Eidgenössische Technische Hochschule Zürich

Linear Algebra I

Autumn Semester 2025

Lecture Notes

Contents

Lineare Algebra I

Fynn Krebser–fkrebser@student.ethz.ch

December 5, 2025

1 Introduction

Lec 1

1.1 Fibonacci Sequences

Definition 1.1:

The fibonacci sequence is defined as follows:

$$a_n = \begin{cases} 0, & n = 0 \\ 1, & n = 1 \\ a_{n-1} + a_{n-2}, & n \geq 2 \end{cases}.$$

The first few terms of the sequence are: 0, 1, 1, 2, 3, 5, 8, 13, ...

Today we try to find a explicit formula for a_n.

From high school we know both the arithmetic and geometric sequences:

Definition 1.2: Arithmetic Sequence

An arithmetic sequence is defined as follows:

$$a_n = \begin{cases} a, & n = 0 \\ a_{n-1} + d, & n \geq 1 \end{cases}.$$

For an arithmetic sequence we have the explicit formula $a_n = a + nd$.

Definition 1.3: Geometric Sequence

A geometric sequence is defined as follows:

$$a_n = \begin{cases} a, & n = 0 \\ a_{n-1} \cdot q, & n \geq 1 \end{cases}.$$

In this case, we aswell have an explicit formula: $a_n = a \cdot q^n$.

However looking at the fibonacci sequence, we see that it is diffrent from theese two. So we will do a mathematical trick to solve it. This trick is generalizing.

Definition 1.4:

Let $a_0, a_1 \in \mathbb{R}$ we define $a_n := a_{n-1} + a_{n-2}$ for $n \geq 2$.

Denote the sequence we get $\mathcal{F}_{a_0,a_1}$.

If $S = (s_0, s_1, \cdots)$ is any sequence of numbers, we say that S is Fibonacci if there exists $a_0, a_1 \in \mathbb{R}$ such that $S = \mathcal{F}_{a_0,a_1}$. Denote by `Fib` the set of all Fibonacci sequences.

`Fib` has some algebraic structures.

Claim 1.5:

If $\mathcal{F}' = (a_0, a_1, \cdots)$ and $\mathcal{F}'' = (b_0, b_1, \cdots)$ are Fibonacci then

$$\mathcal{F}' + \mathcal{F}'' := (a_0 + b_0, a_1 + b_1, \cdots, a_n + b_n, \cdots)$$

is also Fibonacci.

Proof. Lets write $c_0 := a_0 + b_0, c_1 := a_1 + b_1$ and $c_n := a_n + b_n$. Then $\mathcal{F}' + \mathcal{F}'' = (c_0, c_1, \cdots)$. To show this is a fibonacci sequence we need to show that $c_n = c_{n-1} + c_{n-2}$ for $n \geq 2$.

Indeed,

$$\begin{aligned} c_{n-1} + c_{n-2} &= (a_{n-1} + b_{n-1}) + (a_{n-2} + b_{n-2}) \\ &= (a_{n-1} + a_{n-2}) + (b_{n-1} + b_{n-2}) \\ &= a_n + b_n \\ &= c_n. \end{aligned}$$

□

This is somewhat special, as not every set of sequences is closed under addition. In fact, we showed that

$$\mathcal{F}_{a_0,a_1} + \mathcal{F}_{b_0,b_1} = \mathcal{F}_{a_0+b_0,a_1+b_1}.$$

But there is more structure.

Claim 1.6:

Let $\mathcal{A} = (a_0, a_1, \cdots)$ be a Fibonacci sequence and let $\alpha \in \mathbb{R}$. Define $\alpha\mathcal{A} := (\alpha a_0, \alpha a_1, \cdots)$. Then $\alpha\mathcal{A}$ is also Fibonacci. In fact, $\alpha\mathcal{F}_{a_0,a_1} = \mathcal{F}_{\alpha a_0,\alpha a_1}$.

Proof. We need to check that $\alpha a_n = \alpha a_{n-1} + \alpha a_{n-2}$ for $n \geq 2$. But this is "obviously" true, because we know that $a_n = a_{n-1} + a_{n-2}$. □

Note: The sequence $(0, 0, \cdots)$ is also Fibonacci, it is actually the sequence $\mathcal{F}_{0,0}$.

In fact, LINEAR COMBINATIONS of Fibonaccis are also Fibonacci.

$$\alpha\mathcal{F}_{a_0,a_1} + \beta\mathcal{F}_{b_0,b_1} = \mathcal{F}_{\alpha a_0+\beta b_0,\alpha a_1+\beta b_1},$$

for $\alpha, \beta \in \mathbb{R}$.

Corollary 1.7:

In order to find a formula for $\mathcal{F}_{a_0,a_1}$ it is enough to find a formula for $\mathcal{F}_{0,1}$ and $\mathcal{F}_{1,0}$.

Proof. This is because $\mathcal{F}_{a_0,a_1} = a_0\mathcal{F}_{1,0} + a_1\mathcal{F}_{0,1}$.

□

Out of curiosity, could it be that for some, very special $a_0, a_1 \neq (0, 0)$, the sequence $\mathcal{F}_{a_0,a_1}$ is arithmetic or geometric?

Since arithmetic sequences are characterized by a constant difference $a_n = a_{n-1} + d$, we would need $a_{n-1} + a_{n-2} = a_{n-1} + d$ for $n \geq 2$. This means that the sequence would need to be constant, i.e. $a_n = d$ for all $n \geq 1$. This cannot happen for any a_0, a_1 except $(0, 0)$.

Perhaps it could be geometric? Let's try to see if $(1, q, \cdots)$ is Fibonacci for some $q \in \mathbb{R}$.

For such a sequence to be Fibonacci we need

$$q^n = q^{n-1} + q^{n-2} \text{ for } n \geq 2.$$

As $q \neq 0$ we get $q^2 = q + 1$. This is a quadratic equation with solutions

$$q = \frac{1 \pm \sqrt{5}}{2}.$$

Denote $\varphi := \frac{1+\sqrt{5}}{2}$ and $\psi := \frac{1-\sqrt{5}}{2}$.

We have found

$$\mathcal{F}_{1,\varphi} = (1, \varphi, \varphi^2, \varphi^3, \cdots)$$
$$\mathcal{F}_{1,\psi} = (1, \psi, \psi^2, \psi^3, \cdots)$$

Perhaps a linear combination of $\mathcal{F}_{1,\varphi}$ and $\mathcal{F}_{1,\psi}$ will give us $\mathcal{F}_{0,1}$?

We need to find $\alpha, \beta \in \mathbb{R}$ such that

$$\begin{cases} \alpha \cdot 1 + \beta \cdot 1 = 0 \\ \alpha\varphi + \beta\psi = 1 \end{cases}.$$

Solving this system of equations gives

$$\alpha = \frac{1}{\varphi - \psi} = \frac{1}{\sqrt{5}}, \quad \beta = \frac{1}{\psi - \varphi} = -\frac{1}{\sqrt{5}}.$$

So we found

$$\mathcal{F}_{0,1} = \frac{1}{\sqrt{5}}\mathcal{F}_{1,\varphi} - \frac{1}{\sqrt{5}}\mathcal{F}_{1,\psi}$$
$$= \frac{1}{\sqrt{5}}\left(\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n\right).$$

Lec 2

1.2 Logic and Math language

What is a mathematical/logical statement? A statement that can be either true or false, but not both.

> **Example 1.8:**
>
> A: "For every real number x, we have $x^2 \geq 0$".
>
> B: "Every prime number $p \geq 3$ must be odd".
>
> C: "Every odd number $p \geq 3$ must be prime".
>
> A is True, B is True, C is False.

Negation of a statement: If A is a statement, then the statement "A is NOT ture" is called the negation of A, denoted by $\neg A$ ("not A").[1]

Such a statement can be represented in a truth table:

A	$\neg A$
T	F
F	T

[1] Sometimes we write $\bar{A}$ instead of $\neg A$.

> **Definition 1.9: And operation**
>
> Let A, B be statements. Then the statement $A \wedge B$ is the statement "A is true AND B is true". Sometimes we write $A \& B$ instead.

The corresponding truth table is:

A	B	$A \wedge B$
T	T	T
T	F	F
F	T	F
F	F	F

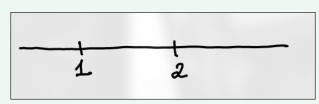
> **Example 1.10:**
>
> For real numbers x, ("x^2=1") $\wedge$ ("$x \geq 0$") is true. This is equivalent to "$x = 1$".

Similarly we can define the "or" operation.

> **Definition 1.11: Or operation**
>
> Let A, B be statements. Then the statement $A \vee B$ is the statement "A is true OR B is true", i.e. "at least one of the statements A or B are true". Sometimes we write $A|B$ instead.

Thus the corresponding truth table is:

A	B	$A \vee B$
T	T	T
T	F	T
F	T	T
F	F	F

> **Example 1.12:**
>
> ("$x > 1$") $\vee$ ("$x < 2$"). We can draw a number line to visualize this situation:
>
> 
>
>
> We see that this statement is true for all $x \in \mathbb{R}$.

> **Example 1.13:**
>
> Consider the statement ("$x < 1$") $\vee$ ("$x > 2$"). Again looking at the number line, we see that this statement is equivalent to "$\neg(1 \leq x \leq 2)$".

> **Definition 1.14: Logical Implications**
>
> Let A, B be statements. We can form a new statement $A \Rightarrow B$ (A implies B). The statement "$A \Rightarrow B$" is "If A is true, then B is also true".

Depicted in a truth table:

A	B	$A \Rightarrow B$
T	T	T
T	F	F
F	T	T
F	F	T

As an exercise: Convice yourself that the statement $A \Rightarrow B$ is equivalent to $(\neg A \vee B)$.

Example 1.15:
("$0 = 1$") $\Rightarrow$ ("Earth is Flat") is a true statement. Notice that in logical terms, implies does not mean because A is true, B is true, no matter if there is no causality.

Similarly there is logical equivalence:

Definition 1.16: Logical Equivalence
$A \Leftrightarrow B$ is the statement $(A \Rightarrow B) \wedge (B \Rightarrow A)$. In diffrent words: A holds if and only if (iff) B holds.

We say that A and B are EQUIVALENT STATEMENTS if $A \Leftrightarrow B$ is true. This happens if A and B have the same truth table.

Example 1.17:

$$x^2 > 0 \Leftrightarrow x \neq 0,$$

$$0 = 1 \Leftrightarrow \text{Earth is flat.}$$

Let A, B be statements. Then there is a statement $A \Rightarrow B$ and $\neg B \Rightarrow \neg A$. These two statements are actually equivalent. This can be shown by truth tables.

As an exercise, state in words, why

$$\neg(\neg A) = A$$

and

$$\neg(A \wedge B) = (\neg A) \vee (\neg B)$$

.

Definition 1.18: Predicate
A PREDICATE is a statement involving some variables taken from a set.

$$P(x), P(x, y).$$

For example, $P(n) =$"$n = n^2$".

"for all integers $n : n = n^2$". This is definelty a false statement. Meanwhile, "There exists an integer $n : n = n^2$" is a true statement.

We denote $\forall$ for "for all/all" and $\exists$ for "there exists". The previous statements can be written as:

$$\forall n \in \mathbb{Z}, P(n) \quad , \quad \exists n \in \mathbb{Z}, P(n).$$

The symbols $\forall$ and $\exists$ are called QUANTIFIERS.

$\forall n \in \mathbb{Z} : n = n^2 \Rightarrow (n = 0) \vee (n = 1)$ is a true statement.
$\forall y \in \mathbb{R} : (y \geq 0) \Rightarrow (\exists x \in \mathbb{R} : x^2 = y)$ is true.

In practice, the second statement would be written as:

$$\forall 0 \leq y \in \mathbb{R}, \exists x \in \mathbb{R} s.t. x^2 = y.$$

Be careful about the order of quantifiers:

$$\forall x \in X, \exists y \in Y : A(x, y),$$

is not the same as

$$\exists y \in Y, \forall x \in X : A(x, y).$$

Example 1.19:
Let X be the set of students at ETH and Y the set of all courses given in HS25. Let $A(x, y)$ be the statement "student x attends course y". Then the first statement means "For every student, there is at least one course that the student attends". Which is true, whilst the second statement means "There is a course that every student attends", which is false.

In practice, instead of writing $\exists x \in X, \exists y \in X$, one rather writes $\exists x, y \in X$, similarly for $\forall$.

When negating quantifiers, we have the following rules:

$$\neg(\forall x \in X : A(x)) \Leftrightarrow \exists x \in X : \neq A(x).$$

$$\neg(\exists x \in X : A(x)) \Leftrightarrow \forall x \in X : \neg A(x).$$

There is also a quantifier "there exists a unique" denoted by $\exists!$. Lec 3

$$\exists! x \in X : A(x).$$

Example 1.20:
"$\exists! x \in \mathbb{R} : x^2 = 4$". This is a false statement as there exist two such x (namely 2 and -2).

"$\exists! x \in \mathbb{R} : x \geq 0 \wedge x^2 = 4$". This is a true statement as there exists exactly one such x (namely 2).

1.3 Set theory

The fundamental object in set theory is the set.

Definition 1.21: Set
A SET is a collection of different objects. These objects are called the ELEMENTS of the set.

For example, $M = \{1, 3, \text{unicorn}\} = \{3, 1, \text{unicorn}\}$ is a set with three elements. As we can see, the order of elements does not matter. Furthermore, repetition of elements does not matter, i.e. $\{1, 1, 1\} = \{1\}$.

If x is an element of a set M, we write $x \in M$. If it is not, we write $x \notin M$.

Another way to write sets is by specifying a property that the elements satisfy.

$$M = \{x : A(x)\} = \{x \mid A(x)\},$$

represents the set of all x such that $A(x)$ holds.

> **Example 1.22:**
>
> $$\{x \mid x \in \mathbb{Z} \wedge x^2 \leq 9\} = \{-3, -2, -1, 0, 1, 2, 3\}.$$

Another important set is the empty set $\{\}$, which contains no elements. It is denoted by $\varnothing$.

A set is allowed to contain other sets as elements one of its elements. For example $M = \{1, \{2, 3\}\}$ is a set with two elements, 1 and $\{2, 3\}$. Similarly, $\{\varnothing\}$ is a set with one element, the empty set.

> **Definition 1.23:**
>
> Let P, Q be sets.
>
> 1. We say that P is a SUBSET of Q if $x \in P \Rightarrow x \in Q$ We write $P \subseteq Q$ (Also: $P \subset Q$).
> 2. $P \subsetneq Q$ means $P \subseteq Q$ but $P \neq Q$.
> 3. $P \nsubseteq Q$ means $\neg(P \subseteq Q)$.

Careful: Every set is a subset of itself. Furthermore, the empty set $\varnothing \subseteq M$ for every set M.

We define the following operations for sets.

$$\begin{aligned}
P \cap Q &:= \{x : x \in P \wedge x \in Q\} && \text{(intersection)}\\
P \cup Q &:= \{x : x \in P \vee x \in Q\} && \text{(union)}\\
P \setminus Q &:= \{x : x \in P \wedge x \notin Q\} && \text{(complement of } Q \text{ in } P)\\
P \Delta Q &:= (P \cup Q) \setminus (P \cap Q) && \text{(symmetric difference)}
\end{aligned}$$

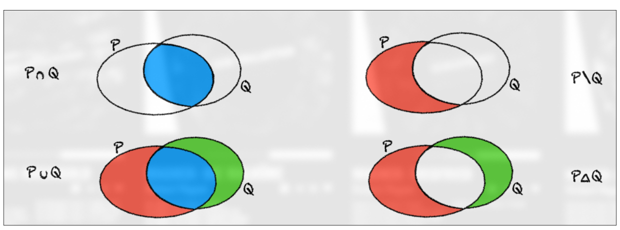


Figure 1: Set operations

Sometimes our sets will be subsets of some ambient/underlying set X.

In this case we define the COMPLEMENT of P in X as

$$P^c := X \setminus P = \{x \in X : x \notin P\}.$$

> **Definition 1.24:**
>
> Let $\mathcal{A}$ be a family of sets, $\mathcal{A} \neq \varnothing$. We can look at the union of the members of $\mathcal{A}$:
>
> $$\bigcup_{A \in \mathcal{A}} A := \{x \mid \exists A \in \mathcal{A} \text{s.t.} x \in A\}.$$
>
> Similarly, we can look at the intersection of the members of $\mathcal{A}$:
>
> $$\bigcap_{A \in \mathcal{A}} A := \{x \mid \forall A \in \mathcal{A} : x \in A\}.$$

When working with big unions and intersections, it is often useful to translate them into set statements involving quantifiers, as this greatly improves intuition.

> **Example 1.25:**
>
> Let $\mathcal{A} = \{2, 3, 4, \cdots\}$.
>
> Let $A_n = \{x \mid x \in \mathbb{Z} x \geq 1, n^2 | x\}$.
>
> Then
>
> $$\bigcup_{n \in \mathcal{A}} A_n = \{a \in \mathbb{Z} \mid \exists k \geq 2, \exists r \in \mathbb{Z} \text{s.t.} r \geq 1 \wedge a = k^2 r\}$$
>
> and
>
> $$\bigcap_{n \in \mathcal{A}} A_n = \{\}.$$

> **Definition 1.26: Cartesian Product**
>
> Let X, Y be sets. The CARTESIAN PRODUCT of X and Y is the set of ordered pairs (x, y).
>
> $$X \times Y := \{(x, y) \mid x \in X, y \in Y\}.$$

> **Definition 1.27: n-Fold Cartesian Product**
>
> Let X be a set, $n \geq 1$. The n-FOLD CARTESIAN PRODUCT of X is defined as
>
> $$X^n := \underbrace{X \times X \times \cdots \times X}_{n \text{ times}}.$$
>
> An element of $X^n = (x_1, x_2, \cdots, x_n)$ is called an N-TUPLE.

The classic example is $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$ – the Cartesian plane.

> **Definition 1.28: Power Set**
>
> Let X be a set. The POWER SET of X is the set of all subsets of X.
>
> $$\mathcal{P}(X) := \{A \mid A \subseteq X\}.$$
>
> Sometimes the notation 2^X is also used.

> **Example 1.29:**
>
> Let $X = \{1, 2, 3\}$. Then
>
> $$\begin{aligned}\mathcal{P}(X) = \{&\varnothing, \{1\}, \{2\}, \{3\},\\ &\{1, 2\}, \{1, 3\}, \{2, 3\},\\ &\{1, 2, 3\}\}\end{aligned}$$
>
> Notice that if X has n elements, then $\mathcal{P}(X)$ has 2^n elements.

The CARDINALITY of a set X is the number of elements in X. For example, if $X = \{1, 2, \cdots, n\}$, then $|X| = n$. This operation is only defined for finite sets at the moment.

1.4 Functions/Maps

Definition 1.30: Maps

Let X, Y be sets. A MAP $f : X \to Y$ is an assignment to *every* element $x \in X$ a uniquely defined element $f(x) \in Y$.

Lec 4

Roughly speaking, for an input $x \in X$, the map f produces an output output $f(x) \in Y$.

Given a map $f : X \to Y$, The set of inputs of a map is called the DOMAIN OF DEFINITION of f, whilst Y is called the TARGET of f.

Given two functions $f : X \to Y$ and $g : X' \to Y'$, we say that f and g are equal if $X = X'$, $Y = Y'$ and $\forall x \in X : f(x) = g(x)$.

Example 1.31:

$$f(x) := x^2.$$

This is lacking some information. Much better would be

$$X := \{x \in \mathbb{R} \mid 0 \le x\}, \quad f : X \to \mathbb{R}, \quad f(x) := x^2.$$

Here is another function:

$$Z := \{x \in \mathbb{R} \mid -12 \le x\}, \quad g : Z \to \mathbb{R}, \quad g(x) := x^2.$$

According to our definition, $f \neq g$ since $X \neq Z$.

Functions can be described by a formula, but do not necessarily have to be.

About notation: $f : X \to Y$ is the most standard notation. Sometimes $X \ni x \to f(x) \in Y$ is used.

Definition 1.32:

Let $f : X \to Y$ be a map. The IMAGE of f is the set

$$f(X) := \{y \in Y \mid \exists x \in X : f(x) = y\}.$$

Sometimes it's also written as $\text{image}(f)$. It holds that $f(X) \subseteq Y$.

Definition 1.33: Restriction

Let $f : X \to Y$, Let $A \subseteq X$. We can define a new map

$$f|_A : A \to Y, \quad f|_A(a) = f(a) \quad \forall a \in A$$

called the RESTRICTION of f to A.

A restriction restricts the domain of definition of a function.

Example 1.34: Characteristic Function

Let X be a set, $A \subseteq X$. The CHARACTERISTIC FUNCTION of A is defined as

$$\mathbb{1}_A : X \to \{0,1\}, \quad \mathbb{1}_A(x) := \begin{cases} 1 & x \in A \\ 0 & x \notin A \end{cases}.$$

Example 1.35: Projection Function

Let X, Y be sets. The PROJECTION FUNCTION on X is defined as

$$\pi_X : X \times Y \to X, \quad \pi_X(x, y) := x.$$

Similarly, we can define $\pi_Y : Y \times X \to Y$.

Functions can have diffrent properties.

Definition 1.36:

Let $f : X \to Y$ be a map.

- f is called INJECTIVE if

  $$\forall x_1, x_2 \in X : f(x_1) = f(x_2) \Rightarrow x_1 = x_2.$$

  Equivalently, if $x_1 \neq x_2$, then $f(x_1) \neq f(x_2)$.
- f is called SURJECTIVE if

  $$\forall y \in Y, \exists x \in X : f(x) = y.$$

  Equivalently, $f(X) = Y$.
- f is called BIJECTIVE if it is both injective and surjective. Equivalently,

  $$\forall y \in Y, \exists! x \in X : f(x) = y.$$

Definition 1.37:

Let $f : X \to Y$ be a bijection. The INVERSE map $f^{-1} : Y \to X$ is the map that assigns to every $y \in Y$ the unique $x \in X$ such that $f(x) = y$.

An example would be the map $f : \mathbb{R}_{\ge 0} \to \mathbb{R}_{\ge 0}, f(x) = x^2$. Its inverse is $f^{-1} : \mathbb{R}_{\ge 0} \to \mathbb{R}_{\ge 0}, f^{-1}(y) = \sqrt{y}$.

Definition 1.38: Composition

Let $f : X \to Y, g : Y \to Z$ be maps. We define $g \circ f : X \to Z$ as the COMPOSITION of f and g:

$$g \circ f(x) := g(f(x)).$$

Somtimes this is written as

$$X \xrightarrow{f} Y \xrightarrow{g} Z.$$

Example 1.39:

Let $f : X \to Y$ be a bijection and $f^{-1} : Y \to X$ its inverse map. We have $f \circ f^{-1} = \text{id}_Y$ and $f^{-1} \circ f = \text{id}_X$, where $\text{id}_X : X \to X, \text{id}_X(x) = x$ is the identity map on X.

The order of composition is important. In general, $f \circ g$ and $g \circ f$ are different maps. Even worse, one of them might not even be defined.

Lemma 1.40:

Let $f : X \to Y, g : Y \to Z$ be maps.

1. If f, g are injective, then $g \circ f$ is injective.
2. If f, g are surjective, then $g \circ f$ is surjective.
3. If f, g are bijective, then $g \circ f$ is bijective and
$$(g \circ f)^{-1} = f^{-1} \circ g^{-1}.$$

Proof.

1. Let $x_1, x_2 \in X$ such that $(g \circ f)(x_1) = (g \circ f)(x_2)$. Then $g(f(x_1)) = g(f(x_2))$. Since g is injective, $f(x_1) = f(x_2)$. Since f is injective, $x_1 = x_2$. Thus $g \circ f$ is injective.
2. Let $z \in Z$. Since g is surjective, there exists $y \in Y$ such that $g(y) = z$. Since f is surjective, there exists $x \in X$ such that $f(x) = y$. Thus $(g \circ f)(x) = g(f(x)) = g(y) = z$. Hence $g \circ f$ is surjective.
3. Follows from 1. and 2.

□

This lemma is useful when showing a certain map is injective/surjective/bijective, as it allows us to work with simpler maps.

Definition 1.41: Image

Let $f : X \to Y$ be a map, $A \subseteq X$. Then
$$f(A) := \{y \in Y \mid \exists x \in A : f(x) = y\}$$
is called the IMAGE of A under f.

Definition 1.42: Inverse Image

Let $B \subseteq Y$. Define the INVERSE IMAGE of B under f as
$$f^{-1}(B) := \{x \in X \mid f(x) \in B\}.$$

Lec 5

1.4.1 Graphs

Consider the following.

Definition 1.43:

Let $f : X \to Y$ be a map. The **graph** of f is the subset $\text{graph}(f) \subseteq X \times Y$ defined by
$$\text{graph}(f) := \{(x, f(x)) \mid x \in X\}.$$

Example 1.44:

If $X \subseteq \mathbb{R}$ and $Y = \mathbb{R}$, then the graph can be visualized in the Cartesian plane $\mathbb{R}^2$.

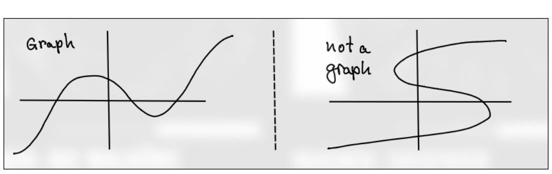


Figure 2: Graphs of some functions $f : \mathbb{R} \to \mathbb{R}$.

Sometimes injectivity and surjectivity can be read off from the graph.

2 Linear Systems of Equations

2.1 Fields

A FIELD (Körper) is a set K with two operations $+$ and $\cdot$. Together they form a commutative group $(K, +)$ with identity 0 and inverses $-a$ for $a \in K$. The operation $\cdot$ is associative and commutative with identity $1 \neq 0$ and inverses a^{-1} for $a \in K \setminus \{0\}$. The operations are linked by the distributive law
$$a \cdot (b + c) = a \cdot b + a \cdot c.$$
Furthermore, we ask that $1 \neq 0$.

Example 2.1: Examples of fields

$$\mathbb{Q} = \left\{a \mid a = \frac{m}{n}, m, n \in \mathbb{Z}, n \neq 0\right\}$$
with the usual addition and multiplication is a field.

$$\mathbb{R} = \text{the real numbers.}$$

$$\mathcal{F}_2 = \{0, 1\},$$
with the follwoing addition and multiplication tables:

+	0	1
0	0	1
1	1	0

·	0	1
0	0	0
1	0	1

Similarly, we can define $\mathcal{F}_3 = \{0, 1, 2\}$ and $\mathcal{F}_5 = \{0, 1, 2, 3, 4\}$, where the operations are defined modulo 3 and 5, respectively.

2.2 Systems of Linear Equations

Consider the following
$$\begin{cases} x + 3y + 4z &= -1 \\ 2x - y + z &= 5 \end{cases}.$$

To solve this, there exists a systematic method which is called ELIMINATION.

$$\begin{cases} x + 3y + 4z = -1 \\ 2x - y + z = 5 \end{cases} \xrightarrow[\to E_2]{-2E_1 + E_2} \begin{cases} x + 3y + 4z = -1 \\ -7y - 7z = 7 \end{cases}$$
$$\xrightarrow[\to E_2]{-\frac{1}{7}E_2} \begin{cases} x + 3y + 4z = -1 \\ y + z = -1 \end{cases}$$
$$\xrightarrow[\to E_1]{-3E_2 + E_1} \begin{cases} x + z = 2 \\ y + z = -1 \end{cases}$$

Staring at this for a while, we can see that z is a free variable. Thus, the set of solutions is given by:

choose z arbitrarily, $z = c \in \mathbb{R}$

then $y = -1 - c$ and $x = 2 - c$

$$(x, y, z) = (2 - c, -1 - c, c), c \in \mathbb{R} \subseteq \mathbb{R}^3.$$

A general system of many equations and matrix notation:

Fix a field $\mathcal{F}$. A system of m equations in n unknowns over $\mathcal{F}$ with unknowns $x_1, \ldots, x_n$ is of the form

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

We call $a_{ij} \in \mathcal{F}$ the coefficients. Furthermore $b_i \in \mathcal{F}$. Such a system is called LINEAR because the unknowns appear only to the first power and are not multiplied together.

We notice the objects can be written as matrices and vectors.

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

This object is called an $m \times n$ matrix with entries in $\mathcal{F}$.

Sometimes we write $A = (a_{ij})_{1 \leq i \leq m}^{1 \leq j \leq n}$.

Further we organize

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}.$$

These are $n \times 1$ VECTORS.

The system can now be written as

$$(S) : A \cdot x = b.$$

And the solution set as

$$L(S) = \{(x_1, x_2, \cdots) \mid x_i \in \mathcal{F} \forall 1 \leq i \leq n \wedge A \cdot x = b\}.$$

We define the matrix multiplication as follows:

$$\begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + \ldots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \ldots + a_{mn}x_n \end{pmatrix}.$$

If we now have $A \cdot x = b$, this is equivalent to the system of equations. We call this the MATRIX NOTATION.

Example 2.2:

The system from the beginning can be written as

$$\begin{pmatrix} 1 & 3 & 4 \\ 2 & -1 & 1 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -1 \\ 5 \end{pmatrix}.$$

We can also write the EXTENDED MATRIX as

$$(A|b) = \left(\begin{array}{ccc|c} 1 & 3 & 4 & -1 \\ 2 & -1 & 1 & 5 \end{array}\right).$$

2.3 Elementary Row Operations

We define the following operations on matrices.

1. Choose $c \neq 0 \in \mathcal{F}$; multiply eq./row i by c.

   $(c \cdot R_i \to R_i)$

2. Choose $c \in \mathcal{F}$, choose $1 \leq i, j \leq m, i \neq j$; replace eq./row j by eq./row $j + c\cdot$ eq./row i.

   $(c \cdot R_i + R_j \to R_j)$

3. Interchange eq./row i and eq./row j.

   $(R_i \leftrightarrow R_j)$

Lec 6

Definition 2.3: Row-Equivalence

Let C and C' be $r \times s$ matrices with entries in $\mathcal{F}$. We say that C' is ROW-EQUIVALENT to C if C' can be obtained from C by a finite sequence of elementary row operations.

If we denote $C = C_0 \to C_1 \to \ldots \to C_k = C'$, then the following holds

Theorem 2.4:

Let $A \cdot x = b$ and $A' \cdot x = b'$ be two systems of m linear equations in n unknowns. Let's call first system (S) and second system (S'). Suppose the extended matrix $(A'|b')$ is row-equivalent to $(A|b)$. Then

$$L(S') = L(S).$$

Proof. We begin with the follwoing claim:

Claim 1: If S' is obtained form S by one elementary row operation, then $L(S) \subseteq L(S')$.

Let $(x_1, \cdots, x_n) \in L(S)$. We need to show that $(x_1, \cdots, x_n) \in L(S')$.

Operation 1: $c(\neq 0) \cdot R_i \to R_i$. In this case all the equations of S' coincide with the equations of S except equation i.

$$\begin{aligned} & a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{in}x_n = b_i \\ \xrightarrow{c \cdot R_i \to R_i} \quad & ca_{i1}x_1 + ca_{i2}x_2 + \ldots + ca_{in}x_n = cb_i. \end{aligned}$$

If $(x_1, \cdots, x_n)$ satisfies the first equation, it also satisfies the second equation, since we can multiply both sides of the first equation by c.

Operation 2: $R_j + c \cdot R_i \to R_j$. Again, all equations of S and S' coincide except equation j. We have

$$\begin{aligned} & a_{j_1}x_1 + \cdots + a_{j_n}x_n = b_j \\ \xrightarrow{R_j + c \cdot R_i \to R_j} \quad & (a_{j1} + ca_{i1})x_1 + \cdots + (a_{jn} + ca_{in})x_n = b_j + cb_i \\ & \Rightarrow (a_{j1}x_1 + \cdots + a_{jn}x_n) + c(a_{i1}x_1 + \cdots + a_{in}x_n) \\ & \quad = b_j + cb_i. \end{aligned}$$

So if $(x_1, \cdots, x_n)$ satisfies the first equation, it also satisfies the second equation, since we can add c times the i-th equation to the j-th equation.

Operation 3: $R_i \leftrightarrow R_j$. In this case, clearly we have that

$$(x_1, \cdots, x_n) \in L(S) \Rightarrow (x_1, \cdots, x_n) \in L(S').$$

This concludes the proof of Claim 1.

Claim 2: If S' is obtained from S by one elementary row operation, then S can be obtained from S' by one elementary row operation.

(1): If $S \xrightarrow[c \neq 0]{c \cdot R_i \to R_i} S'$, then $S' \xrightarrow{\frac{1}{c} \cdot R_i \to R_i} S$.

(2): If $S \xrightarrow[i \neq j]{R_j + c \cdot R_i \to R_j} S'$, then $S' \xrightarrow{R_j - c \cdot R_i \to R_j} S$.

(3): If $S \xrightarrow{R_i \leftrightarrow R_j} S'$, then $S' \xrightarrow{R_i \leftrightarrow R_j} S$.

This concludes the proof of Claim 2.

Claim 3: If S' is obtained from S by one elementary row operation, then $L(S) = L(S')$.

By Claim 1, we have $L(S) \subseteq L(S')$. By Claim 2, S is obtained by S' by one elementary row operation. Hence by Claim 1 again (with the roles of S and S' interchanged), we have $L(S') \subseteq L(S)$.

It follows that $L(S) = L(S')$.

Proof of Theorem: We are now in position to prove the theorem. By assumption, there is a finite sequence of elementary row operations

$$S = S_0 \to S_1 \to \cdots \to S_k = S'.$$

By Claim 3, we have $L(S_0) = L(S_1) = \cdots = L(S_k)$, i.e.

$$L(S) = L(S').$$

□

Definition 2.5: Row-Reduced Matrix

A $m \times n$ matrix A is called ROW-REDUCED if the following two conditions hold:

1. The first non-zero entry in each non-zero row of A is 1. This entry is called the LEADING ENTRY/PIVOT of that row.
2. Each column of A which contains the leading non-zero entry of some row has all its other entries equal to 0.

Example 2.6:

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

This matrix is row-reduced.

$$B = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

This matrix is not row-reduced, because the second condition is violated in column 3.

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

This matrix is not row-reduced, because the first condition is violated in row 1.

It might not be clear yet, why row-reduced matrices are important, however, we will see this soon. However, we will first prove the following theorem.

Theorem 2.7:

Every $m \times n$ matrix A is row-equivalent to a row-reduced $m \times n$ matrix.

Proof. Let $A = (a_{ij})_{1 \le i \le m,\, 1 \le j \le n}$, with $a_{ij} \in \mathcal{F}$.

If all entries in the 1st row of A are 0, the condition (1) is satisfied for row 1.

If row 1 does have a non-zero entry, let k be the smallest index j for which $a_{1j} \neq 0$. Multiply row 1 by a_{1k}^{-1} to make the leading entry 1. Now, condition (1) is satisfied for row 1.

For each row $2 \le i \le m$ add $(-a_{ik})$ times row 1 to row i. Formally

$$R_i + (-a_{ik}) \cdot R_1 \to R_i.$$

The result of these operations is a matrix A' which looks like

$$\begin{pmatrix} 0 & \cdots & 0 & 1 & \cdots \\ \cdots & \cdots & \cdots & 0 & \cdots \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ a_{i1} & \cdots & a_{ik-1} & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix}.$$

Notice that in row i of resulting matrix, the elements standing to the right of entry k (which is now 0) are unchanged. This is because to the left of the pivot in row 1, there are only 0s.

Summary: Condition (2) is now satisfied for the column of the pivot of row 1. Lec 7

We now turn to row number 2. If all elements in row 2 are 0, we just leave it as it is. In case not all elements are 0, we find the pivot in that row, say at entry $2, k'$.

Note that $k' \neq k$, because in column k of row 2, we have 0. Divide row 2 by the element lying at $(2, k')$ to make the pivot equal to 1.

Now we take row two and add suitable multiples of it to every other zero. Notice that for the critical row 1, this operation does not change the pivot, because $(2, k) = 0$. To proof that the pivot stays in row 1, we distinguish two cases:

If $k' < k$, the entry in the first row is already 0, so the multiple is $c = 0$.

If $k' > k$, by definition of k', all entries in row 2 to the left of k' and thus also k are 0. Thus they won't change.

We can continue this process for rows $3, 4, \ldots, m$. In the end, we arrive at a row-reduced matrix. □

Tip 2.8:

Only divide by pivot elements at the end of the algorithm, to avoid fractions.

Definition 2.9:

An $m \times n$ matrix is called ROW-REDUCED ECHELON if the following holds:

(a) It is row-reduced.

(b) Every row which has only 0 entries appears below every row which has a non-zero entry.

(c) If rows $1, \ldots, r$ are the non-zero rows, and if the pivot of row i occurs in column k_i, then

$$k_1 < k_2 < \cdots < k_r.$$

A row-reduced echelon matrix looks as follows:

$$\begin{pmatrix} 0 & \cdots & 0 & 1 & * & 0 & * & 0 & * \\ 0 & \cdots & 0 & 0 & 0 & 1 & * & 0 & * \\ 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 1 & * \\ 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Theorem 2.10:

Every $m \times n$ matrix is row equivalent to a row reduced echelon matrix.

Proof. Apply the previous theorem + permutation between the rows. □

This is useful, for solving $A \cdot x = b$. Take $(A|b)$ and apply a sequence of row operations to bring it to row-reduced echelon form $(A'|b')$. Then $L(A|b) = L(A'|b')$.

Example 2.11:

We solve the following system of equations:

$$\begin{aligned} -9x_2 + 3x_3 + 4x_4 &= 9 \\ x_1 + 4x_2 - x_4 &= 5 \\ 2x_1 + 6x_2 - x_3 + 5x_4 &= -5. \end{aligned}$$

We transform this into an extended matrix

$$(A|b) = \begin{pmatrix} 0 & -9 & 3 & 4 & | & 9 \\ 1 & 4 & 0 & -1 & | & 5 \\ 2 & 6 & -1 & 5 & | & -5 \end{pmatrix}.$$

The corrseponding row reduced matrix is

$$(A'|b') = \begin{pmatrix} 0 & 1 & 0 & -\frac{5}{3} & | & \frac{12}{5} \\ 1 & 0 & 0 & \frac{17}{3} & | & \frac{-23}{5} \\ 0 & 0 & 1 & -\frac{11}{3} & | & \frac{51}{5} \end{pmatrix}.$$

Thus the row reduced echelon is

$$\begin{pmatrix} 1 & 0 & 0 & \frac{17}{3} & | & \frac{-23}{5} \\ 0 & 1 & 0 & -\frac{5}{3} & | & \frac{12}{5} \\ 0 & 0 & 1 & -\frac{11}{3} & | & \frac{51}{5} \end{pmatrix}.$$

We can see that we can take $x_4 \coloneqq a \in \mathbb{R}$ as a free variable. Then we have

$$x_3 = \frac{11}{3}a + \frac{51}{5}, \quad x_2 = \frac{5}{3}a + \frac{12}{5}, \quad x_1 = -\frac{17}{3}a - \frac{23}{5}.$$

It may happen that the system has a row with only zeros. For example

$$0 = b_3 - b_2 + 2b_1.$$

Then if $b_3 - b_2 + 2b_1 \neq 0$, the system has no solution. Otherwise, the system is equivalent to the system without that equation.

3 Vector Spaces

Vector spaces are essentially the playground of linear algebra, the following definition is thus very central for the subject.

Definition 3.1: Vector Space

A Vector Space over a field $\mathcal{F}$ (scalars) is a set V (vecotrs) endowed with two operations:

$$\begin{aligned} + &: V \times V \to V, (v_1, v_2) \mapsto v_1 + v_2 \\ \cdot &: \mathcal{F} \times V \to V, (\lambda, v) \mapsto \lambda \cdot v. \end{aligned}$$

such that the following axioms are satisfied:

V1-4) $(V, +)$ is an abelian group.

V5) $\forall a, b \in \mathcal{F}, \forall v \in V : a \cdot (b \cdot v) = (ab) \cdot v$

V6) $\forall v \in V, 1 \cdot v = v$

V7) $\forall a \in F, v_1, v_2 \in V : a \cdot (v_1 + v_2) = a \cdot v_1 + a \cdot v_2$

V8) $\forall a_1, a_2 \in \mathcal{F}, v \in V : (a_1 + a_2) \cdot v = a_1 \cdot v + a_2 \cdot v$

Tip 3.2:

Vector spaces are essentially ordered lists with component-wise operations.

Theese lists don't have to be finite.

Example 3.3: Coordinate Space

The coordinate space K^n for $n \in \mathbb{N}_0$ then

$$\underbrace{K^n}_{K \times K \cdots \times K} = \{(a_1, \ldots, a_n) \mid a_i \in K, 1 \leq i \leq n\}.$$

We turn $V = K^n$ into a vector space. Let $v = (a_1, \ldots, a_n), w = (b_1, \ldots, b_n) \in V$. Then

$$v + w = (a_1 + b_1, a_2 + b_2, \ldots, a_n + b_n) \in V.$$

Let $a \in K$, then

$$a \cdot v \coloneqq (a \cdot a_1, a \cdot a_2, \ldots, a \cdot a_n) \in V.$$

$$0 \coloneqq (0, 0, \ldots, 0) \in V.$$

As an exercise, show that K^n with theese operations is a vector space.

Example 3.4: Matrix Space

Consider $M_{m \times n}$. The elemnts are

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

This is basiclly the same as $K^{m \cdot n}$.

Lemma 3.5:

Let V be a vector space over K. Then:

(a) The zero vector is unique.

(b) Let $v \in V$. The element $v' \in V$ with $v + v' = 0$ is unique. It is denoted by $-v$.

(c) $\forall v \in V : 0 \cdot v = 0_V$

(d) $\forall a \in K : a \cdot 0_V = 0_V$

(e) $\forall v \in V : (-1) \cdot v = -v$

(f) $\forall v \in V, -(-v) = v$

(g) If $a \cdot v = 0$ for some $a \in K, v \in V$, then either $a = 0$ or $v = 0_V$.

Lec 8

Proof. **(a)**: Since $(V, +)$ is an abelian group, there is an unique identity element, which we denote by 0_V.

(b): Again, since $(V, +)$ is an abelian group, every element has an unique inverse.

(c): Let $v \in V$. We have to show that $0 \cdot v = 0_V$. Indeed,

$$0 \cdot v = (0 + 0) \cdot v \overset{\text{V8}}{=} 0 \cdot v + 0 \cdot v.$$

Add $(-0 \cdot v)$ to both sides. This yields

$$0_V = 0 \cdot v + (0 \cdot v + (-0 \cdot v)) = 0 \cdot v + 0_V = 0 \cdot v.$$

(g): Let $a \in K, v \in V$ such that $a \cdot v = 0_V$. We need to prove that either $a = 0$ or $v = 0_V$. Indeed, assume that $a \neq 0$. By the axioms of a field, $\exists$ an element $a^{-1} \in K$ such that $a^{-1}a = 1 = aa^{-1}$.

Now:

$$v \overset{\text{V6}}{=} 1 \cdot v = (a^{-1}a) \cdot v \overset{\text{V5}}{=} a^{-1} \cdot (a \cdot v) = a^{-1} \cdot 0_V \overset{\text{(d)}}{=} 0_V.$$

□

Definition 3.6: Linear Subspace

Let V be a vector space over K. A subset $W \subseteq V$ is called a LINEAR SUBSPACE (Unterraum) of V if the following holds:

LSS1) $W \neq \varnothing$,

LSS2) $\forall w_1, w_2 \in W : w_1 + w_2 \in W$,

LSS3) $\forall w \in W, \forall a \in K : a \cdot w \in W$.

Lemma 3.7:

Let V be a vector space over K and $W \subseteq V$ a subspace. Then W is a vector space on its own when endowed with the operations coming from V.

Proof. From LSS2 and LSS3, it follows that the operations $+$ and $\cdot$ from V indeed define such operations on W.

Now we have to check the axioms V1-V8. This can be done as an exercise.

For example, for V2, Pick any $w \in W$ (Possible by LSS1). We have

$$0_V \overset{(c)}{=} 0_K \cdot w \overset{\text{LSS3}}{\in} W.$$

Thus $0_V \in W$. Take now $0_W := 0_V$. Clearly, $\forall w \in W$ we have

$$0_W + w = 0_V + w \overset{\text{V2}}{=} w.$$

□

The above definition is often cumbersome to use. The following lemma gives a more handy criterion to check if a subset is a subspace.

Lemma 3.8: Criterion for Subspaces

Let V be a vector space ove K and $W \subseteq V$ a subset. Then W is a subspace of V iff the following holds:

(1) $0_V \in W$,

(2) $\forall a_1, a_2 \in K, \forall w_1, w_2 \in W : a_1 \cdot w_1 + a_2 \cdot w_2 \in W$.

Proof. $\Rightarrow$: Assume that W is a subspace of V. This follow at once from the fact that W is in itself a vector space.

$\Leftarrow$: Assume that (1) and (2) hold. We have to show that W is a subspace. Clearly, LSS1 holds by (1). To show LSS2, pick $w_1, w_2 \in W$. Then by (2) (with $a_1 = a_2 = 1$), we have

$$w_1 + w_2 = 1 \cdot w_1 + 1 \cdot w_2 \in W.$$

Similarly for LSS3, pick $w \in W, a \in K$. Then by (2) (with $a_1 = a, a_2 = 0, w_1 = w, w_2 = 0_V$), we have

$$a \cdot w = a \cdot w + 0 \cdot 0_V \in W.$$

□

Example 3.9: Trivial Subspace

Let V be a vector space over K. Then

$$\{0_V\} \subseteq V,$$

is a subspace of V. It is called the ZERO SUBSPACE of V.

Example 3.10:

Fix $b \in K$. Consider the subset

$$\mathcal{U}_b := \{(x_1, x_2, x_3) \in K^3 \mid x_1 + x_2 + x_3 = b\} \subseteq K^3.$$

Claim: $\mathcal{U}_b$ is a subspace of K^3 iff $b = 0$.

Proof. $\Rightarrow$: Suppose $\mathcal{U}_b$ is a linear subspace of K^3. If (x_1, x_2, x_3) and $(y_1, y_2, y_3) \in \mathcal{U}_b$, then

$$x_1 + x_2 + x_3 = b = y_1 + y_2 + y_3.$$

Since $\mathcal{U}_b$ is a linear subspace, we have

$$(x_1 + y_1, x_2 + y_2, x_3 + y_3) \in \mathcal{U}_b.$$

And thus it follows that

$$(x_1 + y_1) + (x_2 + y_2) + (x_3 + y_3) = b.$$

But from before, we get that this sum is equal to $2b$. Thus $2b = b$, which implies that $b = 0$.

The other direction is left as an exercise. □

Let A be an $m \times n$ matrix with entries in K. From now on we abbreviate this with

$$A \in \mathrm{Mat}_{m \times n}(K) \text{ or } A \in M_{m \times n}(K).$$

Fix $b \in K^m$ such that $b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$ is viewed as a column vector. Consider the subset of K^n defined by

$$L := \{x \in K^n | A \cdot x = b\}.$$

Claim 3.11:
$L \subseteq K^n$ is a linear subspace iff $b = 0_V$.

For the proof we need preparations.

Lemma 3.12:
$\forall a \in K, u, v \in K^n$ we have

(1) $A \cdot (u + v) = A \cdot u + A \cdot v$,

(2) $A \cdot (a \cdot u) = a \cdot (A \cdot u)$.

Proof. Write $A = (a_{ij})_{1 \le i \le m, 1 \le j \le n}$,

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}, u = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix}, v = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}.$$

$$A \cdot u + A \cdot v = \begin{pmatrix} \vdots \\ a_{i1}u_1 + a_{i2}u_2 + \cdots + a_{in}u_n \\ \vdots \end{pmatrix} + \begin{pmatrix} \vdots \\ a_{i1}v_1 + a_{i2}v_2 + \cdots + a_{in}v_n \\ \vdots \end{pmatrix}$$

$$= \begin{pmatrix} \vdots \\ a_{i1}(u_1 + v_1) + a_{i2}(u_2 + v_2) + \cdots + a_{in}(u_n + v_n) \\ \vdots \end{pmatrix}$$

$$= A \cdot \begin{pmatrix} u_1 + v_1 \\ u_2 + v_2 \\ \vdots \\ u_n + v_n \end{pmatrix}$$

$$= A \cdot (u + v).$$

The second statement is left as an exercise. □

Lec 9

Proof. [Claim] $\Leftarrow$: Assume $b = 0$, we want to show that L is a subspace. Indeed $0 \in L$ because $A \cdot 0 = 0$.

If $x = (x_1, \ldots, x_n), y = (y_1, \ldots, y_n) \in L$, then

$$A \cdot (x + y) \overset{(1)}{=} A \cdot x + A \cdot y = 0 + 0 = 0.$$

Thus $x + y \in L$.

Let $x \in L, a \in K$. Then

$$A \cdot (a \cdot x) \overset{(2)}{=} a \cdot (A \cdot x) = a \cdot 0 = 0.$$

Thus $a \cdot x \in L$.

Together theese three properties show that $L \subseteq K^n$ is a subspace.

$\Rightarrow$: Assume that L is a subspace, and we we'll show that $b = 0$. Indeed since $L \subseteq K^n$ is a subspace, $0_V \in L$. Thus $A \cdot 0_V = b$. But $A \cdot 0_V = 0_V$. Thus $b = 0_V$.

□

Example 3.13:
Let K be a field. Define

$$K^\infty := \{(a_1, a_2, \ldots) \mid a_i \in K, i \in \mathbb{N}\}.$$

We turn K^∞ into a vector space as follows:

$$(a_1, a_2, \ldots) + (b_1, b_2, \ldots) := (a_1 + b_1, a_2 + b_2, \ldots)$$
$$c \cdot (a_1, a_2, \ldots) := (c \cdot a_1, c \cdot a_2, \ldots)$$
$$.$$

Recalling the first lecture, taking $K = \mathbb{R}$, we have

$$\text{Fib} := \{(a_1, a_2, \ldots) \mid a_k = a_{k-1} + a_{k-2}, k \ge 3\}.$$

As an exercise, show that Fib is a subspace of $\mathbb{R}^\infty$. More generally, fix $\alpha, \beta \in \mathbb{R}$. Then

$$\mathcal{U}_{\alpha,\beta} := \{(a_1, a_2, \ldots) \in \mathbb{R}^\infty \mid a_k = \alpha a_{k-1} + \beta a_{k-2}, k \ge 3\},$$

is a subspace of $\mathbb{R}^\infty$.

3.1 Spaces of Functions

Fix a field K and S is a non-empty set. Define

$$K^S := \{f : S \to K\}.$$

This denotes the set of all functions from S to K. We define the following operations on K^S:

$$(f + g)(s) := f(s) + g(s)$$
$$(a \cdot f)(s) := a \cdot (f(s)).$$

With these operations, K^S is a vector space over K. (Exercise)

Exercise 3.14:
Let $S = [0, 1]$. Define $C(S) := \{f \in \mathbb{R}^S \mid f\}$ such that f is continuous. Then $C(S)$ is a subspace of $\mathbb{R}^S$.

3.2 Polynomials

Let K be a field. Fix a letter called FORMAL VARIABLE x. A POLYNOMIAL over K (or polynomial with coefficients in K) is an expression of the form

$$f(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n.$$

where $n \in \mathbb{N}_0, a_i \in K \, \forall 0 \le i \le n$. We allow to omit 0-terms like $0 \cdot x^i$.

We call $f(x) = 0$ the ZERO POLYNOMIAL.

We can add and multiply polynomials as follows: Consider polynomials a and b we take $r = \max(n, m)$ and define

$$(a_0 + a_1 x + \cdots + a_n x^n) + (b_0 + b_1 x + \cdots + b_m x^m) :=$$
$$(a_0 + b_0) + (a_1 + b_1)x + \cdots + (a_r + b_r)x^r.$$

The multiplication is defined as

$$c \cdot (a_0 + a_1 x + \cdots + a_n x^n) :=$$
$$(c \cdot a_0) + (c \cdot a_1)x + \cdots + (c \cdot a_n)x^n.$$

We write $K[x]$ for the set of all polynomials with coefficients in K.

Claim 3.15:
$K[x]$ is a vector space over K with the above operations.

Every $f(x) \in K[x]$ defines also a function $\tilde{f} : K \to K$. The subset of K^K which comes from polynomials is a subspace of K^K.

But: from $\tilde{f}$ we cannot always recover f.

Example 3.16:

Take $K = \mathcal{F}_2 = \{0,1\}$. Consider $f(x) = x$ and $g(x) = x^2$. Then theese are diffrent polynomials, but $\tilde{f} = \tilde{g}$.

Definition 3.17: Degree of a polynomial

Let $f(x) = a_0 + a_1 x + \cdots + a_n x^n \in K[x]$ with $a_n \neq 0$. Then we define the DEGREE of f as $\deg(f) = n$. We define $\deg(0) = -\infty$.

Fix $d \in \mathbb{N}_0 \cup \{-\infty\}$. Define $K[x]_d$ to be all the polynomials of degree at most d. Then $K[x]_d$ is a subspace of $K[x]$.

Question 3.18:

Is $\{f(x) \in k[x] \mid \deg(f) = d\}$ also a subspace of $K[x]$?

Cosnider $\{p \in k[x]_5 \mid \sum_{n=0}^{\infty} a_n = 0\}$. Is this a subspace of $K[x]_5$?

Spaces of matrices: Consider $M_{m \times n}(K)$ to be the set of all $m \times n$ matrices with entries in K. Define $+$ and $\cdot$ as

$$(a_{ij}) + (b_{ij}) := (a_{ij} + b_{ij})$$
$$c \cdot (a_{ij}) := (c \cdot a_{ij}).$$

With theese operations, $M_{m \times n}(K)$ is a vector space over K.

As a motivation for the next few lectures: Let V be a vector space, let $S \subseteq V$ be a subset. What is the smallest subspace of V containing S?

Lec 10

3.3 Span

Consider the following lemma

Lemma 3.19:

Let V be a vector space and let $\{W_i\}_{i \in I}$ be a family of subspaces of V (i.e. $W_i \subseteq V$ is a subspace for all $i \in I$). Then the intersection

$$W := \bigcap_{i \in I} W_i$$

is a subspace of V.

Proof. Note that $0_V \in W_i$ for all $i \in I$, thus $0_V \in W$.

Let now $a, b \in K, v, w \in W$. We'll show that $a \cdot v + b \cdot w \in W$.

Let $j \in I$ be any index in I. Since $v, w \in W$, they must also be in W_j. Since W_j is a subspace of V, it follows that

$$a \cdot v + b \cdot w \in W_j.$$

But this holds $\forall j \in I$, therefore,

$$a \cdot v + b \cdot w \in \bigcap_{i \in I} W_i = W.$$

By a previous lemma, this shows that W is a subspace of V. □

Let $S \subseteq V$ be a subset.

Definition 3.20: Span

Define the SPAN of S as

$$\mathrm{Sp}(S) := \bigcap_{W \in \mathcal{N}} W,$$

where

$$\mathcal{N} := \{W \subseteq V \mid W \text{ is a subspace of } V, S \subseteq W\}.$$

Note that $\mathcal{N} \neq \varnothing$ since $V \in \mathcal{N}$.

By the previous lemma, $\mathrm{Sp}(S)$ is a subspace of V. Sometimes, we also call it the subspace of V generated by S.

Lemma 3.21:

Among all subspaces of V that contain S, $\mathrm{Sp}(S)$ is the smallest one, i.e.,

1. $\mathrm{Sp}(S) \subseteq V$ is a subspace of V.
2. If $W \subseteq V$ is a subspace and $W \supseteq S$, then $W \supseteq \mathrm{Sp}(S)$.

Proof. 1. This was shown above.

2. Let $W \subseteq V$ be a subspace with $W \supseteq S$. By definition of $\mathcal{N}$, we have $W \in \mathcal{N}$. Thus by definition of $\mathrm{Sp}(S)$, we have

$$\mathrm{Sp}(S) = \bigcap_{U \in \mathcal{N}} U \subseteq W.$$

□

3.3.1 Linear Combinations

This definition is good from a theoretical point of view, but not very handy when we actually want to compute $\mathrm{Sp}(S)$. We thus introduce the following concept.

Definition 3.22: Linear Combination

Let $n \in \mathbb{Z}_{\geq 1}$, $a_1, \ldots, a_n \in K$ and $v_1, \ldots, v_n \in V$. A LINEAR COMBINATION of the vectors $v_1, \ldots, v_n$ with coefficients $a_1, \ldots, a_n$ is the vector

$$a_1 \cdot v_1 + a_2 \cdot v_2 + \cdots + a_n \cdot v_n.$$

We only look at finite linear combinations, i.e., n is always finite.

Lemma 3.23:

Let $\varnothing \neq S \subseteq V$. Then

$$\mathrm{Sp}(S) = \{a_1 v_1 + \cdots + a_n v_n \mid n \in \mathbb{Z}_{\geq 1}, v_i \in S, a_i \in K\}.$$

i.e. $\mathrm{Sp}(S)$ is the subset of V obtained by taking all possible linear combinations of vectors in S.

Proof. Denote $\widetilde{\mathrm{Sp}}(S) := \{a_1 v_1 + \cdots + a_n v_n \mid n \in \mathbb{Z}_{\geq 1}, v_i \in S, a_i \in K\}$. We will show that $\widetilde{\mathrm{Sp}}(S) = \mathrm{Sp}(S)$.

We first show that 1) $S \subseteq \widetilde{\mathrm{Sp}}(S)$, 2) $\widetilde{\mathrm{Sp}}(S)$ is a subspace of V and 3) $\forall$ subspace $W \subseteq V$ with $S \subseteq W$ we have $\widetilde{\mathrm{Sp}}(S) \subseteq W$.

Note that this is useful since 1)+2) would imply that $\mathrm{Sp}(S) \subseteq \widetilde{\mathrm{Sp}}(S)$ by the previous lemma, and 3) would imply $\widetilde{\mathrm{Sp}}(S) \subseteq \bigcap_{W \in \mathcal{N}} W = \mathrm{Sp}(S)$. Together, this implies $\widetilde{\mathrm{Sp}}(S) = \mathrm{Sp}(S)$.

It remains to prove 1), 2) and 3).

1) If $v \in S$, then $v = 1 \cdot v \in \widetilde{\mathrm{Sp}}(S)$.

2) We'll show that $0_V \in \widetilde{\mathrm{Sp}}(S)$ and that $\forall \alpha, \beta \in K, \forall v, w \in \widetilde{\mathrm{Sp}}(S)$: $\alpha \cdot v + \beta \cdot w \in \widetilde{\mathrm{Sp}}(S)$.

Clearly, $0_V = 0 \cdot v \in \widetilde{\mathrm{Sp}}(S)$ for any $v \in S$.

Write $v = a_1 v_1 + \cdots + a_n v_n$ and $w = b_1 w_1 + \cdots + b_m w_m$. Then

$$\begin{aligned} \alpha \cdot v + \beta \cdot w &= \alpha(a_1 v_1 + \cdots + a_n v_n) + \beta(b_1 w_1 + \cdots + b_m w_m) \\ &= (\alpha a_1) v_1 + \cdots + (\alpha a_n) v_n + (\beta b_1) w_1 + \cdots + (\beta b_m) w_m. \end{aligned}$$

This is a linear combination of vectors in S, thus $\alpha \cdot v + \beta \cdot w \in \widetilde{\mathrm{Sp}}(S)$.

3) Let $W \subseteq V$ be a subspace with $S \subseteq W$. Let v be in $\widetilde{\mathrm{Sp}}(S)$. We need to prove that $v \in W$. Indeed, write $v = a_1 v_1 + \cdots + a_n v_n$ with $v_i \in S$. By assumption, $S \subseteq W$, thus $v_i \in W$. As W is a linear subspace, it follows that $v \in W$. □

We have shown that the span of a subset S of a vector space V can be written in two ways. The first one is more theoretical, the second one more practical.

The span of the empty set is $\{0_V\}$. From now on, let's agree that linear combinations of elements of the empty set is just 0_V.

As a notation: Many times $S = \{v_1, \dots, v_n\}$ is a finite set. We write $\mathrm{Sp}\{v_1, \dots, v_n\}$ or $\mathrm{Sp}(v_1, \dots, v_n)$ instead of $\mathrm{Sp}(\{v_1, \dots, v_n\})$.

Definition 3.24: Generating/Spanning Set

Let V be a vector space and $S \subseteq V$ a subset. We say that S GENERATES or SPANS V if $\mathrm{Sp}(S) = V$.

If $W \subseteq V$ is a subspace, we say that S generates or spans W if $\mathrm{Sp}(S) = W$.

Definition 3.25: Finite Dimensional Vector Space

A vector space V is FINITE DIMENSIONAL if there exists a finite set $S \subseteq V$ that generates V.

If such a finite S does not exist, then V is called INFINITE DIMENSIONAL.

Example 3.26: Subspaces of $\mathbb{R}^2$

Let $V = \mathbb{R}^2, K = \mathbb{R}$. Some subspaces of V are:

1. $\{0_{\mathbb{R}^2}\} = \{(0,0)\} = \mathrm{Sp}(\varnothing) = \mathrm{Sp}\{0_{\mathbb{R}^2}\}$
2. Let $0 \neq v = (a,b) \in \mathbb{R}^2$. Then

$$\mathrm{Sp}\{v\} = \{\alpha \cdot (a,b) \mid \alpha \in \mathbb{R}\} = \{(\alpha a, \alpha b) \mid \alpha \in \mathbb{R}\}.$$

Geometrically speaking, this is the line through the origin and v.

3. Let $0 \neq v_1, w \in \mathbb{R}^2$ If $\mathrm{Sp}\{v\} = \mathrm{Sp}\{w\}$, this is equivalent to saying that $\exists \alpha \in \mathbb{R} \setminus \{0\}$ such that $w = \alpha v$.
4. Let $W \subseteq \mathbb{R}^2$ be a subspace and assume that $W = \mathrm{Sp}\{v, w\}$ where $v \neq 0, w \notin \mathrm{Sp}\{v\}$. Then $W = \mathbb{R}^2$.

Together this tells us that there are only three types of subspaces of $\mathbb{R}^2$: the zero subspace, lines through the origin, and the whole space $\mathbb{R}^2$.

Question: Let $v_1, \dots, v_n \in V$ and $w \in K^m$. How can we determine whether or not $w \in \mathrm{Sp}\{v_1, \dots, v_n\}$?

e.g. let $v_1 = (1,3), v_2 = (7,73)$ and $w = (11,137)$. In this case yes, $w \in \mathrm{Sp}\{v_1, v_2\}$ as $-3 \cdot v_1 + 2 \cdot v_2 = w$. Lec 11

Solution: Take the column vectors $v_1, \dots, v_n$ and make a $m \times n$ matrix out of them:

$$A = \begin{pmatrix} | & | & & | \\ v_1 & v_2 & \cdots & v_n \\ | & | & & | \end{pmatrix}.$$

If $x = (x_1, \dots, x_n) \in K^n$, then

$$A \cdot x = x_1 v_1 + x_2 v_2 + \cdots + x_n v_n = \sum_{i=1}^n x_i v_i.$$

Given $w \in K^m$, the question is wheter or not the system of equations $A \cdot x = w$ has a solution. If a solution exists, then $w \in \mathrm{Sp}\{v_1, \dots, v_n\}$. If a solution does not exist, then $w \notin \mathrm{Sp}\{v_1, \dots, v_n\}$.

In the first case, any solution x will give us a choice of coefficients $x_1, \dots, x_n$ such that $w = \sum_{i=1}^n x_i v_i$.

Example 3.27:

1) Let $V = K^n$. Denote by $e_1 := (1,0,0,\dots,0), e_2 := (0,1,0,\dots,0), \dots, e_n := (0,0,0,\dots,1)$. Then

$$\mathrm{Sp}\{e_1, e_2, \dots, e_n\} = K^n.$$

2) Let $W = k[x]_d$ (polynomials in x of degree at most d). Then

$$W = \mathrm{Sp}\{1, x, x^2, \dots, x^d\}.$$

Similarly, take $V = K[x]$. Then

$$V = \mathrm{Sp}\{1, x, x^2, x^3, \dots\} \text{ (infinite set) }.$$

3) Let $M = M_{m \times n}(K)$ be the set of all $m \times n$ matrices. For every $1 \leq i \leq m, 1 \leq j \leq n$, define the matrix

$$E_{ij} := \begin{pmatrix} 0 & \cdots & 0 & \cdots & 0 \\ \vdots & & \vdots & & \vdots \\ 0 & \cdots & 1 & \cdots & 0 \\ \vdots & & \vdots & & \vdots \\ 0 & \cdots & 0 & \cdots & 0 \end{pmatrix}.$$

i.e. the matrix with a 1 in the (i,j)-th position and 0 elsewhere.

Then

$$M = \mathrm{Sp}\{E_{ij} \mid 1 \leq i \leq m, 1 \leq j \leq n\}.$$

All of these examples are finite dimensional vector spaces except $K[x]$. As an exercise, show that $K[x]$ is infinite dimensional.

3.4 Linear Independence

Let V be a vector space over K

Definition 3.28:

Let $v_1, \dots, v_n$ be a list of n vectors in V. We say that that $v_1, \dots, v_n$ are LINEARLY INDEPENDENT if the only linear combination of $v_1, \dots, v_n$ that equals 0_V is the one with all coefficients equal to zero.

If the vectors are not linearly independent, they are called LINEARLY DEPENDENT.

Alternatively, $v_1, \dots, v_n$ are linearly independent iff

$\forall a_1, \dots, a_n \in K$ for which $a_1 v_1 + \cdots + a_n v_n = 0_V$, we must have $a_1 = a_2 = \cdots = a_n = 0$.

Let's agree that $\varnothing$ is linearly independent.

Remark 3.29:

1) 0_V can always be written as a linear combination of any list of vectors as $0_V = 0 \cdot v_1 + 0 \cdot v_2 + \cdots + 0 \cdot v_n$. This is called the TRIVIAL linear combination.

2) Suppose that in our list there is a repetition, say $v_i = v_j$ for some $i \neq j$. Then the list is linearly dependent, since

$$0_V = 1 \cdot v_i + (-1) \cdot v_j + 0 \cdot v_1 + \cdots + 0 \cdot v_n.$$

3) The order of the elements in the list does not matter.

4) If 0 is one of the elements in the list, then this list cannot be linearly independent.

As a generalization for families of vectors consider the following definition.

Definition 3.30:

Let $\mathcal{F} = \{v_i\}_{i \in \mathcal{I}}$ be a family of vectors in V. We say that $\mathcal{F}$ is LINEARLY INDEPENDENT if $\forall n \in \mathbb{Z}_{\geq 1}$ and any sequence of distinct indices $i_1, \dots, i_n \in \mathcal{I}$, the list of vectors $v_{i_1}, \dots, v_{i_n}$ is linearly independent.

Let $\varnothing \neq S \subseteq V$ be a subset. We say that S is linearly independent if every finite list of distinct vectors in S is linearly independent.

Lemma 3.31:

If $\varnothing \neq S \subseteq V$, is linearly independent, then every subset of it is also linearly independent.

Proof. Assume $\exists a_1, \dots, a_n \in K$ and $v_1, \dots, v_n \in W \subseteq S$ such that $a_1 v_1 + \cdots + a_n v_n = 0_V$ with not all a_i equal to zero. Since $W \subseteq S$, this contradicts the linear independence of S. $\square$

Example 3.32:

1) $e_1, \dots, e_n \in K^n$ are linearly independent.

Solution. Suppose $a_1 e_1 + \cdots + a_n e_n = 0$. This means that

$$a_1 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + a_2 \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} + \cdots + a_n \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

But since we are working component-wise, this means that

$$\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

In particular, $a_i = 0$ for all $1 \leq i \leq n$.

2) The set $\{e_1, e_2, \cdots\} \subseteq K^\infty$ is linearly independent.

3) The set $\{1, x, x^2, x^3, \dots, x^d\} \subseteq K[x]_d$ is linearly independent.

4) The set $\{1, x, x^2, x^3, \dots\} \subseteq K[x]$ is linearly independent.

5) Let V be a vector space, let $v \in V$. Show that $\{v\}$ is linearly independent iff $v \neq 0_V$.

3.5 Basis

The following is one of the most central definitions in linear algebra.

Definition 3.33: Basis

A subset $S \subseteq V$ is called a BASIS of V if the following holds:

1) S is linearly independent.

2) S spans V, i.e. $\mathrm{Sp}(S) = V$.

Proposition 3.34:

A subset $S \subseteq V$ is a basis of V iff every $v \in V$ can be written in a *unique* way as a linear combination of vectors from S. Lec 12

The meaning of "unique" in the above proposition is the following: suppose $v \in V$ is written as $v =$ linear combination #1 with vectors from S, and also as $v =$ linear combination #2 with vectors from S.

Define C_1 to be all the vectors that appear in linear combination #1 without zero coefficient. Similarly, define C_2.

Then $C_1 = C_2$ and $\forall u \in C_1$, the coefficient of u in #1 equals the coefficient of u in #2.

In case, S is a finite set, say $S = \{u_1, \dots, u_n\}$, then uniqueness means that if $v = a_1 u_1 + \cdots + a_n u_n = b_1 u_1 + \cdots + b_n u_n$, then $a_i = b_i$ for all $1 \leq i \leq n$.

Proof. [of proposition] For simplicity assume that S is finite. (The general case is an exercise.)

$\Leftarrow$ Assume every $v \in V$ can be written in a unique way. Because every $v \in V$ can be written, this implies that $\mathrm{Sp}(S) = V$. We'll show now that S is linearly independent.

Write $S = \{v_1, \dots, v_n\}$. Consider the 0_V. Clearly, $0 = 0 \cdot v_1 + \dots + 0 \cdot v_n$. But by assumption, there is a unique linear combination of $v_1, \dots, v_n$ that equals 0_V. It follows that if

$$a_1 v_1 + \dots + a_n v_n = 0_V.$$

Then all a_i must be zero. Thus, S is linearly independent. Thus, S is a basis of V.

$\Rightarrow$ Assume S is a basis of V. Therefore, $\mathrm{Sp}(S) = V$. This implies that every $v \in V$ can be written as a linear combination of vectors from S. It remains to show uniqueness.

Assume by contradiction that $\exists v \in V$ can be written as

$$v = a_1 v_1 + \dots + a_n v_n = b_1 v_1 + \dots + b_n v_n.$$

Where $a_k \neq b_k$ for some $1 \leq k \leq n$.

Now, subtracting both expressions, we get

$$0_V = (a_1 - b_1) v_1 + \dots + \underbrace{(a_k - b_k)}_{\neq 0} v_k + \dots + (a_n - b_n) v_n.$$

Thus 0_V can be written as a non-trivial linear combination of vectors from S. Thus S is not linearly independent, which contradicts our assumption that S is a basis. □

Example 3.35:

1) The set $\{e_1, \dots, e_n\}$ is a basis of K^n.

2) The set $\{1, x, x^2, \dots, x^d\}$ is a basis of $K[x]_d$.

3) The set $\{E_{ij} \mid 1 \leq i \leq m, 1 \leq j \leq n\}$ is a basis of $M_{m \times n}(K)$.

In case of K^n, we call the basis $\{e_1, \dots, e_n\}$ the STANDARD BASIS or CANONICAL BASIS of K^n.

Lemma 3.36:

Let $v_1, \dots, v_m \in V$ be a list of linearly dependent vectors. Then $\exists 1 \leq j \leq m$ such that

1) $v_j \in \mathrm{Sp}(v_1, \dots, v_{j-1})$

2) $\mathrm{Sp}(v_1, \dots, v_m) = \mathrm{Sp}(v_1, \dots, v_{j-1}, v_{j+1}, \dots, v_m)$

Proof. Since the list is linearly dependent, $\exists a_1, \dots, a_m \in K$ not all of them zero, such that

$$a_1 v_1 + \dots + a_m v_m = 0_V.$$

Define $j := \max\{i \mid a_i \neq 0\}$ (i.e. $a_j \neq 0$ and $a_{j+1} = \dots = a_m = 0$).

Then $a_1 v_1 + \dots + a_{j-1} v_{j-1} + a_j v_j = 0_V$. Since $a_j \neq 0$, we can solve for v_j:

$$v_j = -\frac{a_1}{a_j} v_1 - \dots - \frac{a_{j-1}}{a_j} v_{j-1}. \tag{3.1}$$

This implies $v_j \in \mathrm{Sp}(v_1, \dots, v_{j-1})$.

To prove 2), let $v \in \mathrm{Sp}(v_1, \dots, v_m)$. Then

$$v = c_1 v_1 + \dots + c_m v_m.$$

Substituting (3.1) into this expression, we obtain, that v is a linear combination of $v_1, \dots, v_{j-1}, v_{j+1}, \dots, v_m$. Hence,

$$\mathrm{Sp}(v_1, \dots, v_m) \subseteq \mathrm{Sp}(v_1, \dots, v_{j-1}, v_{j+1}, \dots, v_m).$$

But clearly, we also have the other inclusion, thus

$$\mathrm{Sp}(v_1, \dots, v_m) = \mathrm{Sp}(v_1, \dots, v_{j-1}, v_{j+1}, \dots, v_m).$$

□

Lemma 3.37:

Let $v_1, \dots, v_m$ be a list of linearly dependent vectors such that $v_1, \cdots, v_k$ are linearly independent for some $1 \leq k < m$. Then j from Lemma 3.36 satisfies $k < j$.

Proof. Assume by contradiction that $j \leq k$. We have:

$$v_j = a_1 v_1 + \dots + a_{j-1} v_{j-1},$$

for some $a_1, \dots, a_{j-1} \in K$.

But this implies that

$$0 = a_1 v_1 + \dots + a_{j-1} v_{j-1} - 1 \cdot v_j.$$

Since there is a -1 coefficient, this is a non-trivial linear combination of $v_1, \dots, v_k$ that equals 0_V. This contradicts the linear independence of $v_1, \dots, v_k$. □

Lemma 3.38:

Let $w_1, \dots, w_n \in V$ such that $\mathrm{Sp}(w_1, \dots, w_n) = V$ and let $v \in V$. Then the list $v, w_1, \dots, w_n$ is linearly dependent.

Proof. Write $v = a_1 w_1 + \dots + a_n w_n$ for some $a_1, \dots, a_n \in K$. Then

$$0_V = (-1) \cdot v + a_1 w_1 + \dots + a_n w_n.$$

This is a non-trivial linear combination of $v, w_1, \dots, w_n$ that equals 0_V. Thus, the list is linearly dependent. □

Lemma 3.39:

Let $v_1, \dots, v_n \in V$ such that $\mathrm{Sp}(v_1, \dots, v_n) = V$. Let $u_1, \dots, u_m$ be a list of linearly independent vectors in V. Then $m \leq n$.

Proof. The proof will have m steps.

Step 1: We will replace u_1 by one of the vectors $v_1, \dots, v_n$. How? Consider the list $u_1, v_1, \dots, v_n$. By Lemma 3.38, this list is linearly dependent. By Lemma 3.36, one of the vectors in this list is in the span of the vectors that appear before it, and if we drop this vector from the list, the overall span does not change. By Lemma 3.37, this vector cannot be u_1 (because u_1 is linearly independent).

We now drop that vector and get a list of n vectors that still span V and the first vector is u_1.

Step j ($2 \leq j \leq m$): From step $j-1$, we should have a list of n vectors that span V and the first $j-1$ vectors are $u_1, \dots, u_{j-1}$ and the other rest $n-(j-1)$ of vectors are taken from the list $v_1, \dots, v_n$.

Let's write this list of n vectors as $u_1, \dots, u_{j-1}, w_1, \dots, w_{n-(j-1)}$. Consider now the list $u_1, \dots, u_{j-1}, u_j, w_1, \dots, w_{n-(j-1)}$. Lec 13

This list with $n+1$ vectors spans V as the span cannot be reduced by adding more vectors. By Lemma 3.38, we know that this list is linearly dependent. By Lemma 3.36, one of the vectors in this list is in the span of the vectors that appear before it, and if we drop this vector from the list, the overall span does not change. By Lemma 3.37, this vector can be not of $u_1, \dots, u_{j-1}$ (because they are linearly independent). Thus, the vector that we drop is in the set $\{w_1, \dots, w_{n-(j-1)}\}$.

So, the new list is $u_1, \dots, u_j, w_1, \dots, w_{n-j}$ and it still spans V.

Assume now, by contradiction, that $m > n$. Then, we can perform the steps $j = 1, \dots, j = n$ and obtain at step $j = n$ a list that looks like

$$u_1, u_2, \dots, u_n, u_n.$$

This list spans V as we have shows. But there are more u's which are not in the list, in particular u_{n+1}. If we now add this vecotr to the list, by lemma 3.38, the list is linearly dependent. But this contradicts the linear independence of $u_1, \dots, u_{n+1}$. □

In fact, we proved a bit more than Lemma 3.39, namely: We can remove m vectors from the list $v_1, \dots, v_n$ such that the remaining $n - m$ vectors $v_{i_1}, \dots, v_{i_{n-m}}$ when put together with $u_1, \dots, u_m$ still span V.

In particular, if $m = n$, then $\mathrm{Sp}(u_1, \dots, u_n) = V$.

Theorem 3.40:

Let V be a finite dimensional vector space over K. Then V has a finite basis. Moreover, every basis of V is finite and has the same number of elements.

Lemma 3.41:

Let $w_1, \dots, w_l \in V$ and assume $\forall 1 \le j \le l$, $w_j \notin \mathrm{Sp}(w_1, \dots, w_{j-1})$. Then $w_1, \dots, w_l$ are linearly independent.

Proof. If $w_1, \dots, w_l$ were linearly dependent, then by Lemma 3.36, $\exists 1 \le j \le l$ such that $w_j \in \mathrm{Sp}(w_1, \dots, w_{j-1})$, contradicting our assumption. □

Lemma 3.42:

Assume $\mathrm{Sp}(v_1, \dots, v_n) = V$. Then $\exists$ a subset of $\{v_1, \dots, v_n\}$ that is a basis of V.

Proof. We'll have n steps, and each step we'll have consider one of the vectors form $v_1, \dots, v_n$ and decide wether or not to drop it.

Step 1: If $v_1 = 0_V$, we drop it. If not, we keep it.

Step $2 \le j \le n$: If $v_j \in \mathrm{Sp}(v_1, \dots, v_{j-1})$, we drop it. Otherwise, we keep it. After performing n steps as above, we get a possibly shorter list $v_{i_1}, \dots, v_{i_m}$ with

$$1 \le i_1 < i_2 < \dots < i_m \le n.$$

We claim that $\mathrm{Sp}(v_{i_1}, \dots, v_{i_m}) = V$. Indeed, in any of the steps $1 \le j \le n$, we dropped the vector v_j only if $v_j \in \mathrm{Sp}(v_1, \dots, v_{j-1})$. From this it follows that dropping v_j does not change the overall span.

$$\mathrm{Sp}(v_{i_1}, \dots, v_{i_m}) = \mathrm{Sp}(v_1, \dots, v_n) = V.$$

We now claim, that $v_{i_1}, \dots, v_{i_m}$ are linearly independent. Indeed, $\forall 1 \le k \le m$ we have $v_{i_k} \notin \mathrm{Sp}(v_{i_1}, \dots, v_{i_k - 1})$. Thus in particular,

$$v_{i_k} \notin \mathrm{Sp}(v_{i_1}, \dots, v_{i_{k-1}}).$$

By Lemma 3.41 applied to the list $v_{i_1}, \dots, v_{i_m}$, we deduce that they are linearly independent. So $v_{i_1}, \dots, v_{i_m}$ is a basis of V. □

At last, we can prove Theorem 3.40:

Proof. [of Theorem 3.40] We need to prove three things:

1) V has a finite basis. Take any finite set $S \subseteq V$, such that $\mathrm{Sp}(S) = V$ (this is possible by definition of finite dimensional). By Lemma 3.42, there is a subset of S' of S that is a basis of V.

2) Every basis of V has finitely many elements. Let $\mathcal{C}$ be any basis for V. We'll show that $\mathcal{C}$ is a finite set. Suppose by contradiction that $\mathcal{C}$ is infinite. Also, take any finite basis, say S, for V (we know such a basis exists by 1), say $S = \{v_1, \dots, v_n\}$. Choose $n + 2025$ vectors, $u_1, \dots, u_{n+2025} \in \mathcal{C}$. By Lemma 3.39, we have that

$$n + 2025 \le n.$$

This is a contradiction. Thus, every basis of V is finite.

3) Any two bases $\mathcal{A}, \mathcal{B}$ of V have the same number of elements. By statement 2), both $\mathcal{A}$ and $\mathcal{B}$ are finite sets. Now $\mathrm{Sp}(\mathcal{B}) = V$ and $\mathcal{A}$ is linearly independent. By Lemma 3.39, we have that

$$|\mathcal{A}| \le |\mathcal{B}|.$$

But interchanging the roles of $\mathcal{A}$ and $\mathcal{B}$, we also have that

$$|\mathcal{B}| \le |\mathcal{A}|.$$

Thus, $|\mathcal{A}| = |\mathcal{B}|$. □

This inspires the following definition:

Definition 3.43: Dimension

Let V be a finite dimensional vector space over K. We define the DIMENSION of V to be the unique number $n \in \mathbb{Z}_{\ge 0}$ of elements in any basis of V. We write

$$\dim(V) = n.$$

Sometimes, people also write $\dim_K(V)$ to emphasize the field K.

Example 3.44:

1) $\dim\{0_V\} = 0$.

2) $\dim(K^n) = n$.

3) $\dim(K[x]_d) = d + 1$.

4) $\dim(M_{m \times n}(K)) = m \cdot n$.

As a summary of the algorithm we developped, we have: Let V be a finite dimensional vector space over K. Then:

1) Every finite list of vectors that spans V will contain a sublist which is a basis of V.

2) Every linearly independent list of vectors in V can be extended to get a basis of V. Lec 14

Theorem 3.45:

Let V be a finite dimensional vector space of dimension $n := \dim(V)$. The follwoing statements are equivalent:

1. $v_1, \dots, v_n \in V$ are linearly independent.
2. $v_1, \dots, v_n$ span V.
3. $v_1, \dots, v_n$ form a basis of V.

Proof. $1 \Rightarrow 3$: By what was done last time, we can extend the list $v_1, \dots, v_n$ and get a basis for V. But this list has already n vectors in it. So there is no extension needed. Thus, $v_1, \dots, v_n$ is a basis of V.

$3 \Rightarrow 1\&2$: By definition of basis.

$2 \Rightarrow 3$: By again, what we did last time, we can possibly drop some vectors from the list $v_1, \dots, v_n$ and get a basis for V. But the length of the list $v_1, \dots, v_n$ is already n. So dropping is impossible. Thus $v_1, \dots, v_n$ was a basis already from the beginning. □

Theorem 3.46:

Suppose V is finite dimensional and $n = \dim(V)$. Let $v_1, \dots, v_k \in V$ be a list. Then:

1. If $k < n$ then $v_1, \dots, v_k$ do not span V.
2. If $k > n$ then $v_1, \dots, v_k$ are linearly dependent.

Proof. 1) Assume $k < n$. Suppose by contradiction that this list spans V. As we have seen, we can thin this list to a sublist which is a basis of V. But any basis of V has n elements. So we have a contradiction.

2) Assume $k > n$. Assume by contradiction that $v_1, \dots, v_k$ are linearly independent. By what we did last time, we can extend this list to a basis of V. But any basis of V has n elements. So we have a contradiction. □

Proposition 3.47: Let V be a finite dimensional vector space. Then every subspace $U \subseteq V$ is also finite dimensional and we have

$$\dim(U) \leq \dim(V).$$

Moreover, $\dim(U) = \dim(V) \Leftrightarrow U = V$.

Proof. Let $n = \dim(V)$. We'll show first that U is finite dimensional.

If $U = \{0_V\}$, then $\dim(U) = 0$ and we are done. Assume now, that $U \neq \{0_V\}$. Take any vector $v_1 \in U \setminus \{0_V\}$. If $U = \mathrm{Sp}(v_1)$, then $\dim(U) = 1$ and we are done.

If $U \neq \mathrm{Sp}(v_1)$, take $v_2 \in U \setminus \mathrm{Sp}(v_1)$. By lemma 3.41, v_1, v_2 are linearly independent.

We continue this process and after j steps, we obtain a list $v_1, \ldots, v_j \in U$ that are linearly independent.

However, this process must stop after at most n steps, since otherwise we would have a list of $n+1$ linearly independent vectors in V, contradicting the fact that $\dim(V) = n$.

For the equality case, conversly suppose $k = n$. Thus the list of vectors $v_1, \ldots, v_n$ is a basis for U, hence these are linearly independent vectors. They continue to be linearly independent even when viewed as vectors in V. By the above theorem, $v_1, \ldots, v_n$ also span V. Thus $U = \mathrm{Sp}(v_1, \ldots, v_n) = V$. □

3.6 Row and Column Spaces

Consider K^m. Let $v_1, \ldots, v_n \in K^m$. The follwoing questions arise:

1. Let $w \in K^m$. How to determine whether or not $w \in \mathrm{Sp}(v_1, \ldots, v_n)$?
2. How do we describe $\mathrm{Sp}(v_1, \ldots, v_n)$?
3. How to determine if $v_1, \ldots, v_n$ are linearly independent?

Let us start with question 1).

For $w \in \mathrm{Sp}(v_1, \ldots, v_n)$, this is equivalent to the existence of $a_1, \ldots, a_n \in K$ such that

$$w = a_1 v_1 + \cdots + a_n v_n.$$

We can write this as

$$\begin{pmatrix} | & | & \cdots & | \\ v_1 & v_2 & \cdots & v_n \\ | & | & \cdots & | \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} | \\ w \\ | \end{pmatrix}. \tag{3.2}$$

So the question is, wether or not the system of equations (3.2) with the unknowns $a_1, \ldots, a_n$ has a solution.

For the second question, apply answer 1) to a general $b = w$. So do elimination to the augmented matrix $(A|b)$, untill we get a reduced row echelon matrix $(A'|b')$. And $A'x = b'$ has a solution iff $Ax = b$ has a solution.

In case A' does not have any 0-rows (i.e each row of A' has a pivot). In this case, $A'x = b'$ will always have a solution, which implies that $\mathrm{Sp}(v_1, \ldots, v_n) = K^m$.

If A' does have some 0-rows, then the entries in b' corresponding to these 0-rows must also be 0 for $A'x = b'$ to have a solution. Thus

$$b \in \mathrm{Sp}(v_1, \ldots, v_n) \Leftrightarrow b'_{r+1} = b'_{r+2} = \cdots = b'_m = 0.$$

Example 3.48: Take the following in $\mathbb{R}^3$:

$$v_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}.$$

The span of v_1, v_2 is given by all vectors $b \in \mathbb{R}^3$ such that

$$\begin{pmatrix} 1 & 1 & b_1 \\ 1 & 0 & b_2 \\ 1 & -1 & b_3 \end{pmatrix} \to \begin{pmatrix} 1 & 1 & b_1 \\ 0 & 1 & b_1 - b_2 \\ 0 & 0 & b_1 - 2b_2 + b_3 \end{pmatrix}.$$

From this, we see that for $\mathrm{Sp}(v_1, v_2)$, we must have

$$\mathrm{Sp}(v_1, v_2) = \left\{ \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} \in \mathbb{R}^3 \mid b_1 - 2b_2 + b_3 = 0 \right\}.$$

For question three, we see that $v_1, \ldots, v_n$ are linearly independent iff the system of equations

$$\begin{pmatrix} | & | & \cdots & | \\ v_1 & v_2 & \cdots & v_n \\ | & | & \cdots & | \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} | \\ 0_V \\ | \end{pmatrix}.$$

Has only the trivial solution $a_1 = a_2 = \cdots = a_n = 0$. So Take $A = (v_1 \cdots v_n)$, do elimination and get A'. Lec 15

In other words, $v_1, \ldots, v_n$ are linearly independent iff the reduced row echelon form A' of A has percisely n pivots (i.e no zero rows). So the matrix is of the form

$$A' = \begin{pmatrix} 1 & * & * & \cdots & * \\ 0 & 1 & * & \cdots & * \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \end{pmatrix}.$$

Let us revisit theorem 3.46. Take a list $v_1, \ldots, v_n \in K^m$, where $n < m$. Form the matrix

$$A = \begin{pmatrix} | & | & \cdots & | \\ v_1 & v_2 & \cdots & v_n \\ | & | & \cdots & | \end{pmatrix}.$$

We can have at most n pivots but $m > n$. When trying to solve $Ax = b$, we will get at least one zero row in the reduced row echelon form of A. Now, the span is given by

$$\mathrm{Sp}(v_1, \ldots, v_n) = \{b \in K^m \mid b'_{n+1} = b'_{n+2} = \cdots = b'_m = 0\}.$$

Assume now that $n > m$. Why is $v_1, \ldots, v_n$ linearly dependent? Well, we need to check if the system

$$\begin{pmatrix} | & | & \cdots & | \\ v_1 & v_2 & \cdots & v_n \\ | & | & \cdots & | \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} | \\ 0_V \\ | \end{pmatrix}$$

has non-trivial solutions. After elimination, we get $A'x = 0$ and A' can at most have m pivots, since there are only m rows. But $n > m$, so there are at least $n - m$ free variables. Thus, there are non-trivial solutions.

Definition 3.49: Row and Column Spaces

Let $A \in M_{m\times n}(K)$. Denote by $u_1, \dots, u_m \in K^n$ the rows of A and by $v_1, \dots, v_n \in K^m$ the columns of A.

$$A = \begin{pmatrix} - & u_1 & - \\ - & u_2 & - \\ & \vdots & \\ - & u_m & - \end{pmatrix} = \begin{pmatrix} | & | & & | \\ v_1 & v_2 & \cdots & v_n \\ | & | & & | \end{pmatrix}.$$

Define two spaces:

$$\mathrm{RowS}(A) := \mathrm{Sp}(u_1, \dots, u_m) \subseteq K^n$$
$$\mathrm{ColS}(A) := \mathrm{Sp}(v_1, \dots, v_n) \subseteq K^m.$$

Where we call $\mathrm{RowS}(A)$ the ROW SPACE of A and $\mathrm{ColS}(A)$ the COLUMN SPACE of A.

Lemma 3.50:

If $A, B \in M_{m\times n}(K)$ are row equivalent, then

$$\mathrm{RowS}(A) = \mathrm{RowS}(B).$$

Notice that it is NOT true that $\mathrm{ColS}(A) = \mathrm{ColS}(B)$

Proof. Let $M \in M_{m\times m}(K)$.

1) When we exchange two rows of M, the span of the rows does not change. So $\mathrm{RowS}(M)$ does not change under this operation.

2) The operation $\lambda R_i \to R_i$ for $\lambda \in K \setminus \{0\}$. Here too, the span of the rows does not change. So $\mathrm{RowS}(M)$ does not change under this operation.

3) The operation $R_i + \lambda R_j \to R_i$ for $\lambda \in K, i \neq j$. We need to show $\mathrm{Sp}(R_1, \dots, R_m) = \mathrm{Sp}(R_1, \dots, R_i + \lambda R_j, \dots, R_m)$.

Indeed, $\mathrm{Sp}(R_1, \dots, R_i + \lambda R_j, \dots, R_m) \subseteq \mathrm{Sp}(R_1, \dots, R_m)$ since $R_i + \lambda R_j \in \mathrm{Sp}(R_1, \dots, R_m)$.

We also have $\mathrm{Sp}(R_1, \dots, R_m) \subseteq \mathrm{Sp}(R_1, \dots, R_i + \lambda R_j, \dots, R_m)$. Indeed, $R_i = (R_i + \lambda R_j) - \lambda R_j$, and since $i \neq j$, R_j is in the span on the right hand side.

So we conclude that after one elementary row operation on M, the $\mathrm{RowS}(M)$ is unchanged.

If A and B are row equivalent, then $\exists$ a finite sequence of matrices $A = A_0, A_1, \dots, A_k = B$ such that each A_{i+1} is obtained from A_i by an elementary row operation. By the above,

$$\mathrm{RowS}(A_0) = \mathrm{RowS}(A_1) = \cdots = \mathrm{RowS}(A_k) = \mathrm{RowS}(B).$$

□

Definition 3.51:

Let $A \in M_{m\times n}(K)$. The ROW-RANK of A is defined as

$$\text{row-rank}(A) := \dim(\mathrm{RowS}(A))$$
$$\text{col-rank}(A) := \dim(\mathrm{ColS}(A)).$$

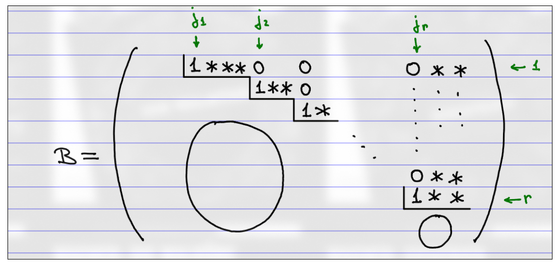
Lemma 3.52:

Let $B \in M_{m\times n}(K)$ which is in row reduced echelon form. Then the rows of B that are not totally 0 form a basis of $\mathrm{RowS}(B)$. In particular, $\text{row-rank}(B)$ is equal to the number of pivots in B.

Also the pivot columns of B form a basis of $\mathrm{ColS}(B)$. In particular, for matrices B that are in row reduced echelon form, we have

$$\text{col-rank}(B) = \text{row-rank}(B).$$

Proof. Consider the following matrix:



Denote by $j_1 < \cdots < j_r$ the column numbers of the pivots. Denote by $u_1, \dots, u_r$ the non-zero rows of B. We claim that $u_1, \dots, u_r$ are linearly independent.

Indeed, assume that $\lambda_1 u_1 + \cdots + \lambda_r u_r = 0$. This is a vector in K^n. Entry number j_k in this vector is precisely λ_k, since u_k has a pivot in column j_k. But if we assume that $\lambda_1 u_1 + \cdots + \lambda_r u_r = 0$, then all entries of this vector are 0. In particular, $\lambda_k = 0$ for all $k = 1, \dots, r$. Thus, $u_1, \dots, u_r$ are linearly independent. This shows our claim.

But $\mathrm{RowS}(B) = \mathrm{Sp}(u_1, \dots, u_r)$ by definition.

By Thm 3.45, $u_1, \dots, u_r$ form a basis of $\mathrm{RowS}(B)$. Also we get that $\text{row-rank}(B) = r$.

For the column space, clearly

$$\mathrm{ColS}(B) \subseteq \{x \in K^m \mid x_{r+1} = \dots x_m = 0\} = K^r \times \{0\}^{m-r}.$$

At the same time, the pivot columns of B are of the form

$$\begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ * \end{pmatrix}.$$

So $e_1, \dots, e_r$ form a basis for $K^r \times \{0\}$. Hence we have, $K^r \times \{0\} = \mathrm{Sp}(e_1, \dots, e_r) \subseteq \mathrm{ColS}(B)$.

Thus, $\mathrm{ColS}(B) = K^r \times \{0\}$ and the pivot columns of B form a basis of $\mathrm{ColS}(B)$. In particular, $\text{col-rank}(B) = r = \text{row-rank}(B)$.

□ Lec 16

We thus have found an algorithm to find a basis for the span of a list of vectors $u_1, \dots, u_m \in K^n$, viewed as rows. Just put them as rows of a matrix $A \in M_{m\times n}(K)$, row reduce A to B, and take the non-zero rows of B as a basis of $\mathrm{Sp}(u_1, \dots, u_m)$.

$$A = \begin{pmatrix} - & u_1 & - \\ - & u_2 & - \\ & \vdots & \\ - & u_m & - \end{pmatrix}.$$

Remark 3.53:

The number of pivots one gets after elimination does NOT depend on the specific elimination process because the number of pivots is equal to the dimension of the row space, which is an invariant of the matrix.

3.6.1 Transpose of a Matrix

Sometimes its useful to switch between rows and columns of a matrix.

Definition 3.54: **Transposed Matrix**

Let $A = (a_{ij}) \in M_{m\times n}(K)$. The TRANSPOSED MATRIX of A, is

$$A^T := (b_{ij}) \in M_{n\times m}(K) \text{ with } b_{ij} = a_{ji}.$$

Sometimes also A^t is written.

Example 3.55:

Let

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

Then

$$A^T = \begin{pmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{pmatrix}.$$

Simple properties of the transpose: $\forall A, B \in M_{m\times n}(K)$ and $\lambda \in K$:

1. $(A+B)^T = A^T + B^T$
2. $(\lambda A)^T = \lambda A^T$
3. $(A^T)^T = A$
4. $\mathrm{RowS}(A^T) = \mathrm{ColS}(A)$ and $\mathrm{ColS}(A^T) = \mathrm{RowS}(A)$

3.7 Sums of Vector Spaces

Let V be a vector space over K.

Definition 3.56: **Sum of Subspaces**

Let $U, W \subseteq V$ be subspaces. We define the SUM as

$$U + W := \{u + w \mid u \in U, w \in W\} \subseteq V.$$

In similar fashion, we can also take sums of more subspaces.

Proposition 3.57:

Let $U, W \subseteq V$ be subspaces. Then

1. $U + W = \mathrm{Sp}(U \cup W)$. In particular, $U + W$ is a subspace of V.
2. Suppose U, W are finite-dimensional. Then $U + W$ is also finite-dimensional and one can write a basis for $U + W$ as follows:
   - Choose a basis $p_1, \dots, p_k$ for $U \cap W$ where $k = \dim(U \cap W)$.
   - Extend this to a basis $p_1, \dots, u_1, \dots, u_{l-k}$ of U, where $l = \dim(U)$.
   - Extend $p_1, \dots, p_k$ to a basis of W say $p_1, \dots, p_k, w_1, \dots, w_{m-k}$, where we have $m = \dim(W)$.
   - $p_1, \dots, p_k, u_1, \dots, u_{l-k}, w_1, \dots, w_{m-k}$ are a basis of $U + W$.
3. In particular

   $$\dim(U+W) = \dim(U) + \dim(W) - \dim(U \cap W).$$

Proof. 1) Clearly, $U + W \subseteq \mathrm{Sp}(U \cup W)$. We'll show that $\mathrm{Sp}(U \cup W) \subseteq U + W$.

Indeed, let $v \in \mathrm{Sp}(U \cup W)$. Then $v = \sum_{i=1}^{s} a_i u_i + \sum_{j=1}^{r} b_j w_j$, where $a_i, b_j \in K$, $u_i \in U$ and $w_j \in W$. But the first sum belongs to U and the second to W because U, W are subspaces. Thus, $v \in U + W$.

Together this shows $U + W = \mathrm{Sp}(U \cup W)$.

2) We first show that

$$S := \{p_1, \dots, p_k, u_1, \dots, u_{l-k}, w_1, \dots, w_{m-k}\}$$

are linearly independent. Indeed, if

$$\sum_{i=1}^{k} a_i p_i + \sum_{j=1}^{l-k} b_j u_j + \sum_{t=1}^{m-k} c_t w_t = 0_V.$$

Write $v := c_1 w_1 + \dots c_{m-k} w_k \in W$. Note that

$$v = -(\sum_{i=1}^{k} a_i p_i + \sum_{j=1}^{l-k} b_j u_j) \in U.$$

So $v \in U \cap W$. Since $p_1 \dots, p_k$ is a basis of $U \cap W$, $\exists$ coefficients $\alpha_1, \dots, \alpha_k \in K$ such that

$$v = \sum_{i=1}^{k} \alpha_i p_i.$$

Furthermore, since $p_1, \dots, p_k, w_1, \dots, w_{m-k}$ is a basis of W, Thus,

$$\sum_{i=1}^{k} \alpha_i p_i - \sum_{t=1}^{m-k} c_t w_t = 0_V$$
$$\Rightarrow \alpha_1 = \alpha_2 = \dots = \alpha_k = 0$$

.

But $p_1, \dots, p_k, u_1, \dots, u_{l-k}$ is a basis of U, so

$$a_1 = \dots = a_k = b_1 = \dots = b_{l-k} = 0.$$

This shows that the p vectors with the u and w vectors are linearly independent. It remains to show that they span $U + W$.

Indeed let $z \in U + W$. Write $z = u + w$. We can write

$$u = \alpha_1 p_1 + \dots + \alpha_k p_k + a_1 u_1 + \dots + a_{l-k} u_{l-k}$$
$$w = \beta_1 p_1 + \dots + \beta_k p_k + b_1 w_1 + \dots + b_{m-k} w_{m-k}.$$

Then,

$$\begin{aligned} z = (\alpha_1+\beta_1)p_1 + \cdots + (\alpha_k+\beta_k)p_k \\ + a_1u_1 + \cdots + a_{l-k}u_{l-k} \\ + b_1w_1 + \cdots + b_{m-k}w_{m-k}. \end{aligned}$$

But this definetly belongs to $\mathrm{Sp}(S)$. Thus S is a basis of $U+W$.

3) We know

$$\begin{aligned} \dim(U+W) &= k + (l-k) + (m-k) \\ &= l + m - k \\ &= \dim(U) + \dim(W) - \dim(U\cap W). \end{aligned}$$

□

Corollary 3.58:

Let $U, W \subseteq V$ be finite-dimensional subspaces. Then the following are equivalent:

1. $\dim(U+W) = \dim(U) + \dim(W)$
2. $\dim(U\cap W) = 0$
3. $U \cap W = \{0_V\}$
4. Every $v \in U+W$ can be written uniquely as $v = u + w$ with $u \in U$ and $w \in W$.
5. If $u + w = 0_V$, then $u = 0_V$ and $w = 0_V$.

Proof. $1 \Leftrightarrow 2$: Follows from the dimesion formula in Proposition 3.57.

$2 \Leftrightarrow 3$: Again using proposition 3.57 and the fact that there is only one space of dimension 0, namely $\{0_V\}$.

$3 \Rightarrow 4$: By the definition 3.56, every $v \in U+W$ *can* be written as $v = u+w$ with $u \in U$ and $w \in W$. We need to show now the uniqueness. Assume $v = u + w = u' + w'$ with $u, u' \in U$ and $w, w' \in W$. Then $u - u' = w' - w$. But the left hand side belongs to U and the right hand side to W. Thus, $u - u' \in U \cap W$. By 3), $u - u' = 0_V$, so $u = u'$. Consequently, $w = w'$.

$4 \Rightarrow 5$: $0 = 0 + 0$, so if $0 = u + v$ and 4 holds, then $u = 0$ and $v = 0$.

$5 \Rightarrow 3$: Let $v \in U \cap W$. Clearly $0 = v + (-v)$ with $v \in U$, $-v \in W$. By 5), $v = 0_V$. □

Definition 3.59: Complement of a Subspace

Let $U \subseteq V$ be a subspace. A subspace $W \subseteq V$ is called a COMPLEMENT of U in V if

$$U + W = V \text{ and } U \cap W = \{0_V\}.$$

Lec 17

Notice there are in general many complements of a given subspace. There is not a canonical choice.

Proposition 3.60:

Let V be a finite-dimensional vector space over K and $U \subseteq V$ a subspace. Then there exists a subspace $W \subseteq V$ which is a complement of U in V.

Proof. Choose a basis $u_1, \ldots, u_l$ of U, where $l = \dim(U)$. Extend this list to a basis of V, say

$$u_1, \ldots, u_l, w_1, \ldots, w_m.$$

So $l+m = \dim(V)$. Take $W := \mathrm{Sp}(w_1, \ldots, w_m)$. Clearly, $U+W = V$ because $U + W = \mathrm{Sp}(u_1, \ldots, u_l, w_1, \ldots, w_m) = V$.

Also $U \cap W = \{0_V\}$ because if $\sum_{i=1}^{l} a_i u_i = \sum_{j=1}^{m} b_j w_j$, then

$$\sum_{i=1}^{l} a_i u_i - \sum_{j=1}^{m} b_j w_j = 0_V.$$

But since $u_1, \ldots, u_l, w_1, \ldots, w_m$ is a basis of V, they are linearly independent, so all coefficients are 0.

Alternatively, by corollary 3.58, we can compute

$$\dim(U\cap W) = \dim(U) + \dim(W) - \dim(U+W) = l + m - (l+m) = 0.$$

□

4 Linear Maps

Let V, W be vector spaces over a field K.

A linear map is a function between vector spaces that respects the structure of the vector spaces. Other names include, linear transformations or homomorphisms of vector spaces.

Definition 4.1: Linear Map

A map $T : V \to W$ is called LINEAR if

1. $\forall u, v \in V : T(u+v) = T(u) + T(v)$
2. $\forall v \in V, \forall a \in K : T(a \cdot v) = a \cdot T(v)$

The set of all linear maps $T : V \to W$ is denoted by $\mathrm{Hom}(V, W)$ or $\mathrm{Hom}_K(V, W)$.

Sometimes, we write $Tv = T(v)$. The reason, V and W are over the same field K is that the definition would not make sense otherwise.

Exercise 4.2:

Show that $T : V \to W$ is linear iff

$$T(au + bv) = aT(u) + bT(v) \quad \forall u, v \in V, \forall a, b \in K.$$

Example 4.3:

1) $\mathrm{id} : V \to V$ is a linear map.

2) $0 : V \to W, v \mapsto 0_W$ is a linear map.

3) $D : K[x] \to K[x]$, the (formal) derivative, i.e.

$$D(a_0 + a_1 x + \cdots + a_n x^n) = a_1 + 2a_2 x + \cdots + n a_n x^{n-1}$$

is a linear map as

$$\begin{aligned} D(ap(x) + bq(x)) &= ap'(x) + bq'(x) \\ &= aD(p(x)) + bD(q(x)). \end{aligned}$$

4) Let $C([a, b])$ be the vector space of continuous functions $f : [a, b] \to \mathbb{R}$. Then the map

$$T : C([a, b]) \to \mathbb{R}, \quad f \mapsto \int_a^b f(x)\, dx,$$

is linear.

5) $S : K^\infty \to K^\infty, (a_1, a_2, \dots) \mapsto (a_2, a_3, \dots)$ called the SHIFT map is linear.

Example 4.4: Important Example

Let $A \in M_{m \times n}(K)$. Define

$$T_A : K^n_{\text{col}} \to K^m_{\text{col}}, \quad x \mapsto Ax.$$

We showed that $A \cdot (a \cdot x) = a \cdot (Ax)$ and $A \cdot (x + y) = Ax + Ay$. Thus, T_A is linear.

Proposition 4.5:

Let $T : V \to W$ be a linear map. Then:

1. $\forall n \in \mathbb{Z}_{\geq 0}, \forall v_1, \dots, v_n \in V, \forall a_1, \dots, a_n \in K$:
$$T\left(\sum_{i=1}^n a_i v_i\right) = \sum_{i=1}^n a_i T(v_i).$$
2. $T(0_V) = 0_W$

Proof. 1) We prove this by induction on n.

Base case: $n = 2$: This is just the definition 4.1. Induction step: Assume true for n. We show for $n + 1$:

$$\begin{aligned} T\left(\sum_{i=1}^{n+1} a_i v_i\right) &= T\left(\sum_{i=1}^n a_i v_i + a_{n+1} v_{n+1}\right) \\ &= T\left(\sum_{i=1}^n a_i v_i\right) + T(a_{n+1} v_{n+1}) \\ &= \sum_{i=1}^n a_i T(v_i) + a_{n+1} T(v_{n+1}) \\ &= \sum_{i=1}^{n+1} a_i T(v_i). \end{aligned}$$

2) Note that

$$T(0_V) = T(0 \cdot v) = 0 \cdot T(v) = 0_W.$$

□

Theorem 4.6:

Let V be finite-dimensional, and let $v_1, \dots, v_n$ be a basis of V. Then $\forall w_1, \dots, w_n \in W$, $\exists!$ linear map $T : V \to W$ such that $T(v_i) = w_i$.

In other words, we can define a linear map by defining where the basis vectors go.

Proof. Existence of T: Let $v \in V$. We want to define $T(v)$. Since $v_1, \dots, v_n$ is a basis of V, $\exists\ a_1, \dots, a_n \in K$ such that

$$v = a_1 v_1 + a_2 v_2 + \cdots + a_n v_n. \tag{4.1}$$

Define $T(v) \coloneqq a_1 w_1 + a_2 w_2 + \cdots + a_n w_n$. Note that T is well-defined because $\forall v \in V$, the coefficients a_i are unique. We claim that T is linear.

Indeed, let $u, v \in V$. Then $\exists a_i, b_i \in K$ such that

$$u = \sum_{i=1}^n a_i v_i, \quad v = \sum_{i=1}^n b_i v_i.$$

Now,

$$u + v = \sum_{i=1}^n (a_i + b_i) v_i.$$

But also by definition of T by (4.1):

$$\begin{aligned} T(u + v) &= \sum_{i=1}^n (a_i + b_i) w_i \\ &= \sum_{i=1}^n a_i w_i + \sum_{i=1}^n b_i w_i \\ &= T(u) + T(v). \end{aligned}$$

Let now $\alpha \in K, v \in V$ with $v = \sum_{i=1}^n a_i v_i$. Then,

$$\begin{aligned} T(\alpha v) &= T\left(\sum_{i=1}^n \alpha a_i v_i\right) \overset{(4.1)}{=} \sum_{i=1}^n \alpha a_i w_i \\ &= \alpha \sum_{i=1}^n a_i w_i = \alpha T(v). \end{aligned}$$

This proves that T is linear. Note that by definition, $T(v_i) = w_i$. We now have to show uniqueness of T.

Suppose $T, S : V \to W$ are linear maps such that $T(v_i) = S(v_i) = w_i$. We'll show that $T = S$. Let $v \in V$. Write $v = \sum_{i=1}^n a_i v_i$. Then,

$$\begin{aligned} T(v) = T\left(\sum_{i=1}^n a_i v_i\right) &\overset{\text{Prop. 4.5}}{=} \sum_{i=1}^n a_i T(v_i) \\ &= \sum_{i=1}^n a_i w_i \\ &= \sum_{i=1}^n a_i S(v_i) \overset{\text{Prop. 4.5}}{=} S\left(\sum_{i=1}^n a_i v_i\right) = S(v). \end{aligned}$$

□

Example 4.7:

Consider K^n with the standard basis $e_1, \dots, e_n$. Let $w_1, \dots, w_n \in K^m$ be n arbitrary vectors. By Theorem 4.6, $\exists!$ linear map $T : K^n \to K^m$ such that $T(e_i) = w_i$. We can describe T as $T = T_A$ where

$$A = \begin{pmatrix} | & | & & | \\ w_1 & w_2 & \cdots & w_n \\ | & | & & | \end{pmatrix}.$$

Recall that $T_A : K^n \to K^m$ is linear. Indeed, for e_i,

$$T_A(e_i) = A \cdot e_i = w_i.$$

So we have constructed a linear map T_A such that $T_A(e_i) = w_i$. By uniqueness in Theorem 4.6, $T = T_A$.

Lemma 4.8:

$\forall$ linear map $T : K^n \to K^m$, $\exists! A \in M_{m\times n}(K)$ such that $T = T_A$.

In fact,

$$A = \begin{pmatrix} | & | & & | \\ T(e_1) & T(e_2) & \cdots & T(e_n) \\ | & | & & | \end{pmatrix}.$$

Lec 18

4.1 Kernel and Image

Let $T : V \to W$ be a linear map.

Definition 4.9: Kernel and Image

The set

$$\mathrm{Ker}(T) := \{v \in V \mid T(v) = 0_W\} \subseteq V$$

is called the KERNEL of T.

The set

$$\mathrm{Im}(T) := \{Tv \mid v \in V\} \subseteq W$$

is called the IMAGE of T.

bWith the notation from set theory, we have

$$\mathrm{Ker}(T) = T^{-1}(\{0_W\}) \quad \text{and} \quad \mathrm{Im}(T) = T(V).$$

Proposition 4.10:

The kernel $\mathrm{Ker}(T)$ is a subspace of V and the image $\mathrm{Im}(T)$ is a subspace of W.

Proof. 1) Let $a, b \in K$ and $u, v \in \mathrm{Ker}(T)$. Then

$$T(av + bu) = aT(v) + bT(u) = a \cdot 0_W + b \cdot 0_W = 0_W.$$

Hence, $av + bu \in \mathrm{Ker}(T)$ and since $0_V \in \mathrm{Ker}(T)$, $\mathrm{Ker}(T)$ is a subspace of V.

2) Let $a, b \in K$ and $w_1, w_2 \in \mathrm{Im}(T)$. Then $\exists v_1, v_2 \in V$ such that

$$T(v_1) = w_1, \quad T(v_2) = w_2.$$

Now $T(av_1 + bv_2) = aT(v_1) + bT(v_2) = aw_1 + bw_2$. Thus $aw_1 + bw_2 \in \mathrm{Im}(T)$. Also $0_W = T(0_V) \in \mathrm{Im}(T)$. Hence, $\mathrm{Im}(T)$ is a subspace of W. □

Example 4.11:

Let $A \in M_{m\times n}(K)$ and $T_A : K^n \to K^m$ be a linear map $T_A(x) = Ax$. Then $\mathrm{Ker}(T_A) = \{x \in K^n \mid Ax = 0\}$ is the solution space of the homogeneous system of linear equations associated to A.

Proposition 4.12:

The following are equivalent for a linear map $T : V \to W$:

1. T is injective $\Leftrightarrow \mathrm{Ker}(t) = \{0_V\}$
2. T is surjective $\Leftrightarrow \mathrm{Im}(T) = W$

Proof. 2) This is true by definition of surjectivity and image. 1) $\Rightarrow$: Clearly $0_V \in \mathrm{Ker}(T)$. So $\{0_V\} \subseteq T^{-1}(\{0_W\})$. Since T is injective, $T^{-1}(\{0_W\})$ contains at most one element. Thus, $\mathrm{Ker}(T) = \{0_V\}$.

$\Leftarrow$: Let $u, v \in V$ such that $T(u) = T(v)$. Then we have

$$T(u) - T(v) = 0_W \Rightarrow T(u - v) = 0_W \Rightarrow u - v \in \mathrm{Ker}(T).$$

But by assumption, $\mathrm{Ker}(T) = \{0_V\}$, so $u - v = 0_V \Rightarrow u = v$ which shows the injectivity. □

Exercise 4.13:

Show that

1. If $V' \subseteq V$ is a subspace, then $T(V') \subseteq W$ is a subspace.
2. If $W' \subseteq W$ is a subspace, then $T^{-1}(W') \subseteq V$ is a subspace.
3. Explain why 1 & 2 are generalisations of Proposition 4.10.

Recall that $f : X \to Y$ is called bijective if it is both injective and surjective. In this case, $\exists!$ inverse map

$$g : Y \to X \quad \text{such that} \quad g \circ f = \mathrm{id}_X, \quad f \circ g = \mathrm{id}_Y.$$

The other direction of this also holds: If $\exists g : Y \to X$ such that $g \circ f = \mathrm{id}_X$ and $f \circ g = \mathrm{id}_Y$, then f is bijective.

Definition 4.14: Isomorphism

A linear map $T : V \to W$ is called an ISOMORPHISM if $\exists$ a linear map $S : W \to V$ such that

$$S \circ T = \mathrm{id}_V, \quad T \circ S = \mathrm{id}_W .$$

We say that V is isomorphic to W if $\exists$ an isomorphism $T : V \to W$ and write $V \cong W$.

Any isomorphism is obviously bijective. The other direction also holds:

Lemma 4.15:

Let $T : V \to W$ be a linear map. Then T is an isomorphism iff T is bijective.

Proof. Let $a, b \in K, w_1, w_2 \in W$. Put $v_1 = T^{-1}(w_1)$ and $v_2 = T^{-1}(w_2)$. Since T is linear we get

$$T(av_1 + bv_2) = aT(v_1) + bT(v_2) = aw_1 + bw_2.$$

But then

$$T^{-1}(aw_1 + bw_2) = av_1 + bv_2 = aT^{-1}(w_1) + bT^{-1}(w_2).$$

Hence, T^{-1} is linear and thus T is an isomorphism. □

In short, for linear maps, isomorphism = bijection.

Lemma 4.16:

Let $T : V \to W, S : W \to U$ be linear maps. Then the composition

$$S \circ T : V \to U, \quad v \mapsto S(T(v))$$

is also a linear map.

Proof. Let $a, b \in K, u, v \in V$. Then

$$\begin{aligned} S \circ T(au + bv) &= S(T(au + bv)) \\ &= S(aT(u) + bT(v)) \\ &= aS(T(u)) + bS(T(v)) \\ &= a(S \circ T)(u) + b(S \circ T)(v). \end{aligned}$$

□

Exercise 4.17:

Show that $\cong$ defines an equivalence relation on the set of vector spaces over K.

Definition 4.18: Endomorphism and Automorphism

A linear map $T : V \to V$ (both domain and target are V) is called an ENDOMORPHISM of V.

We denote the set of all endomorphisms of V by

$$\mathrm{End}(V) := \mathrm{Hom}(V, V).$$

A linear map $T : V \to V$ which is an isomorphism is called an AUTOMORPHISM of V.

Sometimes, if T is injective, we call it a MONOMORPHISM, and if T is surjective, we call it an EPIMORPHISM.

Lemma 4.19:

Let $v_1, \dots, v_n$ be a basis of V. Then $\mathrm{Im}(T) = \mathrm{Sp}(T(v_1), \dots, T(v_n))$.

Proof. We first show $Im(T) \supseteq \mathrm{Sp}(T(v_1), \dots, T(v_n))$. Let $a_1, \dots, a_n \in K$. Then

$$\sum_{i=1}^{n} a_i T(v_i) = T\left(\sum_{i=1}^{n} a_i v_i\right) \in \Im(T).$$

Thus $\mathrm{Sp}(T(v_1), \dots, T(v_n)) \subseteq \Im(T)$.

For the other inclusion, let $w \in \Im(T)$. Then $\exists v \in V$ such that $T(v) = w$. Since $v_1, \dots, v_n$ is a basis of V, $\exists a_1, \dots, a_n \in K$

$$v = \sum_{i=1}^{n} a_i v_i \Rightarrow w = T(v) = T\left(\sum_{i=1}^{n} a_i v_i\right) = \sum_{i=1}^{n} a_i T(v_i).$$

Thus $w \in \mathrm{Sp}(T(v_1), \dots, T(v_n))$ and hence this inclusion holds as well. □

Theorem 4.20: Rank Theorem

Let V be finite-dimensional and $T : V \to W$ a linear map. Then

$$\dim(V) = \dim(\mathrm{Ker}(T)) + \dim(\mathrm{Im}(T)).$$

Proof. Let $n := \dim V$ Let $u_1, \dots, u_k$ be a basis of $\mathrm{Ker}(T)$, where $k = \dim(\mathrm{Ker}(T))$. Extend this to a basis of V (Since the kernel is a subspace of V its also finite-dimensional). We extend it to a basis of V:

$$u_1, \dots, u_k, v_1, \dots, v_{n-k}.$$

By Lemma 4.19, we have

$$\begin{aligned} \mathrm{Im}(T) &= \mathrm{Sp}(\underbrace{T(u_1)}_{=0}, \dots, \underbrace{T(u_k)}_{=0}, T(v_1), \dots, T(v_{n-k})) \\ &= \mathrm{Sp}(T(v_1), \dots, T(v_{n-k})). \end{aligned}$$

We claim that $T(v_1), \dots, T(v_{n-k})$ form a basis for $\mathrm{Im}(T)$. Indeed we just proved that theese vectors span $\mathrm{Im}(T)$. We now show linear independence.

Let $a_1, \dots, a_{n-k} \in K$ and assume $\sum_{i=1}^{n-k} a_i T(v_i) = 0$. Since T is linear,

$$T(\sum_{i=1}^{n-k} a_i v_i) = \sum_{i=1}^{n-k} a_i T(v_i) = 0.$$

Thus, $\sum_{i=1}^{n-k} a_i v_i \in \mathrm{Ker}(T)$. But $u_1, \dots, u_k$ is a basis of $\mathrm{Ker}(T)$, so $\exists b_1, \dots, b_k \in K$ such that

$$a_1 v_1 + \dots + a_{n-k} v_{n-k} = b_1 u_1 + \dots + b_k u_k$$
$$\Rightarrow a_1 v_1 + \dots + a_{n-k} v_{n-k} - b_1 u_1 - \dots - b_k u_k = 0.$$

But since $u_1, \dots, u_k, v_1, \dots, v_{n-k}$ is a basis of V, they must be linearly independent. Thus, all coefficients are 0, which proves that $T(v_1), \dots, T(v_{n-k})$ are linearly independent.

We can see that $\dim(\mathrm{Im}(T)) = n - k = \dim(V) - \dim(\mathrm{Ker}(T))$. which proves the theorem. □ Lec 19

Corollary 4.21:

Let $T : V \to W$ be a linear map between finite-dimensional vector spaces. Then the following holds

- If $\dim W < \dim V$, then T is not injective.
- If $\dim V < \dim W$, then T is not surjective.
- If $\dim V = \dim W$, then

$$T \text{ injective } \Leftrightarrow T \text{ surjective } \Leftrightarrow T \text{ bijective}.$$

Proof. 1) Since $\mathrm{Im}(T) \subseteq W$, we have

$$\dim(\mathrm{Im}(T)) \leq \dim(W) < \dim(V).$$

By the rank theorem (Theorem 4.20),

$$\dim(\mathrm{Ker}(T)) = \dim(V) - \dim(\mathrm{Im}(T)) > 0.$$

But this means that $\mathrm{Ker}(T) \neq \{0_V\}$, so by Proposition 4.12, T is not injective.

2) Again by the rank theorem,

$$\dim(\mathrm{Im}(T)) = \dim(V) - \dim(\mathrm{Ker}(T)) \leq \dim(V) < \dim(W).$$

Since the image has strictly smaller dimension than W, we have $\mathrm{Im}(T) \subsetneq W$ and thus T is not surjective.

3) T is injective $\Leftrightarrow$ $\dim(\mathrm{Ker}(T)) = 0$. But this is equivalent to $\dim(\mathrm{Ker}(T)) = 0$. By the rank theorem, this is equivalent to $\dim(\mathrm{Im}(T)) = \dim(W)$, which is equivalent to T being surjective.

□

Corollary 4.22:

Two finite-dimensional vector spaces V and W are isomorphic iff $\dim(V) = \dim(W)$.

Proof. $\Rightarrow$: Suppose $T : V \to W$ is an isomorphism. Then T is bijective, so by Proposition 4.12,

$$\mathrm{Ker}(T) = \{0_V\} \Rightarrow \dim(\mathrm{Ker}(T)) = 0,$$

and

$$\mathrm{Im}(T) = W \Rightarrow \dim(\mathrm{Im}(T)) = \dim(W).$$

By the rank theorem, we have

$$\dim(W) = \dim(V) - \dim(\mathrm{Ker}(T)) = \dim(V) - 0 = \dim(V).$$

$\Leftarrow$: Assume $\dim(V) = \dim(W)$. Denote by n their common dimension.

Let $v_1, \ldots, v_n$ be a basis for V and $w_1, \ldots, w_n$ be a basis for W.

Define a linear map $T : V \to W$ by $T(v_i) = w_i$ for $i = 1, \ldots, n$. By Theorem 4.6, such a linear map exists and is unique. By Lemma 4.19, we have

$$\mathrm{Im}(T) = \mathrm{Sp}(T(v_1), \ldots, T(v_n)) = \mathrm{Sp}(w_1, \ldots, w_n) = W.$$

Hence, T is surjective. By Corollary 4.21, we have that T is also injective, and thus bijective, i.e. isomorphic. □

Theorem 4.23:

Let T be an isomorphism. Let S be a set of vectors from V. Write

$$T(s) := \{T(v) \mid v \in S\}.$$

Then

1) S is linearly independent iff $T(S)$ is linearly independent.

2) S spans V iff $T(S)$ spans W.

3) S is a basis of V iff $T(S)$ is a basis of W.

Given $V \cong W$, then V and W have the same algebraic properties. We can think of V and W as the same space represented in two different ways. For all practical purposes, V and W are identical.

Now, unfortunately, there is no *canonical* isomorphism between two vector spaces and none is preferred over another in general. Since this is the case, it is therefore better to look at all of them.

So for example, $K[x]_3 \cong K^4$, but there is no preferred isomorphism between them. A more extreme example is the vector space

$$\left\{ \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \mid a_{11} + a_{22} = 0 \right\}.$$

This is a 3-dimensional vector space over K, so it is isomorphic to K^3, but there is no obvious isomorphism between them.

Definition 4.24: Rank of a Linear Map

We define the RANK of a linear map $T : V \to W$ as

$$\mathrm{rank}(T) := \dim(\mathrm{Im}(T)).$$

Exercise 4.25:

Let $T : V \to W$, $S : W \to U$ be linear maps where U, V, W are finite-dimensional. Show that

1. $\mathrm{rank}(S \circ T) \leq \min(\mathrm{rank}(S), \mathrm{rank}(T))$.
2. If S is injective, then $\mathrm{rank}(S \circ T) = \mathrm{rank}(T)$.
3. If T is surjective, then $\mathrm{rank}(S \circ T) = \mathrm{rank}(S)$.

4.2 Linear Maps and coordinates

In this section, we will always assume that all vector spaces are finite-dimensional. We will write a basis as an ordered tuple $\mathcal{B} = (v_1, \ldots, v_n)$.

Definition 4.26: Coordinate Vector

Let V be a vector space over K and $\mathcal{B} = (v_1, \ldots, v_n)$ a basis for V. Let $v \in V$. Define the COORDINATE VECTOR of v with respect to the basis $\mathcal{B}$ as

$$[v]_{\mathcal{B}} := \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} \in K^n.$$

Where $a_1, \ldots, a_n \in K$ are the unique scalars such that

$$v = a_1 v_1 + a_2 v_2 + \cdots + a_n v_n.$$

Define a map $\Phi_{\mathcal{B}} : V \to K^n$ by

$$\Phi_{\mathcal{B}}(v) = [v]_{\mathcal{B}} \quad \forall v \in V.$$

Proposition 4.27:

The map $\Phi_{\mathcal{B}} : V \to K^n$ is an isomorphism.

Proof. Linearity of $\Phi_{\mathcal{B}}$. Let $a, b \in K$ and $u, v \in V$. Now write v and u as a linear combination of the basis vectors:

$$v = \sum_{i=1}^{n} a_i v_i, \quad a_i \in K \Rightarrow [v]_{\mathcal{B}} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$$

$$u = \sum_{i=1}^{n} b_i v_i, \quad b_i \in K.$$

Now it holds that

$$\begin{aligned}
av+bu &= a\sum_{i=1}^{n} a_i v_i + b\sum_{i=1}^{n} b_i v_i \\
&= \sum_{i=1}^{n} (aa_i + bb_i) v_i \\
\Rightarrow [av+bu]_{\mathcal{B}} &= \begin{pmatrix} aa_1+bb_1 \\ aa_2+bb_2 \\ \vdots \\ aa_n+bb_n \end{pmatrix} = a\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} + b\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} \\
&= a[v]_{\mathcal{B}} + b[u]_{\mathcal{B}}.
\end{aligned}$$

Now in the $\Phi_{\mathcal{B}}$ definition, we have

$$\Phi_{\mathcal{B}}(av+bu) = [av+bu]_{\mathcal{B}} = a[v]_{\mathcal{B}} + b[u]_{\mathcal{B}} = a\Phi_{\mathcal{B}}(v) + b\Phi_{\mathcal{B}}(u).$$

This shows linearity. We now show that $\Phi_{\mathcal{B}}$ is an isomorphism.

$\Phi_{\mathcal{B}}$ is a map between two vector spaces of the same dimension n. Hence, we will show that $\mathrm{Ker}(\Phi_{\mathcal{B}}) = \{0_V\}$ to prove injectivity.

Indeed, if $v \in \mathrm{Ker}(\Phi_{\mathcal{B}})$, then

$$[v]_{\mathcal{B}} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

But this means that $v = 0_V$. Thus, $\mathrm{Ker}(\Phi_{\mathcal{B}}) = \{0_V\}$ which implies injectivity. By Corollary 4.21, $\Phi_{\mathcal{B}}$ is also surjective, and thus an isomorphism. □

Lec 20

Remark 4.28:
The inverse $\Phi_{\mathcal{B}}^{-1} : K^n \to V$ of $\Phi_{\mathcal{B}}$ can be written as:

$$\Phi_{\mathcal{B}}^{-1}\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = a_1 v_1 + a_2 v_2 + \cdots + a_n v_n.$$

Example 4.29:
Consider $V = \mathbb{R}^2$ and let $\xi = \left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}\right)$ be the standard basis of $\mathbb{R}^2$. Let $\mathcal{B} = \left(\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix}\right)$ be another basis. Let $v = \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2$. Then

$$[v]_{\xi} = \begin{pmatrix} x \\ y \end{pmatrix}, \quad [v]_{\mathcal{B}} = \begin{pmatrix} \frac{x+y}{2} \\ \frac{x-y}{2} \end{pmatrix}.$$

But how can we do the second equality quickly? We can consider the first the basis vectors of ξ as linear combinations of the basis vectors of $\mathcal{B}$: Write $[v_1]_{\mathcal{B}} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$ Hence, by a system of equations, we have

$$a_1\begin{pmatrix} 1 \\ 1 \end{pmatrix} + a_2\begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \Rightarrow a_1 = \frac{1}{2}, a_2 = \frac{1}{2}.$$

Similarly, for v_2 we have

$$b_1\begin{pmatrix} 1 \\ 1 \end{pmatrix} + b_2\begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \Rightarrow b_1 = \frac{1}{2}, b_2 = -\frac{1}{2}.$$

Since the map is linear, we find the desired result.

Let V, W be finite-dimensional vector spaces over K. Let $n := \dim V, m := \dim W$. Let T be a linear map from V to W. Fix a basis $\mathcal{B}$ for V and a basis $\mathcal{C}$ for W.

$$\begin{array}{ccc}
V & \xrightarrow{\ T\ } & W \\
{\scriptstyle \Phi_{\mathcal{B}}^{-1}}\uparrow & & \downarrow{\scriptstyle \Phi_{\mathcal{C}}} \\
K^n & \xrightarrow{\ \Phi_{\mathcal{C}}\cdot T\cdot \Phi_{\mathcal{B}}^{-1}\ } & K^m
\end{array}$$

Since it does not matter which path we take from K^n to W, we say that the diagram COMMUTES.

Recall that composition of linear maps is linear. Hence, $\Phi_{\mathcal{C}} \cdot T \cdot \Phi_{\mathcal{B}}^{-1} : K^n \to K^m$ is a linear map.

By previous knowledge (4.7), $\exists! A \in M_{m\times n}(K)$ such that

$$\Phi_{\mathcal{C}} \cdot T \cdot \Phi_{\mathcal{B}}^{-1}(x) = T_A.$$

Where $T_A : K^n \to K^m$ is the linear map defined by $T_A x := A \cdot x$.

The matrix A is called the REPRESENTATION of T with respect to the bases $\mathcal{B}$ and $\mathcal{C}$. It is denoted by $[T]_{\mathcal{C}}^{\mathcal{B}}$.

Now, how do we calculate $[T]_{\mathcal{C}}^{\mathcal{B}}$? Put$A := [T]_{\mathcal{C}}^{\mathcal{B}}$.

$$\begin{array}{ccc}
V & \xrightarrow{\ T\ } & W \\
{\scriptstyle \Phi_{\mathcal{B}}^{-1}}\uparrow & & \downarrow{\scriptstyle \Phi_{\mathcal{C}}} \\
K^n & \xrightarrow{\ T_A\ } & K^m
\end{array}$$

Let now $e_i \in K^n$ then under the map $\Phi_{\mathcal{B}}^{-1}$, it becomes v_i, the i-th basis vector of $\mathcal{B}$. Applying T to v_i gives $T(v_i) \in W$. Under $\Phi_{\mathcal{B}}$, this becomes $[T(v_i)]_{\mathcal{C}} \in K^m$.

In other words,

$$A = [T]_{\mathcal{C}}^{\mathcal{B}} = \begin{pmatrix} | & & | \\ [T(v_1)]_{\mathcal{C}} & \cdots & [T(v_n)]_{\mathcal{C}} \\ | & & | \end{pmatrix}.$$

Another way to describe this $[T]_{\mathcal{C}}^{\mathcal{B}}$ is the following: Write $[T]_{\mathcal{C}}^{\mathcal{B}} = (a_{ij})$. The entries a_{ij} are uniquely defined by:

$$T(v_j) = \sum_{i=1}^{m} a_{ij} w_i \quad \text{for } j = 1, \ldots, n.$$

(Here, $\mathcal{C} = (w_1, \ldots, w_m)$.)

Proposition 4.30:
Let $T : V \to W$ be a linear map and let $\mathcal{B}$, $\mathcal{C}$ be bases for V and W, respectively. Put $A = [T]_{\mathcal{C}}^{\mathcal{B}}$. Then $\forall v \in V$,

$$[T(v)]_{\mathcal{C}} = A \cdot [v]_{\mathcal{B}}.$$

Equivalently: $[Tv]_{\mathcal{C}} = [T]_{\mathcal{C}}^{\mathcal{B}} \cdot [v]_{\mathcal{B}}$.

Proof.

$$\begin{array}{ccc}
V & \xrightarrow{\ T\ } & W \\
{\scriptstyle \Phi_{\mathcal{B}}}\downarrow & & \downarrow{\scriptstyle \Phi_{\mathcal{C}}} \\
K^n & \xrightarrow{\ T_A\ } & K^m
\end{array}$$

This diagram commutes again. Going around the top right gives us

$$\Phi_{\mathcal{C}}(T(v)) = [T(v)]_{\mathcal{C}}.$$

Which is equal to

$$[T(v)]_{\mathcal{C}} = T_A([v]_{\mathcal{B}}) = A \cdot [v]_{\mathcal{B}}.$$

□

Example 4.31:

Let $V = K^n, W = K^m$, $T : V \to W$ be a linear map with $T = T_Q$. Take $\mathcal{B}$ and $\mathcal{C}$ to be the standard bases of K^n and K^m, then $[T]^{\mathcal{B}}_{\mathcal{C}} = Q$.

Example 4.32:

Let $V = K[x]_3$, $W = K[x]_2$ and let $D : V \to W$ with $D(f) = f'$ be the differentiation map. Take $\mathcal{B} = (1, x, x^2, x^3)$ and $\mathcal{C} = (1, x, x^2)$ as bases of V and W, respectively. We now want to find $[D]^{\mathcal{B}}_{\mathcal{C}}$.

We have

$$\begin{aligned} D(1) &= 0 = 0 \cdot 1 + 0 \cdot x + 0 \cdot x^2 \\ D(x) &= 1 = 1 \cdot 1 + 0 \cdot x + 0 \cdot x^2 \\ D(x^2) &= 2x = 0 \cdot 1 + 2 \cdot x + 0 \cdot x^2 \\ D(x^3) &= 3x^2 = 0 \cdot 1 + 0 \cdot x + 3 \cdot x^2. \end{aligned}$$

Hence we find

$$[D]^{\mathcal{B}}_{\mathcal{C}} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}.$$

4.3 Matrices

Let $T : V \to W$ and $S : W \to U$ be linear maps between finite-dimensional vector spaces. Let $\mathcal{A} = (v_1, \ldots, v_n)$ be a basis for V, $\mathcal{B} = (w_1, \ldots, w_m)$ a basis for W and $\mathcal{C} = (u_1, \ldots, u_p)$ a basis for U.

Consider $S \circ T : V \to U$. We want to find the relation between $[S \circ T]^{\mathcal{A}}_{\mathcal{C}}$, $[T]^{\mathcal{A}}_{\mathcal{B}}$ and $[S]^{\mathcal{B}}_{\mathcal{C}}$.

Write $[T]^{\mathcal{A}}_{\mathcal{B}} = A = (a_{ij})$ and $[S]^{\mathcal{B}}_{\mathcal{C}} = B = (b_{ij})$. Furthermore, write $[S \circ T]^{\mathcal{A}}_{\mathcal{C}} = C = (c_{ij})$.

We know that

$$T(v_j) = a_{1j}w_1 + a_{2j}w_2 + \cdots + a_{mj}w_m.$$

Furthermore

$$S(w_j) = b_{1j}u_1 + b_{2j}u_2 + \cdots + b_{pj}u_p.$$

Lastly,

$$S \circ T(v_j) = c_{1j}u_1 + c_{2j}u_2 + \cdots + c_{pj}u_p.$$

But $(S \circ T)(v_j) = S(T(v_j))$, so we have

$$\begin{aligned} S(T(v_j)) = S\left(\sum_{i=1}^m a_{ij}w_i\right) &= \sum_{i=1}^m a_{ij}S(w_i) \\ = \sum_{i=1}^m a_{ij}\left(\sum_{k=1}^p b_{ki}u_k\right) &= \sum_{k=1}^p \left(\sum_{i=1}^m b_{ki}a_{ij}\right) u_k. \end{aligned}$$

Lec 21 So the coefficient $c_{kj} = \sum_{i=1}^m b_{ki}a_{ij}$.

So

$$[S \circ T]^{\mathcal{A}}_{\mathcal{C}} = \begin{pmatrix} \cdots & c_{1j} & \cdots \\ \cdots & c_{2j} & \cdots \\ & \vdots & \\ \cdots & c_{pj} & \cdots \end{pmatrix}.$$

If we write the matrices A and B we get

$$[S \circ T]^{\mathcal{A}}_{\mathcal{C}} = \begin{pmatrix} \cdots & \cdots & \cdots \\ b_{i1} & \cdots & b_{im} \\ \cdots & \cdots & \cdots \end{pmatrix} \begin{pmatrix} \cdots & a_{1j} & \cdots \\ \cdots & a_{2j} & \cdots \\ & \vdots & \\ \cdots & a_{mj} & \cdots \end{pmatrix}.$$

Definition 4.33: **Matrix Multiplication**

Let $A \in M_{m\times n}(K)$ and $B \in M_{n\times p}(K)$. Write $A = (a_{ij})$ and $B = (b_{ij})$. Define a new matrix $C \in M_{m\times p}(K)$ with entries (c_{ij}), where

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

We denote this new matrix C by $A \cdot B$ and call it the MULTIPLICATION or PRODUCT of A and B.

It is important that the number of columns of A equals the number of rows of B.

Another way to describe $A \cdot B$ is the following:

$$(A) \cdot \begin{pmatrix} | & \cdots & | \\ z_1 & \cdots & z_p \\ | & \cdots & | \end{pmatrix} = \begin{pmatrix} | & \cdots & | \\ Az_1 & \cdots & Az_p \\ | & \cdots & | \end{pmatrix}.$$

Example 4.34:

Compute

$$\begin{pmatrix} 3 & 2 & 1 \\ 1 & 0 & 2 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 \\ 0 & 1 \\ 4 & 0 \end{pmatrix}.$$

We have

$$A \cdot B = \begin{pmatrix} 7 & 8 \\ 9 & 2 \end{pmatrix}.$$

In this example, we can also compute the other way around and get:

$$B \cdot A = \begin{pmatrix} 5 & 2 & 5 \\ 1 & 0 & 2 \\ 12 & 8 & 4 \end{pmatrix}.$$

This example tells us that in general, matrix multiplication is not commutative, i.e. $A \cdot B \neq B \cdot A$ in general.

Let $A \in M_{n\times n}(K)$. Let $v \in K^n_{\text{col}}$. We define Av previously. Viewing v as a $n \times 1$ matrix, we can see that this is just a special case of matrix multiplication.

Using the same idea, we can define multiplication of a $1 \times n$ matrix with a $n \times n$ matrix. Let $w \in K^{1\times n}$ and $A \in M_{n\times n}(K)$. Then we define

$$w \cdot A = \begin{pmatrix} \sum_{k=1}^n w_{1k}a_{k1} & \cdots & \sum_{k=1}^n w_{1k}a_{kn} \end{pmatrix}.$$

which is a $1 \times n$ matrix.

So matrix multiplication can also be described like

$$\begin{pmatrix} \cdots & \cdots & \cdots \\ - & v_i & - \\ \cdots & \cdots & \cdots \end{pmatrix} \cdot (B) = \begin{pmatrix} \cdots & \cdots & \cdots \\ - & v_i \cdot B & - \\ \cdots & \cdots & \cdots \end{pmatrix}.$$

Conclusion: $[S \circ T]^{\mathcal{A}}_{\mathcal{C}} = [S]^{\mathcal{B}}_{\mathcal{C}} \cdot [T]^{\mathcal{A}}_{\mathcal{B}}$.

As a notation, let I be an index set. $\forall i, j \in I$, define

$$\delta_{ij} := \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}.$$

This is called the KRONECKER DELTA.

Define the identity matrix $I_n \in M_{n\times n}(K)$ as

$$I_n = (\delta_{ij}) = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}.$$

Remark 4.35:

Let v be a finite dimensional vector space with dimension n. Let $\mathcal{B}$ be any basis of V. Consider $\mathrm{id}_V : V \to V$. Then $[\mathrm{id}_V]^{\mathcal{B}}_{\mathcal{B}} = I_n$.

Proposition 4.36:

1) $\forall A \in M_{m\times n}(K)$, $B \in M_{n\times p}(K)$, $C \in M_{p\times q}(K)$, we have

$$A \cdot (B \cdot C) = (A \cdot B) \cdot C.$$

Hence, there is a meaning to write $A \cdot B \cdot C$ without parentheses.

2) $\forall A, B \in M_{m\times n}, C \in M_{n\times p}, C' \in M_{q\times m}$, we have

$$(A + B) \cdot C = A \cdot C + B \cdot C$$
$$C' \cdot (A + B) = C' \cdot A + C' \cdot B.$$

3) $\forall A \in M_{m\times n}(K)$, we have

$$I_m \cdot A = A = A \cdot I_n.$$

4) $\forall \alpha \in K, A \in M_{m\times n}(K), B \in M_{n\times p}(K)$, we have

$$(\alpha A) \cdot B = \alpha(A \cdot B) = A \cdot (\alpha B).$$

Proof. 1) Let $T_A : K^n \to K^m$, $T_B : K^p \to K^n$ and $T_C : K^q \to K^p$ be the linear maps defined by A, B and C, respectively. $\forall l$, denote by E_l the standard basis for K^l.

$$[T_A]^{E_n}_{E_m} = A, \quad [T_B]^{E_p}_{E_n} = B, \quad [T_C]^{E_q}_{E_p} = C.$$

We have

$$(T_A \circ T_B) \circ T_C = T_A \circ (T_B \circ T_C),$$

since composition of functions is associative. Hence

$$[(T_A \circ T_B) \circ T_C]^{E_q}_{E_m} = [T_A \circ (T_B \circ T_C)]^{E_q}_{E_m}.$$

By what we did earlier,

$$[T_A \circ T_B]^{E_q}_{E_m} \cdot [T_C]^{E_q}_{E_p} = [T_A]^{E_n}_{E_m} \cdot [T_B \circ T_C]^{E_q}_{E_n}.$$

But we can expand this further:

$$([T_A]^{E_n}_{E_m} \cdot [T_B]^{E_p}_{E_n}) \cdot [T_C]^{E_q}_{E_p} = [T_A]^{E_n}_{E_m} \cdot ([T_B]^{E_p}_{E_n} \cdot [T_C]^{E_q}_{E_p}).$$

But this is exactly what we wanted to prove.

2) If $T, S : V \to W$ are linear maps, $\mathcal{B}$ is a basis for V and $\mathcal{C}$ is a basis for W, then

$$[T + S]^{\mathcal{B}}_{\mathcal{C}} = [T]^{\mathcal{B}}_{\mathcal{C}} + [S]^{\mathcal{B}}_{\mathcal{C}}.$$

From this, the result follows directly. 3) Look at the definition of I_n in the standard basis. From this, the result follows directly.

4) The outline of the proof is to let $A = (a_ij)$ and $B = (b_ij)$ and compute both sides. □ Lec 22

Remark 4.37:

Let V be a vector space, take $\alpha \in K$. And let $Q(v) = \alpha v$. Then

$$[Q]^{\mathcal{B}}_{\mathcal{B}} = \begin{pmatrix} \alpha & 0 & \cdots & 0 \\ 0 & \alpha & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \alpha \end{pmatrix}.$$

Definition 4.38: Invertible Matrix

Let A_n be a $n \times n$ matrix. We say that A is INVERTIBLE if $\exists B_n$ such that

$$A \cdot B = I_n B \cdot A.$$

Remark 4.39:

If A is invertible, then B st. $A \cdot B = B \cdot A = I_n$ is unique.

Proof. Assume that B, C are both inverses of A. Then

$$B = B \cdot I_n = B \cdot (A \cdot C) = (B \cdot A) \cdot C = I_n \cdot C = C.$$

□

Definition 4.40:

Let A be an invertible matrix. We denote by A^{-1} the unique matrix such that

$$A \cdot A^{-1} = A^{-1} \cdot A = I_n.$$

A^{-1} is called the INVERSE of A.

We denote by $\mathrm{GL}_n(K)$ the set of all invertible $n \times n$ matrices over K. It is called the GENERAL LINEAR GROUP of degree n over K.

Notice that $\mathrm{GL}_n(K) \neq \varnothing$ since $I_n \in \mathrm{GL}_n(K)$. Also, $0_n \notin \mathrm{GL}_n(K)$. Hence, $\mathrm{GL}_n(K) \subsetneq M_{n\times n}(K)$.

Proposition 4.41:

Let $A, B \in \mathrm{GL}_n(K)$. Then $A \cdot B \in \mathrm{GL}_n(K)$ and also $A^{-1} \in \mathrm{GL}_n(K)$. Moreover, $(A \cdot B)^{-1} = B^{-1} \cdot A^{-1}$ and $(A^{-1})^{-1} = A$.

Proof. It holds that

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

Also

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

Also

$$A \cdot A^{-1} = I_n = A^{-1} \cdot A.$$

So A is the inverse of A^{-1}. □

Corollary 4.42:

Let $A \in \mathrm{GL}_n(K)$. Then $\forall b \in K^n$, the linear system of equations $Ax = b$ has a unique solution $x \in K^n$, which is given by $x = A^{-1}b$.

Proof. If $Ax = b$, then multiplying both sides by A^{-1} gives

$$A^{-1}Ax = A^{-1}b \Rightarrow I_n x = A^{-1}b \Rightarrow x = A^{-1}b.$$

Also, if $x = A^{-1}b$, then $Ax = A(A^{-1}b) = (AA^{-1})b = I_n b = b$. □

Proposition 4.43:

Let $A \in M_{n\times n}(K)$. A is invertible if and only if $T_A : K^n \to K^n$ is an isomorphism. Moreover, $(T_A)^{-1} = T_{A^{-1}}$.

Proof. Assume first that A is invertible. We claim that $T_{A^{-1}} \circ T_A = \mathrm{id}_{K^n}$ and same for the other way around. This will show that T_A is an isomorphism with inverse $T_{A^{-1}}$.

Indeed, $\forall v \in K^n$, $T_{A^{-1}}(T_A(v)) = T_{A^{-1}}(Av) = A^{-1}(Av) = (A^{-1}A)v = I_n v = v$. Hence, $T_{A^{-1}} \circ T_A = \mathrm{id}_{K^n}$. Similarly, $T_A \circ T_{A^{-1}} = \mathrm{id}_{K^n}$.

Conversely, assume that T_A is an isomorphism. Denote $S := (T_A)^{-1}$. S is a linear map from K^n to K^n. Recall that $\forall F : K^n \to K^n$ linear, $\exists! M \in M_{n\times n}(K)$ such that $F = T_M$. So, $\exists B \in M_{n\times n}(K)$ such that $S = T_B$. Now $\forall v \in K^n$ we have:

$$v = S \circ T_A(v) = T_B(T_A(v)) = T_B(Av) = B(Av) = (BA)v.$$

Applying this to $v = e_1, v = e_2, \dots, v = e_n$ shows that $BA = I_n$. In a similar way, we show that $AB = I_n$. Hence, A is invertible with inverse B. □

Proposition 4.44:

1) Let $A \in M_{m\times n}(K), B \in M_{n\times p}(K)$. Then

$$(A \cdot B)^T = B^T \cdot A^T.$$

2) If $A \in \mathrm{GL}_n(K)$, then $A^T \in \mathrm{GL}_n(K)$ and

$$(A^T)^{-1} = (A^{-1})^T.$$

Definition 4.45: Triangular and Diagonal Matrices

A matrix $A \in M_{n\times n}(K)$ is called:

- UPPER TRIANGULAR if $a_{ij} = 0$ for all $i > j$.
- LOWER TRIANGULAR if $a_{ij} = 0$ for all $i < j$.
- DIAGONAL if $a_{ij} = 0$ for all $i \neq j$.

Lemma 4.46:

Let $A, B \in M_{n\times n}(K)$ be upper triangular(or lower triangular or diagonal) matrices. Then $A \cdot B$ is also upper triangular(or lower triangular or diagonal, respectively).

We can ask ourselves how $[T]^{\mathcal{B}}_{\mathcal{C}}$ depends on the choice of bases $\mathcal{B}$ and $\mathcal{C}$.

Corollary 4.47:

Let V, W be finite-dimensional vector spaces over K of dimensions n and m, respectively. Let $\mathcal{B}, \mathcal{B}'$ be bases for V and $\mathcal{C}, \mathcal{C}'$ be bases for W. Let $T : V \to W$ be a linear map. Then

$$[T]^{\mathcal{B}'}_{\mathcal{C}'} = [\mathrm{id}_W]^{\mathcal{C}}_{\mathcal{C}'} \cdot [T]^{\mathcal{B}}_{\mathcal{C}} \cdot [\mathrm{id}_V]^{\mathcal{B}'}_{\mathcal{B}}.$$

Proof. $T = \mathrm{id}_W \circ T \circ \mathrm{id}_V$. But this we can write as

$$\mathrm{id}_W \circ (T \circ \mathrm{id}_V).$$

This implies that

$$[T]^{\mathcal{B}'}_{\mathcal{C}'} = [\mathrm{id}_W]^{\mathcal{C}}_{\mathcal{C}'} \cdot [T \circ \mathrm{id}_V]^{\mathcal{B}'}_{\mathcal{C}} = [\mathrm{id}_W]^{\mathcal{C}}_{\mathcal{C}'} \cdot [T]^{\mathcal{B}}_{\mathcal{C}} \cdot [\mathrm{id}_V]^{\mathcal{B}'}_{\mathcal{B}}.$$

□

Corollary 4.48:

Furthermore $[id_V]^{\mathcal{B}'}_{\mathcal{B}} \in \mathrm{GL}_n(K)$ and $[\mathrm{id}_W]^{\mathcal{C}}_{\mathcal{C}'} \in \mathrm{GL}_m(K)$ and

$$([\mathrm{id}_V]^{\mathcal{B}'}_{\mathcal{B}})^{-1} = [\mathrm{id}_V]^{\mathcal{B}}_{\mathcal{B}'}, \quad ([\mathrm{id}_W]^{\mathcal{C}}_{\mathcal{C}'})^{-1} = [\mathrm{id}_W]^{\mathcal{C}'}_{\mathcal{C}}.$$

Proof. We know that $\mathrm{id}_V = \mathrm{id}_V \circ \mathrm{id}_V$. Hence,

$$I_n = [\mathrm{id}_V]^{\mathcal{B}}_{\mathcal{B}} = [\mathrm{id}_V]^{\mathcal{B}'}_{\mathcal{B}} \cdot [\mathrm{id}_V]^{\mathcal{B}}_{\mathcal{B}'}.$$

Similarly,

$$I_n = [\mathrm{id}_V]^{\mathcal{B}'}_{\mathcal{B}'} = [\mathrm{id}_V]^{\mathcal{B}}_{\mathcal{B}'} \cdot [\mathrm{id}_V]^{\mathcal{B}'}_{\mathcal{B}}.$$

Hence, $[\mathrm{id}_V]^{\mathcal{B}}_{\mathcal{B}'}$ is invertible with inverse $[\mathrm{id}_V]^{\mathcal{B}'}_{\mathcal{B}}$. □ Lec 23

Definition 4.49: Change of Basis Matrix

Let V be a finite-dimensional vector space with dimension n. Let $\mathcal{B}, \mathcal{B}'$ be bases for V. The matrix $[\mathrm{id}_V]^{\mathcal{B}}_{\mathcal{B}'} \in \mathrm{GL}_n(K)$ is called the CHANGE OF BASIS MATRIX between $\mathcal{B}$ and $\mathcal{B}'$, or the TRANSITION MATRIX between $\mathcal{B}$ and $\mathcal{B}'$.

If $\mathcal{B} = (v_1, \dots, v_n)$ and $\mathcal{B}' = (v'_1, \dots, v'_n)$, then

$$[\mathrm{id}_V]^{\mathcal{B}}_{\mathcal{B}'} = \begin{pmatrix} | & & | \\ [v_1]_{\mathcal{B}'} & \cdots & [v_n]_{\mathcal{B}'} \\ | & & | \end{pmatrix}.$$

Proposition 4.50:

Let $T : V \to W$ be an isomorphism between finite-dimensional vector spaces. Let $\mathcal{B}, \mathcal{C}$ be bases for V and W, respectively. Then $[T]^{\mathcal{B}}_{\mathcal{C}} \in \mathrm{GL}_n(K)$ and

$$\left([T]^{\mathcal{B}}_{\mathcal{C}}\right)^{-1} = \left[T^{-1}\right]^{\mathcal{C}}_{\mathcal{B}}.$$

Proof. Let $n := \dim V = \dim W$. We have

$$[T^{-1}]^{\mathcal{C}}_{\mathcal{B}} \cdot [T]^{\mathcal{B}}_{\mathcal{C}} = [T^{-1} \circ T]^{\mathcal{B}}_{\mathcal{B}} = [\mathrm{id}_V]^{\mathcal{B}}_{\mathcal{B}} = I_n.$$

Similarly,

$$[T]^{\mathcal{B}}_{\mathcal{C}} \cdot [T^{-1}]^{\mathcal{C}}_{\mathcal{B}} = [T \circ T^{-1}]^{\mathcal{C}}_{\mathcal{C}} = [\mathrm{id}_W]^{\mathcal{C}}_{\mathcal{C}} = I_n.$$

□

Consider now $T : V \to V$. Let $\mathcal{B}, \mathcal{B}'$ be bases for V. The question is, what's the relation between $[T]^{\mathcal{B}}_{\mathcal{B}}$ and $[T]^{\mathcal{B}'}_{\mathcal{B}'}$?

Corollary 4.51:

Let $T : V \to V$ be a linear map and let $\mathcal{B}, \mathcal{B}'$ be bases for V. Then

$$[T]^{\mathcal{B}'}_{\mathcal{B}'} = [\mathrm{id}_V]^{\mathcal{B}}_{\mathcal{B}'} \cdot [T]^{\mathcal{B}}_{\mathcal{B}} \cdot [\mathrm{id}_V]^{\mathcal{B}'}_{\mathcal{B}}.$$

where $[\mathrm{id}_V]^{\mathcal{B}}_{\mathcal{B}'} = ([\mathrm{id}_V]^{\mathcal{B}'}_{\mathcal{B}})^{-1}$.

Definition 4.52: Matrix Equivalence

1) Let $A, B \in M_{m\times n}(K)$. We say that A and B are EQUIVALENT if $\exists P \in \mathrm{GL}_m(K)$ and $Q \in \mathrm{GL}_n(K)$ such that
$$B = P \cdot A \cdot Q.$$

2) Let $A, B \in M_{n\times n}(K)$. We say that A and B are SIMILAR if $\exists P \in \mathrm{GL}_n(K)$ such that
$$B = P^{-1} \cdot A \cdot P.$$

Equivalence between matrices is an equivalence relation on the set of all $m \times n$ matrices over K.

Furthermore, $A, B \in M_{n\times n}(K)$ are equivalent if and only if $\exists$ two vector spaces V, W of dimension n, m, and bases $\mathcal{B}, \mathcal{B}'$ of V and $\mathcal{C}, \mathcal{C}'$ of W and a linear map $T : V \to W$ such that
$$A = [T]^{\mathcal{B}}_{\mathcal{C}}, \quad B = [T]^{\mathcal{B}'}_{\mathcal{C}'}.$$

Similarity enjoys the same properties, with $V = W$.

As a notation, the zero $p \times q$ matrix is denoted by $0_{p,q}$.

Proposition 4.53:

Let V, W be finite-dimensional vector spaces over K of dimensions n and m, respectively. Let $T : V \to W$ be a linear map with $\mathrm{rank}(T) = r$. Then $\exists$ bases $\mathcal{B}$ of V and $\mathcal{C}$ of W such that
$$[T]^{\mathcal{B}}_{\mathcal{C}} = \left(\begin{array}{c|c} I_r & 0_{r,n-r} \\ \hline 0_{m-r,r} & 0_{m-r,n-r} \end{array}\right) \in M_{m\times n}(K).$$

Proof. Put $l = \dim(\ker(T))$. By the rank theorem (4.20), we have $r = n - l$. Let $v_1, \ldots, v_l$ be a basis for $\ker(T)$. We can extend it to a basis $(v_1, \ldots, v_n)$ of V. We'll take
$$\mathcal{B} = (v_{l+1}, \ldots, v_n, v_1, \ldots, v_l).$$
In the proof of the rank theorem, we saw that
$$(T(v_{l+1}), \ldots, T(v_n))$$
is a basis for $\mathrm{Im}(T)$. Define $w_i := T(v_{l+i})$ for $i = 1, \ldots, r$. (note that $l + r = n$). We can extend $(w_1, \ldots, w_r)$ to a basis
$$\mathcal{C} = (w_1, \ldots, w_r, w_{r+1}, \ldots, w_m)$$
of W.

So now, we can construct the matrix $[T]^{\mathcal{B}}_{\mathcal{C}}$.
$$[T]^{\mathcal{B}}_{\mathcal{C}} = \begin{pmatrix} | & & | & | & & | \\ [T(v_{l+1})]_{\mathcal{C}} & \cdots & [T(v_n)]_{\mathcal{C}} & [T(v_1)]_{\mathcal{C}} & \cdots & [T(v_l)]_{\mathcal{C}} \\ | & & | & | & & | \end{pmatrix}.$$
Now $[T(v_i)]_{\mathcal{C}} = 0$ for $i = 1, \ldots, l$ since $v_i \in \ker(T)$. Also, for $j = 1, \ldots, r$, we have
$$T(v_{l+j}) = w_j = 1 \cdot w_j + 0 \cdot w_{r+1} + \cdots + 0 \cdot w_m.$$
So we can write $[T(v_{l+j})]_{\mathcal{C}}$ as desired. $\square$

Corollary 4.54:

Let $A \in M_{m\times n}(K)$. Then $\exists P \in \mathrm{GL}_m(K)$ and $Q \in \mathrm{GL}_n(K)$ such that
$$P \cdot A \cdot Q = \left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right).$$
where $r = \text{col-rank}(A)$.

Proof. Consier $T_A : K^n \to K^m$. Then $r := \text{col-rank}(A)$. From the definition of column rank, we have
$$\mathrm{ColS}(A) = \mathrm{Im}(T_A).$$
So also $r = \mathrm{rank}(T_A)$. By 4.53, $\exists$ bases $\mathcal{B}$ of K^n and $\mathcal{C}$ of K^m such that
$$[T_A]^{\mathcal{B}}_{\mathcal{C}} = \left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right).$$
We also know that
$$[T_A]^{\mathcal{B}}_{\mathcal{C}} = [\mathrm{id}_{K^m}]^{E_m}_{\mathcal{C}} \cdot A \cdot [\mathrm{id}_{K^n}]^{\mathcal{B}}_{E_n}.$$
If we now call $P := [\mathrm{id}_{K^m}]^{E_m}_{\mathcal{C}} \in \mathrm{GL}_m(K)$ and $Q := [\mathrm{id}_{K^n}]^{\mathcal{B}}_{E_n} \in \mathrm{GL}_n(K)$, we are done. $\square$

A consequence for this corollary is that $\text{col-rank}(A) \leq \min(m, n)$. This is because $r = \text{col-rank}(A)$ and A is $m \times n$.

Corollary 4.55:

Let $A, B \in M_{m\times n}(K)$. Then A and B are equivalent if and only if $\text{col-rank}(A) = \text{col-rank}(B)$.

Proof. Assume $\text{col-rank}(A) = \text{col-rank}(B) = r$. By 4.54,
$$A \sim \left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right) \sim B.$$
By transitivity of equivalence, we have $A \sim B$.

Assume now that $A \sim B$. Then $\exists P \in \mathrm{GL}_m(K)$ and $Q \in \mathrm{GL}_n(K)$ such that $B = P \cdot A \cdot Q$. Hence $T_B = T_P \circ T_A \circ T_Q$. Then
$$\begin{aligned} \mathrm{Im}(T_B) &= T_B(K^n) = T_P(T_A(T_Q(K^n))) \\ &= T_P(T_A(K^n)) = T_P(\mathrm{Im}(T_A)) \\ &= \mathrm{Im}(T_A). \end{aligned}$$
Hence, $\mathrm{rank}(T_B) = \mathrm{rank}(T_A)$, so $\text{col-rank}(B) = \text{col-rank}(A)$. $\square$

Theorem 4.56: Row Rank = Column Rank

Let $A \in M_{m\times n}(K)$. Then
$$\text{row-rank}(A) = \text{col-rank}(A).$$
This common number is denoted by $\mathrm{rank}(A)$ and is called the RANK of A.