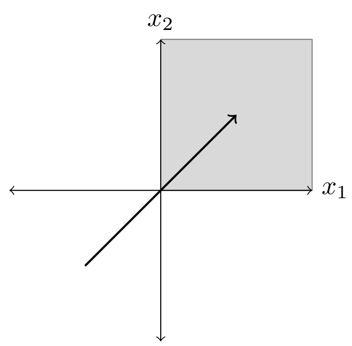
1. $\underline{v} + \underline{w} = \alpha\underline{u} + \beta\underline{u} = (\alpha + \beta)\underline{u} \in \text{span}\{\underline{u}\}$
2. $\gamma \in \mathbb{R}, \gamma\underline{v} = \gamma \cdot (\alpha\underline{u}) = (\gamma \cdot \alpha)\underline{u} \in \text{span}\{\underline{u}\}$

**Example**

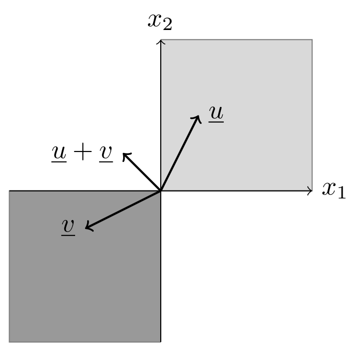
1.




2.




3.



## 4.2 Linear Independence

**Definition**

Let us consider vector space $V$ and $\underline{v}_1, \dots, \underline{v}_n \in V$. $\underline{v}_1, \dots, \underline{v}_n$ are linearly dependent if there exists scalars $\alpha_1, \dots, \alpha_n$ not all equal to zero, such that $\alpha_1 \underline{v}_1 + \cdots + \alpha_n \underline{v}_n = \underline{0}$

If no such scalars exist, the vectors $\underline{v}_1, \dots, \underline{v}_n$ are linearly independent.

**Definition**

Vectors $\underline{v}_1, \dots, \underline{v}_n \in V$ are linearly independent if the following is true:

$$\alpha_1 \underline{v}_1 + \cdots + \alpha_n \underline{v}_n = 0 \Rightarrow \text{ all } \alpha_i = 0, i = 1, \dots, n$$

**Example**

1. Let us consider $\mathbb{R}^n$ and vectors

$$\underline{E}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \underline{E}_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \dots, \underline{E}_i = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \underline{E}_n = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}$$

$\underline{E}_1, \dots, \underline{E}_n$ are linearly independent.

2. Let us consider $\mathbb{R}^2$, $\underline{u}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $\underline{u}_2 = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$. Are they linearly independent? See proof 2.

**Proof**

1. Assume that

$$\alpha_1 \underline{E}_1 + \cdots + \alpha_n \underline{E}_n = 0 \Rightarrow \alpha_1 \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \cdots + \alpha_n \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

$\Rightarrow$ then all $\alpha_i = 0$ for $i = 1, \dots, n$,then based on the definition $\underline{E}_1, \dots, \underline{E}_n$ are linearly independent.

2. Let us consider $\alpha_1 \underline{u}_1 + \alpha_2 \underline{u}_2 = \underline{0} \Rightarrow \alpha_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \alpha_2 \begin{pmatrix} 3 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

$$\begin{cases} \alpha_1 + 3\alpha_2 = 0 \\ \alpha_1 + \alpha_2 = 0 \end{cases} \rightarrow \begin{cases} 2\alpha_2 = 0 \\ \alpha_1 + \alpha_2 = 0 \end{cases} \rightarrow \begin{cases} \alpha_2 = 0 \\ \alpha_1 = 0 \end{cases}$$

If we assume $\alpha_1\underline{u}_1 + \alpha_2\underline{u}_2 = \underline{0}$, we have to show that all $\alpha_i$ are zeroes $\Rightarrow$ vectors are linearly independent.

$\square$

**Example**

Let us consider $\mathbb{R}^2$, $\underline{u}_1 = \begin{pmatrix}1\\1\end{pmatrix}$, $\underline{u}_2 = \begin{pmatrix}2\\2\end{pmatrix}$. Let us assume that

$$\alpha_1\underline{u}_1 + \alpha_2\underline{u}_2 = \underline{0} \Rightarrow \alpha_1\begin{pmatrix}1\\1\end{pmatrix} + \alpha_2\begin{pmatrix}2\\2\end{pmatrix} = \begin{pmatrix}0\\0\end{pmatrix}$$

$$\begin{cases}\alpha_1 + 2\alpha_2 = 0\\ \alpha_1 + 2\alpha_2 = 0\end{cases} \rightarrow \begin{cases}\alpha_1 + 2\alpha_2 = 0\\ 0 = 0\end{cases}$$

One possible solution:

$$\begin{cases}\alpha_1 = -2\\ \alpha_2 = 1\end{cases}$$

Linearly dependent.

**Recap**

If we consider vectors $\underline{v}_1, \ldots, \underline{v}_n \in V$, then

$$\textbf{span}\{\underline{v}_1, \ldots, \underline{v}_n\} = \{\alpha_1\underline{v}_1, \ldots, \alpha_n\underline{v}_n \mid \text{ for all possible } \alpha_1, \ldots, \alpha_n \in \mathbb{R}\}$$

**Definition**

If vector space $v$ is generated by $\{\underline{v}_1, \ldots, \underline{v}_n\}$ (in other words, $V = \text{span}\{\underline{v}_1, \ldots, \underline{v}_n\}$) and $\underline{v}_1, \ldots, \underline{v}_n$ are linearly independent, then $\{\underline{v}_1, \ldots, \underline{v}_n\}$ is called basis of $V$

**Example**

Let us consider $\mathbb{R}^n$ and $\underline{E}_1, \ldots, \underline{E}_n$. They form basis of $\mathbb{R}^n$.

**Proof**

1. "$V$ is generated by $\underline{v}_1, \ldots, \underline{v}_n$". Let us consider any vector $\underline{u} \in \mathbb{R}^n$

$$\underline{u} = \begin{pmatrix}\underline{u}_1\\ \vdots\\ \underline{u}_n\end{pmatrix}, \text{ we have}$$

$$\underline{u} = \begin{pmatrix}\underline{u}_1\\ \vdots\\ \underline{u}_n\end{pmatrix} = \underline{u}_1\underline{E}_1 + \cdots + \underline{u}_n\underline{E}_n \Rightarrow \mathbb{R}^n = \text{span}\{\underline{E}_1, \ldots, \underline{E}_n\}$$

2. "Linear independence" already proven before.

$\square$

**Example**

Let us consider $\mathbb{R}^2$ and $\underline{u}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $\underline{u}_2 = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$, is it a basis?

1. Is $\mathbb{R}^2 = \text{span}\left\{\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 3 \\ 1 \end{pmatrix}\right\}$? Let us consider an arbitrary vector $\underline{v} = \begin{pmatrix} \underline{v}_1 \\ \underline{v}_2 \end{pmatrix} \in \mathbb{R}^2$. We should check that there exists scalars $\alpha_1, \alpha_2$ such that

$$\underline{v} = \alpha_1 \underline{u}_1 + \alpha_2 \underline{u}_2 \rightarrow \underline{v} = \alpha_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \alpha_2 \begin{pmatrix} 3 \\ 1 \end{pmatrix} = \begin{pmatrix} \underline{v}_1 \\ \underline{v}_2 \end{pmatrix}$$

$$\begin{cases} \alpha_1 + 3\alpha_2 = \underline{v}_1 \\ \alpha_1 + \alpha_2 = \underline{v}_2 \end{cases} \rightarrow \begin{cases} 2\alpha_2 = \underline{v}_1 - \underline{v}_2 \\ \alpha_1 + \alpha_2 = \underline{v}_2 \end{cases} \rightarrow \begin{cases} \alpha_2 = \dfrac{\underline{v}_1 - \underline{v}_2}{2} \\ \alpha_1 = \underline{v}_2 - \dfrac{\underline{v}_1 - \underline{v}_2}{2} = \dfrac{3\underline{v}_2 - \underline{v}_1}{2} \end{cases}$$

2. $\underline{u}_1, \underline{u}_2$ = linearly independent (We showed it before).

**Definition**

Let us consider vector space $V$ and vectors $\underline{v}_1, \dots, \underline{v}_n$ that form a basis of $V$. If vector $\underline{x} \in V$ can be written as $\underline{x} = x_1 \underline{v}_1 + \dots + x_n \underline{v}_n$ then $(x_1, \dots, x_n)$ are called the coordinates of $\underline{x}$ with respect to basis $\{v_1, \dots, v_n\}$

**Theorem**

Let us consider vector space $V$ and $v_1, \dots, v_n$ that are linearly independent. Let us assume that $\underline{x} = \alpha_1 v_1 + \dots + \alpha_n v_n$ and $\underline{x} = \beta_1 \underline{v}_1 + \dots + \beta_n \underline{v}_n$, then

$$\alpha_i = \beta_i \quad \forall i = 1, \dots, n$$

**Proof**

We have

$$x = \alpha_1 \underline{v}_1 + \dots + \alpha_n \underline{v}_n = \beta_1 \underline{v}_1 + \dots + \beta_n \underline{v}_n \rightarrow (\alpha_1 - \beta_1)\underline{v}_1 + \dots + (\alpha_n - \beta_n)\underline{v}_n = \underline{0}$$

Since $v_1, \dots, v_n$ are linearly independent $\Rightarrow \alpha_i = \beta_i$, $\forall i = 1, \dots, n$

$\square$

**<u>Remark:</u>**

The coordinates of any vector $\underline{x}$ with respect to given basis $\{\underline{v}_1, \dots, \underline{v}_n\}$ are unique.

**Theorem**

Let us consider vector space $V$. The number of vectors in any basis of $V$ is always the same.

**Remark:**

The number of vectors in the basis of vector space $V$ is called the dimension of vector space $V$.

## 4.3 Rank of Matrix

**Definition**

The row rank of matrix A is a maximum number of linearly independent rows of matrix $A$.

**Definition**

The column rank of matrix $A$ is a maximum number of linearly independent columns of matrix $A$.

**Remark:**

For any matrix $A \in \mathbb{R}^{m,n}$, the row rank is equal to the column rank. Therefore the row rank and column rank are sometimes called rank of matrix $A$, $\text{rank}(A)$.

**Example**

1.
$$A = \begin{pmatrix} 1 & 3 \\ 1 & 1 \end{pmatrix}$$
We have shown before that $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $\begin{pmatrix} 3 \\ 1 \end{pmatrix}$ are linearly independent, therefore $\text{rank}(A) = 2$.

2.
$$A = \begin{pmatrix} 1 & 0 \\ 7 & 0 \\ 3 & 0 \\ -1 & 0 \end{pmatrix} \in \mathbb{R}^{4,2}$$
The column vectors $\begin{pmatrix} 1 \\ 7 \\ 3 \\ -1 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$ are linearly dependent, thus $\text{rank}(A) = 1$ (i.e. the maximum number of linearly independent columns is 1).

**Remark:**

Two vectors are orthogonal if $\langle \underline{u}, \underline{v} \rangle = \underline{u}^T \underline{v} = 0$ (they basically must be perpendicular, i.e. the angle between $\underline{u}$ and $\underline{v}$ is 90 degrees).

**Definition**

Two subspaces $U$ and $W$ of vector space $V$ are orthogonal, if $\forall \underline{u} \in U$ and $\forall \underline{w} \in W$, we have $\langle \underline{u}, \underline{w} \rangle = 0$

**Definition**

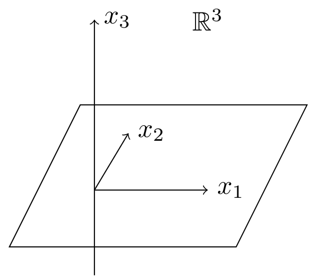
Orthogonal complement of subspace $M$ of vector space $V$ contains every vector orthogonal to $M$. This subspace is usually denoted by $M^{\perp}$

**Remark:**

$\dim M + \dim M^{\perp} = \dim V$

**Example**

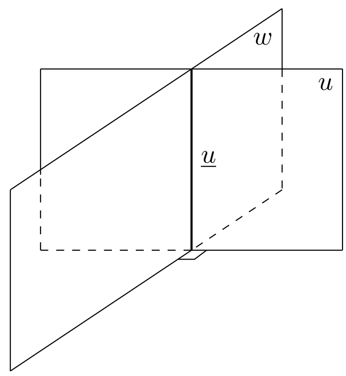
Consider $\mathbb{R}^3$




line $\alpha$ plane - orthogonal subspace. Orthogonal complement of each other

**Example**

Not orthogonal subspace!
$\underline{u} \neq 0$
$\underline{u} \in I$ & $\underline{u} \in W$
$\langle \underline{u} \in U, \underline{u} \in W \rangle = 0$



**Note**

If vector $\underline{u}$ belongs to 2 orthogonal subspaces, this vector is necessarily a zero vector, $\underline{u} = 0$ because we should have

$$\langle \underline{u}, \underline{u} \rangle = \underline{u}^T \underline{u} = 0 \Rightarrow \underline{u} = \underline{0}$$

# Chapter 5

# Orthogonality

## 5.1 Linear Mapping

**Definition**

Let us consider two vector spaces $V$ and $W$. A function (or mapping) $L : V \to W$ is called a *linear mapping* if the following two conditions are satisfied:

1. For any $\underline{v} \in V$ and $\underline{v}' \in V$, $L(\underline{v} + \underline{v}') = L(\underline{v}) + L(\underline{v}')$
2. For any $\underline{v} \in V$ and any scalar $\alpha$, $L(\alpha\underline{v}) = \alpha \cdot L(\underline{v})$

#### Example

Let us consider matrix $A \in \mathbb{R}^{n,m}$. We define a linear mapping $L_A$ as follows:

$$L_A(\underline{v}) = A\underline{v} \qquad L_A : \mathbb{R}^m \to \mathbb{R}^n$$

Is $L_A$ a linear mapping? Yes. A linear mapping or transformation can always be represented as a matrix-vector product and vice-versa.

#### Proof

1. $\forall \underline{v}, \underline{v}' \in \mathbb{R}^m$, we have:

$$L_A(\underline{v} + \underline{v}') = A(\underline{v} + \underline{v}') = A\underline{v} + A\underline{v}' = L_A(\underline{v}) + L_A(\underline{v}')$$

2. $\forall \underline{v} \in \mathbb{R}^m, \forall \alpha$ ($\alpha$ is scalar), we have:

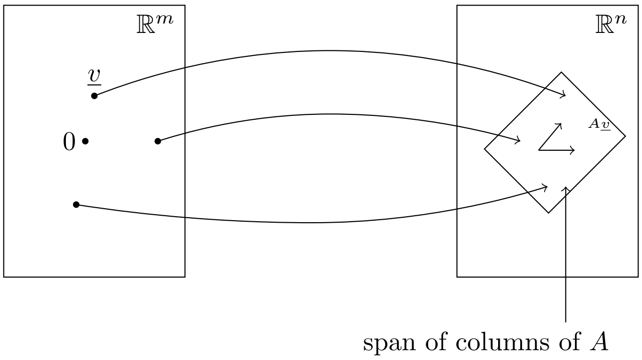
$$L_A(\alpha\underline{v}) = A(\alpha\underline{v}) = \alpha \cdot A\underline{v} = \alpha L_A(\underline{v})$$

$\square$

## 5.2 The 4 Fundamental Subspaces of a Matrix

Let us consider matrix $A \in \mathbb{R}^{n,m}$, $A : \mathbb{R}^m \to \mathbb{R}^n$. Let us consider the vector $\underline{v} \in \mathbb{R}^m$.

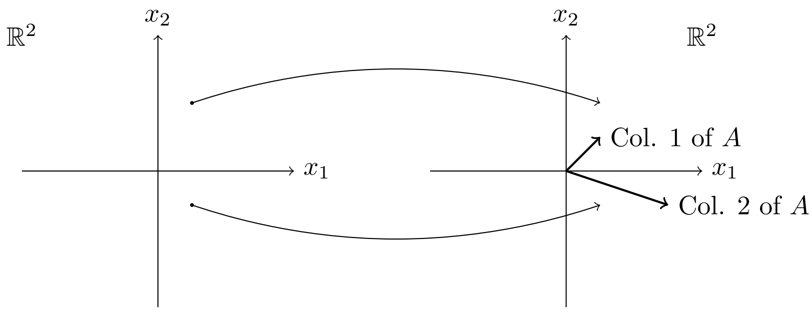
$$A\underline{v} = \underbrace{v_1 \cdot \begin{pmatrix} a_{11} \\ \vdots \\ a_{n1} \end{pmatrix} + v_2 \cdot \begin{pmatrix} a_{12} \\ \vdots \\ a_{n2} \end{pmatrix} + ... + v_m \cdot \begin{pmatrix} a_{1m} \\ \vdots \\ a_{nm} \end{pmatrix}}_{\text{Linear combination of columns of } A}$$




**Example**

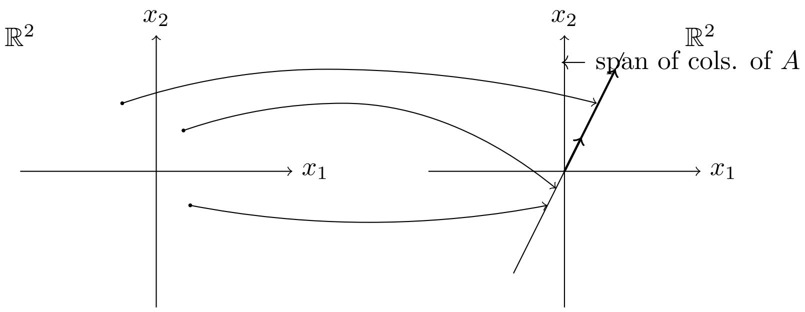
1.

$$A \in \mathbb{R}^{2,2} = \begin{pmatrix} 1 & 3 \\ 1 & -1 \end{pmatrix}$$




2.

$$A = \begin{pmatrix} 1 & 3 \\ 2 & 6 \end{pmatrix}$$




**Note**

In order for a solution of $A\underline{x} = \underline{b}$ to exist, $\underline{b}$ should belong to a *span of the columns* of the matrix $A$.

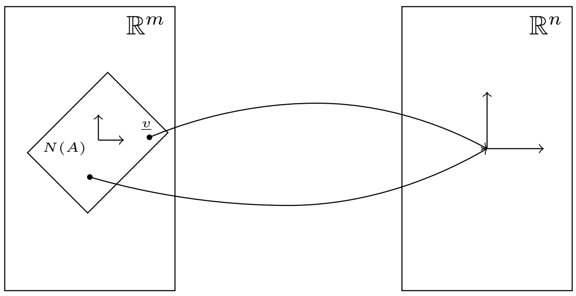
**Definition**

The span (or all possibile linear combinations) of columns of the matrix $A \in \mathbb{R}^{n,m}$ is called a *column space* of $A$, denoted by $C(A)$, where $C(A) \subset R^n$.

**Definition**

Let us consider the matrix $A \in \mathbb{R}^{n,m}, A : \mathbb{R}^m \to \mathbb{R}^n$. The null space of $A$ is defined as

$$N(A) = \{\underline{v} \in \mathbb{R}^m \mid A\underline{v} = \underline{0}\}, N(A) \subset \mathbb{R}^m$$




**Example**

Consider the following matrix

$$A = \begin{pmatrix} 1 & 3 \\ 2 & 6 \end{pmatrix}$$

*What is the nullspace of $A$?*

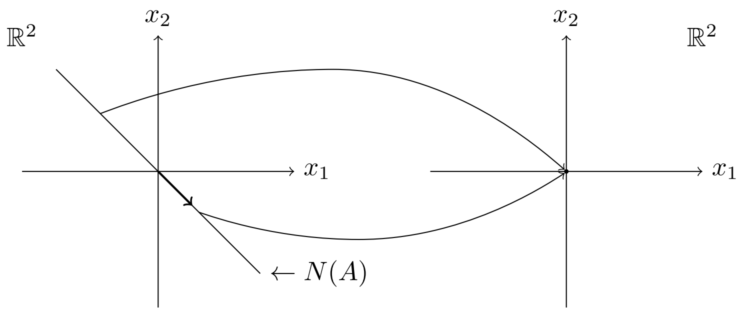
We should find all solutions for the following equation

$$A\underline{x} = \underline{0}$$

$$\begin{cases} x_1 + 3x_2 = 0 \\ 2x_1 + 6x_2 = 0 \end{cases} \rightarrow \begin{cases} x_1 + 3x_2 = 0 \\ 0 = 0 \end{cases} \rightarrow \begin{cases} x_1 = -3x_2 \\ 0 = 0 \end{cases}$$

The nullspace of this matrix will be a line formed by a linear combination of the vector $\begin{pmatrix} -3 \\ 1 \end{pmatrix}$, or in other words $\alpha \cdot \begin{pmatrix} -3 \\ 1 \end{pmatrix}$, for all possible $\alpha$, or in other words it will be the *span*$(\begin{pmatrix} -3 \\ 1 \end{pmatrix})$.

$$x_1 = -3x_2 = -3\alpha, x_2 = \alpha \rightarrow \alpha \begin{pmatrix} -3 \\ 1 \end{pmatrix}, \alpha \begin{pmatrix} -6 \\ 2 \end{pmatrix}$$




**Theorem**

The nullspace, $N(A)$, of $A \in \mathbb{R}^{n,m}$ is a subspace of $\mathbb{R}^m$.

**Proof**

Let us assume that $\underline{x}, \underline{x}' \in N(A)$ and $\alpha$ is arbitrary scalar.

1. $A(\underline{x} + \underline{x}') = A\underline{x} + A\underline{x}' = \underline{0} + \underline{0} = \underline{0} \Rightarrow (\underline{x} + \underline{x}') \in N(A)$

2. $A(\alpha\underline{x}) = \alpha(A\underline{x}) = \alpha \cdot \underline{0} = \underline{0} \Rightarrow \alpha\underline{x} \in N(A)$

$\square$

**Theorem**

The column space, $C(A)$, of $A \in \mathbb{R}^{n,m}$ is a subspace of $\mathbb{R}^n$.

**Definition**

The row space of a matrix $A \in R^{n,m}$ is the span of the rows of $A$. Clearly, $R(A) = C(A^T)$ and $R(A) \subset \mathbb{R}^m$.

**Definition**

The left nullspace of $A$ is defined as $N(A^T)$. $N(A^T) \subset \mathbb{R}^n$.

**Theorem**

$R(A)$ is a subspace of $\mathbb{R}^m$.

**Proof**

Same as for the proof that $C(A)$ is a subspace of $\mathbb{R}^n$, but for $A^T$.

$\square$

**Theorem**

$N(A^T)$ is a subspace of $\mathbb{R}^n$.

**Proof**

Same as for $N(A)$ but replace $A$ with $A^T$.

$\square$

**Theorem**

$R(A)$ and $N(A)$ are *orthogonal subspaces* in $\mathbb{R}^m$ for $A \in \mathbb{R}^{n,m}$.

**Proof**

Let us consider $\forall \underline{x} \in N(A), A\underline{x} = \underline{0}$

$$A\underline{x} = \begin{pmatrix} - \text{ row 1 of } A \rightarrow \\ \vdots \\ - \text{ row } n \text{ of } A \rightarrow \end{pmatrix} \cdot \begin{pmatrix} | \\ x \\ \downarrow \end{pmatrix} = \begin{pmatrix} < \text{ row 1 of } A, \underline{x} > \\ \vdots \\ < \text{ row } n \text{ of } A, \underline{x} > \end{pmatrix} \overset{\underline{x} \in N(A)}{=} \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

$\underline{x}$ is orthogonal to every row of $A$. $\underline{x}$ is orthogonal to every linear combination of rows of $A$. $\underline{x}$ is orthogonal to $R(A)$. In fact, what we just showed is that $N(A)$ and $R(A)$ are *orthogonal complements*.

$\square$

**Theorem**

$N(A^T)$ and $C(A) = R(A^T)$ are orthogonal complements in $\mathbb{R}^n$

Suppose we have the following matrix $A \in \mathbb{R}^{n,m} : \mathbb{R}^m \rightarrow \mathbb{R}^n$.
The row rank of $A = \text{rank}(A) = \dim(R(A)) = \dim(C(A)) =$ column rank of $A$.

$$N(A) : A\underline{x} = \underline{0} \quad \forall x \in \mathbb{R}^m$$

$$\begin{aligned} C(A) : A\underline{v} = & \text{ Linear combinations of columns of } A \\ & = v_1 \cdot \text{col 1 of } A + \cdots + v_n \cdot \text{col } n \text{ of } A \in \mathbb{R}^n \end{aligned}$$

**Theorem**

$N(A)$ is an orthogonal complement of $R(A)$ in $\mathbb{R}^m$,

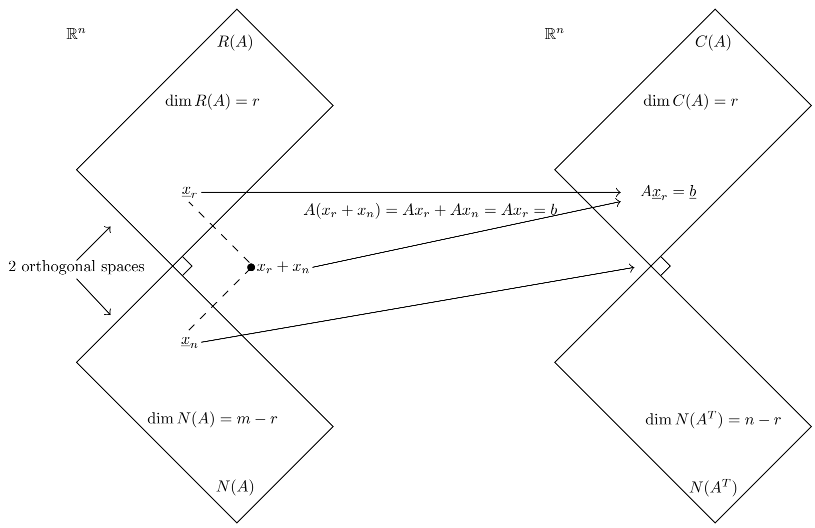
$$\dim N(A) + \underbrace{\dim R(A)}_{=\text{rank}(A)} = m$$

**Theorem**

$N(A^T)$ is an orthogonal complement of $R(A^T) = C(A)$ in $\mathbb{R}^n$,

$$\dim N(A^T) + \underbrace{\dim C(A)}_{=\text{rank}(A)} = n$$

Let us consider $A \in \mathbb{R}^{n,m}, A : \mathbb{R}^m \to \mathbb{R}^n$ and $\text{rank}(A) = r$




**Lemma**

For any vector $\underline{b}$ in $C(A)$, there exists *one and only one* vector $\underline{x}_r \in R(A)$ such that

$$A\underline{x}_r = b$$

**Proof**

Let us assume that $\underline{x}_r$ and $\underline{x}'_r$ are in the row space, $R(A)$. Let us assume that $A\underline{x}_r = A\underline{x}'_r$. We have

$$\underline{x}_r - \underline{x}'_r \in R(A)$$

But we also have

$$A\underline{x}_r - A\underline{x}'_r = \underbrace{A(\underline{x}_r - \underline{x}'_r)}_{\in N(A)} = \underline{0}$$

It means that $(\underline{x}_r - \underline{x}'_r)$ is in $R(A)$ and $N(A)$, but they are orthogonal subspaces, therefore

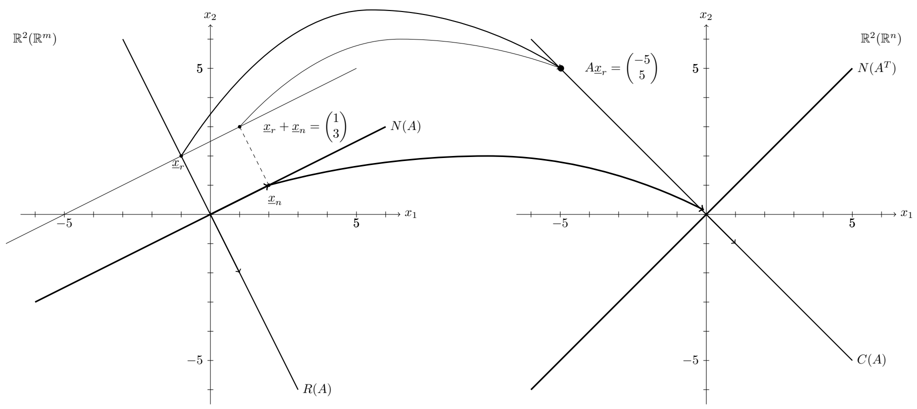
$$\underline{x}_r - \underline{x}'_r = \underline{0} \Rightarrow \underline{x}_r = \underline{x}'_r$$

$\square$

**Example**

Let us consider

$$A = \begin{pmatrix} 1 & -2 \\ -1 & 2 \end{pmatrix} \in \mathbb{R}^{2,2}$$




*What is the row space of A?*

It is basically all possible linear combinations of the row vectors of $A$:

$$R(A) = \text{span}\left\{ \begin{pmatrix} 1 \\ -2 \end{pmatrix}, \begin{pmatrix} -1 \\ 2 \end{pmatrix} \right\} = \text{span}\left\{ \begin{pmatrix} 1 \\ -2 \end{pmatrix} \right\}$$

Note that the vectors are linearly dependent. Observe $\text{rank}(A) = 1$, then $\dim(R(A)) = 1$.

*What is the nullspace of A?*

We need basically to find the solutions to the following equation

$$A\underline{x} = \underline{0}$$

$$\Rightarrow \begin{cases} x_1 - 2x_2 = 0 \\ -x_1 + 2x_2 = 0 \end{cases} \Rightarrow \begin{cases} x_1 - 2x_2 = 0 \\ 0 = 0 \end{cases} \Rightarrow x_1 = 2x_2 \text{ (which represents a line)}$$

The dimention of $N(A)$ is equal to the number of columns minus the dimension of the row space of $A$, so $\dim(N(A)) = 2 - 1 = 1$.

*What is the column space of A?*

The dimension of $C(A)$ is equal to the dimension of $R(A)$. The column space is defined as a linear combination of the column vectors of $A$:

$$C(A) = \text{span}\left\{\begin{pmatrix}1\\-1\end{pmatrix}, \begin{pmatrix}-2\\2\end{pmatrix}\right\} = \text{span}\left\{\begin{pmatrix}1\\-1\end{pmatrix}\right\}$$

*What is the left null space of $A$?*

First, $\dim N(A^T) = 2 - 1 = 1$.

Now, consider

$$\begin{aligned}
\underline{x}_r &= \begin{pmatrix}-1\\2\end{pmatrix} \Rightarrow A\underline{x}_r = \begin{pmatrix}1 & -2\\-1 & 2\end{pmatrix}\begin{pmatrix}-1\\2\end{pmatrix} = \begin{pmatrix}-5\\5\end{pmatrix}\\
\underline{x}_n &= \begin{pmatrix}2\\1\end{pmatrix} \Rightarrow A\underline{x}_n = \begin{pmatrix}1 & -2\\-1 & 2\end{pmatrix}\begin{pmatrix}2\\1\end{pmatrix} = \begin{pmatrix}0\\0\end{pmatrix}
\end{aligned}$$

## 5.3 Orthogonal Basis and Gram-Schmidt process

### 5.3.1 Orthogonal and Orthonormal

**Definition**

Vectors $\underline{q}_1, \ldots, \underline{q}_m$ are *orthogonal* if:

$$\langle \underline{q}_i, \underline{q}_j \rangle = \underline{q}_i^T \underline{q}_j = 0 \quad \text{if } i \neq j$$

In other words, two vectors $\underline{q}_i$ and $\underline{q}_j$ are orthogonal if their dot product is zero, so two orthogonal vectors must be perpendicular.

**Definition**

Vectors $\underline{q}_1, \ldots, \underline{q}_m$ are *orthonormal* if:

$$\langle \underline{q}_i, \underline{q}_j \rangle = \underline{q}_i^T \underline{q}_j = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j \end{cases}$$

In other words, two vectors $\underline{q}_i$ and $\underline{q}_j$ are orthonormal if they are orthogonal and they both have length of 1 (they are *unit vectors*).

If the columns of the matrix are orthonormal vectors, then this matrix is usually denoted by $Q$. In this case, we have

$$Q^T Q = I$$

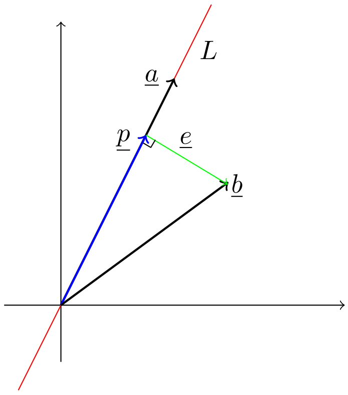
Note that if $Q$ is *not* a square matrix, then $QQ^T$ is not necessarily $I$.

**Definition**

A square matrix is called orthogonal (if its columns are orthonormal vectors) if $Q^TQ = I$. In this case, since it is a square matrix, $QQ^T = I$

### 5.3.2 Projection onto a Line

Let us assume that we have a line $L$ which is the span of the vector $\underline{a} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \in \mathbb{R}^n$, and suppose we have another vector $\underline{b} \in \mathbb{R}^n$. We want to find the vector $\underline{p}$ belonging to the line $L$ (spanned by $\underline{a}$), closest to vector $\underline{b}$. In other words, we are looking for the vector $\underline{p}$, which is the orthogonal projection of $\underline{b}$ onto the line given by $\underline{a}$.




$\underline{p}$ is vector multiple of the vector $\underline{a}$ (it has the same direction, but possibly not the same length)

$$\underline{p} = c\underline{a}$$

where $c$ is some scalar.

Now, let us define the so-called *error vector* $\underline{e}$ in the following way

$$\underline{e} = \underline{b} - \underline{p} = \underline{b} - c\underline{a}$$

As you can see from the picture, the error vector $\underline{e}$ is orthogonal to the line, therefore

$$\begin{aligned}
&\langle \underline{a}, \underline{e} \rangle = 0 \\
&\langle \underline{a}, \underline{e} \rangle = \underline{a}^T \cdot (\underline{b} - c\underline{a}) = \underline{a}^T \cdot \underline{b} - c(\underline{a}^T \cdot \underline{a}) = 0 \\
&\Rightarrow c = \frac{\underline{a}^T \underline{b}}{\underline{a}^T \underline{a}}
\end{aligned}$$

Note that $c$ is the scalar that you multiply by $\underline{a}$ to obtain $\underline{p}$, therefore

$$\underline{p} = c\underline{a} = \underline{a}c = \underline{a}\frac{\underline{a}^T\underline{b}}{\underline{a}^T\underline{a}} = \underbrace{\frac{\underline{a}\underline{a}^T}{\underline{a}^T\underline{a}}}_{P\in\mathbb{R}^{n,n}\ \text{(projection matrix)}} \cdot\underline{b}$$

**Example**

Let us consider $\underline{a} = \begin{pmatrix}1\\2\\2\end{pmatrix} \in \mathbb{R}^3$. We first want to find the projection matrix $P$:

$$P = \frac{\underline{a}\cdot\underline{a}^T}{\underline{a}^T\underline{a}} = \begin{pmatrix}1\\2\\2\end{pmatrix}\cdot\begin{pmatrix}1 & 2 & 2\end{pmatrix}\cdot\frac{1}{9} = \frac{1}{9}\cdot\begin{pmatrix}1 & 2 & 2\\2 & 4 & 4\\2 & 4 & 4\end{pmatrix}$$

Note that the numerator produces a matrix, but the denominator produces a number which is equivalent to the dot product of $a$ with itself.

Let us project the following vector $\underline{b} = \begin{pmatrix}1\\1\\1\end{pmatrix}$ onto the line spanned by $\underline{a}$. We just need to multiply $P$ by $\underline{b}$:

$$\underline{p} = P\underline{b} = \frac{1}{9}\begin{pmatrix}1 & 2 & 2\\2 & 4 & 4\\2 & 4 & 4\end{pmatrix}\begin{pmatrix}1\\1\\1\end{pmatrix} = \frac{1}{9}\begin{pmatrix}5\\10\\10\end{pmatrix}$$

**Note**

$\underline{p}^2 = \underline{p}$

**Note**

$(I-P)-$ projection onto subspace orthogonal to the line given by $\underline{a}$

### 5.3.3 Gram-Schmidt Process

Given linearly independent vectors $\underline{a}, \underline{b}, \underline{c}, \ldots$ (note that linearly indepence does not necessarily mean that the vectors are perpendicular to each other!), the Gram-Schmidt process allows us to construct an *orthogonal basis* of span $\underline{a}, \underline{b}, \underline{c}, \ldots \in \mathbb{R}^n$: we first want to find orthogonal vectors $\underline{a}', \underline{b}', \underline{c}', \ldots$ which span the same subspace as $\underline{a}, \underline{b}, \underline{c}, \ldots$, and then we normalise them (we make them unit vectors by dividing each of them by their respective length).

The general process is the following:

1. Choose $\underline{a}' = \underline{a}$
2. It is likely that $\underline{b}$ is not orthogonal to $\underline{a}'$, so we need to subtract the projection of $\underline{b}$ onto $\underline{a}'$ from $\underline{b}$:

$$\underline{b}' = \underline{b} - \frac{\underline{a}'^T\underline{b}}{\underline{a}'^T\underline{a}'}\underline{a}'$$

Note that $\underline{a} = \underline{a}'$, so it does not matter if we calculate the projection of $\underline{b}$ onto $\underline{a}'$ or $\underline{a}$ (in this case). If you are still not familiar how projections work, try to draw the situation on a paper, and you will definitely understand better.

3. $\underline{c}'$ is likely not orthogonal to $\underline{a}'$ and $\underline{b}'$. Again, subtract its projections

$$\underline{c}' = \underline{c} - \frac{\underline{a}'^T \underline{c}}{\underline{a}'^T \underline{a}'}\underline{a}' - \frac{\underline{b}'^T \underline{c}}{\underline{b}'^T \underline{b}'}\underline{b}'$$

and so on.

4. Finally, normalise $\underline{a}'$, $\underline{b}'$, $\underline{c}'$, ...

$$\underline{\hat{q}}_1 = \frac{\underline{a}'}{\|\underline{a}'\|}, \quad \underline{\hat{q}}_2 = \frac{\underline{b}'}{\|\underline{b}'\|}, \quad \underline{\hat{q}}_3 = \frac{\underline{c}'}{\|\underline{c}'\|}, \dots$$

Note that $\underline{\hat{q}}_1$, $\underline{\hat{q}}_2$, $\underline{\hat{q}}_3$, ... are the orthogonal unit vectors, which form a orthogonal basis. Vectors that are denoted with ˆ over them are often intended to be *unit vectors*.

### Example

Suppose we have the following linearly independent vectors

$$\underline{a} = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \underline{b} = \begin{pmatrix} 2 \\ 0 \\ -2 \end{pmatrix}, \underline{c} = \begin{pmatrix} 3 \\ -3 \\ 3 \end{pmatrix}$$

We first need to find $\underline{a}', \underline{b}', \underline{c}'$, and then the unit vectors $\underline{\hat{q}}_1, \underline{\hat{q}}_2, \underline{\hat{q}}_3$.

1.

$$\underline{a}' = \underline{a} = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$$

2.

$$\underline{b}' = \underline{b} - \frac{\underline{a}'^T \underline{b}}{\underline{a}'^T \underline{a}'}\underline{a}' = \begin{pmatrix} 2 \\ 0 \\ -2 \end{pmatrix} - \frac{\left\langle \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ -2 \end{pmatrix} \right\rangle}{\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}} = \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}$$

3.

$$\underline{c}' = \underline{c} - \frac{\underline{a}'^T \underline{c}}{\underline{a}'^T \underline{a}'}\underline{a}' - \frac{\underline{b}'^T \underline{c}}{\underline{b}'^T \underline{b}'}\underline{b}' = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

We can try to check if $\underline{a}'$, $\underline{b}'$ and $\underline{c}'$ are perpendicular to each other by checking if their dot product is equals to zero:

$$\langle \underline{a}', \underline{b}' \rangle = 0$$

$$\langle \underline{a}', \underline{c}' \rangle = 0$$

$$\langle \underline{b}', \underline{c}' \rangle = 0$$

4. Finally, we normalise $\underline{a}'$, $\underline{b}'$ and $\underline{c}'$:

$$\underline{\hat{q}}_1 = \frac{\underline{a}'}{\|\underline{a}'\|} = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\\0\end{pmatrix}$$

$$\underline{\hat{q}}_2 = \frac{\underline{b}'}{\|\underline{b}'\|} = \frac{1}{\sqrt{6}}\begin{pmatrix}1\\1\\-2\end{pmatrix}$$

$$\underline{\hat{q}}_3 = \frac{\underline{c}'}{\|\underline{c}'\|} = \frac{1}{\sqrt{3}}\begin{pmatrix}1\\1\\1\end{pmatrix}$$

### 5.3.4 Projection onto a Subspace

Assume we have linearly independent vectors $a_1, \ldots, a_m \in \mathbb{R}^n$. We want to project the vector $\underline{b} \in \mathbb{R}^n$ onto the subspace $V$ spanned by $a_1, \ldots, a_m$. Subspace consists of all linear combinations

$$x_1 a_1 + \cdots + x_m a_m = \underbrace{\begin{pmatrix}| & & |\\ a_1 & \ldots & a_m\\ \downarrow & & \downarrow\end{pmatrix}}_{A \in \mathbb{R}^{n,m}} \cdot \underbrace{\hat{x}}_{\in \mathbb{R}^m}$$

We are looking for the projection $\underline{p}$ of $\underline{b}$ onto subspace $V$. We can define $\underline{e} = \underline{b} - \underline{p}$, where $\underline{e}$ should be orthogonal to all $a_1, \ldots, a_m$

$$\left.\begin{aligned}\langle a_1, \underline{e}\rangle &= \underline{a_1}^T \cdot (\underline{b} - A\hat{x}) = 0\\ &\vdots\\ \langle a_m, \underline{e}\rangle &= \underline{a_m}^T \cdot (\underline{b} - A\hat{x}) = 0\end{aligned}\right\} \Rightarrow \underbrace{\begin{pmatrix}-\underline{a}_1^T \rightarrow\\ \vdots\\ -\underline{a}_m^T \rightarrow\end{pmatrix}}_{A^T}(\underline{b} - A\hat{x}) = 0$$

$$\begin{aligned}A^T(\underline{b} - A\hat{x}) &= 0\\ A^T\underline{b} - A^T A\hat{x} &= 0\end{aligned}$$

**Theorem**

$A$ has linearly independent columns. Then $A^T A$ is:

- Square
- Symmetric
- Invertible

$$\begin{aligned}\underline{\hat{x}} &= A(A^T A)^{-1} A^T \underline{b}\\ \underline{p} &= A\hat{x} = \underbrace{A(A^T A)^{-1} A^T}_{P \text{ - Proj. matrix}} \cdot \underline{b} \text{ - Projection vector}\end{aligned}$$

# Chapter 6

# Determinant

Let us consider matrix $A \in \mathbb{R}^{n,n}$, a square matrix. The determinant of $A$ is a number, usually written as $\det(A)$ or $|A|$. Let us consider

$$\begin{aligned} A &= \begin{pmatrix} a & b \\ c & d \end{pmatrix} \\ \det(A) &= ad - bc \\ A^{-1} &= \frac{1}{ad-bc} \cdot \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \Rightarrow \text{ Inverse of } A \end{aligned}$$

In order for $A^{-1}$ to exist, $\det(A)$ should not be equal to 0. If $\det(A) = 0$, then $A^{-1}$ does not exist, and $A$ is not invertible. $A$ is a triangular matrix

$$\begin{aligned} A^{-1} &= \begin{pmatrix} \frac{d}{ad-bc} & -\frac{b}{ad-bc} \\ -\frac{c}{ad-bc} & \frac{a}{ad-bc} \end{pmatrix} \\ \det(A^{-1}) &= \frac{d}{ad-bc} \cdot \frac{a}{ad-bc} - \frac{-b}{ad-bc} \cdot \frac{-c}{ad-bc} \\ &= \frac{da-bc}{(ad-bc)^2} = \frac{1}{ad-bc} = \frac{1}{\det(A)} \end{aligned}$$

Consider

$$A^* = \begin{pmatrix} a & b \\ a & b \end{pmatrix}$$

Then $\det(A^*) = ab - ab = 0$. Let us now consider

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

where the rows are switched. Then

$$\det(A') = cb - ad = -\det(A)$$

**Properties**

The following properties are true for any $n \times n$ matrix.

1. Determinant of $I \in \mathbb{R}^{n,n}$ (identity matrix) is equal to 1

2. If 2 rows of matrix $A \in \mathbb{R}^{n,n}$ are exchanged, the determinant changes its sign
3. The determinant is a linear function of each row, all other rows stay the same
4. If $A \in \mathbb{R}^{n,n}$ has at least 2 equal rows, then $\det(A) = 0$ (i.e. 2 or more equal rows)
5. If we add a multiple of one row to another row, the determinant does not change.
6. If $A$ has row of zeroes, then $\det(A) = 0$
7. Let us consider $A$ and upper or lower triangular matrix

$$A = \begin{pmatrix} a_{11} & & * \\ \vdots & \ddots & \\ 0 & \dots & a_{nn} \end{pmatrix} \text{ or } A = \begin{pmatrix} a_{11} & & 0 \\ \vdots & \ddots & \\ * & \dots & a_{nn} \end{pmatrix}$$

   Then $\det(A) = a_{11} \cdot \dots \cdot a_{nn}$

8. If $A$ is singular then $\det(A) = 0$. If $A$ is non singular, then $\det(A) \neq 0$.
9. $\det(A \cdot B) = \det(A) \cdot \det(B)$
10. $\det(A^T) = \det(A)$

**Example**

P2 Permutation matrix $P$ - identity matrix with rows exchanged

- If rows are exchanged an odd number of times, then $\det(P) = -1$
- If rows are exchanged an even number of times, then $\det(P) = 1$

P3

- If we multiply a row, say the top row, by a number $t$ we get that:

$$\begin{pmatrix} ta & tb \\ c & d \end{pmatrix}$$

  which has determinant of $tad - tbc = t(ad - bd)$. So, multiplying a row by $t$ multiplies the determinant by $t$, or visually:

$$\begin{vmatrix} ta & tb \\ c & d \end{vmatrix} = t \begin{vmatrix} a & b \\ c & d \end{vmatrix}$$

- $$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = \begin{vmatrix} a & 0 \\ c & d \end{vmatrix} + \begin{vmatrix} 0 & b \\ c & d \end{vmatrix}$$

- $$\begin{vmatrix} 4 & 8 & 8 \\ 3 & 7 & 9 \\ 2 & 1 & 4 \end{vmatrix} = 4 \cdot \begin{vmatrix} 1 & 2 & 2 \\ 3 & 7 & 9 \\ 2 & 1 & 4 \end{vmatrix} = \begin{vmatrix} 4 & 0 & 0 \\ 3 & 7 & 9 \\ 2 & 1 & 4 \end{vmatrix} + \begin{vmatrix} 0 & 8 & 8 \\ 3 & 7 & 9 \\ 2 & 1 & 4 \end{vmatrix}$$

**Proof**

P4 Let us assume that rows $i$ and $j$ are equal we can exchange these rows. The resulting matrix $A'$ is in fact equal to $A$. But due to P2

$$\begin{aligned} &\det(A') = -\det(A) \\ &A' = A \Rightarrow \det(A') = \det(A) \\ \Rightarrow &\det(A) = -\det(A) = 0 \end{aligned}$$

P5

$$A = \begin{pmatrix} \text{row } 1 \\ \vdots \\ \text{row } n \end{pmatrix}$$

$$\begin{vmatrix} \text{row } 1 \rightarrow \\ \text{row } i \rightarrow \\ \text{row } j + 2\text{row } i \rightarrow \\ \text{row } n \rightarrow \end{vmatrix} \stackrel{P3}{=} \begin{vmatrix} \text{row } 1 \rightarrow \\ \text{row } i \rightarrow \\ \text{row } j \rightarrow \\ \text{row } n \rightarrow \end{vmatrix} + 2 \begin{vmatrix} \text{row } 1 \rightarrow \\ \text{row } i \rightarrow \\ \text{row } i \rightarrow \\ \text{row } n \rightarrow \end{vmatrix} = \det(A)$$

Note that $2 \begin{vmatrix} \text{row } 1 \rightarrow \\ \text{row } i \rightarrow \\ \text{row } i \rightarrow \\ \text{row } n \rightarrow \end{vmatrix} = 0$, because of property 4.

Remark: Our standard row operation in gaussian elimination do not change the determinant. The only exception is the exchange of rows.

P6 Add any other row to the zero row and get a matrix with 2 equal rows. From P4 $\Rightarrow \det(A) = 0$

P7 Let us first assume that all $a_{ii}$ are not equal to zeroes, $\forall i = 1, \dots, n$. Then by adding rows, we an bring the matrix to the diagonal form. Then we will set matrix

$$\begin{aligned} \begin{vmatrix} a_{11} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & a_{nn} \end{vmatrix} &= a_{11} \cdot \begin{vmatrix} 1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & a_{nn} \end{vmatrix} = a_{11} \cdot a_{22} \cdot \begin{vmatrix} 1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & a_{nn} \end{vmatrix} \\ &= a_{11} \cdot \dots \cdot a_{nn} \underbrace{\begin{vmatrix} 1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 1 \end{vmatrix}}_{1} \\ &= a_{11} \cdot \dots \cdot a_{nn} \end{aligned}$$

If $a_{ii}$ is equal to zero, we can use all other diagonal elements that are not zero, and we can eliminate all non-zero elements from row $i$ using gaussian elimination. At the end, we will get a matrix with row of zeroes, for which the determinant is equal to 0 by propriety 6 and therefore

$$\det(A) = 0 = a_{11} \cdot \dots \cdot a_{nn}$$

Remark: When we use the gaussian elimination, we bring the matrix to an upper triangular form. At the end we have pivot elements on the diagonal. If all pivot elements are non-zero elements, the determinant is not equal to zero since it is a product of pivot elements.

If some elements on the diagonal are zero, then the matrix determinant is equal to zero, the matrix does not have an inverse, the matrix is singular.

P8 We use gaussian elimination to reduce our matrix to an upper triangular matrix. If all pivot elements are non-zero (non-singular matrix), then $\det(A) \neq 0$, according to property 7. Otherwise $\det(A) = 0$.

P9

$$\begin{aligned} \det(A \cdot A^{-1}) &= \det(I) = 1 \\ \det(A) \cdot \det(A^{-1}) &= 1 \\ \Rightarrow \det(A^{-1}) &= \frac{1}{\det(A)} \end{aligned}$$

□

<u>**Remark:**</u>

Everything we just said about rows is also valid for columns.

## 6.1 Compute the Determinant

1. The first method to compute the determinant of a matrix is by using *gaussian elimination* to bring the matrix to its *upper triangular form*, and then the determinant is the product of the diagonal elements. Whenever we have to exchange 2 rows, the determinant changes its sign. Note that you should count how many times you exchange 2 rows.

2. The second method consists of calculating the matrix cofactors.
   Let us consider the following matrix

$$A = \begin{pmatrix} a_{11} & \dots & a_{1j} & \dots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{i1} & \dots & a_{ij} & \dots & a_{in} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \dots & a_{nj} & \dots & a_{nn} \end{pmatrix} \in \mathbb{R}^{n,n}$$

   We can construct a matrix $M_{ij}$ by throwing out row $i$ and column $j$ of $A$.
   For example $M_{11}$ (without column and row 1) would look like:

$$M_{11} = \begin{pmatrix} a_{22} & \dots & a_{2j} & \dots & a_{2n} \\ \vdots & & \vdots & & \vdots \\ a_{i2} & \dots & a_{ij} & \dots & a_{in} \\ \vdots & & \vdots & & \vdots \\ a_{n2} & \dots & a_{nj} & \dots & a_{nn} \end{pmatrix} \in \mathbb{R}^{n-1,n-1}$$

Or in general:

$$M_{ij} = \begin{pmatrix} a_{11} & \dots & a_{1j-1} & \dots & a_{1j+1} & \dots & a_{1n} \\ \vdots & & \vdots & & \vdots & & \vdots \\ a_{i-11} & \dots & a_{i-1j-1} & \dots & a_{i-1j+1} & \dots & a_{i-1n} \\ \vdots & & \vdots & & \vdots & & \vdots \\ a_{i+11} & \dots & a_{i+1j-1} & \dots & a_{i+1j+1} & \dots & a_{i+1n} \\ \vdots & & \vdots & & \vdots & & \vdots \\ a_{n1} & \dots & a_{nj-1} & \dots & a_{nj+1} & \dots & a_{nn} \end{pmatrix} \in \mathbb{R}^{n-1,n-1}$$

The determinant of $M_{ij}$ is usually called *minor*.
We define the cofactor

$$C_{ij} = (-1)^{i+j} \cdot \det(M_{ij})$$

.

The determinant of $A$ can be written as the *sum of the cofactors* of any row or column of the matrix multiplied by the *entries* that generated them. In other words, the cofactor expansion along row $i$ gives:

$$det(A) = a_{i1} \cdot C_{i1} + a_{i2} \cdot C_{i2} + \dots + a_{in} \cdot C_{in}$$

And the cofactor expansion along column $j$ gives:

$$det(A) = a_{1j} \cdot C_{1j} + a_{2j} \cdot C_{2j} + \dots + a_{nj} \cdot C_{nj}$$

**Example**

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \in \mathbb{R}^{2,2}$$

we will use expansion by cofactors using row 1

$$\begin{aligned} \det(A) &= a_{11}c_{11} + a_{12}c_{12} \\ &= a_{11} \cdot (-1)^{1+1} \cdot \det(a_{22}) + a_{12} \cdot (-1)^{1+2} \cdot \det(a_{21}) \\ &= a_{11} \cdot a_{22} - a_{12} \cdot a_{21} \end{aligned}$$

**Example**

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \in \mathbb{R}^{3,3}$$

Let us again use the expansion by row 1. Then

$$\begin{aligned} \det(A) &= a_{11}c_{11} + a_{12}c_{12} + a_{13}c_{13} \\ &= a_{11} \cdot (-1)^{1+1} \cdot \left| \begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix} \right| + a_{12} \cdot (-1)^{1+2} \cdot \left| \begin{pmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{pmatrix} \right| \\ &\qquad + a_{13} \cdot (-1)^{1+3} \cdot \left| \begin{pmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix} \right| \\ &= a_{11}(a_{22}a_{33} - a_{32}a_{23}) - a_{12}(a_{21}a_{33} - a_{31}a_{23}) + a_{13}(a_{21}a_{32} - a_{31}a_{22}) \\ &= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{32}a_{23} - a_{12}a_{21}a_{33} - a_{13}a_{31}a_{22} \end{aligned}$$

$$\begin{array}{ccc|cc}
a_{11} & a_{12} & a_{13} & a_{11} & a_{12} \\
a_{21} & a_{22} & a_{23} & a_{21} & a_{22} \\
a_{31} & a_{32} & a_{33} & a_{31} & a_{32} \\
- & - & - \quad + & + & +
\end{array}$$

In this case the determinant is given by the product of the diagonals left to right, minus the product of the diagonals right to left. This is called the *rule of Sarrus*. Note that *rule of Sarrus* does not work for matrices with size greater than 3.

Most of the time, we use gaussian elimination to compute the determinant. We can use the cofactor formula mostly when $A$ has many zeroes.

## 6.2 Cramer's Rule

*Cramer's rule* is an explicit formula for the solution of a linear system of equations with *as many equations as unknowns*, valid whenever the system has a *unique* solution.
Let us consider the equation

$$A\underline{x} = \underline{b}$$

where $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$, $\underline{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ and $\underline{b} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$
So, our equation is

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$$

We can then substitute the first column of $A$ with $\underline{b}$, and we call this new matrix $B_1$:

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \cdot \begin{pmatrix} x_1 & 0 & 0 \\ x_2 & 1 & 0 \\ x_3 & 0 & 1 \end{pmatrix} = \begin{pmatrix} b_1 & a_{12} & a_{13} \\ b_2 & a_{22} & a_{23} \\ b_3 & a_{32} & a_{33} \end{pmatrix} = B_1$$

According to the properties of calculating the determinant of a matrix, we know that:

$$\det(A \cdot B) = \det(A) \cdot \det(B)$$

We can apply this rule to calculate the determinant of

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \cdot \begin{pmatrix} x_1 & 0 & 0 \\ x_2 & 1 & 0 \\ x_3 & 0 & 1 \end{pmatrix}$$

So we have:

$$\det(A) \cdot \det \begin{pmatrix} x_1 & 0 & 0 \\ x_2 & 1 & 0 \\ x_3 & 0 & 1 \end{pmatrix} = \det(B_1)$$

Note that $\begin{pmatrix} x_1 & 0 & 0 \\ x_2 & 1 & 0 \\ x_3 & 0 & 1 \end{pmatrix}$ is a lower triangular matrix, thus its determinant is the product of the diagonal entries, thus its determinant is $x_1$:

$$\begin{aligned} \det(A) \cdot x_1 &= \det(B_1) \\ x_1 &= \frac{\det(B_1)}{\det(A)} \end{aligned}$$

Similarly, we can again substitute the second row of $A$ with $\underline{b}$ to form $B_2$:

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \cdot \begin{pmatrix} 1 & x_1 & 0 \\ 0 & x_2 & 0 \\ 0 & x_3 & 1 \end{pmatrix} = \begin{pmatrix} a_{11} & b_1 & a_{13} \\ a_{21} & b_2 & a_{23} \\ a_{23} & b_3 & a_{33} \end{pmatrix} = B_2$$

$$\begin{aligned} \det(A) \cdot x_2 &= \det(B_2) \\ x_2 &= \frac{\det(B_2)}{\det(A)} \end{aligned}$$

In general

$$x_i = \frac{\det(B_i)}{\det(A)}, i = 1, \dots, n$$

where $B_i$ is $A$ with column $i$ replaced by $\underline{b}$, where $\det(A) \neq 0$ (division by zero is not defined).

**Example**

Suppose we want to find the solutions to the following linear system of equations:

$$\begin{cases} 3x + 2y + 4z = 1 \\ 2x - y + z = 0 \\ x + 2y + 3z = 1 \end{cases}$$

We first create the matrix $A$ with the coefficients of the addends that are on the left side of the equals sign of each equation:

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

and similarly we create the vector $\underline{b}$ using the numbers on the right side of the equals sign of each equation:

$$\underline{b} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$$

Now, we just need to find $B_1$, $B_2$ and $B_3$ using the process described above:

$$B_1 = \begin{pmatrix} 1 & 2 & 4 \\ 0 & -1 & 1 \\ 1 & 2 & 3 \end{pmatrix}$$

$$B_2 = \begin{pmatrix} 3 & 1 & 4 \\ 2 & 0 & 1 \\ 1 & 1 & 3 \end{pmatrix}$$

$$B_3 = \begin{pmatrix} 3 & 2 & 1 \\ 2 & -1 & 0 \\ 1 & 2 & 1 \end{pmatrix}$$

Finally, we calculate the determinant of each $B_i$ (where $i = 1...3$), and we divide it by the determinant of $A$ (which must be different from 0) to find respectively $x$, $y$ and $z$:

$$x = \frac{\begin{vmatrix} 1 & 2 & 4 \\ 0 & -1 & 1 \\ 1 & 2 & 3 \end{vmatrix}}{\begin{vmatrix} 3 & 2 & 4 \\ 2 & -1 & 1 \\ 1 & 2 & 3 \end{vmatrix}} = -\frac{1}{5}$$

$$y = \frac{\begin{vmatrix} 3 & 1 & 4 \\ 2 & 0 & 1 \\ 1 & 1 & 3 \end{vmatrix}}{\begin{vmatrix} 3 & 2 & 4 \\ 2 & -1 & 1 \\ 1 & 2 & 3 \end{vmatrix}} = 0$$

$$z = \frac{\begin{vmatrix} 3 & 2 & 1 \\ 2 & -1 & 0 \\ 1 & 2 & 1 \end{vmatrix}}{\begin{vmatrix} 3 & 2 & 4 \\ 2 & -1 & 1 \\ 1 & 2 & 3 \end{vmatrix}} = \frac{2}{5}$$

## 6.3 Inverse of a Matrix

We can use Cramer's rule also to find the inverse of a matrix.

We know that if

$$AX = XA = I$$

then $X$ is called the inverse of $A$, also denoted as $A^{-1}$.

So we have that:

$$\begin{pmatrix} a_{11} & \dots & a_{13} \\ \vdots & & \vdots \\ a_{31} & \dots & a_{33} \end{pmatrix} \cdot \begin{pmatrix} x_{11} & \dots & x_{13} \\ \vdots & & \vdots \\ x_{31} & \dots & x_{33} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

which implies that:

$$\begin{pmatrix} a_{11} & \dots & a_{13} \\ \vdots & & \vdots \\ a_{31} & \dots & a_{33} \end{pmatrix} \cdot \begin{pmatrix} x_{11} \\ x_{21} \\ x_{31} \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

Now, we can apply the *Cramer's rule* to find $x_{11}$, $x_{21}$ and $x_{31}$:

$$x_{11} = \frac{\begin{vmatrix} 1 & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & a_{32} & a_{33} \end{vmatrix}}{\det(A)} = \frac{c_{11}}{\det(A)}$$

$$x_{21} = \frac{c_{12}}{\det(A)}$$

$$x_{31} = \frac{c_{13}}{\det(A)}$$

In general, element $ij$ of $A^{-1}$ can be computed as:

$$\left(A^{-1}\right)_{ij} = \frac{c_{ji}}{\det(A)}$$

Note that it is really $c_{ji}$ and **not** $c_{ij}$!

# Chapter 7

# Linear Mappings

**Definition**

A mapping $L : \mathbb{R}^m \to \mathbb{R}^n$ is said to be a *linear mapping* (or linear transformation or linear map) if for any $\underline{u}, \underline{v} \in \mathbb{R}^m$ and any scalar $\alpha$

$$\begin{aligned} L(\underline{u} + \underline{v}) &= L(\underline{u}) + L(\underline{v}) \\ L(\alpha \underline{u}) &= \alpha L(\underline{u}) \end{aligned}$$

For any matrix $A \in \mathbb{R}^{m,n}$ we can associate with it a linear mapping $L_A$ as

$$L_A(\underline{u}) = A\underline{u} \qquad \forall \underline{u} \in \mathbb{R}^n$$

where $L_A$ is a linear mapping from $\mathbb{R}^n$ to $\mathbb{R}^m$.

In principle, any linear mapping is completely defined by its values on the basis vectors.

**Example**

Let us consider the linear mapping $L : \mathbb{R}^2 \to \mathbb{R}^2$ and basis in $\mathbb{R}^2$ consisting of the following vectors:

$$\underline{b}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \underline{b}_2 = \begin{pmatrix} -1 \\ 2 \end{pmatrix}$$

Lets assume that

$$L \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 7 \\ 3 \end{pmatrix}$$

and

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

How do we find $L \begin{pmatrix} 1 \\ 4 \end{pmatrix}$?

Since $\underline{b}_1$ and $\underline{b}_2$ form a basis for $\mathbb{R}^2$, the vector $\begin{pmatrix}1\\4\end{pmatrix}$ can be represented as a linear combination of basis vectors:

$$\begin{pmatrix}1\\4\end{pmatrix} = \alpha_1 \begin{pmatrix}1\\1\end{pmatrix} + \alpha_2 \begin{pmatrix}-1\\2\end{pmatrix}$$

We first find $\alpha_1$ and $\alpha_2$, in this case we have found $\alpha_1 = 2$ and $\alpha_2 = 1$.

Then, using the first condition to be a linear mapping

$$L(\underline{u} + \underline{v}) = L(\underline{u}) + L(\underline{v})$$

we know that:

$$\begin{aligned} L\begin{pmatrix}1\\4\end{pmatrix} &= L\left(2\begin{pmatrix}1\\1\end{pmatrix} + 1\begin{pmatrix}-1\\2\end{pmatrix}\right) \\ &= 2L\begin{pmatrix}1\\1\end{pmatrix} + L\begin{pmatrix}-1\\2\end{pmatrix} \\ &= 2\begin{pmatrix}7\\3\end{pmatrix} + \begin{pmatrix}0\\1\end{pmatrix} \\ &= \begin{pmatrix}14\\7\end{pmatrix} \end{aligned}$$

For any matrix, we can associate with it a *linear mapping*.

For any linear mapping, we can associate with it a matrix (usually called *transformation matrix*).

**Proof**

Consider the linear mapping $L : \mathbb{R}^m \to \mathbb{R}^n$. Consider also the standard basis $B$ consisting of the following vectors $E_i$:

$$E_1 = \begin{pmatrix}1\\0\\0\\\vdots\\0\end{pmatrix}, E_2 = \begin{pmatrix}0\\1\\0\\\vdots\\0\end{pmatrix}, \dots, E_m = \begin{pmatrix}0\\0\\0\\\vdots\\1\end{pmatrix}$$

where each vector has $m$ components and, generally, all the components of $E_i$ are zero except for the $i^{th}$ component which is one.

Let us denote by

$$A_1 = L(E_1), A_2 = L(E_2), \dots, A_m = L(E_m) \in \mathbb{R}^n$$

Now, if we consider arbitrary vector $x \in \mathbb{R}^m$ as a linear combination of the vectors in our basis $B$:

$$\underline{x} = x_1E_1 + \cdots + x_mE_m$$

Let us use the definition of being a linear mapping:

$$\begin{aligned} L(\underline{x}) &= L(x_1E_1 + \cdots + x_mE_m) \\ &= x_1L(E_1) + \cdots + x_mL(E_m) \\ &= x_1A_1 + \cdots + x_mA_m \\ L(\underline{x}) &= A\underline{x} \end{aligned}$$

where $A$ is a matrix whose columns are $A_1, A_2, \ldots, A_m$. We found matrix $A$ associated with the linear mapping $L$.
Note that $A \in \mathbb{R}^{n,m}$ and $x \in \mathbb{R}^m$. Note also that $L : \mathbb{R}^m \to \mathbb{R}^n$.

$\square$

**Example**

Consider the linear mapping:

$$L\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} - \text{ projection from } \mathbb{R}^3 \text{ to } \mathbb{R}^2$$

What will be the matrix $A$ associated with $L$?

Let us consider:

$$L(E_1) = L\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = A_1$$

$$L(E_2) = L\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} = A_2$$

$$L(E_3) = L\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} = A_3$$

Now we can construct our matrix $A$ using the vectors $A_1$, $A_2$ and $A_3$ as the column vectors of the matrix $A$:

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

$A$ is therefore the matrix associated with the linear mapping $L$ in the particular basis $B$, which in this case is the standard basis for $\mathbb{R}^3$.

We can then conclude that

$$L(\underline{x}) = A\underline{x}$$

$$L\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

## 7.1 Matrix Associated with Linear Mapping in a Particular Basis

From now on we will focus primarily on linear mappings $L : V \to V$ (from a space $V$ to the same space $V$).

Assume $\underline{b}_1, \ldots, \underline{b}_n$ form a basis $B$ for the space $V$, then any vector $\underline{u} \in V$ can be written as a linear combination of the vectors in the basis $B$:

$$\underline{u} = u_1\underline{b}_1 + \cdots + u_n\underline{b}_n$$

We can call $\begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} \in \mathbb{R}^n$, the coordinates of $\underline{u}$ with respect to basis $B$.

Yes, as we will see later, coordinates of a vector can change depending on the basis in which they are represented!

Now, lets consider the same linear mapping $L : V \to V$ and the same basis $B$ seen above.

*How does the matrix associated with the linear mapping $L$ look with respect to the basis $B$?*

Since $B$ is a basis for the vector space $V$, then we can write:

$$\begin{aligned} L(\underline{b}_1) &= c_{11}\underline{b}_1 + c_{12}\underline{b}_2 + \cdots + c_{1n}\underline{b}_n \\ &\vdots \\ L(\underline{b}_n) &= c_{n1}\underline{b}_1 + c_{n2}\underline{b}_2 + \cdots + c_{nn}\underline{b}_n \end{aligned}$$

which basically means that the transformation $L$ is applied to our basis vectors, and the result is a linear combination of the basis vectors.

Now, if we take an arbitrary vector $\underline{u} \in V$, we can represent it as a linear combination of our basis vectors $\underline{b}_1$, $\underline{b}_2$, ... , $\underline{b}_n$:

$$\underline{u} = u_1\underline{b}_1 + u_2\underline{b}_2 + \cdots + u_n\underline{b}_n = \sum_{i=1}^{n} u_i\underline{b}_i$$

then

$$\begin{aligned} L(\underline{u}) = L\left(\sum_{i=1}^{n} u_i\underline{b}_i\right) &= \sum_{i=1}^{n} u_i L(\underline{b}_i) = \sum_{i=1}^{n} u_i \sum_{j=1}^{n} c_{ij}\underline{b}_j \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} u_i c_{ij}\underline{b}_j = \sum_{j=1}^{n} \underline{b}_j \sum_{i=1}^{n} u_i c_{ij} \\ &= \sum_{i=1}^{n} c_{i1}u_i \times \underline{b}_1 + \sum_{i=1}^{n} c_{i2}u_i \times \underline{b}_2 + \cdots + \sum_{i=1}^{n} c_{in}u_i \times \underline{b}_n \end{aligned}$$

Therefore, we get

$$L(\underline{u}) = \begin{pmatrix} \sum_{i=1}^{n} c_{i1} u_i \\ \vdots \\ \sum_{i=1}^{n} c_{in} u_i \end{pmatrix} = C^T \underline{u}$$

On coordinate vectors our linear mapping is represented by $L(\underline{u}) = C^T \underline{u}$ for a given basis $\underline{b}_1, \ldots, \underline{b}_n$. Note that $C^T$ is the transformation matrix associated with the linear transformation $L$.

### Note

For a different basis we will have different coordinates of vectors as well as different associated matrix.

### Example

Consider a linear transformation $L : \mathbb{R}^3 \to \mathbb{R}^3$ and basis $B$ composed of the following vectors $b_1, b_2, b_3$.

Now, assume that:

$$\begin{aligned} L(b_1) &= b_1 + b_2 \\ L(b_2) &= 5 \cdot b_1 - b_2 + 3 \cdot b_3 \\ L(b_3) &= -b_1 + 4b_2 - 7b_3 \end{aligned}$$

The matrix associated with this linear mapping $L$ with respect to the basis $B$ is:

$$C^T = \begin{pmatrix} 1 & 5 & -1 \\ 1 & -1 & 4 \\ 0 & 3 & -7 \end{pmatrix}$$

Now, let us say we have a vector $\underline{u}$ whose coordinates with respect to the basis $B$ are $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$. To obtain $L(\underline{u})$, we can simply multiply $C^T$ by $\underline{u}$:

$$L \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 & 5 & -1 \\ 1 & -1 & 4 \\ 0 & 3 & -7 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = 1 \cdot b_1 + 1 \cdot b_2 + 0 \cdot b_3$$

Note that the entries of the vector $\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$ represent the coordinates of the transformation of $\underline{u}$, $L(\underline{u})$, with respect to the basis $B$.

## 7.2 Change of Basis (Matrix)

The matrix representations of linear mappings (transformation matrices) are also determined by the chosen basis. Since it is often desirable to *work with more than one basis* for a vector space, it is of fundamental importance in linear algebra to be able to easily *transform coordinate representations of vectors taken with respect to one basis to their equivalent representations with respect to another basis.* Such a transformation is called a *change of basis* (matrix).

Let us first look at how coordinates of vectors change when we change the basis.

Assume we have a vector space $V$. Let us also assume we have basis $B$ for $V$ consisting of the vectors $\underline{b}_1, \underline{b}_2, \ldots, \underline{b}_n$ and another basis $D = \{\underline{d}_1, \underline{d}_2, \ldots, \underline{d}_n\}$.

Consider now a vector $\underline{v} \in V$, and let $\begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}$ be the coordinates of vector $\underline{v}$ with respect to basis $B$:

$$\underline{v} = u_1\underline{b}_1 + \cdots + u_n\underline{b}_n = (u_1, \ldots, u_n) \cdot \begin{pmatrix} \underline{b}_1 \\ \vdots \\ \underline{b}_n \end{pmatrix} = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}^T \begin{pmatrix} \underline{b}_1 \\ \vdots \\ \underline{b}_n \end{pmatrix}$$

Note that $\underline{v}$ is here represented as linear combination of the vectors that form the basis $B$, and $u_1, \ldots u_2$ are the coordinates relative to the basis $B$ (as we said above).

And let $\begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix}$ be the coordinates of $\underline{v}$ with respect to basis $D$:

$$\underline{v} = w_1\underline{d}_1 + \cdots + w_n\underline{d}_n = (w_1, \ldots, w_n) \cdot \begin{pmatrix} \underline{d}_1 \\ \vdots \\ \underline{d}_n \end{pmatrix} = \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix}^T \begin{pmatrix} \underline{d}_1 \\ \vdots \\ \underline{d}_n \end{pmatrix}$$

Since $B$ is a basis for $V$, we can express each vector of the basis $D$ in terms of $\underline{b}_1, \ldots, \underline{b}_n$:

$$\begin{aligned} \underline{d}_1 &= S_{11}\underline{b}_1 + S_{12}\underline{b}_2 + \cdots + S_{1n}\underline{b}_n \\ &\vdots \\ \underline{d}_n &= S_{n1}\underline{b}_1 + S_{n2}\underline{b}_2 + \cdots + S_{nn}\underline{b}_n \end{aligned}$$

Our more compactly:

$$\begin{pmatrix} \underline{d}_1 \\ \vdots \\ \underline{d}_n \end{pmatrix} = S \begin{pmatrix} \underline{b}_1 \\ \vdots \\ \underline{b}_n \end{pmatrix}$$

So $\underline{v}$ can be represented as:

$$\underline{v} = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}^T \begin{pmatrix} \underline{b}_1 \\ \vdots \\ \underline{b}_n \end{pmatrix} = \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix}^T \begin{pmatrix} \underline{d}_1 \\ \vdots \\ \underline{d}_n \end{pmatrix} = \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix}^T S \begin{pmatrix} \underline{b}_1 \\ \vdots \\ \underline{b}_n \end{pmatrix}$$

It follows that:

$$\begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}^T = \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix}^T S$$

If we take the transpose (remembering that $(AB)^T = B^T A^T$ and $(A^T)^T = A$) of both sides, we obtain:

$$\underbrace{\begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}}_{N1} = \underbrace{S^T}_{N2} \underbrace{\begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix}}_{N3}$$

- $N1$: Coordinates of $\underline{v}$ relative to the basis $B$.
- Matrix $S$ describes the vectors $\underline{d}_1, \ldots, \underline{d}_n$ with respect to basis $B = \{\underline{b}_1, \ldots, \underline{b}_n\}$.
- $N3$: Coordinates of $\underline{v}$ relative to the basis $D$.

**Lemma**

$S^T$ is invertible (i.e. $(S^T)^{-1}$ exists). We expressed $\begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}$ as $S^T \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix}$. We could do the same procedure, but exchanging $\underline{b}_1, \ldots, \underline{b}_n$ with $\underline{d}_1, \ldots, \underline{d}_n$ and we would arrive to:

$$\begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} = R^T \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}$$

Now we have

$$\underline{u} = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} = S^T \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} = S^T R^T \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} \Rightarrow S^T R^T = RS = I$$

and

$$\underline{w} = \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} = R^T \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} = R^T S^T \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} \Rightarrow R^T S^T = SR = I$$

It follows that $R^T = (S^T)^{-1}$, so we have found the inverse of $S^T$, which is what we wanted to do. Note also that $R = S^{-1}$. We can also conclude that

$\underline{w} = (S^T)^{-1}\underline{u}$.

Now, consider the linear mapping $L : V \to V$. Assume that $L$ is represented by the *transformation matrix* $A$ with respect to the basis $B = \{\underline{b}_1, \ldots, \underline{b}_n\}$ and by transformation matrix $A'$ relative to the basis $D$ formed by the following vectors $\underline{d}_1, \ldots, \underline{d}_n$. Consider a vector $\underline{u} \in V$.

Then $L(\underline{u})$ is represented with respected to the basis $B$ as:

$$L_B(\underline{u}) = A \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}$$

And with respect to the basis $D$ as:

$$L_D(\underline{u}) = A' \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix}$$

We have:

$$A \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} = S^T A' \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix}$$

We can now use the fact that $\begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} = S^T \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix}$ to conclude:

$$\Rightarrow AS^T \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} = S^T A' \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix}$$

Since $\underline{w} = \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix}$ is an ordinary vector:

$$\Rightarrow AS^T = S^T A$$

We multiply from the left side by $(S^T)^{-1}$ both sides:

$$(S^T)^{-1} AS^T = (S^T)^{-1} S^T A'$$

$$(S^T)^{-1} AS^T = IA'$$

We know just change the positions:

$$\underbrace{A'}_{N1} = (S^T)^{-1} \cdot \underbrace{A}_{N2} \cdot S^T$$

- $N1$: Transformation matrix with respect to the basis $\underline{d}_1, \ldots, \underline{d}_n$
- $N2$: Transformation matrix with respect to the basis $\underline{b}_1, \ldots, \underline{b}_n$

We have just found how the matrix representing a linear mapping changes when we change the basis. In the example above, the transformation matrix changed from $A$ to $A' = (S^T)^{-1}AS^T$, when we changed from the basis $B$ to $D$.

**Definition**

Assume that $N \in \mathbb{R}^{n,n}$, $N^{-1}$ exists. $A' = N^{-1}AN$ is called similarity transformation.

**Definition**

Matrices $A'$ and $A$ are called similar matrices, if $\exists N$ such that

$$A' = N^{-1}AN$$

**Example**

Assume that linear mapping $L$ is represented with matrix

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

With respect to basis $B$:

$$B = \left\{ \underline{b}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \ \underline{b}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$$

Now, consider a new basis $D$:

$$D = \left\{ \underline{d}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \ \underline{d}_2 = \begin{pmatrix} 1 \\ -\frac{1}{2} \end{pmatrix} \right\}$$

*How is L represented with respect to the basis $D$, or in other words what is the transformation matrix with respect to $D$?*

We first represent the vectors in the basis $D$ as a linear combination of the vectors in the basis $B$:

$$\begin{cases} d_1 = 1 \cdot b_1 + 1 \cdot b_2 \\ d_2 = 1 \cdot b_1 - \dfrac{1}{2} \cdot b_2 \end{cases}$$

Then we need to find $S$, $S^T$ and $(S^T)^{-1}$:

$$\Rightarrow S = \begin{pmatrix} 1 & 1 \\ 1 & -\frac{1}{2} \end{pmatrix} \Rightarrow S^T = \begin{pmatrix} 1 & 1 \\ 1 & -\frac{1}{2} \end{pmatrix} \Rightarrow (S^T)^{-1} = -\frac{2}{3} \begin{pmatrix} -\frac{1}{2} & -1 \\ -1 & 1 \end{pmatrix}$$

Then we can find the transformation matrix $A'$ with respect to the new basis $D$ as follows:

$$A' = (S^T)^{-1}AS^T = -\frac{2}{3} \begin{pmatrix} -\frac{1}{2} & -1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & -\frac{1}{2} \end{pmatrix} = \begin{pmatrix} 5 & 0 \\ 0 & -1 \end{pmatrix}$$

Note that, in the new basis $D$, our linear mapping is represented with a simpler transformation matrix $A'$, which can ease the work when doing calculations.

# Chapter 8

# Eigenvalues and Eigenvectors

Consider a vector space $V$ and a linear mapping $A : V \to V$

**Definition**

A vector $\underline{v} \in V$, $\underline{v} \neq 0$ is called an eigenvector of $A$, if there exists scalar $\lambda$, such that $A\underline{v} = \lambda\underline{v}$. This scalar $\lambda$ is called an eigenvalue, corresponding to eigenvector $\underline{v}$.

Sometimes, eigenvectors are called characetristic vectors, and eigenvalues are called characteristic values.

**Example**

Consider

$$A \in \mathbb{R}^{n,n} = \begin{pmatrix} a_{11} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & a_{nn} \end{pmatrix}$$

Then

$$E_i = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} - i$$

is an eigenvector with eigenvalue $a_{ii}$, because

$$AE_i = a_{ii}E_i$$

**Lemma**

Consider $A : V \to V$ and $\underline{v} \neq 0$. An eigenvector with $\lambda$ - eigenvalue. Then, for any scalar $\alpha \neq 0$, $(\alpha\lambda)$ is also an eigenvector with the same eigenvalue $\lambda$

**Proof**

$$A\underline{v} = \lambda\underline{v}$$
$$A(\alpha\underline{v}) = \alpha(\lambda\underline{v}) = \lambda(\alpha\underline{v})$$

□

**Theorem**

Consider the linear mapping $A : V \to V$ and eigenvalue $\lambda$. Assume that there exists $v_1, \dots, v_n$ eigenvectors corresponding to the eigenvalue. THen any vector from the span of $v_1, \dots, v_n$ (any linear combination of $v_m, \dots, v_m$) is also an eigenvector of $A$, with the same eigenvalue $\lambda$

**Proof**

Take any linear combination of $v_1, \dots, v_m$

$$\begin{aligned}
&\alpha_1 v_1 + \dots + \alpha_m v_m \\
&A(\alpha_1 v_1 + \dots + \alpha_m v_m) = \alpha_1 A v_1 + \dots + \alpha_m A v_m \\
=&\alpha_1 \lambda v_1 + \dots + \alpha_m \lambda v_m = \lambda(\alpha_1 v_1 + \dots + \alpha_m v_m) \\
\Rightarrow&\alpha_1 v_1 + \dots + \alpha_m v_m \Rightarrow \text{ is indeed an eigenvector with eigenvalue } \lambda
\end{aligned}$$

□

**Remark:**

It means that the span of $v_1, \dots, v_m$ forms a subspace in $V$ and any non - zero vector from this subspace is an eigenvector of $A$ with eigenvalue $\lambda$

This subspace is called an eigenspace of $A$ with eigenvalue $\lambda$

**Theorem**

Consider $A : V \to V$ - linear mapping. Assume that there exists eigenvectors $v_1, \dots, v_m$ with corresponding eigenvalues $\lambda_1, \dots, \lambda_m$. Let us also assume that all eigenvalues are distinct, $\lambda_i \neq \lambda_j$ for $i \neq j$. Then $v_1, \dots, v_m$ are linearly independent

**Proof**

By induction on $m$.

- $m = 1$:
  $v_i$ - eigenvector, $\lambda_1$ - eigenvalue. By definition, $v_1 \neq 0$, therefore $v_1$ is linearly independent.
- $m > 1$:
  Assume that the theorem holds for any $n - 1$ eigenvector and eigenvalue. Let us assume $v_1, \dots, v_m$ are linearly dependent.

$$\alpha_1 v_1 + \dots + \alpha_m v_m \qquad (*)$$

Let us multiply $(*)$ by $\lambda_m$:

$$\alpha_1\lambda_m v_1 + \cdots + \alpha_m\lambda_m v_m = 0$$

let us apply $A$ to $(*)$:

$$\begin{aligned} A(\alpha_1 v_1 + \cdots + \alpha_m v_m) &= \alpha_1 A v_1 + \cdots + \alpha_m A v_m \\ &= \alpha_1\lambda_1 v_1 + \cdots + \alpha_m\lambda_m v_m = 0 \end{aligned}$$

Subtract $1^{st}$ from $2^{nd}$:

$$\alpha_1(\lambda_1 - \lambda_m)v_1 + \cdots + \alpha_{m-1}(\lambda_{m-1} - \lambda_{m-1})v_{m-1} = 0$$

(There are $m-1$ eigenvectors)

$\Rightarrow$ They are linearly independent by the induction hypothesis

$\Rightarrow$ $\alpha_1 \underbrace{(\lambda_1 - \lambda_m)}_{\neq 0} = 0, \ldots, \alpha_{m-1}\underbrace{(\lambda_{m-1} - \lambda_m)}_{\neq 0} = 0$

Since $\lambda_i \neq \lambda_j \Rightarrow \alpha_1 = 0, \ldots, \alpha_{m-1} = 0$ and then from $(*) \Rightarrow \alpha_m = 0$, since $\underline{v}_m \neq 0$

$\Rightarrow$ $v_1, \ldots, v_m$ are linearly independent

□

**Remark:**

If $A : V \to V$ is a linear mapping and $V$ is a $n-$ dimensional space. If we have $v_1, \ldots, v_n$ eigenvectors of $A$ with all distinct $\lambda_1, \ldots, \lambda_n$, then $v_1, \ldots, v_n$ from a basis of $V$.

## 8.1 Characteristic Polynomial

How to find eigenvalues and eigenvectors?

**Recall**

1. Let us consider the linear mapping $A : V \to V$. $A$ is invertible $\Leftrightarrow$ the nullspace of $A$ is $\{\underline{0}\}$

$$N(A) = \{\underline{x} \in V, A\underline{x} = \underline{0}\}$$

2. $A$ is invertible $\Leftrightarrow \det(A) \neq 0$

**Theorem**

Let us consider $A \in \mathbb{R}^{n,n}$. $\lambda$ is an eigenvalue of $A$ iff $(\lambda I - A)$ is not invertible.

**Proof**

$\lambda$ is an eigenvalue of $A$, $\exists \underline{v} \neq \underline{0}$ such that

$$A\underline{v} = \lambda\underline{v} \Rightarrow -\lambda\underline{v} + A\underline{v} = -(\lambda I - A)\underline{v} = 0$$

therefore $\lambda I - A$ has a non - zero vector $\underline{v}$ in its null space, and thanks to 1. $\lambda I - A$ is not invertible

$\square$

**Definition**

Consider $A \in \mathbb{R}^{n,n}$. The characteristic polynomial is defined as

$$p_a(t) = \det(tI - A)$$

**Example**

$$A = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}, p_a(t) = ?$$

$$\begin{aligned} p_a(t) &= \left| \begin{pmatrix} t-1 & 0 & 2 \\ 0 & t-1 & 1 \\ 1 & 0 & t-1 \end{pmatrix} \right| \\ &= (t-1)^3 - 2(t-1) \\ &= (t-1)((t-1)^2 - 2) \\ &= (t-1)(t^2 - 2t - 1) \end{aligned}$$

**Theorem**

Consider $A \in \mathbb{R}^{n,n}$. $\lambda$ is an eigenvalue of $A$ iff $\lambda$ is a root of the characteristic polynomial $p_a(t)$

**Proof**

$\Rightarrow$: $\lambda$ is an eigenvalue of $A$. Then from previous theorem $(\lambda I - A)$ is non invertible.
$\det(\lambda I - A) = 0$ but $\det(\lambda I - A) = p_a(t = \lambda) \Rightarrow \lambda$ is a root of $p_a(t)$

$\Leftarrow$: if $\lambda$ is a root of $p_a(t)$, then $p_a(\lambda) = \det(\lambda I - A) = 0$
$\lambda I - A$ is not invertible, from previous theorem. $\lambda$ is an eigenvalue.

$\square$

**Example**

Consider

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

find the Eigenvalues and Eigenvectors.

$$\begin{aligned} p_a(t) = \det(tI - A) = \begin{vmatrix} t-1 & -4 \\ -2 & t-3 \end{vmatrix} &= (t-1)(t-3) - 8 \\ &= t^2 - 4t - 5 \\ &= (t+1)(t-5) \end{aligned}$$

The two roots of $p_a(t)$ are $\lambda_1 = -1, \lambda_2 = 5$, which are also the eigenvalues. We can now find the eigenvector for $\lambda_1 = -1$:

$$\begin{aligned} \begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = -1 \cdot \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} &\Rightarrow \begin{cases} v_1 + 4v_2 = -v_1 \\ 2v_1 + 3v_2 = -v_2 \end{cases} \\ &\Rightarrow \begin{cases} 2v_1 + 4v_2 = 0 \\ 2v_1 + 4v_2 = 0 \end{cases} \\ &\Rightarrow \begin{cases} v_1 + 2v_2 = 0 \\ 0 = 0 \end{cases} \\ &\Rightarrow v_1 = 1, v_2 = -\frac{1}{2} \\ &\Rightarrow \text{ Eigenvector is } \begin{pmatrix} 1 \\ -\frac{1}{2} \end{pmatrix} \end{aligned}$$

and the eigenvector for $\lambda_2 = 5$:

$$\begin{aligned} \begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = 5 \cdot \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} &\Rightarrow \begin{cases} -4v_1 + 4v_2 = 0 \\ 2v_1 - 2v_2 = 0 \end{cases} \\ &\Rightarrow \begin{cases} -v_1 + v_2 = 0 \\ 0 = 0 \end{cases} \\ &\Rightarrow v_1 = 1, v_2 = 1 \\ &\Rightarrow \text{ Eigenvector is } \begin{pmatrix} 1 \\ 1 \end{pmatrix} \end{aligned}$$

**Example**

Consider

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

What are the eigenvalues and eigenvectors?

$$p_a(t) = \det(tI - A) = \begin{vmatrix} t-2 & -1 & 0 \\ 0 & t-1 & 1 \\ 0 & -2 & t-4 \end{vmatrix} = (t-2)^2(t-3)$$

Then the eigenvalues are $\lambda_1 = 3, \lambda_2 = 2, \lambda_3 = 2$. Eigenvector for $\lambda_1 = 3$:

$$
\begin{aligned}
A\underline{x} = \lambda_1 \underline{x} &\Rightarrow \begin{cases} 2x_1 + x_2 = 3x_1 \\ x_2 - x_3 = 3x_2 \\ 2x_2 + 4x_3 = 3x_3 \end{cases} \\
&\Rightarrow \begin{cases} -x_1 + x_2 = 0 \\ -2x_2 - x_3 = 0 \\ 2x_2 + x_3 = 0 \end{cases} \\
&\Rightarrow \begin{cases} -x_1 + x_2 = 0 \\ 2x_2 + x_3 = 0 \\ 0 = 0 \end{cases} \\
&\Rightarrow \text{Let us choose } x_1 = 1, \text{ then } x_2 = 1, x_3 = -2 \\
&\Rightarrow \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}
\end{aligned}
$$

Eigenvector for $\lambda_{2,3} = 2$:

$$
\begin{aligned}
A\underline{x} = 2\underline{x} &\Rightarrow \begin{cases} 2x_1 + x_2 = 2x_1 \\ x_2 - x_3 = 2x_2 \\ 2x_2 + 4x_3 = 2x_3 \end{cases} \\
&\Rightarrow \begin{cases} x_2 = 0 \\ -x_2 - x_3 = 0 \\ 2x_2 + 2x_3 = 0 \end{cases} \\
&\Rightarrow \begin{cases} x_2 = 0 \\ x_3 = 0 \\ 0 = 0 \end{cases} \\
&\Rightarrow \text{We can choose } x_1 = 1, \text{ then } x_2 = 0, x_3 = 0 \\
&\Rightarrow \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}
\end{aligned}
$$

**Example**

Consider

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

What are the eigenvalues and eigenvectors?

$$p_a(t) = \det(tI - A) = \begin{vmatrix} t-3 & 0 & 0 \\ 0 & t-2 & 0 \\ 0 & 0 & t-2 \end{vmatrix} = (t-2)^2(t-3)$$

Then the eigenvalues are $\lambda_1 = 3, \lambda_2 = 2, \lambda_3 = 2$. Eigenvector for $\lambda_{2,3} = 2$:

$$A\underline{x} = 2\underline{x} \Rightarrow \begin{cases} 3x_1 = 2x_1 \\ 2x_2 = 2x_2 \\ 2x_3 = 2x_3 \end{cases} \Rightarrow \begin{cases} x_1 = 0 \\ 0 = 0 \\ 0 = 0 \end{cases}$$

We have three equations, one variable is determined ($x_1 = 0$) and we have two independent variables ($x_2, x_3$) which we can choose arbitrarily. Therefore:

$$x_2 = 1, x_3 = 0 \Rightarrow \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

$$x_2 = 0, x_3 = 1 \Rightarrow \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

We can get two linearly independent eigenvectors. Each eigenvalue can have 0 or 1 corresponding eigenvectors. $k-$eigenvalues$\rightarrow k$ corresponding eigenvectors at most.

# Chapter 9

# Change of Basis

Old basis $\underline{b}_1, \dots, \underline{b}_n$, new basis $\underline{d}_1, \dots, \underline{d}_n$

$$\begin{pmatrix} \underline{d}_1 \\ \vdots \\ \underline{d}_n \end{pmatrix} = S \begin{pmatrix} \underline{b}_1 \\ \vdots \\ \underline{b}_n \end{pmatrix}$$

If $\underline{v}$ are the coordinates of a vector in the old basis $b_1, \dots, b_n$. $v' = \left(S^T\right)^{-1} \underline{v}$ are the coordinates of the same vector in the new basis. If $A$ is a matrix in the old basis, $A' = \left(S^T\right)^{-1} AS^T$ is the same matrix in the new basis.

**Theorem**

The characteristic polynomial of $\left(S^T\right)^{-1} AS^T$ is the same as of $A$

**Proof**

$$\begin{aligned} \det\left(tI - \left(S^T\right)^{-1} AS^T\right) &= \det\left(t\left(S^T\right)^{-1} IS^T - \left(S^T\right)^{-1} AS^T\right) \\ &= \det\left(\left(S^T\right)^{-1}\right) \det\left(tI - A\right) \det\left(S^T\right) \\ &= \det\left(tI - A\right) \\ &\Rightarrow \det\left(B^{-1}\right) \cdot \det\left(B\right) = 1, \text{ if } B^{-1} \text{ exists} \end{aligned}$$

□

It means that the eigenvalues do not change when we change the basis. Let us assume $A\underline{v} = \lambda \underline{v}$:

$$\underbrace{\left(S^T\right)^{-1} AS^T}_{A'} \cdot \underbrace{\left(S^T\right)^{-1} \underline{v}}_{v'} = \left(S^T\right)^{-1} A\underline{v} = \left(S^T\right)^{-1} \lambda v = \lambda \underbrace{\left(S^T\right)^{-1} v}_{v'}$$

$A'v' = \lambda v'-$ in the new basis it means that the eigenvectors of linear mapping do not change, when we change the basis, only coordinates change.

**Definition**

A sot of all eigenvalues of matrix $A \in \mathbb{R}^{n,n}$ is called spectrum of $A$

Let us consider $A \in \mathbb{R}^{n,n}$. Let us assume that $A$ has $\lambda_1, \ldots, \lambda_n$ eigenvalues and linearly independent eigenvectors $\underline{s}_1, \ldots, \underline{s}_n$.

If we consider $\underline{b}_1, \ldots, \underline{b}_n$ (old basis) to be a standard basis $\underline{E}_1, \ldots, \underline{E}_n$ and $\underline{s}_1, \ldots, \underline{s}_n$ as a new basis. Then

$$\begin{pmatrix} \underline{s}_1 \\ \vdots \\ \underline{s}_n \end{pmatrix} = S \begin{pmatrix} \underline{b}_1 \\ \vdots \\ \underline{b}_n \end{pmatrix}, S = \begin{pmatrix} -s_1^t \rightarrow \\ \vdots \\ -s_n^t \rightarrow \end{pmatrix}$$

$A$ in the new basis, $A' = \left(S^T\right)^{-1} AS^T$

$$A\underline{s}_1 = \lambda_1 \underline{s}_1, A\underline{s}_2 = \lambda_2 \underline{s}_2, \ldots, A\underline{s}_n = \lambda_n \underline{s}_n$$

$$A \begin{pmatrix} | & & | \\ s_1 & \ldots & s_n \\ \downarrow & & \downarrow \end{pmatrix} = \underbrace{\begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}}_{\Lambda-\text{ diagonal matrix}} \begin{pmatrix} | & & | \\ s_1 & \ldots & s_n \\ \downarrow & & \downarrow \end{pmatrix}$$

$$AS^T = \Lambda S^T \quad \Big(\text{multiply by } \left(S^T\right)^{-1} \text{ from the left}\Big)$$

$$\underbrace{\left(S^T\right)^{-1} AS^T}_{A'} = \Lambda$$

If there exists $n$ linearly independent eigenvectors of $A$, the $A$ can be brought to a diagonal by changing the basis.