Math 20250
Abstract Linear Algebra

Cong Hung Le Tran

April 24, 2023

Course: MATH 20250: Abstract Linear Algebra
Section: 44
Professor: Zijian Yao
At: The University of Chicago
Quarter: Spring 2023
Course materials: Linear Algebra by Hoffman and Kunze (2nd Edition), Linear Algebra Done Wrong by Treil

Disclaimer: This document will inevitably contain some mistakes, both simple typos and serious logical and mathematical errors. Take what you read with a grain of salt as it is made by an undergraduate student going through the learning process himself. If you do find any error, I would really appreciate it if you can let me know by email at conghungletran@gmail.com.

Contents

Lecture 1

Abelian Group, Field, Equivalence

21 Mar 2023

Goal

Vector spaces and maps between vector spaces (linear transformations)

1.1 Abelian Group

Definition 1.1 (Abelian Group)

A pair $(A, *)$ is an **Abelian group** if A is a set and $*$ is a map: $A \times A \mapsto A$ (closure is implied) with the following properties:

1. (Additive Associativity)
$$(x * y) * z = x * (y * z),\ \forall\, x, y, z \in A$$
2. (Additive Commutativity)
$$x * y = y * x,\ \forall\, x, y \in A$$
3. (Additive Identity)
$$\exists\, 0 \in A : 0 * x = x * 0 = x,\ \forall\, x \in A$$
4. (Additive Inverse)
$$\forall\, x \in A,\ \exists\, (-x) \in A : x * (-x) = (-x) * x = 0$$

Remark

($*$ is just a symbol, soon to be $+$). Typically write as $(A, +)$ or simply A

Example

1. $(\mathbb{Z}, +)$ is an Abelian group
2. $(\mathbb{Q}, +)$ is an Abelian group
3. $(\mathbb{Z}, \times)$ is **NOT** an Abelian group (because identity $= 1$, and 0 does not have a multiplicative inverse)
4. $(\mathbb{Q}, \times)$ is also not an Abelian group (0 does not have a multiplicative inverse)
5. $(\mathbb{Q}\backslash\{0\}, \times)$ is an Abelian group (identity is 1)
6. $(\mathbb{N}, \times)$ is NOT a group

Remark

A crucial difference between $\mathbb{Z}$ and $\mathbb{Q}\backslash\{0\}$ is that $\mathbb{Q}\backslash\{0\}$ has both $+$ and $\times$ while $\mathbb{Z}$ only has $+$. This gives us inspiration for the definition of a field!

Definition 1.2 (Field)
A **field** is a triple $(F, +, \cdot)$ s.t.

1. $(F, +)$ is an Abelian group with identity 0
2. (Multiplicative Associativity)
$$(x \cdot y) \cdot z = x \cdot (y \cdot z),\ \forall\, x, y, z \in F$$
3. (Multiplicative Commutativity)
$$x \cdot y = y \cdot x,\ \forall\, x, y \in F$$
4. (Distributivity) ($+$ and $\cdot$ talking in the following way)
$$x \cdot (y + z) = (x \cdot y) + (x \cdot z),\ \forall\, x, y, z \in F$$
5. (Multiplicative Identity)
$$\exists\, 1 \in F : 1 \cdot x = x,\ \forall\, x \in F$$
6. (Multiplicative Inverse)
$$\forall\, x \in F \backslash \{0\},\ \exists\, y \in F : x \cdot y = 1$$

Remark
In a field $(F, +, \cdot)$, assume that $1 \neq 0$

Example

1. $(\mathbb{Z}, +, \cdot)$ is not a field (because property 6 failed)
2. $(\mathbb{Q}, +, \cdot)$ is a field
3. $(\mathbb{R}, +, \cdot)$ and $(\mathbb{C}, +, \cdot)$ are fields.

1.2 Finite Fields

Recall
$p \in \mathbb{Z}$ is a prime if $\forall\, m \in \mathbb{N} : m \mid p \Rightarrow m = 1$ or $m = p$

Definition 1.3 ($\mathbb{F}_p$ for p prime)
$$\mathbb{F}_p = \{[0], [1], \ldots, [p-1]\}$$
Then define the operations for $[a], [b] \in \mathbb{F}_p$
$$[a] + [b] = [a + b \mod p]; [a] \cdot [b] = [a \cdot b \mod p]$$
Then $\mathbb{F}_p$ is a field, but this is not trivial.

Lemma 1.1

1. $(\mathbb{F}_p, +)$ is an Abelian group
2. $(\mathbb{F}_p, +, \cdot)$ is a field

Example

$\mathbb{F}_5 = \{[0], [1], [2], [3], [4]\}$

$$[1] + [2] = [3], [2] + [4] = [1], [4] + [4] = [3], [2] + [3] = [0]$$

Then it is trivial that [0] is additive identity, and every element has additive inverse. [1] is multiplicative identity, and every element except [0] has multiplicative inverse. Therefore $\mathbb{F}_5$ is indeed a field.

1.3 Vector Spaces in brief

Intuition

The motivation for vector spaces and maps between them (linear transformations) is essentially to solve linear equations. Let $(\mathbb{K}, +, \cdot)$ be a field. We are then interested in systems of linear equations / $\mathbb{K}$; if there are solutions, and if there are how many.

We then inspect a system of linear equations of n unknowns, m relations:

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

where $a_{ij}, b_k \in \mathbb{K}$.

Example

$$2x_1 - x_2 + x_3 = 0 \tag{1}$$
$$x_1 + 3x_2 + 4x_3 = 0 \tag{2}$$

over some field $\mathbb{K}$.

Explanation

Then, $3 \times (1) + (2)$ (carrying out the operations in $\mathbb{K}$) yields

$$\begin{aligned} 7x_1 + 7x_3 &= 0 \\ 7 \cdot (x_1 + x_3) &= 0 \end{aligned} \tag{3}$$

Then, we have 2 cases.
Case 1: $7 \neq 0$ in $\mathbb{K}$, then $\exists\, 7^{-1} \in \mathbb{K} : 7^{-1} \cdot 7 = 1$.

Then (3) $\Rightarrow 7^{-1} \cdot (7 \cdot (x_1 + x_3)) = 0$

$$\begin{aligned} ((7^{-1}) \cdot 7) \cdot (x_1 + x_3) &= 0 \\ 1 \cdot (x_1 + x_3) &= 0 \\ \Rightarrow x_1 + x_3 &= 0 \\ \Rightarrow x_1 &= -x_3 \end{aligned}$$

Let $x_3 = a \Rightarrow x_1 = -a \Rightarrow x_2 = 2x_1 + x_3 = -a$.
$\Rightarrow \{(-a, -a, a) \mid a \in \mathbb{K}\}$ are solutions.

Case 2: $7 = 0$ in $\mathbb{K}$ (e.g. in $\mathbb{F}_7$) then (3) is automatically true.
Let $x_1 = a, x_3 = b \Rightarrow x_2 = 2x_1 + x_3 = 2a + b$
$\Rightarrow \{(a, 2a + b, b) \mid a, b \in \mathbb{K}\}$ are solutions. □

Remark

When doing $3 \times (1) + (2)$, how do we know if we're gaining or losing information? e.g in $\mathbb{F}_7$ we can just multiply by 7 and get nothing new! Therefore some kind of "equivalence" concept must be introduced!

Definition 1.4 (Linear combination)
Suppose $S = \{\sum a_{ij}x_j = b_i\}_{1\le i\le m, 1\le j\le n}$ is a system of linear equations over $\mathbb{K}$. $S' = \{\sum a'_{ij}x_j = b_i\}_{1\le i\le m, 1\le j\le n}$ is another system of linear equations (not too important how many equations there are in S'). Then, S' is a **linear combination** of S if every linear equations $\sum a'_{ij}x_j = b_i$ in S' can be obtained as linear combinations of equations in S, i.e. $\sum a'_{ij}x_j = b'_i$ is obtained through

$$\sum c_i \left(\sum a_{ij}x_j\right) = \sum c_i b_i, 1 \le i \le m, \text{ for some } c_i \in \mathbb{K}$$

Definition 1.5 (Equivalance)
2 systems S, S' are **equivalent** if S' is a linear combination of S and vice versa. Denote $\mathbf{S} \sim \mathbf{S'}$

Example

In previous example, $S = \{(1), (2)\}, S' = \{(1), (3)\}, S'' = \{(2), (3)\}, S''' = \{(3)\}$.
Then, $S \not\sim S'', S \sim S'$ always , $S \sim S''$ only if 3 is invertible

Explanation

From S', (1) = (1), (2) = (3) - $3 \cdot$ (1). Therefore S is a linear combination of S'. $\Rightarrow S \sim S'$.
From S'', (2) = (2), $3 \cdot$ (1) = (3) - (2). If $3^{-1} \in \mathbb{K}$(i.e. $3 \neq 0$) then (1) $= 3^{-1}$((3) - (2)) is thus recoverable from S'', then $S \sim S''$. Otherwise, no. □

Lecture 2

Matrices

28 Mar 2023

Proposition 2.1

If 2 systems of linear equations are equivalent, $S \sim S'$ then they have the same set of solutions

Remark

Why is this important? This becomes important if we have a complicated system and want to transform into a simpler system to solve.

Proof (Proposition 2.1)

If $(x_1 = \alpha_1, x_2 = \alpha_2, \ldots, x_n = \alpha_n)$ is a solution ot S then we claim that it's also a solution of S' and vice versa. This is trivial because $S \sim S'$. $\square$

Definition 2.1 (Matrix)

Let $\mathbb{K}$ be a field. Then an $\mathbf{m} \times \mathbf{n}$ **matrix** with coefficients in $\mathbb{K}$, is an ordered tuple of elements in $\mathbb{K}$, typically written as

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \in \mathbb{M}_{m\times n}(\mathbb{K})$$

Definition 2.2 (Matrix Multiplication)

If $T_1 \in \mathbb{M}_{m\times n}(\mathbb{K}), T_2 \in \mathbb{M}_{n\times l}(\mathbb{K})$ then $T_1 \cdot T_2 \in \mathbb{M}_{m\times l}(\mathbb{K})$ (where $m, n, l \in \mathbb{N}$). Specifically,

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & \cdots & a_{mn} \end{bmatrix} \cdot \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1l} \\ b_{21} & b_{22} & \cdots & b_{2l} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nl} \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} & \cdots & \cdots & c_{1l} \\ c_{21} & c_{22} & \cdots & \cdots & c_{2l} \\ \vdots & \vdots & \ddots & \cdots & \vdots \\ c_{m1} & c_{m2} & \cdots & \cdots & c_{ml} \end{bmatrix}$$

where

$$\begin{aligned} c_{ij} &= \text{the “inner product” of i-th row of } T_1 \text{ and j-th row of } T_2 \\ &= \sum_{t=1}^{n} a_{it} b_{tj} \\ &\forall\, (i, j), 1 \leq i \leq m, 1 \leq j \leq l \end{aligned}$$

In particular, if $T_1, T_2 \in \mathbb{M}_n := \mathbb{M}_{n\times n}(\mathbb{K})$ then $T_1 \cdot T_2$ and $T_2 \cdot T_1$ are both valid. In general, they're often not equal.

Observe

We can write system of linear equations as

$$T \cdot \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

where

$$T \in \mathbb{M}_{m\times n}(\mathbb{K}), \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{M}_{n\times 1}(\text{indeterminants}), \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} \in \mathbb{M}_{m\times 1}(\mathbb{K})$$

Then, finding solutions to S is equivalent to finding $(\alpha_1, \alpha_2, \dots, \alpha_n) \in \mathbb{K}$ s.t.

$$T \cdot \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

Exercise 2.1

If $T_1, T_2, T_3 \in \mathbb{M}_n(\mathbb{K})$ then $(T_1 \cdot T_2) \cdot T_3 = T_1 \cdot (T_2 \cdot T_3)$. This is by no means obvious.

Definition 2.3 (Identity Matrix)

$$I_n = id_n = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \ddots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \vdots & \cdots & \ddots & 1 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{bmatrix} \in \mathbb{M}_n(\mathbb{K})$$

Observe

$$I_n \cdot T = T \cdot I_n, \ \forall\, T \in \mathbb{M}_n(\mathbb{K})$$

Thus, $(\mathbb{M}_n(\mathbb{K}), \cdot)$ is “trying” to be a group, but it’s not.

Definition 2.4 (Invertible Matrix)
A matrix $T \in \mathbb{M}_n(\mathbb{K})$ is **invertible** if $\exists\, T' \in \mathbb{M}_n(\mathbb{K})$ s.t.

$$T \cdot T' = I_n$$

Exercise 2.2
If $T \cdot T' = I_n \Rightarrow T' \cdot T = I_n$

Definition 2.5 (General Linear Group $GL_n(\mathbb{K})$)

$$GL_n(\mathbb{K}) = \{T \in \mathbb{M}_n(\mathbb{K}) \mid T \text{ is invertible}\}$$

Remark
Then $(GL_n(\mathbb{K}), \cdot)$ is a group.

Definition 2.6 (Elementary Row operations)
Let S be the system of equations:

$$\sum a_{1j}x_j = b_1 \quad (1)$$
$$\sum a_{2j}x_j = b_2 \quad (2)$$
$$\vdots = \vdots$$
$$\sum a_{mj}x_j = b_m \quad (\text{m})$$

then there are 3 **elementary row operations**:

1. Switching 2 of the equations
2. Replace (i) with $c \cdot$ (i) where $c \neq 0$
3. Replace (i) by (i) + d(j) where $i \neq j$

Proposition 2.2
If S' can be obtained from S via a finite sequence of elementary row operations then $S \sim S'$.

Corollary 2.1
S can also be obtained from S' via a finite sequence of elementary row operations.

Corollary 2.2
If S' can be obtained from S via a finite sequence of elementary row operations then they have the same solutions.

Lecture 3

Vector Spaces

30 Mar 2023

3.1 Elementary Row Operations and Systems of Linear Equations

Question: What are we doing to the matrices $A, B(Ax = B)$ (A of size $m \times n$, B of size $n \times 1$) when elementary row operations are carried out?
Answer: The row operations operate on the **rows** of A (switching rows, multiplying by scalar, adding other rows)

Example

$$A_0 = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 1 & 1 \end{bmatrix} \overset{(1')=(1)+-2(3)}{\sim} A_1 = \begin{bmatrix} 0 & -1 & -1 \\ 1 & 2 & 3 \\ 1 & 1 & 1 \end{bmatrix} \sim \cdots \sim A_7 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} b\dots \\ b\dots \\ b\dots \end{bmatrix}$$

We eventually arrived $LHS = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ itself, due to the properties of I_3. By "simplifying" rows this way, we can therefore solve systems of linear equations.

Definition 3.1 (Row-reduced Matrix)
The **row-reduced** form of a matrix has 1 as the leading non-zero coefficient for each of its rows (0-padded on the left). Furthermore, each column which contains the leading non-zero entry of some row has all its other entries as 0. By convention, the leading coefficient of a row of higher row index also has a higher column index.

Proof (Proposition 2.2)
We only provide a sketch of the proof. We re-enumerate the types of operations:

1. $(i) \leftrightarrow (j)$
2. $(i) \to c(i), c \neq 0$
3. $(i) \to (i) + d(j), j \neq i$

Explanations:

1. Trivial
2. Clearly S' is obtainable from S, and trivially all other equations except for (i) of S are obtainable from S'. However, $(i) = c^{-1}(c(i)) = c^{-1}(i')$. Therefore $S \sim S'$.

3. Similarly, S' is clearly obtainable from S, while $(i) = (i') - d(j) = (i') - d(j')$. Therefore $S \sim S'$.

□

3.2 Vector Spaces

Definition 3.2 (Vector Space)

Let $\mathbb{K}$ be a field. A **vector space over** $\mathbb{K}$ ("$\mathbb{K}$-vector space")("k-vs") is an Abelian group V with a map: $\mathbb{K} \times V \to V$ ($\mathbb{K}$-action on V). An element in V is called a **vector**. They have to satisfy $\forall\, a, b \in \mathbb{K}$; $\forall\, v, v_1, v_2 \in V$:

1. $0 \cdot v = 0$
   $1 \cdot v = v$
2. $(a + b) \cdot v = (a \cdot v) + (b \cdot v)$
   $(a \cdot b) \cdot v = a \cdot (b \cdot v)$
3. $a \cdot (v_1 + v_2) = (a \cdot v_1) + (a \cdot v_2)$

Essentially, $\mathbb{K}, V$ with operations:

1. $+ : \mathbb{K} \times \mathbb{K} \to \mathbb{K}, \cdot : \mathbb{K} \times \mathbb{K} \to \mathbb{K}$ (Field)
2. $+ : V \times V \to V$ (Abelian group)
3. $\cdot : \mathbb{K} \times V \to V$ (Action)

Example

Field $\mathbb{K} = \mathbb{R}$, $V = \mathbb{R}^n \doteq \{(x_1, x_2, \ldots, x_n) \mid x_i \in \mathbb{R}\}$. Indeed, $\mathbb{R}^n$ is an Abelian group.

Definition 3.3 (Linear Combination)

Let V be a k-vs. If $v_1, v_2, \ldots, v_r \in V; r \in \mathbb{N}$ then a **linear combination** of $\{v_1, v_2, \ldots, v_r\}$ is a vector of the form

$$c_1 \cdot v_1 + c_2 \cdot v_2 + \cdots + c_r \cdot v_r \text{ where } c_i \in \mathbb{K}$$

Definition 3.4 (Linear Span)

Then the **linear span** of $v_1, v_2, \ldots, v_r$ in V is the set of all such linear combinations.

Lecture 4

Linear Transformation, Homomorphism, Kernel, Image

04 Apr 2023

4.1 Vector Subspace

Definition 4.1 (Vector Subspace)

Let V be a $\mathbb{K}$-vector space. A **subspace** (or **sub-vector space**) of V is a subset $W \subseteq V$ s.t. W is itself a $\mathbb{K}$-vector space under addition and scaling induced from V. A priori, we know that

$$+ : W \times W \to V, \cdot : W \times W \to V$$

but this subspace requirement implies that

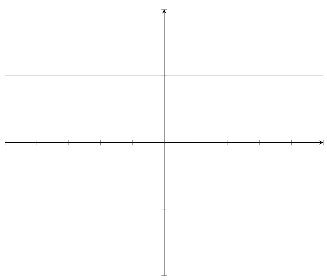
$$\forall\, x, y \in W, x + y \in W$$

$$\forall\, \alpha \in \mathbb{K}, x \in W, \alpha \cdot x \in W$$

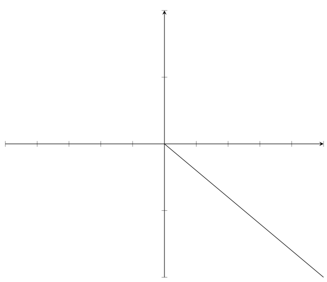
In other words, the subspace is closed under addition and scaling.

Example

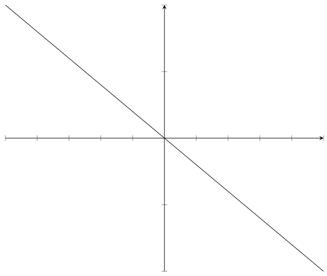
Take $\mathbb{K} = \mathbb{R}, V = \mathbb{R}^2$, with ordinary addition and scaling.
Consider the subset represented by line $y = 1$.



This is not a subspace because there exists no 0 element. This kinda implies that any subspace of $\mathbb{R}^2$ must pass through the origin $(0, 0)$.
Consider another instance, this time the following ray:



This is also not a subspace, since there's no additive inverse. Therefore a subspace shall look something like this:



4.2 Mapping

Motivation

A map from sets to sets can be anything. e.g. $x : \mathbb{Z} \mapsto x^2 : \mathbb{Z}$ doesn't preserve the "group" structure $(x+y)^2 \neq x^2 + y^2$ most of the time.

Definition 4.2 (Group Homomorphism)

Let A, B be Abelian groups. Map $\psi : A \to B$ is called a **group homomorphism** if:

$$\psi(x+y) = \psi(x) + \psi(y)$$

Then $x : \mathbb{Z} \mapsto x^2 : \mathbb{Z}$ is not a group homomorphism, but $x : \mathbb{Z} \mapsto nx : \mathbb{Z}$ for fixed n is a group homomorphism.
Here, a natural question arises: If given 2 vector spaces, what maps are allowed between them? What structures do we have to preserve?

Definition 4.3 (Linear Transformation)

Let V, W be $\mathbb{K}$-vector spaces. Then a **vector space homomorphism** is also called a **linear transformation**, a map $\psi : V \to W$ s.t.

1. $\psi(v_1 + v_2) = \psi(v_1) + \psi(v_2)\ \forall\ v_1, v_2 \in V$
2. $\psi(\alpha \cdot v) = \alpha \cdot \psi(v)\ \forall\ \alpha \in \mathbb{K}, v \in V$

Denote $\mathbf{Hom}_{\mathbb{K}}(\mathbf{V}, \mathbf{W})$ as the set of all linear transformations $V \to W$.

Example

$\mathbb{K} = \mathbb{R}, V = W = \mathbb{R}$
$\mathrm{Hom}_{\mathbb{R}}(V, W) = \{\psi : \mathbb{R} \to \mathbb{R} \mid (1), (2) \text{ are satisfied } \}$
We claim that $\psi(1)$ uniquely determines the map ψ, because

$$\psi(\alpha) = \alpha \cdot \psi(1)$$

Essentially, there exists a bijection between $\mathrm{Hom}_{\mathbb{R}}(V, W)$ and $\mathbb{R}$:

$$\begin{aligned} \mathrm{Hom}_{\mathbb{R}}(V, W) &\to \mathbb{R} \\ \psi &\to \psi(1) \\ (\psi_\beta : x \mapsto x \cdot \beta) &\leftarrow \beta \end{aligned}$$

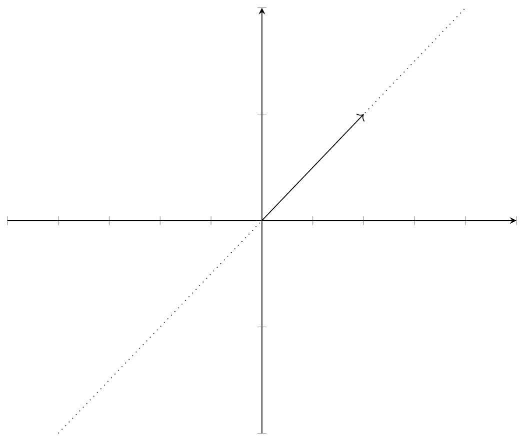
Example

$\mathbb{K} = \mathbb{R}, V = \mathbb{R}, W =$ any $\mathbb{K}$-vector space
We, similarly, claim that there is a bijection between $\mathrm{Hom}_{\mathbb{R}}(V, W)$ and W. With the same reasoning, ψ is determined by $\psi(1)$, though this time $\psi(1) \in W$.

$$\begin{aligned} \mathrm{Hom}_{\mathbb{R}}(V, W) &\to W \\ \psi &\to \psi(1) \in W \\ (\psi_\beta : x \mapsto x \cdot w) &\leftarrow w \end{aligned}$$

Example

As a sub-example of the example above, consider $W = \mathbb{R}^2$:



Then if $\psi(1) = (4, 5)$ as above (and $\psi(0) = (0, 0)$ implicit), then ψ would map the rest of $V = \mathbb{R}$ onto the dotted line above.
An interesting point to note is that if $\psi(1) = (0, 0)$, then the entire real line would get sent (and compressed) to $(0, 0)$. $\psi_{(0,0)}$ therefore contracts $\mathbb{R}$ into one point (the origin (0, 0)) while others output a subspace of $\mathbb{R}^2$.

Example

$\mathbb{K} = \mathbb{R}, V = \mathbb{R}^2, W =$ any $\mathbb{R}$-vector space
We claim that there exists a bijection between $\mathrm{Hom}_{\mathbb{R}}(\mathbb{R}^2, W)$ and $W \oplus W$; as each ψ is determined by $\psi((1, 0))$ and $\psi((0, 1))$.
The notation $\oplus$ is defined as: If V, W are $\mathbb{K}$-vector spaces then

$$V \oplus W = \{(v, w) \mid v \in V, w \in W\}$$

e.g. $\mathbb{R}^2 = \mathbb{R} \oplus \mathbb{R}$
Then $V \oplus W$ would also be a $\mathbb{K}$-vector space with operations $+, \cdot$ defined intuitively:

$$(v_1, w_1) + (v_2, w_2) = (v_1 + v_2, w_1 + w_2)$$
$$\alpha \cdot (v, w) = (\alpha \cdot v, \alpha \cdot w)$$

Back to the example, $\forall\, v = (x, y) \in V, v = x(1, 0) + y(0, 1)$, therefore

$$\psi(v) = \psi((x, y)) = x \cdot \psi((1, 0)) + y \cdot \psi((0, 1))$$

ψ is therefore uniquely defined by $\psi((1,0))$ and $\psi((0,1))$.

Example

$\mathbb{K} = \mathbb{R}, V = \mathbb{R}^m, W =$ any $\mathbb{R}$-vector space
Think about $W = \mathbb{R}^n$ with similar reasoning.
Hint: We want to show there exists a bijection between $\mathrm{Hom}_{\mathbb{R}}(\mathbb{R}^m, \mathbb{R}^n)$ and $\mathbb{R}^{m \cdot n}$, but this is often rewritten as $\mathbb{M}_{m \times n}(\mathbb{R})$

4.3 Isomorphism, Kernel, Image

Every linear transformation is just a map, and we can therefore question if it is injective, surjective or bijective. In all cases, these concepts simply deal with the sets (vector spaces) as simply sets.

Definition 4.4 (Isomorphism)
A $\mathbb{K}$-linear transformation $\psi : V \to W$ is an **isomorphism** if it is bijective.

Definition 4.5 (Kernel, Image)
Let $\psi : V \to W$ be a linear transformation over $\mathbb{K}$. Then:

1. **Kernel**: $\ker(\psi) := \{v \in V \mid \psi(v) = 0\} \subseteq V$
2. **Image**: $\mathrm{im}(\psi) := \{w \in W \mid \exists\, v \in V \text{ s.t. } \psi(v) = w\}$

Lemma 4.1

1. $\ker(\psi)$ is a $\mathbb{K}$-vector subspace of V
2. $\mathrm{im}(\psi)$ is a $\mathbb{K}$-vector subspace of W

Proof (Lemma)
We want to show that if $x, y \in \ker(\psi)$ then $x + y \in \ker(\psi)$.

$$\begin{aligned}\psi(x+y) &= \psi(x) + \psi(y) (\text{ since } \psi \text{ is a linear transformation }) \\ &= 0 + 0 \\ &= 0\end{aligned}$$

Therefore $x + y \in \ker(\psi)$
Furthermore, $\forall\, \alpha \in \mathbb{K}, x \in \ker(\psi)$ then

$$\psi(\alpha, x) = \alpha \cdot \psi(x) = \alpha \cdot 0 = 0 \Rightarrow \alpha \cdot x \in \ker(\psi)$$

Therefore $\ker(\psi)$ is a subspace.
Similarly, $\mathrm{im}(\psi)$ is a subspace. □

Definition 4.6 (Finite Dimensional, Dimension)

1. Let V be a $\mathbb{K}$-vector space. V is called **finite dimensional** if there exists a surjective linear transformation $\mathbb{K}^r \to V$ where $r \in \mathbb{Z}_{\geq 0}$. As a consequence, $\mathbb{K}^r$ is also finite dimensional, with an identity mapping.

2. If V is finite dimensional then **dimension** of V is defined as

$$\dim V \coloneqq \min\{k \in \mathbb{Z}_{\geq 0} \mid \ \exists \text{ linear transformation } \mathbb{K}^r \to V\}$$

Lecture 5

Span, Linear Independence, Basis

06 Apr 2023

Recall

Linear Combination: Let $V = \mathbb{K}$-vector space with $v_1, v_2, \ldots, v_r \in V$ then

$$\mathbb{K}\langle v_1, v_2, \ldots, v_r \rangle \coloneqq \{w \in W \mid w = a_1 v_1 + \cdots + a_r v_r; a_i \in \mathbb{K}\} \subseteq V (\text{ is a subspace of } V)$$

Definition 5.1 (Span)
$\{v_1, v_2, \ldots, v_r\}$ span V if

$$\mathbb{K}\langle v_1, v_2, \ldots, v_r \rangle = V$$

i.e. equality is achieved: every vector in V can be written as linear combinations of $\{v_1, v_2, \ldots, v_r\}$

Connecting to the previous lecture, let $\psi : \mathbb{K}^r \to V$ then $\psi \in \mathrm{Hom}_{\mathbb{K}}(\mathbb{K}^r, V) \xrightarrow{\sim} V^{\oplus r}$, i.e. ψ corresponds to $(v_1, v_2, \ldots, v_r)$ in V.
In particular, $(v_1, v_2, \ldots, v_r) \in V^{\oplus r}$ determines the map:

$$\begin{aligned} \psi : (1, 0, \ldots, 0) \in \mathbb{K}^r &\to v_1 \\ (0, 1, \ldots, 0) \in \mathbb{K}^r &\to v_2 \\ &\vdots \\ (0, 0, \ldots, 1) \in \mathbb{K}^r &\to v_r \\ (\alpha_1, \alpha_2, \ldots, \alpha_r) \in \mathbb{K}^r &\to \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_r v_r \end{aligned}$$

Lemma 5.1

1. Let $\psi : \mathbb{K}^r \to V$ be a linear transformation determined by $v_1, v_2, \ldots, v_r \in V$, i.e. $\psi(\alpha_1, \alpha_2, \ldots, \alpha_r) \coloneqq \sum_{i=1}^r \alpha_i v_i$, then

   $$\mathrm{im}(\psi) = \mathbb{K}\langle v_1, v_2, \ldots, v_r \rangle$$

   is a subspace of V

2. $\{v_1, v_2, \ldots, v_r\}$ span $V \Leftrightarrow \psi$ is surjective
   i.e. a surjection $\mathbb{K}^r \to V$ corresponds to r vectors $v_1, v_2, \ldots, v_r \in V$ that span V

Remark

V is finite dimensional when $\exists$ surjection $\mathbb{K}^d \to V$
$\Leftrightarrow \exists\, d$ vectors $v_1, v_2, \ldots, v_r$ that span V.
Recall: $\dim V = \min\{r \in \mathbb{Z}_{\geq 0} \text{ s.t. } \exists \text{ surjective } \mathbb{K}^r \to V\}$.
Next, what does it mean for ψ to be injective?

Definition 5.2 (Linear Independence)
$v_1, v_2, \ldots, v_r \in V$ are **linearly independent** if

$$a_1 v_1 + a_2 v_2 + \cdots + a_r v_r = 0; a_i \in \mathbb{K} \Rightarrow a_1 = a_2 = \cdots = a_r = 0$$

i.e. there doesn't exist non-trivial relations between the vectors.

Example

In $\mathbb{R}^2$, (0, 1) and (0, 2) are not linearly independent because

$$(-2)(0,1) + (0,2) = (0,0)$$

But (0, 1) and (1,0) are linearly independent.

Consequentially, they are **linearly dependent** otherwise, i.e.

$$\exists\, a_i \text{ not all } 0 \text{ s.t. } \sum a_i v_i = 0$$

Lemma 5.2

Given $\psi : \mathbb{K}^r \to V$ corresponds to $v_1, v_2, \ldots, v_r$ then $v_1, v_2, \ldots, v_r$ are linearly independent if and only if ψ is injective

In order to prove the lemma above, we shall make use of a more convenient test for whether a map $\varphi : \mathbb{K}^r \to V$ is injective.

Lemma 5.3

Let $\varphi : V \to W$ be a linear transformation then φ is injective if and only if

$$\ker(\varphi) = \{0\} \subseteq V$$

Proof (Lemma 5.3)

$\boxed{\Rightarrow}$ We assume that φ is injective, want to show that $\ker(\varphi) = \{0\}$.
We know that $\varphi(0) = 0 \Rightarrow 0 \in \ker(\varphi)$ but since φ is injective, $\nexists v \neq 0 \in V$ s.t. $\varphi(v) = 0$.
It follows that $\ker(\varphi) = 0$
$\boxed{\Leftarrow}$ We want to show that $x, y \in V$ s.t. $\varphi(x) = \varphi(y) \Rightarrow x = y$
Since $\varphi(x - y) = \varphi(x + (-y)) = \varphi(x) - \varphi(y) = 0$, combined with $\ker(\varphi) = 0$

$$\Rightarrow x - y = 0 \Rightarrow x = y$$

$\square$

Proof (Lemma 5.2)

Applying Lemma 5.3, we want to show: $\ker(\varphi) = 0$ iff $v_1, v_2, \ldots, v_r$ are linearly independent.
$\boxed{\Rightarrow}$ Suppose $\ker(\varphi) = \{0\}$ then want to show

$$a_1 v_1 + a_2 v_2 + \cdots + a_r v_r = 0 \Rightarrow a_i = 0\ \forall\, i$$

But $LHS = \varphi((a_1, a_2, \ldots, a_r)) \Rightarrow (a_1, a_2, \ldots, a_r) \in \ker(\varphi) \Rightarrow (a_1, a_2, \ldots, a_r) = 0$.
Therefore $a_i = 0\ \forall\, i$.
$\boxed{\Leftarrow}$ Suppose that $v_1, v_2, \ldots, v_r$ are linearly independent.
Then for $v \in \ker(\varphi) \Rightarrow \varphi(v) = 0$, with $v = (a_1, a_2, \ldots, a_r)$

$$\begin{aligned} \Rightarrow 0 &= \varphi(v) \\ &= \varphi((a_1, a_2, \ldots, a_r)) \\ &= a_1 v_1 + a_2 v_2 + \cdots + a_r v_r \end{aligned}$$

But since $v_1, v_2, \ldots, v_r$ are linearly independent

$$\Rightarrow a_i = 0 \; \forall \; i \Rightarrow v = 0 \Rightarrow \ker(\varphi) = 0$$

$\square$

Corollary 5.1

If V has dimension d over $\mathbb{K}$ then there exists isomorphic $\varphi : \mathbb{K}^d \xrightarrow{\sim} V$
i.e. φ is a bijective linear transformation

Proof (Corollary)

Since $d = \dim V$, by definition there exists surjective linear transformation $\pi : \mathbb{K}^d \twoheadrightarrow V$
We then claim that π is also injective.
Proving by contradiction, we suppose that π is not injective.
let $v_1, v_2, \ldots, v_d$ be the d vectors that correspond to π, i.e.

$$\pi((a_1, a_2, \ldots, a_d)) = a_1 v_1 + \cdots + a_d v_d$$

By Lemma 5.2, π being not injective implies that $v_1, v_2, \ldots, v_d$ are linearly dependent.
i.e. there exists $b_1, b_2, \ldots, b_d \in \mathbb{K}$ not identically 0 s.t.

$$b_1 v_1 + b_2 v_2 + \cdots + b_d v_d = 0$$

WLOG, assume $b_1 \neq 0$.

$$\begin{aligned}
\Rightarrow b_1 v_1 &= -(b_2 v_2 \ldots b_d v_d) \\
\Rightarrow v_1 &= -b^{-1}(b_2 v_2 \ldots b_d v_d)(\,\exists\, b^{-1} \because b_1 \neq 0) \\
&= c_2 v_2 + c_3 v_3 + \cdots + c_d v_d
\end{aligned}$$

We already know that since π is surjective, thus $v_1, v_2, \ldots, v_d$ span V. However, the above equality implies that $v_2, \ldots, v_d$ already span V!
It follows that there must exist a surjective linear transformation $\pi' : \mathbb{K}^{d-1} \twoheadrightarrow V$
$\Rightarrow\Leftarrow$, since $d = \min\{r \mid \exists \text{ surjective } \pi^r : \mathbb{K}^r \twoheadrightarrow V\}$
Therefore π is injective. It is already surjective, and therefore bijective, making it an isomorphism. $\square$

Recall

$\psi : \mathbb{K}^d \to V$ as determined by $v_1, v_2, \ldots, v_d$ is

1. **injective** when $v_1, v_2, \ldots, v_d$ are linearly independent
2. **surjective** when $v_1, v_2, \ldots, v_d$ span V

This naturally leads to our next definition.

Definition 5.3 (Basis)

$\{v_1, v_2, \ldots, v_r\}$ is called a **basis** of V if they span V and are linearly independent,
i.e. $\psi_{(v_1, v_2, \ldots, v_r)} : \mathbb{K}^r \to V$ is an isomorphism.

Corollary 5.2

$\dim_{\mathbb{K}} V = d \Leftrightarrow \exists$ basis $\{v_1, v_2, \ldots, v_d\}$ for V

Corollary 5.3

If $\{v_1, v_2, \ldots, v_d\}$ and $\{w_1, w_2, \ldots, v_{d'}\}$ are basis for V then $d = d'$.

Lecture 6

Vector Space as Direct Sums of Subspaces

13 Apr 2023

Lemma 6.1

Let V, W be vector spaces over $\mathbb{K}$. If $\dim_{\mathbb{K}} V = d_1, \dim_{\mathbb{K}} W = d_2$ then $V \oplus W$ is finite dimensional and $\dim_{\mathbb{K}}(V \oplus W) = d_1 + d_2$

Proof (Lemma)

We claim that: If $\{v_1, v_2, \ldots, v_{d_1}\}$ is a basis for V, $\{w_1, w_2, \ldots, w_{d_2}\}$ is a basis for W then

$$\{(v_1, 0), (v_2, 0), \ldots, (v_{d_1}, 0), (0, w_1), (0, w_2), \ldots, (0, w_{d_2})\}$$

is a basis for $V \oplus W$.

Span

If $x \in V \oplus W$ then $x = (v, w)$ for some $v \in V, w \in W$.

Therefore

$$\begin{aligned} x &= (v, 0) + (0, w) \\ &= \sum_{i=1}^{d_1} \alpha_i (v_i, 0) + \sum_{j=1}^{d_2} \beta_j (0, w_j) \end{aligned}$$

for some $\alpha_i, \beta_j \in \mathbb{K}$, since $\{v_i\}, \{w_j\}$ are bases.

$\{(v_1, 0), (v_2, 0), \ldots, (v_{d_1}, 0), (0, w_1), (0, w_2), \ldots, (0, w_{d_2})\}$ indeed spans $V \oplus W$.

Linearly Independent

Suppose there exists $\sum_{i=1}^{d_1} \alpha_i (v_i, 0) + \sum_{j=1}^{d_2} \beta_j (0, w_j) = (0, 0)$

By comparing the 2 "coordinates", $\sum_{i=1}^{d_1} \alpha_i v_i = 0 \in V$ and $\sum_{j=1}^{d_2} \beta_j w_j = 0 \in W$.

But since $\{v_i\}, \{w_j\}$ are bases $\Rightarrow \alpha_i = \beta_j = 0 \in \mathbb{K}$.

It follows that $\{(v_1, 0), (v_2, 0), \ldots, (v_{d_1}, 0), (0, w_1), (0, w_2), \ldots, (0, w_{d_2})\}$ are indeed linearly independent.

Dimension as size of basis:

$$\Rightarrow \dim_{\mathbb{K}}(V \oplus W) = d_1 + d_2 = \dim_{\mathbb{K}} V + \dim_{\mathbb{K}} W$$

$\square$

Example

$\mathbb{R} \oplus \mathbb{R} = \mathbb{R}^2$.

We can view $\mathbb{R}$ as a "subspace" of $\mathbb{R}^2$, by prescribing the other coordinate. Some ways are described as follows:

1. $L_0 : \mathbb{R} \to \mathbb{R}^2, a \to (0, 0)$
2. $L_1 : \mathbb{R} \to \mathbb{R}^2, x \to (x, 0)$
3. $L_2 : \mathbb{R} \to \mathbb{R}^2, y \to (0, y)$
4. $L_3 : \mathbb{R} \to \mathbb{R}^2, z \to (z, z)$

Then, when are these direct sums of subspaces either lacking/redundant to get $\mathbb{R}^2$? For example, $L_0 \oplus L_1$ is lacking, while $L_1 \oplus \mathbb{R}^2$ is redundant. We thus investigate the relationship between a vector space and its subspaces.

Let W be a vector space over $\mathbb{K}$. V_1, V_2 are subspaces of W. Consider

$$\begin{aligned} V_1 \oplus V_2 &\xrightarrow{\pi} W \\ (v_1, v_2) &\to v_1 + v_2 \end{aligned}$$

We then inspect the injectivity and surjectivity of this mapping π.

Lemma 6.2

π as above is injective $\Leftrightarrow V_1 \cap V_2 = \{0\} \subseteq W$

Proof (Lemma)

$\boxed{\Rightarrow}$ Suppose π is injective.
Let $x \in V_1 \cap V_2$ then $x \in V_1, x \in V_2 \Rightarrow (-x) \in V_2$.
It follows that $(x, -x) \in V_1 \oplus V_2$ and $\pi(x, -x) = x + (-x) = 0$.
Therefore, for π to be injective, $x = 0 \Rightarrow V_1 \cap V_2 = \{0\}$
$\boxed{\Leftarrow}$ Suppose $V_1 \cap V_2 = \{0\}$. To prove that π is injective, we prove that $\ker(\pi) = 0$
Let $y = (v_1, v_2) \in \ker(\pi)$, i.e. $v_1 \in V_1, v_2 \in V_2, 0 = \pi(y) = \pi((v_1, v_2)) = v_1 + v_2 \in W$
It follows that $v_1 = -v_2 \in V_2 \Rightarrow v_1 \in V_1 \Rightarrow v_1 \in V_1 \cap V_2 \Rightarrow v_1 = 0 \Rightarrow v_2 = -v_1 = 0$
Thus $y = (0, 0) = 0_{V \oplus W}$. Therefore $\ker(\pi) = \{0\}$ □

Corollary 6.1

Suppose V_1, V_2 are subspaces of W s.t.

1. (surjective) every $w \in W$ can be written as $w = v_1 + v_2$ for some $v_1 \in V_1, v_2 \in V_2$
2. (injective) $V_1 \cap V_2 = \{0\}$

then we have a (natural) isomorphism:

$$\begin{aligned} V_1 \oplus V_2 &\xrightarrow{\sim} W \\ (x, y) &\to x + y \end{aligned}$$

Remark

Essentially, this answers the question: when can we write a vector space as direct sum of 2 subspaces?

Proposition 6.3

Let V, W be finite dimensional vector spaces over $\mathbb{K}$. Let $\psi : V \to W$ be a linear transformation over $\mathbb{K}$ then there exists isomorphism

$$\ker(\psi) \oplus \mathrm{im}(\psi) \xrightarrow{\sim} V$$

Consequentially, $\dim_{\mathbb{K}} V = \dim_{\mathbb{K}}(\ker(\psi)) + \dim_{\mathbb{K}}(\mathrm{im}(\psi))$
Warning: $\ker(\psi)$ is a subspace of V, but $\mathrm{im}(\psi)$ is only a subspace of W! We therefore can't straightaway apply the results of the previous corollary, but can do that by constructing a subspace of V that is isomorphic to $\mathrm{im}(\psi)$.

Remark

$\dim_{\mathbb{K}}(\ker(\psi))$ is called the **nullity of** ψ.
$\dim_{\mathbb{K}}(\mathrm{im}(\psi))$ is called the **rank of** ψ

Proof (Proposition)

Since W is finite dimensional, $\mathrm{im}(\psi) \subseteq W$ is therefore finite dimensional.
Let $\{e_1, e_2, \dots, e_r\}$ be a basis for $\mathrm{im}(\psi) \subseteq W$.
Since $e_i \in \mathrm{im}(\psi) \Rightarrow \ \exists\ \psi^{-1}(e_i) = \{v \in V \mid \psi(v) = e_i\} \neq \emptyset$
Pick some $e_i' \in \psi^{-1}(e_i)$ for each i then let

$$U :- \mathbb{K}\langle e_1', e_2', \dots, e_r'\rangle \subseteq V$$

be the subspace spanned by $\{e_i'\}$.
Claim 1: ψ induces an isomorphism

$$U \xrightarrow{\sim} \mathrm{im}(\psi)$$
$$\sum_{i=1}^{r} \alpha_i e_i' \to \sum_{i=1}^{r} \alpha_i e_i$$

Claim 2: $\ker(\psi)$ and U satisfy the conditions in the above corollary as subspaces of V.
Before proving the details, we show that the 2 claims give us QED:
Claim 1: $\Rightarrow \ker(\psi) \oplus U \xrightarrow{\sim} \ker(\psi) \oplus \mathrm{im}(\psi)$
Claim 2: $\Rightarrow \ker(\psi) \oplus U \xrightarrow{\sim} V$ □
Proving Claim 1: From construction,

$$U \xrightarrow{\varphi} \mathrm{im}(\psi)$$
$$\sum_{i=1}^{r} \alpha_i e_i' \to \sum_{i=1}^{r} \alpha_i e_i$$

is surjective. It remains for us to show that it is injective $\Leftrightarrow \ker(\varphi) = \{0\}$
Suppose $\sum_{i=1}^{r} \alpha_i e_i' \in \ker(\varphi)$ then

$$\mathrm{im}(\psi) \ni 0 = \varphi\left(\sum_{i=1}^{r} \alpha_i e_i'\right) = \sum_{i=1}^{r} \alpha_i e_i$$

But since $\{e_i\}$ forms a basis for $\mathrm{im}(\psi) \Rightarrow \alpha_i = 0 \in \mathbb{K} \Rightarrow \sum_{i=1}^{r} \alpha_i e_i' = 0 \in U \Rightarrow \ker(\varphi) = \{0\}$
φ is therefore injective.
Proving Claim 2: Let $v \in V$, we want to write v as sum of an element from U and an element from $\ker(\psi)$.
Let $w = \psi(v) \in \mathrm{im}(\psi) = \sum \alpha_i e_i$
Let $v' = \sum \alpha_i e_i' \in U$, then

$$\psi(v - v') = \psi(v) - \psi(v') = w - w = 0$$

Therefore $v - v' \in \ker(\psi)$, and we can write

$$v = (v - v')(\in \ker(\psi)) + v'(\in U)$$

It remains for us to show that $\ker(\psi) \cap U = \{0\}$.
Let any $x \in \ker(\psi) \cap U$ then $\psi(x) = 0 \in \mathrm{im}(\psi)$.
But from claim 1, it follows that $x = 0 \Rightarrow \ker(\psi) \cap U = \{0\}$

Lecture 7

Linear Transformation and Matrices

23 Apr 2023

Recall

1. $\psi \in \mathrm{Hom}_{\mathbb{K}}(\mathbb{K}^r, V)$ corresponds to r vectors: $v_1, v_2, \ldots, v_r$:

$$(\psi : \mathbb{K}^r \to V) \to \{v_i\} = \{\psi(0, \ldots, 1, \ldots, 0)\} \text{ (1 in } i\text{−th position)}$$
$$(\psi : (a_1, a_2, \ldots, a_r) \to \sum a_i v_i) \leftarrow \{v_i\}$$

2. V has dimension $d \Leftrightarrow V$ has basis $\{v_1, v_2, \ldots, v_d\}$

3. $\psi : V \xrightarrow{\sim} W$ then ψ sends a set of basis $\{v_i\}_{1\leq i\leq d}$ to a set of basis $\psi(v_i)$ of W

Proof (Recall 3)

Approach 1

One might first prove this statement from first principles, that is to show that:

1. $\{w_i = \psi(v_i)\}$ span W
2. $\{w_i = \psi(v_i)\}$ are linearly independent

This approach is doable, though a little bit tedious.

Approach 2

Observe that $\{v_i\}$ corresponds to a map:

$$\mathbb{K}^d \xrightarrow{\sim} V$$

while

$$V \underset{\psi}{\xrightarrow{\sim}} W$$

by assumption.

It then follows that $\mathbb{K}^d \xrightarrow{\sim} W$, following the function composition, it would yield that this mapping corresponds to $\{w_i = \psi(v_i)\}$. Therefore $\{w_i\}$ forms a basis of W. □

7.1 Linear Transformation as Matrix Multiplication

Claim 7.1

Let V, W be vector spaces over $\mathbb{K}$ of dimensions n, m respectively. Let $\psi : V \to W$ be a linear transformation. Then once we've fixed bases $\{v_i\}_{1\leq i\leq n}$ of V and $\{w_j\}_{1\leq j\leq m}$ of W, ψ corresponds to $T_\psi \in \mathbb{M}_{m\times n}(\mathbb{K})$

In other words,

$$\psi \in \mathrm{Hom}_{\mathbb{K}}(V, W) \leftrightarrow T_\psi \in \mathbb{M}_{m\times n}(\mathbb{K})$$

Specifically,

$$T_\psi = (\alpha_{ji}) = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{m1} & \alpha_{m2} & \cdots & \alpha_{mn} \end{pmatrix}$$

corresponds to

$$\psi : v_i \mapsto \alpha_{1i}w_1 + \alpha_{2i}w_2 + \cdots + \alpha_{mi}w_m = \sum_{j=1}^{m} \alpha_{ji}w_j \text{ for } 1 \leq i \leq n$$

For any $v = \sum_{i=1}^{n} \beta_i v_i \in V$ then

$$\begin{aligned} w = \psi(v) &= \sum_{i=1}^{n} \beta_i \psi(v_i) \\ &= \sum_{i=1}^{n} \beta_i \left(\sum_{j=1}^{m} \alpha_{ji} w_j \right) \\ &= \sum_{i=1}^{n} \sum_{j=1}^{m} \alpha_{ji} \beta_i w_j \end{aligned}$$

An alternative perspective is that $v = \sum_{i=1}^{n} \beta_i v_i$ can be thought of as a "matrix" multiplication:

$$\begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} (v_1 \dots v_n)$$

where $\begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} \in \mathbb{M}_{n\times 1}(\mathbb{K})$ and $(v_1 \dots v_n)$ is just the basis in the row vector form.
(Warning: It is not a matrix, since $v_i \notin \mathbb{K}$)
Upshot: If we fix basis $v_1, v_2, \dots, v_n$ then any $v \in V$ would be uniquely expressed as $v = \beta_i v_i$.

The fixed basis would then correspond to unique matrices $\begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} \in \mathbb{M}_{n\times 1}(\mathbb{K})$

Note that if we change the basis to another $\{v_i'\}$ then

$$v = \sum \beta_i v_i = \sum \beta_i' v_i' \text{ where } \begin{pmatrix} \beta_1' \\ \vdots \\ \beta_n' \end{pmatrix} \in \mathbb{M}_{n\times 1}(\mathbb{K})$$

Now, if $T_\psi = (a_{ji})_{1\leq j\leq m, 1\leq i\leq n}$ then the map ψ sends $v \leftrightarrow \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix}$ to

$$\begin{pmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{m1} & \alpha_{m2} & \cdots & \alpha_{mn} \end{pmatrix} \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{n} \alpha_{1i}\beta_i = \gamma_1 \\ \sum_{i=1}^{n} \alpha_{2i}\beta_i = \gamma_2 \\ \vdots \\ \sum_{i=1}^{n} \alpha_{mi}\beta_i = \gamma_m \end{pmatrix} \in \mathbb{M}_{m\times 1}(\mathbb{K})$$

which corresponds to writing $w \in W$ under $\{w_j\}$ as

$$\begin{aligned} w &= \gamma_1 w_1 + \cdots + \gamma_m w_m \\ &= \sum_{j=1}^{m} \gamma_j w_j = \sum_{j=1}^{m} \sum_{i=1}^{n} \alpha_{ji} \beta_i w_j \end{aligned}$$

which is similar to the expression above.
Therefore, once we choose basis $\{v_i\}, \{w_j\}$ of V, W respectively then $\psi \leftrightarrow T_\psi \in \mathbb{M}_{m\times n}(\mathbb{K})$:

$$v = \sum_{i=1}^{n} \beta_i v_i \to \psi(v) = \sum_{i=1}^{n} \sum_{j=1}^{m} \alpha_{ji} \beta_i v_i$$

$$(\alpha_{ji}) \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} \leftrightarrow \begin{pmatrix} \gamma_1 \\ \vdots \\ \gamma_m \end{pmatrix}$$

7.2 Going from linear transformation to matrix

We've successfully represented linear transformation $\psi \in \mathrm{Hom}_{\mathbb{K}}(V, W)$ from T_ψ. How about the other way around, i.e. we know ψ and want to find its corresponding matrix T_ψ?
Consider $\psi : v_i \to \psi(v_i) \in W = c_1 w_1 + \cdots + c_m w_m$ then we can define $a_{ji} = c_j$ in this expression. Iterating over $1 \leq i \leq n$ would yield us $T_\psi = (a_{ji})$.

7.2.1 Standard $\mathbb{K}^n \to \mathbb{K}^m$

We have $\mathbb{K}^n, \mathbb{K}^m (\mathbb{K}^n = \mathbb{K}^{\oplus n} = \{x_1, x_2, \ldots, x_n \mid x_i \in \mathbb{K}\})$ then there's a preferred basis $\{e_i\}_{1\leq i\leq n}$:

$$\begin{aligned} e_1 &= (1, 0, \ldots, 0) \in \mathbb{K}^n \\ e_i &= (0, 0, \ldots, 1, \ldots, 0) \in \mathbb{K}^n \ (i\text{−th position}) \\ &\vdots \\ e_n &= (0, 0, \ldots, 1) \in \mathbb{K}^n \end{aligned}$$

and similarly for $e'_j \in \mathbb{K}^m$.

Under this basis, $(x_1, x_2, \ldots, x_n)$ corresponds to $\begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{M}_{n\times 1}(\mathbb{K})$

It follows that any linear transformation $\psi \in \mathrm{Hom}_{\mathbb{K}}(\mathbb{K}^n, \mathbb{K}^m)$ corresponds to

$$T_\psi = (\alpha_{ji}) = \begin{pmatrix} \alpha_{11} & \cdots & \alpha_{1n} \\ \vdots & \ddots & \vdots \\ \alpha_{m1} & \cdots & \alpha_{mn} \end{pmatrix}$$

with ψ sending:

$$\begin{pmatrix} \alpha_{11} & \cdots & \alpha_{1n} \\ \vdots & \ddots & \vdots \\ \alpha_{m1} & \cdots & \alpha_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}$$

7.2.2 General case $V \to W$

With $\psi \in \mathrm{Hom}_{\mathbb{K}}(V,W)$, and isomorphisms $\psi_1 : \mathbb{K}^n \xrightarrow{\sim} V, \psi_2 : \mathbb{K}^m \xrightarrow{\sim} W$ with corresopnding bases $\{v_i\}, \{w_j\}$:

$$\begin{tikzcd}
\mathbb{K}^n \arrow[r, dotted, "\tilde{\psi}"] \arrow[d, "\psi_1"'] & \mathbb{K}^m \arrow[d, "\psi_2"] \\
V \arrow[r, "\psi"'] & W \arrow[u, bend left, "\psi_2^{-1}"]
\end{tikzcd}$$

then $\psi \in \mathrm{Hom}_{\mathbb{K}}(V,W)$ corresponds to $\tilde{\psi} \in \mathrm{Hom}_{\mathbb{K}}(\mathbb{K}^n, \mathbb{K}^m)$ (through ψ_1, ψ_2), and this $\tilde{\psi}$ corresponds to $T_{\tilde{\psi}}$!

Exercise 7.1

Given linear transformation $\psi : \mathbb{K}^n \to \mathbb{K}^n$ that corresponds to $T_\psi \in \mathbb{M}_{n\times n}(\mathbb{K})$.
Show that ψ is isomorphism $\Leftrightarrow T_\psi$ is invertible.

Remark

Consider $\psi : \mathbb{K}^n \to \mathbb{K}^m$ that corresponds to matrix $T_\psi = A = (\alpha_{ji})$. Then,

$$\ker(\psi) = \{v \in \mathbb{K}^n \mid \psi(v) = 0\} = \left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{K}^n \,\middle|\, A \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = 0 \right\}$$
$$= \text{ null space of } A$$

$$\mathrm{im}(\psi) = \{w \in \mathbb{K}^m \mid w = \psi(v) \text{ for some } v\} = \left\{ \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = A \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \text{ for some } \{x_1, \ldots, x_n\} \right\}$$
$$= \text{ range of } A$$

Recall

Relating the this with a previous dimensional equality:

$$\begin{aligned}
\dim_{\mathbb{K}} \mathbb{K}^n &= n \\
&= \dim_{\mathbb{K}}(\mathrm{im}(\psi)) + \dim_{\mathbb{K}}(\ker(\psi)) \\
&= \text{ rank of } A + \text{ nullity of } A
\end{aligned}$$

7.3 Determinant

Determinant is simply a function $D : \mathbb{M}_{n\times n}(\mathbb{K}) \to \mathbb{K}$

Definition 7.1 (Multilinearity and Alternating)

A function $f : \mathbb{M}_{n\times n}(\mathbb{K}) \to \mathbb{K}$ is called **multilinear** if the following holds:

Given $A = \begin{pmatrix} r_1 \\ \vdots \\ r_n \end{pmatrix}$ where row $r_i = (a_{i1}\, a_{i2}\, \dots\, a_{in})$,

$$f\begin{pmatrix} r_1 \\ \vdots \\ \alpha r_i + \beta r_i' \\ \vdots \\ r_n \end{pmatrix} = \alpha f\begin{pmatrix} r_1 \\ \vdots \\ r_i \\ \vdots \\ r_n \end{pmatrix} + \beta f\begin{pmatrix} r_1 \\ \vdots \\ r_i' \\ \vdots \\ r_n \end{pmatrix} \text{ where } \alpha, \beta \in \mathbb{K}$$

f is **alternating** if the following holds:

1. $f\begin{pmatrix} r_1 \\ \vdots \\ r_n \end{pmatrix} = 0$ whenever $\exists\, r_i = r_j, i \neq j$

2. $f\begin{pmatrix} r_1 \\ \vdots \\ r_i \\ r_{i+1} \\ \vdots \\ r_n \end{pmatrix} = -f\begin{pmatrix} r_1 \\ \vdots \\ r_{i+1} \\ r_i \\ \vdots \\ r_n \end{pmatrix}$

Remark

If $2 \neq 0$ in $\mathbb{K}$ then the second condition for alternating implies the first one.

Definition 7.2 (Determinant)

A **determinant** function $\mathbb{M}_{n\times n}(\mathbb{K})$ is a multilinear and alternating function $D : \mathbb{M}_{n\times n}(\mathbb{K}) \to \mathbb{K}$ s.t. $D(I_n) = 1$

Remark

For each n there is a unique determinant function $\mathbb{M}_{n\times n}(\mathbb{K})$, usually written as det. To be discussed further next lecture.

Lecture 8

Determinant

23 Apr 2023

Motivation

The motivation for representing matrices in such a manner now becomes clearer for us. Let $\psi_1 : \mathbb{K}^l \to \mathbb{K}^n, \psi_2 : \mathbb{K}^n \to \mathbb{K}^m$ be linear transformations with corresponding $T_1 \in \mathbb{M}_{n\times l}(\mathbb{K}), T_2 \in \mathbb{M}_{m\times n}(\mathbb{K})$:

$$\mathbb{K}^l \xrightarrow{\psi_1, T_1} \mathbb{K}^n \xrightarrow{\psi_2, T_2} \mathbb{K}^m$$

then it is also an exercise to show that $\psi_2 \circ \psi_1$ is also a linear transformation, that corresponds to $T_2 \cdot T_1 \in \mathbb{M}_{m\times l}(\mathbb{K})$.
Matrix multiplication is therefore built in such a way that $T_2 \cdot T_1$ results in an $m \times l$ matrix. It makes sense to multiply in such a way to fit the shape requirements: $i-$th row by $j-$th column.

Recall

$D : \mathbb{M}_n(\mathbb{K}) \to \mathbb{K}$ is a function that is multilinear, alternating and satisfies: $D(I_n) = 1$. As of now, we don't know if this function exists at all!

Remark

Assuming that D is multilinear, then the first condition for alternating implies the second. When $2 \neq 0$, the second condition imlpies the first one.

Proof (Remark)

$\boxed{\Rightarrow}$ We want to show that

$$D\begin{pmatrix} r_1 \\ \vdots \\ r_n \end{pmatrix} = 0 \text{ whenever } \exists\, i \neq j : r_i = r_j \Rightarrow D\begin{pmatrix} \vdots \\ r_i \\ r_{i+1} \\ \vdots \end{pmatrix} = -D\begin{pmatrix} \vdots \\ r_{i+1} \\ r_i \\ \vdots \end{pmatrix}$$

We have:

$$LHS = D\begin{pmatrix} \vdots \\ r_i \\ r_{i+1} \\ \vdots \end{pmatrix} + 0 = D\begin{pmatrix} \vdots \\ r_i \\ r_{i+1} \\ \vdots \end{pmatrix} + D\begin{pmatrix} \vdots \\ r_{i+1} \\ r_{i+1} \\ \vdots \end{pmatrix}$$

$$= D\begin{pmatrix} \vdots \\ r_i + r_{i+1} \\ r_{i+1} \\ \vdots \end{pmatrix}$$

Similarly,

$$RHS = -D\begin{pmatrix} \vdots \\ r_i + r_{i+1} \\ r_i \\ \vdots \end{pmatrix}$$

Thus,

$$LHS - RHS = D\begin{pmatrix} \vdots \\ r_i + r_{i+1} \\ r_{i+1} \\ \vdots \end{pmatrix} + D\begin{pmatrix} \vdots \\ r_i + r_{i+1} \\ r_i \\ \vdots \end{pmatrix} = D\begin{pmatrix} \vdots \\ r_i + r_{i+1} \\ r_i + r_{i+1} \\ \vdots \end{pmatrix} = 0$$

$\boxed{\Leftarrow}$ The proof backward is similar, only with the requirement that $2 \neq 0$ in $\mathbb{K}$. □

Proposition 8.4

$\forall\, n,\ \exists\,!$ such a function D.

Proof (Proposition)

If $n = 1, D : \mathbb{K}^1 \to \mathbb{K}$, since D must be multilinear (in this case, simply linear):

$$D(\alpha) = D(\alpha \cdot 1) = \alpha \cdot D(1) = \alpha$$

It is trivial that this D satisfies all conditions (2nd condition is satisfied as there are no 2 rows to swap) and is indeed unique.

If $n = 2$:

$$
\begin{aligned}
D\left(\begin{pmatrix} a & b \\ c & d \end{pmatrix}\right) &= D\left(\begin{pmatrix} a & 0 \\ c & d \end{pmatrix} + \begin{pmatrix} 0 & b \\ c & d \end{pmatrix}\right) \\
&= aD\left(\begin{pmatrix} 1 & 0 \\ c & d \end{pmatrix}\right) + bD\left(\begin{pmatrix} 0 & 1 \\ c & d \end{pmatrix}\right) \\
&= a\left[D\left(\begin{pmatrix} 1 & 0 \\ c & d \end{pmatrix}\right) - cD\left(\begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}\right)\right] \\
&+ b\left[D\left(\begin{pmatrix} 0 & 1 \\ c & d \end{pmatrix}\right) - dD\left(\begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}\right)\right] \\
&= aD\left(\begin{pmatrix} 1 & 0 \\ 0 & d \end{pmatrix}\right) + bD\left(\begin{pmatrix} 0 & 1 \\ c & 0 \end{pmatrix}\right) \\
&= adD\left(\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right) + bcD\left(\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\right) \\
&= adD(I_2) - bcD(I_2) \\
&= ad - bc
\end{aligned}
$$

$\square$