Math 20250
Abstract Linear Algebra

Hung C. Le Tran

June 18, 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 $S = \{\sum a_{ij}x_j = b_i\}_{1 \leq i \leq m, 1 \leq j \leq n}$ is a system of linear equations over $\mathbb{K}$. $S' = \{\sum a'_{ij}x_j = b_i\}_{1 \leq i \leq m, 1 \leq j \leq n}$ 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

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

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. $\square$

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 $S \sim S'$. □

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 $(\alpha_1, \alpha_2, \ldots, \alpha_n) \in \mathbb{K}$ 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:

$$\sum a_{1j}x_j = b_1 \tag{1}$$
$$\sum a_{2j}x_j = b_2 \tag{2}$$
$$\vdots = \vdots$$
$$\sum a_{mj}x_j = b_m \tag{m}$$

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

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

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

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

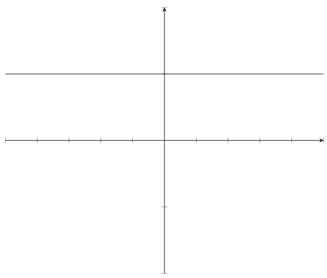
$$\forall\, x, y \in W, x + y \in W$$

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

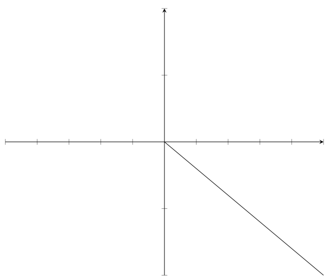
In other words, the subspace is closed under addition and scaling.

Example

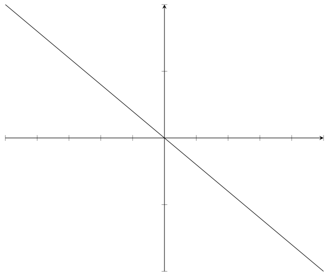
Take $\mathbb{K} = \mathbb{R}, V = \mathbb{R}^2$, with ordinary addition and scaling.
Consider the subset represented by line $y = 1$.



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

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

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

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

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

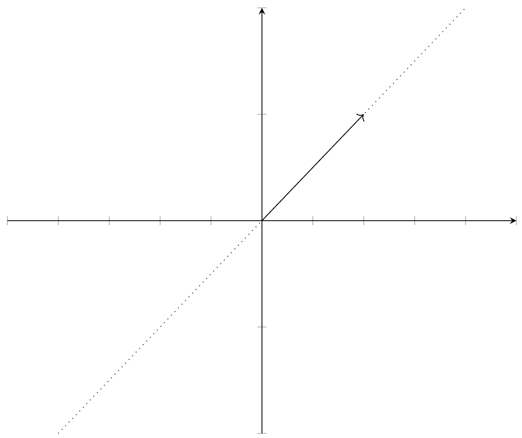
Example

$\mathbb{K} = \mathbb{R}, V = \mathbb{R}, W =$ any $\mathbb{K}$-vector space
We, similarly, claim that there is a bijection between $\text{Hom}_{\mathbb{R}}(V, W)$ 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:

$$\begin{aligned} (v_1, w_1) + (v_2, w_2) &= (v_1 + v_2, w_1 + w_2) \\ \alpha \cdot (v, w) &= (\alpha \cdot v, \alpha \cdot w) \end{aligned}$$

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

$$\dim V := \min\{k \in \mathbb{Z}_{\geq 0} \mid \exists \text{ surjective linear transformation } \mathbb{K}^r \twoheadrightarrow V\}$$

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

$$\mathbb{K}\langle v_1, v_2, \ldots, v_r\rangle \coloneqq \{w \in W \mid w = a_1v_1 + \cdots + a_rv_r; a_i \in \mathbb{K}\} \subseteq V(\text{ is a subspace of } V)$$

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 $\psi \in \text{Hom}_{\mathbb{K}}(\mathbb{K}^r, V) \xrightarrow{\sim} V^{\oplus r}$, 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:

$$\begin{aligned}
\psi : (1, 0, \ldots, 0) \in \mathbb{K}^r &\to v_1 \\
(0, 1, \ldots, 0) \in \mathbb{K}^r &\to v_2 \\
&\vdots \\
(0, 0, \ldots, 1) \in \mathbb{K}^r &\to v_r \\
(\alpha_1, \alpha_2, \ldots, \alpha_r) \in \mathbb{K}^r &\to \alpha_1v_1 + \alpha_2v_2 + \cdots + \alpha_rv_r
\end{aligned}$$

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. $\psi(\alpha_1, \alpha_2, \ldots, \alpha_r) \coloneqq \sum_{i=1}^{r} \alpha_iv_i$, then
$$\text{im}(\psi) = \mathbb{K}\langle v_1, v_2, \ldots, v_r\rangle$$
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 $\mathbb{K}^r \to V$
$\Leftrightarrow$ $\exists\, r$ vectors $v_1, v_2, \ldots, v_r$ that span V.
Recall: $\dim V = \min\{r \in \mathbb{Z}_{\geq 0}$ s.t. $\exists$ 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

$$a_1v_1 + a_2v_2 + \cdots + a_rv_r = 0; a_i \in \mathbb{K} \Rightarrow a_1 = a_2 = \cdots = a_r = 0$$

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 $v_1, v_2, \dots, v_r$ then $v_1, v_2, \dots, v_r$ 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 $\ker(\varphi) = 0$

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

□

Proof (Lemma 5.2)

Applying Lemma 5.3, we want to show: $\ker(\varphi) = 0$ iff $v_1, v_2, \dots, v_r$ 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 $LHS = \varphi((a_1, a_2, \dots, a_r)) \Rightarrow (a_1, a_2, \dots, a_r) \in \ker(\varphi) \Rightarrow (a_1, a_2, \dots, a_r) = 0$.
Therefore $a_i = 0\ \forall\, i$.
$\boxed{\Leftarrow}$ Suppose that $v_1, v_2, \dots, v_r$ are linearly independent.
Then for $v \in \ker(\varphi) \Rightarrow \varphi(v) = 0$, with $v = (a_1, a_2, \dots, a_r)$

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

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

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

□

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 $v_1, v_2, \dots, v_d$ be the d vectors that correspond to π, i.e.

$$\pi((a_1, a_2, \dots, a_d)) = a_1v_1 + \cdots + a_dv_d$$

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

$$b_1v_1 + b_2v_2 + \cdots + b_dv_d = 0$$

WLOG, assume $b_1 \neq 0$.

$$\begin{aligned} &\Rightarrow b_1v_1 = -(b_2v_2 \dots b_dv_d) \\ &\Rightarrow v_1 = -b^{-1}(b_2v_2 \dots b_dv_d)(\,\exists\, b^{-1} \because b_1 \neq 0) \\ &\qquad = c_2v_2 + c_3v_3 + \cdots + c_dv_d \end{aligned}$$

We already know that since π is surjective, thus $v_1, v_2, \dots, v_d$ span V. However, the above equality implies that $v_2, \dots, v_d$ 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. □

Recall

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

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

This naturally leads to our next definition.

Definition 5.3 (Basis)

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

Corollary 5.2

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

Corollary 5.3

If $\{v_1, v_2, \dots, v_d\}$ and $\{w_1, w_2, \dots, v_{d'}\}$ 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:

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

□

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

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

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 6.1

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

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

Consequentially, $\dim_{\mathbb{K}} V = \dim_{\mathbb{K}}(\ker(\psi)) + \dim_{\mathbb{K}}(\operatorname{im}(\psi))$
Warning: $\ker(\psi)$ is a subspace of V, but $\operatorname{im}(\psi)$ 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 $\operatorname{im}(\psi)$.

Remark

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

Proof (Proposition)

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

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

be the subspace spanned by $\{e_i'\}$.

Claim 1: ψ induces an isomorphism

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

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: $\Rightarrow \ker(\psi) \oplus U \xrightarrow{\sim} \ker(\psi) \oplus \operatorname{im}(\psi)$
Claim 2: $\Rightarrow \ker(\psi) \oplus U \xrightarrow{\sim} V$ $\square$
Proving Claim 1: From construction,

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

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

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

But since $\{e_i\}$ forms a basis for $\operatorname{im}(\psi) \Rightarrow \alpha_i = 0 \in \mathbb{K} \Rightarrow \sum_{i=1}^{r} \alpha_i e_i' = 0 \in U \Rightarrow \ker(\varphi) = \{0\}$
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 $w = \psi(v) \in \operatorname{im}(\psi) = \sum \alpha_i e_i$
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 $\psi(x) = 0 \in \operatorname{im}(\psi)$.
But from claim 1, it follows that $x = 0 \Rightarrow \ker(\psi) \cap U = \{0\}$

Lecture 7

Linear Transformation and Matrices

18 Apr 2023 (Zoom)

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
$$V \xrightarrow[\psi]{\sim} W$$
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 $T_\psi = (a_{ji})_{1\le j\le m, 1\le i\le n}$ 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 $\psi : v_i \to \psi(v_i) \in W = c_1w_1 + \cdots + c_mw_m$ then we can define $a_{ji} = c_j$ in this expression. Iterating over $1 \le i \le n$ 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 $\{e_i\}_{1\le i\le n}$:

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

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

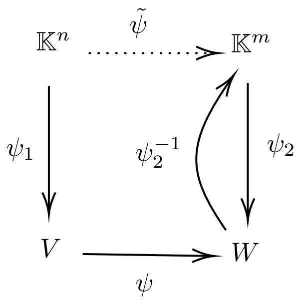
$$T_\psi = (\alpha_{ji}) = \begin{pmatrix}\alpha_{11} & \cdots & \alpha_{1n}\\ \vdots & \ddots & \vdots\\ \alpha_{m1} & \cdots & \alpha_{mn}\end{pmatrix}$$

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 $\{v_i\}, \{w_j\}$:



then $\psi \in \text{Hom}_{\mathbb{K}}(V, W)$ corresponds to $\tilde{\psi} \in \text{Hom}_{\mathbb{K}}(\mathbb{K}^n, \mathbb{K}^m)$ (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,

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

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

Recall

Relating the this with a previous dimensional equality:

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

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

20 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:

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

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 8.1

$\forall\, n,\ \exists\,!$ 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 $n = 2$:

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

□

Lecture 9

The Uniqueness of the Determinant

25 Apr 2023

Goal

For all n, there exists a unique function $\det : \mathbb{M}_n(\mathbb{K}) \to \mathbb{K}$

9.1 Uniqueness of Determinant Function

Proposition 9.1

If a determinant function $D : \mathbb{M}_n(\mathbb{K}) \to \mathbb{K}$ exists then it is unique

Proof (Proposition)

We first prove the proposition for $n = 3$. Suppose

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

Then, if a determinant function D exists,

$$D(A) = a_{11} D \begin{pmatrix} 1 & 0 & 0 \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} + a_{12} D \begin{pmatrix} 0 & 1 & 0 \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} + a_{13} D \begin{pmatrix} 0 & 0 & 1 \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

Now, observe that

$$D \begin{pmatrix} 1 & 0 & 0 \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = a_{21} D \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ a_{31} & a_{32} & a_{33} \end{pmatrix} + D \begin{pmatrix} 1 & 0 & 0 \\ 0 & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

due to the linearity of the second row.

Moreover, $a_{21} D \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = 0$ since rows 1 and 2 are equal. Therefore,

$$D \begin{pmatrix} 1 & 0 & 0 \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = D \begin{pmatrix} 1 & 0 & 0 \\ 0 & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

Consequently,

$$D \begin{pmatrix} 1 & 0 & 0 \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = D \begin{pmatrix} 1 & 0 & 0 \\ 0 & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = D \begin{pmatrix} 1 & 0 & 0 \\ 0 & a_{22} & a_{23} \\ 0 & a_{32} & a_{33} \end{pmatrix}$$

for similar reasons. It follows that we can write $D(A)$:

$$\begin{aligned}
D(A) &= a_{11} D\begin{pmatrix} 1 & 0 & 0 \\ 0 & a_{22} & a_{23} \\ 0 & a_{32} & a_{33} \end{pmatrix} + a_{12} D\begin{pmatrix} 0 & 1 & 0 \\ a_{21} & 0 & a_{23} \\ a_{31} & 0 & a_{33} \end{pmatrix} + a_{13} D\begin{pmatrix} 0 & 0 & 1 \\ a_{21} & a_{22} & 0 \\ a_{31} & a_{32} & 0 \end{pmatrix} \\
&= a_{11}\left[a_{22} D\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & a_{32} & a_{33} \end{pmatrix} + a_{23} D\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & a_{32} & a_{33} \end{pmatrix}\right] + \cdots \\
&= a_{11}a_{22}a_{33} D\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + a_{11}a_{23}a_{32} D\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} + \cdots
\end{aligned}$$

Note that every $D(I')$ in the last expression evaluates to ± 1, since every I' is some row-swapping of the identity matrix, making $D(I') = \pm D(I) = \pm 1$. It naturally follows that D is unique for $n = 3$. Though a tedious procedure, we can carry out the same simplifying steps for all other values of n - therefore we can conclude that if D exists in the first place (if not, we can't even evaluate the last steps), then it must be unique. □

In the next section, we shall inductively construct a determinant function to prove its existence.

9.2 Inductive Construction of Determinant Function

Proposition 9.2

Suppose $\forall\ m \leq n-1, \det_m : \mathbb{M}_m(\mathbb{K}) \to \mathbb{K}$ exists (and is therefore unique). Then a construction of $\det_n : \mathbb{M}_n(\mathbb{K}) \to \mathbb{K}$ is:

$$\det_n(A) = a_{11} \det_{n-1}(A_{11}) - a_{12} \det_{n-1}(A_{12}) + \cdots + (-1)^{n+1} a_{1n} \det_{n-1}(A_{1n}) = \sum_{j=1}^{n} (-1)^{1+j} a_{1j} \det_{n-1}(A_{1j})$$

where $A_{ij} \in \mathbb{M}_{n-1}(\mathbb{K})$ in this case has a meaning that is different from normal usage, being matrix A with its $i-$th row and $j-$th column removed.

Proof (Proposition)

It suffices for us to show that above-constructed $\det_n$ is a determinant function, i.e. multilinear and alternating.

Step 1: Multilinearity

Denote $\det = \det_n$. Let $A = \begin{pmatrix} \alpha_1 \\ \vdots \\ c\alpha_i + d\alpha_i' \\ \vdots \\ \alpha_n \end{pmatrix}, B = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_i \\ \vdots \\ \alpha_n \end{pmatrix}, B' = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_i' \\ \vdots \\ \alpha_n \end{pmatrix}$, where α_i are row vectors.

We want to show

$$\det A = c \det B + d \det B'$$

If $i \neq 1$, the coefficients of the $\det_{n-1}$ terms in all 3 expansions are the same (a_{1j}). Furthermore, since $\det_{n-1}$ is multilinear,

$$\det_{n-1}(A_{1j}) = c \det_{n-1}(B_{1j}) + d \det_{n-1}(B'_{1j})$$

If $i = 1$, the $\det_{n-1}$ terms are all equal, while the coefficients adhere to the multilinearity (row vectors $A_1 = cB_1 + dB_1'$). From these 2 cases, it is clear that det is multilinear.

Step 2: Alternating
We want to show that if the i-th row is the same as the (i+1)-th row, then $\det(A) = 0$.
We first consider the harder case $i = 1$. Let A be as follows

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} = a_{11} & a_{22} = a_{12} & \cdots & a_{2n} = a1n \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

Define T_{ij} as matrix A with rows 1 and 2, columns i, j removed (clearly, $T_{ij} = T_{ji}$). We recall that

$$\det A = a_{11} \det_{n-1}(A_{11}) - a_{12} \det_{n-1}(A_{12}) + \cdots$$

Expanding each component,

$$\begin{aligned}
\det_{n-1}(A_{11}) &= a_{22} \det_{n-2}(T_{12}) - a_{23} \det_{n-2}(T_{13}) + \cdots \\
&= a_{12} \det_{n-2}(T_{12}) - a_{13} \det_{n-2}(T_{13}) + \cdots \\
\det_{n-1}(A_{12}) &= a_{21} \det_{n-2}(T_{21}) - a_{23} \det_{n-2}(T_{23}) + \cdots \\
&= a_{11} \det_{n-2}(T_{21}) - a_{13} \det_{n-2}(T_{23}) + \cdots \\
&\cdots \\
\det_{n-1}(A_{1m}) &= \pm a_{21} \det_{n-2}(T_{m1}) \pm a_{22} \det_{n-2}(T_{m2}) \pm \cdots \pm a_{2n} \det_{n-2}(T_{mn}) \\
&= \pm a_{11} \det_{n-2}(T_{m1}) \pm a_{12} \det_{n-2}(T_{m2}) \pm \cdots \pm a_{1n} \det_{n-2}(T_{mn}) \text{ (without } a_{1m} \text{ term)}
\end{aligned}$$

As a high-level explanation, to calculate $\det_{n-1}(A_{1m})$, we have already removed the first row and m-th column. Therefore, the coefficients are going to be from the second row (which, in this case, is the same as the first row), and we have to remove one more column other than the m-th (which is why there's no a_{1m} term). It is a good reminder that in each expression, the signs alternate. Combining these steps into the original expression:

$$\begin{aligned}
\det_n(A) &= a_{11} \det_{n-1}(A_{11}) - a_{12} \det_{n-1}(A_{12}) + \cdots \\
&= a_{11}(a_{12} \det_{n-2}(T_{12}) - a_{13} \det_{n-2}(T_{13}) + \cdots) + \cdots
\end{aligned}$$

In this expansion, the term $a_{1i}a_{1j} \det_{n-2}(T_{ij})$ appear twice with opposite signs, making $\det A = 0$.
The easier case is when $i \neq 1$. Since the i-th and (i+1)-th rows are both included in the matrix onto which $\det_{n-1}$ is applied, and since $\det_{n-1}$ is alternating, it follows that all $\det_{n-1}$ terms are 0 in the expansion, resulting in $\det A = 0$. det is therefore also alternating. □

Lecture 10

Properties of the Determinant

27 Apr 2023

Recall

The determinant function $D : \mathbb{M}_n(\mathbb{K}) \to \mathbb{K}$ is unique and is defined by:

$$D(A) = \sum_{j=1}^{n} (-1)^{j+1} a_{1j} \det_{n-1}(A_{1j})$$

Proposition 10.1

The proof previously sketched shows that if $F : \mathbb{M}_n(\mathbb{K}) \to \mathbb{K}$ is a multilinear, alternating map then

$$F(A) = \det_n(A) F(I_n)$$

In particular, if $F(I_n) = 1$ then $F(A) = \det(A)$. It is quite clear why $F(I_n)$ should appear at the final expression, as according to what we did to prove uniqueness from existence of the determinant function, the final expression involves the determinant function applied to row permutations of I_n, the value of which evaluates to $\pm F(I_n)$ due to the alternating nature of the function.

Definition 10.1 (Row-varying Determinant Function)

$$D_i(A) := \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det_{n-1}(A_{ij})$$

Then $D_i : \mathbb{M}_n(\mathbb{K}) \to \mathbb{K}$ is also multilinear and alternating, and $D_i(A) = \det A$

10.1 Multiplicativity of Determinant

Let $B \in \mathbb{M}_n(\mathbb{K})$ be any matrix. Consider:

$$\begin{aligned} H : \mathbb{M}_n(\mathbb{K}) &\to \mathbb{K} \\ A &\to \det(A \cdot B) \end{aligned}$$

Claim 10.1

H is multilinear and alternating. In order to prove this, we shall introduce some notations:

- If $S \in \mathbb{M}_{m\times n}(\mathbb{K})$ then $S^T \in \mathbb{M}_{n\times m}(\mathbb{K}), S^T_{ij} := S_{ji}$. This is the transpose matrix.
- If $v = (v_1, v_2, \ldots, v_n) \in \mathbb{M}_{1\times n}(\mathbb{K})$, $v^T = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \in \mathbb{M}_{n\times 1}(\mathbb{K})$
- If $w, v \in \mathbb{M}_{1\times n}(\mathbb{K})$, we define the dot product of v, w as

$$v \cdot w = vw^T = \sum_{i=1}^{n} v_i w_i$$

Proof (Claim)

Multilinearity

Write $A = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}, B = \begin{pmatrix} w_1^T \\ \vdots \\ w_n^T \end{pmatrix}$ where each α_i are row vectors of A, w_i are column vectors of B. Then,

$$A \cdot B = \begin{pmatrix} \alpha_1 w_1^T & \alpha_1 w_2^T & \cdots & \alpha_1 w_n^T \\ \alpha_2 w_1^T & \alpha_2 w_2^T & \cdots & \alpha_2 w_n^T \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_n w_1^T & \alpha_n w_2^T & \cdots & \alpha_n w_n^T \end{pmatrix}$$

If $A = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_i + c\alpha_i' \\ \vdots \\ \alpha_n \end{pmatrix}, A_1 = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_i \\ \vdots \\ \alpha_n \end{pmatrix}, A_2 = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_i' \\ \vdots \\ \alpha_n \end{pmatrix}$ then we want to show $H(A) = H(A_1) + cH(A_2)$

$$\begin{aligned} H(A) = \det(AB) &= \det \begin{pmatrix} \alpha_1 w_1^T & \alpha_1 w_2^T & \cdots & \alpha_1 w_n^T \\ \vdots & \vdots & \ddots & \vdots \\ (\alpha_i + c\alpha_i')w_1^T & (\alpha_i + c\alpha_i')w_2^T & \cdots & (\alpha_i + c\alpha_i')w_n^T \\ \alpha_n w_1^T & \alpha_n w_2^T & \cdots & \alpha_n w_n^T \end{pmatrix} \\ &= \det(A_1 B) + c \det(A_2 B) \\ &= H(A_1) + cH(A_2) \end{aligned}$$

Proof that H is alternating is left as an exercise. □

Corollary 10.1

$$\begin{aligned} H(A) &= \det(A)H(I_n) \text{ (since } H \text{ is multilinear, alternating)} \\ \Rightarrow \det(AB) &= \det A \det B \\ \Rightarrow \det(AB) &= \det A \det B = \det(BA) \end{aligned}$$

Corollary 10.2
If A has either a left or right inverse, then $\det A \neq 0$ in $\mathbb{K}$.

Proof (Corollary)
Suppose $\exists A'$ s.t. $AA' = I_n \Rightarrow 1 = \det(I_n) = \det A \det A' \Rightarrow \det A \neq 0$ □

Proposition 10.2
Given $A \in \mathbb{M}_n(\mathbb{K})$ then

$$\det A = \det A^T$$

Proof (Proposition)
Suppose this is true up to $n-1$. Then,

$$\det A = \sum_1^n (-1)^{1+j} a_{1j} \det_{n-1}(A_{1j})$$

Consider $D_{vert} : \mathbb{M}_n(\mathbb{K}) \to \mathbb{K}$, where

$$D_{vert}(B) = \sum_{i=1}^{n} (-1)^{1+i} b_{i1} \det_{n-1}(B_{i1})$$

We claim that D_{vert} is multilinear and alternating, and $D_{vert}(I_n) = 1$, which would therefore imply that $D_{vert} = \det$. This is trivial.
So,

$$\begin{aligned}
\det(A^T) &= D_{vert}(A^T) \\
&= \sum_{i=1}^{n} (-1)^{1+i} a_{i1}^T \det_{n-1}(A_{i1}^T) \\
&= \sum_{i=1}^{n} (-1)^{1+i} a_{1i} \det_{n-1}((A_{1i})^T) \\
&= \sum_{i=1}^{n} (-1)^{i+1} a_{1i} \det_{n-1}(A_{1i}) \\
&= \sum_{j=1}^{n} (-1)^{j+1} a_{1j} \det_{n-1}(A_{1j}) \\
&= \det A
\end{aligned}$$

□

Definition 10.2 (Adjunct Matrix?)
Let $A \in \mathbb{M}_n(\mathbb{K})$. Define

$$A' = \begin{pmatrix} (-1)^2 \det A_{11} & (-1)^3 \det A_{12} & \cdots & (-1)^{1+n} \det A_{n1} \\ \vdots & \vdots & (-1)^{i+j} \det A_{ij} & \vdots \\ (-1)^{n+1} \det A_{n1} & \cdots & \cdots & (-1)^{2n} \det A_{nn} \end{pmatrix}$$

Then the **adjunct matrix** of A is $(A')^T$

Claim 10.2

$$A(A')^T = (A')^T A = \det A I_n$$

We will prove this in the next lecture, with the upshot being that if $\det A \neq 0$ then

$$A((\det A)^{-1}(A')^T) = I_n \Rightarrow A^{-1} = (\det A)^{-1}(A')^T$$

We can therefore use this formula to find the inverse of A!

Corollary 10.3
A is invertible iff $\det A \neq 0$

Proof (Corollary)
$\boxed{\Rightarrow}$ As proven above in Corollary 10.2
$\boxed{\Leftarrow}$ The above claim provides a constructive way of finding the inverse of A through $\det A$. □

Lecture 11

Spectral Theory: Eigenvalues, Eigenvectors, Eigenspaces

02 May 2023 (Zoom)

Recall

We claimed last lecture that

$$A(A')^T = (A')^T A = (\det A)I_n$$

Proof (Claim)

First recall that

$$A'_{ij} = (-1)^{i+j}\det(A_{ij})$$
$$\Rightarrow (A'^T)_{ij} = A_{ji} = (-1)^{i+j}\det(A_{ji})$$

Let

$$\begin{aligned} P &= A(A')^T \\ &= \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} \det(A_{11}) & \cdots & (-1)^{n+1}\det(A_{n1}) \\ (-1)\det(A_{12}) & \cdots & \vdots \\ \cdots & (A'^T)_{kj} = (-1)^{j+k}\det(A_{jk}) & \vdots \\ (-1)^{n+1}\det(A_{1n}) & \cdots & (-1)^{2n}\det(A_{nn}) \end{pmatrix} \end{aligned}$$

then $P_{ii} = \sum_{j=1}^n a_{ij}(A'^T)_{ji} = \sum_{j=1}^n a_{ij}(-1)^{i+j}\det(A_{ij}) = \det A$, so indeed the diagonal entries are $\det A$. It remains for us to prove that $P_{ij} = 0$ for $i \neq j$.

$$P_{ij} = \sum_{k=1}^n a_{ik}(-1)^{j+k}\det A_{jk}$$

We will change A a little bit by replacing its $j-$th row with its $i-$th row, and denote this matrix B. We know that $\det B = 0$ (having 2 equal rows), and can write its expression:

$$\det B = \sum_{k=1}^n (-1)^{i+k} a_{ik}\det B_{ik}$$

and we can see that

$$\det B_{ik} = (-1)^{i+j}\det A_{jk}$$

(LHS has $i-$th row removed, but the $j-$th row is just the former $i-$th row, while RHS has the original $j-$th row removed. The sign change is to compensate for the alternating nature.)
Therefore,

$$\begin{aligned} 0 = \det B &= \sum_{k=1}^n (-1)^{i+k} a_{ik}\det B_{ik} \\ &= \sum_{k=1}^n (-1)^{i+j+i+k} a_{ik}\det A_{jk} \\ &= \sum_{k=1}^n (-1)^{j+k} a_{ik}\det A_{jk} \\ &= P_{ij} \end{aligned}$$

Just a note, that this entire maneuver could have only been accomplished if $i \neq j$, with the row-changing in the first step. □

11.1 Spectral Theory

Motivation

Spectral theory is the study of eigenvalues and eigenvectors.
Let $T : V \to V$ be a linear transformation over $\mathbb{K}$, otherwise known as a **linear operator** since it acts $V \to V$. We know that if we pick a basis $\mathcal{B} = \{e_1, e_2, \ldots, e_n\}$ then T corresponds to some matrix $[T]_{\mathcal{B}} \in \mathbb{M}_n(\mathbb{K})$. Then the question arises: does there exist a basis such that $[T]_{\mathcal{B}}$ is "simple", whatever that measure of "simple" is: maybe it can have lots of zeroes, triangular etc. However, for this case, we opt to ask if there exists a basis such that $[T]_{\mathcal{B}}$ is diagonal.

Remark

Why are diagonal matrices good? Well, matrix multiplication is much easier to perform on diagonal matrices!
If $A = \begin{pmatrix} \alpha_1 & 0 & \cdots \\ 0 & \alpha_2 & \cdots \\ \vdots & \ddots & 0 \\ \cdots & 0 & \alpha_n \end{pmatrix}, v = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}$ then $Av = \begin{pmatrix} \alpha_1 v_1 \\ \alpha_2 v_2 \\ \vdots \\ \alpha_n v_n \end{pmatrix}$, in particular:

$$A \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} \alpha_1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \alpha_1 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

so A is sending a vector to a multiple of itself!

Definition 11.1 (Eigenvalue, Eigenvector)
Let $T : V \to V$ be a linear operator over $\mathbb{K}$, V is finite dimensional. Then

1. $a \in \mathbb{K}$ is called an **eigenvalue** of T if $\exists\, v \neq 0 \in V$ s.t.

$$T \cdot v = a \cdot v$$

2. If a is eigenvalue of T, then a vector v s.t.

$$Tv = av$$

is called an **eigenvector** of T for the eigenvalue a.

Example

$\begin{pmatrix} 2 & \\ & 1 \end{pmatrix}$ has eigenvalue 2, 1; with $\begin{pmatrix} x \\ 0 \end{pmatrix}$ being an eigenvector of eigenvalue 2, $\begin{pmatrix} 0 \\ y \end{pmatrix}$ being an eigenvector of eigenvalue 1.
$\begin{pmatrix} 1 & \\ & 1 \end{pmatrix}$ has eigenvalue 1, and since $id \cdot v = v = 1 \cdot v \ \forall\ v \in \mathbb{K}^2$, every $v \in \mathbb{K}^2$ is an eigenvector of eigenvalue 1.

Therefore, if $\mathcal{B} = \{e_1, e_2, \ldots, e_n\}$ is a basis of V, and $[T]_\mathcal{B} = \begin{pmatrix} c_1 & & & \\ & c_2 & & \\ & & \ddots & \\ & & & c_n \end{pmatrix}$ then $c_1, c_2, \ldots, c_n$ are eigenvalues and each e_i is an eigenvector of c_i (Its coordinates would be all zeroes, except for the $i-$th position, which would be multiplied by c_i)

Definition 11.2 (Eigenspace)
If a is an eigenvalue of T, define

$$V_a \coloneqq \{v \in V \mid T(v) = a \cdot v\}$$

to be the **eigenspace** of the eigenvalue a.

Claim 11.1
V_a is a vector subspace of T, the proof of which is left as exercise.

Example

If $T : \mathbb{K}^2 \to \mathbb{K}^2, [T]_\mathcal{B} = \begin{pmatrix} 2 & \\ & 1 \end{pmatrix}$ then

$$V_2 = \{(x, 0) \mid x \in \mathbb{K}\}$$
$$V_1 = \{(0, y) \mid y \in \mathbb{K}\}$$

Then a natural question arises, that given $T : V \to V$, how do we find its eigenvalues and eigenvectors?

Proposition 11.1
Let $T : V \to V$ be a linear operator over $\mathbb{K}$, then the following are equivalent:

1. $a \in \mathbb{K}$ is an eigenvalue for T
2. $T - a \cdot id : V \to V$ is not an isomorphism
3. There exists a basis $\mathcal{B} = \{e_1, e_2, \ldots, e_n\}$ for V s.t. $M = [T]_\mathcal{B}$ satisfies $\det(M - a \cdot I_n) = 0$

Remark
In (3), the term $\det(M - aI_n)$ is independent of basis, i.e. if we have a different basis $\mathcal{B}'$ then

$$\det(M' - aI_n) = \det(M - aI_n) = 0$$

This is because $M = CM'C^{-1}$ through change of basis, which implies

$$\begin{aligned} M - aI_n &= C(M' - aI_n)C^{-1} \\ \Rightarrow \det(M - aI_n) &= \det C \det(M' - aI_n) \det C^{-1} \\ &= \det(M' - aI_n) \end{aligned}$$

Proof (Proposition)
We first prove (1) $\Rightarrow$ (2). If $a \in \mathbb{K}$ is an eigenvalue for $T \Rightarrow$ There exists eigenvector v for a.
Therefore, $T - aI_n$ is not an isomorphism since $\exists\, v \neq 0$ s.t. $(T - aI_n)(v) = 0$.
In fact, $V_a = \{v \mid Tv = av = aI_nv \Leftrightarrow (T - aI_n)v = 0\} = \ker(T - aI_n)$
We now prove (2) $\Rightarrow$ (1). We know that $T - aI_n$ is not an isomorphism, and is either not injective or not surjective.
If it is not injective, then $\ker(T - aI_n) \neq \{0\}$.

If it is not surjective, then $\dim \operatorname{im}(T - aI_n) < \dim V$, and since

$$\dim \ker(T - aI_n) + \dim \operatorname{im}(T - aI_n) = \dim V \Rightarrow \dim \ker(T - aI_n) > 0 \Rightarrow \ker(T - aI_n) \neq \{0\}$$

In either case, $\ker(T - aI_n) \neq \{0\} \Rightarrow \ \exists\, v \neq 0$ s.t. $T(v) = aI_n v = av$.
We now prove (2) $\Leftrightarrow$ (3).
Let $\mathcal{B} = \{e_1, \ldots, e_n\}$ be any basis. Let $M = [T]_{\mathcal{B}}$, then $T - aI_n$ corresponds to $M - aI_n$ under $\mathcal{B}$.
Then, $T - aI_n$ is (not) an isomorphism $\Leftrightarrow$ $M - aI_n$ is (not) invertible $\Leftrightarrow \det(M - aI_n) = (\neq)0$ □

Definition 11.3 (Characteristic Polynomial of Matrix)
Let $M \in \mathbb{M}_n(\mathbb{K})$ the **characteristic polynomial** of M is

$$f(x) \coloneqq \det(xI_n - M)$$

Example
$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ then

$$\begin{aligned} f(x) &= \begin{vmatrix} x-a & -b \\ -c & x-d \end{vmatrix} \\ &= (x-a)(x-d) - bc \\ &= x^2 - (a+d)x + (ad - bc) \end{aligned}$$

Definition 11.4 (Chacteristic Polynomial of Linear Operator)
$T : V \to V$ is a linear operator over $\mathbb{K}$. Let $\mathcal{B} = \{e_1, e_2, \cdots, e_n\}$ be any basis. Then the **characteristic polynomial** of T is the characteristic polynomial of $[T]_{\mathcal{B}}$.

Remark
Note, that this is independent of the choice of basis! For the same reason above: that the change of basis matrices can always be applied and not alter these properties.

Corollary 11.1
Eigenvalues of T are precisely the roots of its characteristic polynomial.

Example
If $A = [T]_{\mathcal{B}} = \begin{pmatrix} a_1 & & \\ & \ddots & \\ & & a_n \end{pmatrix}$ then

$$\begin{aligned} f_T(x) &= \det(xI_n - A) \\ &= \begin{vmatrix} x - a_1 & & \\ & \ddots & \\ & & x - a_n \end{vmatrix} \\ &= (x - a_1) \cdots (x - a_n) \end{aligned}$$

Example

If $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ on $\mathbb{K}^2 \to \mathbb{K}^2$ $(\mathbb{K} = \mathbb{R}, \mathbb{C})$ then

$$P_A(x) = \det \begin{pmatrix} x & -1 \\ 1 & x \end{pmatrix} = x^2 + 1$$

Then, if $\mathbb{K} = \mathbb{R}$, then there does not exist any solutions, i.e. no eigenvalues for A over $\mathbb{R}$. However, if $\mathbb{K} = \mathbb{C}$, then $\pm i$ are eigenvalues for A.

Lecture 12

Review before Midterm 2

02 May 2023

12.1 Change of Basis

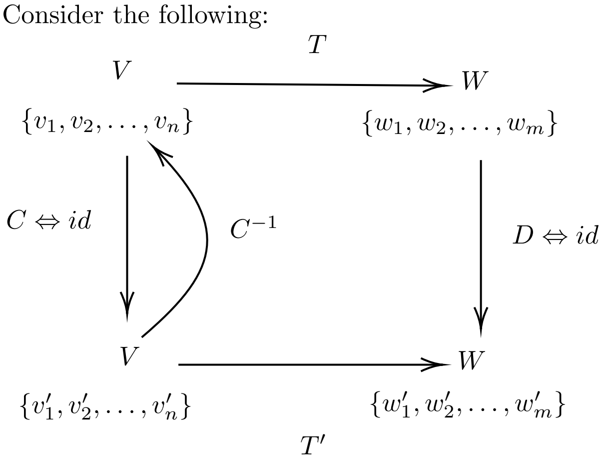
Consider the **identity map**, that is

$$\begin{array}{ccc} V & \xrightarrow{id} & V \\ \mathcal{B} = \{v_1, v_2, \dots, v_n\} & & \mathcal{B}' = \{v'_1, v'_2, \dots, v'_n\} \\ v_i & \mapsto & v_i \end{array}$$

Note that this is not the map **represented** by the identity matrix, which would essentially just map the coordinates to be the same between 2 bases.
The matrix $C = [id]_{\mathcal{B},\mathcal{B}'}$ is the **change of basis matrix**. Essentially all vectors are kept the same (it's an identity mapping!), but just the coordinates are changed (because the basis changed!).
In particular, this mapping would map:

$$v_i \mapsto C_{i1}v'_1 + C_{i2}v'_2 + \cdots C_{in}v'_n$$
$$\begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \mapsto C \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} C_{i1} \\ \vdots \\ C_{ij} \\ \vdots \\ C_{in} \end{pmatrix}$$

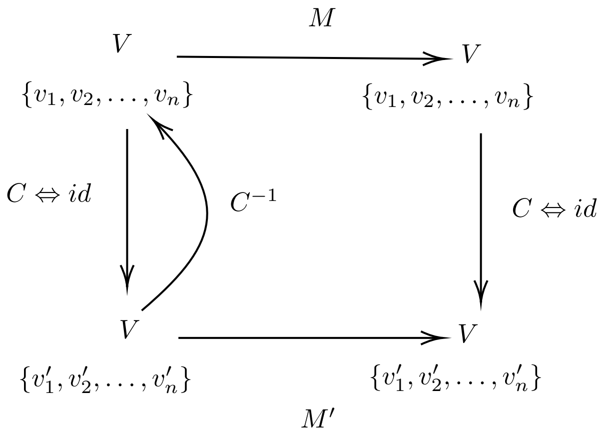
Consider the following:



then

$$T' = DTC^{-1}$$

simply tracing back!
In a specific case,



then

$$M' = CMC^{-1}$$

12.2 Eigenspaces

If α is an eigenvalue of T, $V_\alpha \coloneqq \{v \mid Tv = \alpha v\} = \ker(T - \alpha I_n)$.

Exercise 12.1

If $\alpha \neq \beta, V_\alpha \cap V_\beta = \{0\}$

Lecture 13

Diagonalizable Linear Operators

09 May 2023

Recall

Linear operator $T : V \to V$, then the char poly is $\det(xI_n - T) = \det(xI_n - [T]_{\mathcal{B}})$, which is consistent regardless of choice of basis $\mathcal{B}$!
Solutions of the char poly are the eigenvalues of T (or of $M = [T]_{\mathcal{B}}$).
For each eigenvalue λ of T, there is

$$V_\lambda := \{v \in V \mid Tv = \lambda v\} = \ker(T - \lambda I_n)$$

Exercise 13.1

If $\lambda \neq \lambda'$ then $V_\lambda \cap V_{\lambda'} = \{0\}$

Definition 13.1 (Diagonalizable)

Let $T : V \to V$ be a linear operator. T is diagonalizable if there exists a basis $\mathcal{B} = \{e_1, e_2, \ldots, e_n\}$ of V s.t. $[T]_{\mathcal{B}}$ is a diagonal matrix, i.e. each e_i is an eigenvector of T.

If $[T]_{\mathcal{B}} = \begin{pmatrix} c_1 & & \\ & \ddots & \\ & & c_n \end{pmatrix}$, it sends $Te_i = c_i e_i \Rightarrow e_i$ is an eigenvector of eigenvalue c_i

Proposition 13.1

Let $T : V \to V$ be a linear operator with $\dim V = d$. TFAE (The Following Are Equivalent):

1. T is diagonalizable
2. $P_T(x) = (x - \lambda_1)^{d_1}(x - \lambda_2)^{d_2} \ldots (x - \lambda_n)^{d_n}$ where $\lambda_i \neq \lambda_j$ if $i \neq j$ and $d_i = \dim V_{\lambda_i}$ (the power has to match the dimension of the eigenspace!)
3. Let $\{\lambda_i\}_{i=1,\ldots,l}$ be distinct eigenvalues of T, then $\dim V = \sum \dim V_{\lambda_i}$

Example

$T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, then $P_T(x) = (x-1)^2 \Rightarrow \lambda_1 = 1$

$$V_1 = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in V \middle| \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix} \right\} = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in V \middle| \begin{pmatrix} x+y \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix} \right\}$$

This implies $y = 0 \Rightarrow \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\}$ forms a basis for $V_1 \Rightarrow \dim V_1 = 1 \neq 2$ (the power of $(x-1)^2$). T is therefore NOT diagonalizable.

Proof (Proposition)

(1) $\Rightarrow$ **(2)**

Let $d = \dim V$. Let $\mathcal{B} = \{e_1, e_2, \ldots, e_d\}$ be the basis s.t. $[T]_\mathcal{B}$ is diagonal:

$$[T]_\mathcal{B} = \begin{pmatrix} c_1 & & \\ & \ddots & \\ & & c_d \end{pmatrix}$$

then the characteristic polynomial is

$$(x - c_1)(x - c_2)\ldots(x - c_d) = \prod_{\lambda_i}(x - \lambda_i)^{d_i}$$

and $\dim V_{\lambda_i} = d_i$ (picking out the d_i vectors in $\mathcal{B}$ that correspond to the eigenvalue λ_i)
(2) $\Rightarrow$ **(3)**
If (2) holds then $\sum \dim V_{\lambda_i} = \sum d_i = \deg(P(x)) = \dim V$
(3) $\Rightarrow$ **(1)**
Given eigenvalues $\{\lambda_i\}_{i=1,2,\ldots,l}$.
Consider the map $\pi : V_{\lambda_1} \oplus V_{\lambda_2} \oplus \cdots \oplus V_{\lambda_l} \to V$, sending $(v_1, v_2, \ldots, v_l) \mapsto v_1 + \cdots + v_l$
We claim that $\ker(\pi) = \{0\}$, which in combination with $\dim LHS = \sum \dim V_{\lambda_i} = \dim V = \dim RHS$, implies that π is isomorphic.
Suppose that $(w_1, w_2, \ldots, w_l) \in \ker(\pi)$. Since it is an element of the direct sum of eigenspaces, $Tw_i = \lambda_i w_i$ and $0 = \pi(w_1, \ldots, w_l) = \sum w_i$
Let $w = \sum w_i (= 0)$ then $Tw = \lambda_1 w_1 + \cdots + \lambda_l w_l$
Then

$$0 = Tw - \lambda_1 w = (\lambda_2 - \lambda_1)w_2 + \cdots + (\lambda_l - \lambda_1)w_l$$

which implies $(w_2' = (\lambda_2 - \lambda_1)w_2, \ldots, w_l' = (\lambda_l - \lambda_1)w_l) \in \ker(\oplus_{i=2,\ldots,l} V_{\lambda_i} \to V)$
Repeating this,

$$(\lambda_l - \lambda_1)(\lambda_l - \lambda_2)\ldots(\lambda_l - \lambda_{l-1})w_l \in \ker(V_{\lambda_l} \to V)$$

But V_{λ_l} is a subspace of $V \Rightarrow \ker(V_{\lambda_l} \to V) = \{0\} \Rightarrow (\lambda_l - \lambda_1)(\lambda_l - \lambda_2)\ldots(\lambda_l - \lambda_{l-1})w_l = 0$
But $\lambda_i \neq \lambda_j \Rightarrow w_l = 0 \Rightarrow w_{l-1} = 0 \Rightarrow \cdots \Rightarrow w_1 = 0$
In particular, if $l = 2$,

$$V_{\lambda_1} \oplus V_{\lambda_2} \xrightarrow{\pi} V$$

$$\begin{aligned} \ker(\pi) &= \{(w_1, w_2) \mid w_1 + w_2 = 0\} \\ &= \{(w_1, -w_1) \mid w_1 \in V_{\lambda_1} \cap V_{\lambda_2}\} = \{0\} \end{aligned}$$

Coming back, if π is indeed an isomorphism, then we can pick out bases from V_{λ_i} to form a basis for V that are all eigenvectors. □

Definition 13.2 (Algebraic and Geometric Multiplicity)
If $P_T(x) = (x - \lambda_1)^{d_1}(x - \lambda_2)^{d_2}\ldots(x - \lambda_l)^{d_l}$ then d_i is the **algebraic multiplicity** of λ_i, while $\dim V_{\lambda_i}$ is the **geometric multiplicity** of λ_i. Generally, $\dim V_{\lambda_i} \leq d_i$

Definition 13.3 (Invariant/Stabilised Subspace)
Let $T : V \to V$ be a linear operator. A subspace $W \subseteq V$ is an **invariant/stabilised subspace under** T if $T(W) \subseteq W$, i.e. $T(w) \in W \ \forall\, w \in W$.

Lemma 13.1
Linear operator $T : V \to V$, stabilised subspace W and let $v_1, v_2, \ldots, v_l$ be eigenvectors corresponding to distinct eigenvalues then if $v_1 + \cdots + v_l \in W$ then $v_i \in W$.

Remark
In particular, if $W = \{0\}$ then the result from the previous proof follows, i.e. if $0 = \pi(v_1, v_2, \ldots, v_n) =$

$v_1 + v_2 + \cdots + v_n$ then $v_i = 0$

Proof (Lemma)

We can induct on l. Base case $l = 1$ is clear.
Suppose the lemma holds for $m \leq l - 1$, then we want to show it is true for l.
Let $v = v_1 + \cdots + v_l \in W$, then

$$\begin{aligned} Tv &= \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_l v_l \\ \lambda_1 v &= \lambda_1 v_1 + \lambda_1 v_2 + \cdots + \lambda_1 v_l \\ Tv - \lambda_1 v &= (\lambda_2 - \lambda_1) v_2 + \cdots + (\lambda_l - \lambda_1) v_l \end{aligned}$$

Since $v \in W \Rightarrow Tv, \lambda_1 v \in W \Rightarrow Tv - \lambda_1 v \in W \Rightarrow RHS \in W$.
The induction hypothesis is true for $(l-1) \Rightarrow v_2, \ldots, v_l \in W \Rightarrow v_1 \in W$ □

Corollary 13.1

Suppose T is diagonalizable on V, $W \subseteq V$ be a stabilised subspace under T then

$$W = \bigoplus_{\lambda_i} (W \cap V_{\lambda_i})$$

Proof (Corollary)

Similar to previous proof, consider

$$\bigoplus_{\lambda_i} (W \cap V_{\lambda_i}) \xrightarrow{\pi} W$$

We want to show that π is surjective. It's more obvious that π is injective, but for demonstration let us show surjectivity.
Let $w \in W \Rightarrow w = v_1 + \cdots + v_l$, simply viewing $w \in V$, since there exists an eigenbasis $\{v_1', v_2', \ldots, v_l'\}$, and v_i has simply absorbed all the coefficients. By Lemma, each $v_i \in W \Rightarrow v_i \in W \cap V_{\lambda_i}$ □

Corollary 13.2

If $T : V \to V$ is diagonalizable on V, and W is a stabilised subspace then $T|_W$ is also diagonalizable.

Lecture 14

Generalized Eigenspaces and Cayley-Hamilton

11 May 2023

We continue from last lecture.

Proposition 14.1
Let $T : V \to V$ be a linear operator, and W be a stabilised subspace under T, then $T|_W$ is diagonalizable.

Proof (Proposition)
Since T is diagonalizable,

$$V = \bigoplus_{\lambda} V_\lambda$$

for distinct λ. From previous lecture, we know that

$$W = \bigoplus_{\lambda}(W \cap V_\lambda)$$

for distinct λ of $T : V \to V$. It therefore suffices to show that $W \cap V_\lambda$ is an eigenspace of $T : W \to W$, i.e. $W_\lambda = W \cap V_\lambda$. However, this is clear from the definition itself:

$$LHS = \{w \in W \mid Tw = \lambda w\} = RHS$$

Therefore, $W = \bigoplus_{\lambda} W_\lambda \Rightarrow T|_W$ is diagonalizable. □

14.1 Simultaneously Diagonalizable Linear Operators

Question
Given linear operators $T_1, T_2 : V \to V$ that are both diagonalizable, then when are they simultaneously diagonalizable? i.e. when does there exist an eigenbasis that is an eigenbasis for both T_1, T_2?

Observe that if there exists a common eigenbasis $\mathcal{B}$ then

$$[T_1]_\mathcal{B} = \begin{pmatrix} a_1 & & \\ & \ddots & \\ & & a_n \end{pmatrix}, [T_2]_\mathcal{B} = \begin{pmatrix} b_1 & & \\ & \ddots & \\ & & b_n \end{pmatrix} \Rightarrow [T_1]_\mathcal{B}[T_2]_\mathcal{B} = [T_2]_\mathcal{B}[T_1]_\mathcal{B}$$

T_1, T_2 are commutative!

Proposition 14.2
Let $T_1, T_2 : V \to V$ be diagonalizable linear operators, then they are simultaneously diagonalizable if and only if $T_1 \circ T_2 = T_2 \circ T_1$

Proof (Proposition)
$\boxed{\Rightarrow}$ From our observation above □
$\boxed{\Leftarrow}$ First, let $V = \bigoplus_{\lambda} V_\lambda$ for λ of $T_1 : V \to V$, i.e. $V_\lambda : \{v \in V \mid T_1 v = \lambda v\}$
We then claim that V_λ is invariant under T_2. This would imply that since T_2 is diagonalizable, $T_2|_{V_\lambda}$ is diagonalizable.
Then each $V_\lambda = \bigoplus_{\alpha} W_{\lambda,\alpha}$ for α of $T_2 : V_\lambda \to V_\lambda$ where $W_{\lambda,\alpha} = \{v \in V_\lambda \mid T_2 v = \alpha v\}$.
It then follows that $V = \bigoplus_{\lambda} V_\lambda = \bigoplus_{\lambda}(\bigoplus_{\alpha} W_{\lambda,\alpha})$. Now, since $W_{\lambda,\alpha} = \{v \in V \mid T_1 v = \lambda v, T_2 v = \alpha v\}$ then we are done.
It remains for us to prove the claim, i.e. $T_2(V_\lambda) \subseteq V_\lambda$.
If $x \in V_\lambda$, we want to show that $T_2(x) \in V_\lambda \Leftrightarrow T_1(T_2(x)) = \lambda(T_2(x))$.

But $T_1 \circ T_2 = T_2 \circ T_1 \Rightarrow T_1(T_2(x)) = T_2(T_1(x)) = T_2(\lambda x) = \lambda T_2(x)$ □

Observe

$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ has eigenvalue 1, but not diagonalizable over $\mathbb{R}, \mathbb{C}$, but

$\begin{pmatrix} & 1 \\ -1 & \end{pmatrix}$ is clearly not diagonalizable over $\mathbb{R}$ (no eigenvalue), but potentially over $\mathbb{C}$. It was therefore the fault of the field (in this case, $\mathbb{R}$) that made it not diagonalizable, not fault of the matrix itself. This prompts us to generate a more general definition of eigenspace.

Definition 14.1 (Generalized Eigenspace)

Let λ be an eigenvalue of $T : V \to V$. Define the **generalized eigenspace** of T:

$$\tilde{V}_\lambda \coloneqq \{v \in V \mid (T - \lambda)^m \cdot v = 0 \text{ for some } m \in \mathbb{N}\}$$

$\forall\, m \geq 1$, we can define:

$$V_\lambda^{(m)} \coloneqq \{v \in V \mid (T - \lambda)^m \cdot v = 0\}$$

In particular, the typical eigenspace

$$V_\lambda = V_\lambda^{(1)}$$

and

$$V_\lambda^{(1)} \subseteq V_\lambda^{(2)} \subseteq \cdots \subseteq V_\lambda^{(m)} \subseteq \tilde{V}_\lambda = \bigcup_{m \geq 1} V_\lambda^{(m)}$$

Example

$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ has eigenvalue 1, $\dim V_1 = 1$ but $\dim \tilde{V}_1 = 2$, since

$$(M - I)^2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}^2 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

Lemma 14.1

Suppose the characteristic polynomial of $T : V \to V$ has the form:

$$P(x) = \prod_{\lambda_i} (x - \lambda_i)^{d_i}$$

then $\dim \tilde{V}_{\lambda_i} = d_i$

Lemma 14.2

If the characteristic polynomial has the same form as above, then $V \cong \bigoplus_\lambda \tilde{V}_\lambda$.

Proof (Lemma 14.2)

Using the same idea, we want to show

$$\bigoplus_\lambda \tilde{V}_\lambda \xrightarrow{\tilde{\pi}} V$$

where $\tilde{\pi}$ is the standard addition map, is an isomorphism.

Then we want $\ker(\tilde{\pi}) = \{0\}$, and $\sum \dim \tilde{V}_\lambda = \dim V$, which would imply isomorphism.
It is also an exercise to prove, during this process, that $\tilde{V}_\lambda \cap \tilde{V}_{\lambda'} = 0$ □

Proof (Lemma 14.1)
We inspect the special case when there only exists 1 eigenvalue λ, i.e. $P(x) = (x - \lambda)^d$.
It is already clear that $\dim \tilde{V_1} \leq d$, since $\tilde{V_1} \subseteq V$.
However, we want to show $\dim \tilde{V_1} = d$, i.e. $\forall\, v \in V, (T - \lambda)^m v = 0$ for some v.
Equivalently, we claim that $(T - \lambda)^d v = 0 \;\forall\, v \in V$, i.e. $(T - \lambda)^d = 0$ □

This claim will be proven by the following theorem, which shall prove the previous 2 lemmas at one go.

Theorem 14.1 (Cayley-Hamilton)
Let $T : V \to V$ be a linear operator, and its characteristic polynomial be

$$P(x) = \det(xI_n - T)$$

Then

$$P(T) = 0$$

i.e. if $P(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_d x^d; a_i \in \mathbb{K}$ then

$$a_0 I + a_1 T + a_2 T^2 + \cdots + a_d T^d = 0$$

Remark
This makes the claims above obvious, since $P(x) = (x - \lambda)^d \Rightarrow (T - \lambda)^d = 0$

Example
$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \Rightarrow P(X) = (x - 1)^2$ then according to theorem,

$$\left(\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} - I_2 \right)^2 = 0$$

Remark
If $P(x) = \prod_{\lambda_i} (x - \lambda_i)^{d_i}$ then we can switch around the order

$$(T - \lambda_1)^{d_1} (T - \lambda_2)^{d_2} \cdots = (T - \lambda_2)^{d_2} (T - \lambda_1)^{d_1} \cdots = 0$$

(we could switch the multiplying order in the polynomial, and thus switch the multiplying order of the matrices)
Then if $(T - \lambda_1)^{d_1} (T - \lambda_2)^{d_2} v = 0$ then it's relatively easy to conclude that $v \in \tilde{V_1}$ or $\tilde{V_2}$

Lecture 15

Proving Cayley-Hamilton

16 May 2023

Goal

To prove Cayley-Hamilton!

Recall

Let $T : V \to V$ be a linear operator, and $P(x)$ be the characteristic polynomial of T, then $P(T) = 0$

Proof (C-H for Diagonalizable Linear Operator)

We first look at a special case, where T is diagonalizable. Then, choosing the representation of T in the eigenbasis (it does not matter which basis we choose!):

$$\begin{aligned} P(x) &= (x-\lambda_1)(x-\lambda_2)\cdots(x-\lambda_n) \\ \Rightarrow P(T) &= (T-\lambda_1 I_n)(T-\lambda_2 I_n)\cdots(T-\lambda_n I_n) \\ &= \begin{pmatrix} 0 & & & \\ & \lambda_2-\lambda_1 & & \\ & & \ddots & \\ & & & \lambda_n-\lambda_1 \end{pmatrix} \begin{pmatrix} \lambda_1-\lambda_2 & & & \\ & 0 & & \\ & & \ddots & \\ & & & \lambda_n-\lambda_2 \end{pmatrix} \cdots = 0 \end{aligned}$$

□

Remark

If $P(A) = 0$ for $A \in \mathbb{M}_n(\mathbb{K}) \Leftrightarrow P(CAC^{-1}) = 0$
Again, it doesn't matter which basis we choose!

The following section should be read with a ton of salt, as I could only understand and jot down the brief ideas, and couldn't capture every subtlety that Professor Yao introduced to us.

Proof (Sketch 1, Jordan Canonical Form)

Only works for $\mathbb{C}$, or any algebraically closed field $\mathbb{K}$.
Jordan: If $A \in \mathbb{M}_n(\mathbb{C})$ then $\exists\, C$ s.t. $A' = CAC^{-1}$ has the form

$$A' = \begin{pmatrix} \lambda_1 & * & & & & & & & \\ & \lambda_1 & * & & & & & & \\ & & \lambda_1 & * & & & & & \\ & & & \lambda_1 & & & & & \\ & & & & \lambda_2 & * & & & \\ & & & & & \lambda_2 & & & \\ & & & & & & \lambda_l & * & \\ & & & & & & & \lambda_l & * \\ & & & & & & & & \lambda_l \end{pmatrix}$$

essentially has entries on the diagonal, and 1 off the diagonal, sectioned off into different "squares" with the same eigenvalue on the diagonal within each "square". Then we can show that eventually the off-diagonal entries will die off after some steps. □

Proof (Sketch 2, Abstractifying)
A wrong approach one might have tried is as follows:

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

This is clearly wrong, since x in this case is a scalar, and can't just be replaced by a matrix A. But what if...?

Definition 15.1 (Commutative Ring)
$(R, +, \cdot)$ is a **commutative ring** if

1. $(R, +)$ is an Abelian group
2. $\cdot$ is associative , $\exists\, 1 \in R$
3. $(a + b) \cdot c = (a \cdot c) + (b \cdot c)$
4. $ab = ba$

and we assume that $1 \neq 0$.
i.e. A commutative ring is simply a field, without the multiplicative inverse requirement.

Definition 15.2 (Module over commutative ring)
A **module** over a commutative ring is similar to a vector space over a field, i.e. with the same requirements and properties.
A module M over R is finitely generated(?) if there exists a surjection $R^d \twoheadrightarrow M$.

Let us now consider the polynomials over a field $\mathbb{K}$:

$$\mathbb{K}[x] := \left\{ \sum_{i=0}^{d} a_i x^i \mid a_i \in \mathbb{K} \right\}$$

then $\mathbb{K}[x]$ is actually a commutative ring, with addition and multiplication well-defined. Of course, we can't (and it is not required of us to) enforce the multiplicative inverse requirement. In this case, x is an indeterminant.
Switching perspective, let's look at linear operators. Let $T : V \to V$, and consider

$$\Sigma_T = \text{"}\mathbb{K}[T]\text{"}$$

a "polynomials with entries T (formerly x), with coefficients in $\mathbb{K}$" of sorts.
Intuitively, each element in $\sum_T$ is a linear operator (scaled powers of T are also linear operators, just applied repeatedly and scaled), i.e.

$$\mathbb{K}[T] \subseteq \mathrm{Hom}_{\mathbb{K}}(V, V)$$

In particular, since $\sum_T$ = "$\mathbb{K}[T]$" is a commutative ring, we can consider matrices with entries in this ring. Consider

$$M_{n \times n}(\Sigma_T) := \{(\alpha_{ij}) \mid \alpha_{ij} \in \Sigma_T\}$$

then this has considerably enlarged what we can put into matrices.
Let's draw a parallel between $P(x)$ and $P(T)$. Fix $\mathcal{B} = \{v_1, v_2, \ldots, v_n\}$ to be a basis of V, and let $A = [T]_{\mathcal{B}} = (a_{ij})$.
Then first, since we know $P(x) = \det(xI_n - A)$ (again, does not matter which basis we take!), let us

actually take the determinant of $(xI_n - A)^T$, for purposes that will make sense later:

$$P(x) = \det \begin{pmatrix} x - a_{11} & -a_{21} & \cdots & -a_{n1} \\ -a_{12} & x - a_{22} & \cdots & -a_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{1n} & -a_{2n} & \cdots & x - a_{nn} \end{pmatrix}$$

where the matrix inside the determinant function $\in \mathbb{M}_n(\mathbb{K})$, then similarly

$$P(T) = \det \begin{pmatrix} T - a_{11}I & -a_{21}I & \cdots & -a_{n1}I \\ -a_{12}I & T - a_{22}I & \cdots & -a_{n2}I \\ \vdots & \vdots & \ddots & \vdots \\ -a_{1n}I & -a_{2n}I & \cdots & T - a_{nn}I \end{pmatrix}$$

where the matrix inside the determinant function $B \in \mathbb{M}_n(\sum_T)$
Keep in mind that $P(T) = \det B \in \sum_T$ are linear operators, therefore we now need to prove that it is the zero mapping. To do that, we shall show that it sends v_i in the basis to 0, i.e.

$$(\det B)v_i = 0 \;\forall\; i$$

Equivalently,

$$\begin{pmatrix} \det B & & \\ & \ddots & \\ & & \det B \end{pmatrix} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = 0$$

but we know that

$$\begin{pmatrix} \det B & & \\ & \ddots & \\ & & \det B \end{pmatrix} = (\det B)I_n \in \mathbb{M}_n(\Sigma_T)$$

and recall that we proved

$$(M')^T M = (\det M)I_n$$

for $M \in \mathbb{M}_n(\mathbb{K})$, so we can also use that here when $B \in \mathbb{M}_n(\Sigma_T)$, i.e.

$$(B')^T B = (\det B)I_n$$

Therefore, we want to show that $(B')^T B$ sends $\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$ to 0. Then,

$$B \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} T - a_{11}I & -a_{21}I & \cdots & -a_{n1}I \\ -a_{12}I & T - a_{22}I & \cdots & -a_{n2}I \\ \vdots & \vdots & \ddots & \vdots \\ -a_{1n}I & -a_{2n}I & \cdots & T - a_{nn}I \end{pmatrix} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$$

The entry on the first row would be:

$$\begin{aligned}&(T - a_{11}I)v_1 - a_{21}Iv_2 - \cdots \\ &= Tv_1 - (a_{11}v_1 + a_{21}v_2 + \cdots) \\ &= 0\end{aligned}$$

by sheer definition of $A = [T]_{\mathcal{B}}$. And this argument works for all rows! □

Remark

It should be noted that the last step clarifies why we took the determinant of the transpose matrix in the first place: simply to make the final computations easier. If we did not, then we would have had to show computations with $(B')^T$, and that's not nice.

Proof (Sketch 3, Analysis)

We can perturb A slightly by introducing δ_{ij} to diagonal entries A_{jj} to form a sequence of perturbed matrices A_i that are very close to A. By magic, we can show that

$$P(A) = \lim_{i\to\infty} P_i(A_i) \in \mathbb{M}_n(\mathbb{C})$$

The proof of which uses some "norm" on $\mathbb{M}_n(\mathbb{C}) \cong C^{n^2}$ to show the convergence.
And note that each $P_i(A_i) = 0 \Rightarrow P(A) = 0$ □

Proof (Sketch 4, General Matrices)

My apologies in advance for butchering the arguments, hopefully one day I'll be able to get this in its full glory.
Consider field $L := \mathbb{K}[x_{ij}]_{i,j\leq n}$ then the a super general matrix would be

$$A^{gen} = \begin{pmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & \ddots & \vdots \\ x_{n1} & \cdots & x_{nn} \end{pmatrix} \in \mathbb{M}_n(L)$$

We then plug in all possible roots of L ($\bar{L}$, the algebraic closure of L).
Then

$$\begin{aligned}P(A) &= \det(xI_n - A) \\ &= P(x) = (x-\lambda_1)\cdots(x-\lambda_n) \in \mathbb{M}_n(L)\end{aligned}$$

And we have $\lambda_1, \cdots, \lambda_n$ all distinct since $P'(x) \neq 0 \Rightarrow P(A^{gen} = 0)$ □

Remark

Understanding $\mathbb{Z}$:
For a long long time no one really understood $\mathbb{Z}$, but it shall be understood as

$$\{(a,b) | a, b \in \mathbb{N}, b \neq 0\}$$

equipped with a notion of equivalence $\sim$ defined as:

$$(a,b) \sim (c,d) \Leftrightarrow a + d = b + c \Leftrightarrow \text{“}a - b = c - d\text{”}$$

Then the pair of natural numbers (a,b) shall represent "$a - b$" in the common sense.
It is mindblowing to observe that we intuitively understand this type of construction for the positive

rationals $\mathbb{Q}_+$, just equipped with a slightly different notion of equivalence, namely:

$$(a, b) \sim (c, d) \Leftrightarrow ad = bc \Leftrightarrow \text{“}\frac{a}{b} = \frac{c}{d}\text{”}$$

but ancient mathematicians failed to realize the similar construction for the integers.
Also, a lot of number systems (Roman, Chinese) did not have the symbol for “0” too, indicating that they might not have fully grasped what its meaning is, not even mentioning the negative numbers.
Somehow it got here when we were trying to turn $\mathbb{Z}$ into a field, but it was cool. Fun class!

Lecture 16

Inner Product and Normed Spaces

18 May 2023

16.1 Motivating Examples on Inner Product and Norm

Remark

For the scope of this lecture, we shall restrict our field $\mathbb{K}$ to either $\mathbb{R}$ or $\mathbb{C}$ (I'll denote $\mathbb{R}/\mathbb{C}$), which are called "normed" fields, that are roughly fields in which we can make sense of the notion of distance.

Example

On $\mathbb{R}, \alpha \in \mathbb{R} \rightsquigarrow |\alpha|$
On $\mathbb{C}, \beta = a + bi \in \mathbb{C} \rightsquigarrow |\beta| = |\beta \cdot \overline{\beta}|^{\frac{1}{2}} = \sqrt{a^2 + b^2}$

Let V be a vector space over $\mathbb{R}/\mathbb{C}$, then we now want to talk about "distancne" in V

Example

If $V = \mathbb{R}^d$. Let

$$x = (x_1, x_2, \dots, x_d) \in \mathbb{R}^d$$
$$y = (y_1, y_2, \dots, y_d) \in \mathbb{R}^d$$

Then the "dot product" is defined over $\mathbb{R}^d$ as follows:

$$\langle x, y \rangle = x \cdot y^T = \sum x_i y_i \in \mathbb{R}$$

Then $\langle x, x \rangle = \sum x_i^2$ and $\|x\| \coloneqq \sqrt{\langle x, x \rangle}$ is the distance to origin.

Example

If $V = \mathbb{C}^d$. Let

$$x = (x_1, x_2, \dots, x_d) \in \mathbb{C}^d$$
$$y = (y_1, y_2, \dots, y_d) \in \mathbb{C}^d$$

Then the "dot product" is defined over $\mathbb{C}^d$ as follows:

$$\langle x, y \rangle = x \cdot \overline{y} = \sum x_i \overline{y_i} \in \mathbb{C}$$

Then $\|x\| \coloneqq \sqrt{\langle x, x \rangle} = \sqrt{x\overline{x}}$, which obviously also applies to $\mathbb{R}$.
If $x = (x_1, x_2) \in \mathbb{C}^2$ then $\|x\| = \sqrt{x_1\overline{x_1} + x_2\overline{x_2}} = \sqrt{|x_1|^2 + |x_2|^2}$

16.2 Inner Product Space

Definition 16.1 (Inner Product Space)
An **inner product space** is a vector space V over $\mathbb{R}/\mathbb{C}$, together with an "inner product mapping":

$$\langle , \rangle : V \times V \to \mathbb{K}(= \mathbb{R}/\mathbb{C}) \text{ s.t.}$$

1. $\langle x, y \rangle = \overline{\langle y, x \rangle}$ (over $\mathbb{R}$, this is simply an equality without the conjugate)
2. $\langle \alpha x + \beta y, z \rangle = \alpha \langle x, z \rangle + \beta \langle y, z \rangle$
3. $\langle x, x \rangle \geq 0 \, \forall \, x \in V$ and
$$\langle x, x \rangle = 0 \Leftrightarrow x = 0$$

This non-negative requirement strengthens our understanding of this as a notion of "distance"!

If $(V, \langle , \rangle)$ is an inner product space as such, we can write

$$\|x\| := \sqrt{\langle x, x \rangle}$$

Observe

$\langle x, y \rangle$ is linear with respect to the first component, and is "conjugate linear" with respect to the second component. While the first part is clue from property (2), the second part can be shown as follows:

$$\begin{aligned}\langle x, \beta_y + z \rangle &= \overline{\beta \langle y, x \rangle + \langle z, x \rangle} \\ &= \overline{\beta \langle y, x \rangle} + \overline{\langle z, x \rangle} \\ &= \overline{\beta} \langle x, y \rangle + \langle x, z \rangle\end{aligned}$$

Lemma 16.1 (Cauchy-Schwarz)
Let $(V, \langle , \rangle)$ be an inner product space. Then $\forall\, x, y \in V$,

$$|\langle x, y \rangle| \leq \|x\| \|y\|$$

Proof (Lemma)
We shall prove it for the case $\mathbb{K} = \mathbb{R}$; when $\mathbb{K} = \mathbb{C}$ the proof is very similar.
Observe than if $y = 0 \Rightarrow \langle x, y \rangle = 0, \|y\| = 0$, this case is trivial.
Therefore let's assume that $y \neq 0$.
Consider

$$\begin{aligned}0 &\leq \langle x - ty, x - ty \rangle \text{ for some } t \in \mathbb{R} \\ &= \langle x, x - ty \rangle - \langle ty, x - ty \rangle \\ &= \langle x, x \rangle - \langle x, ty \rangle - \langle ty, x \rangle + \langle ty, ty \rangle \\ &= \|x\|^2 - 2t \langle x, y \rangle + t^2 \|y\|^2 (\langle x, y \rangle = \langle y, x \rangle \text{ over } \mathbb{R})\end{aligned}$$

Therefore,

$$\|x\|^2 + t^2 \|y\|^2 \geq 2t \langle x, y \rangle$$

In particular, we can choose a convenient value for $t = \frac{\langle x, y \rangle}{\|y\|^2}$, then:

$$\begin{aligned}\|x\|^2 + \frac{|\langle x, y \rangle|^2}{\|y\|^2} &\geq \frac{2|\langle x, y \rangle|^2}{\|y\|^2} \\ \Rightarrow \|x\|^2 \|y\|^2 &\geq |\langle x, y \rangle|^2 \\ \Rightarrow \|x\| \|y\| &\geq |\langle x, y \rangle|\end{aligned}$$

□

Corollary 16.1 (Triangle Inequality)
If $x, y \in V$ where $(V, \langle , \rangle)$ is an inner product space over $\mathbb{R}/\mathbb{C}$ then

$$\|x + y\| \leq \|x\| + \|y\|$$

Also,

$$\|x - y\| \leq \|x\| + \|y\|$$

Proof (Corollary)
WTS:

$$\|x+y\| \leq \|x\| + \|y\| \Leftrightarrow \langle x+y, x+y \rangle \leq \|x\|^2 + \|y\|^2 + 2\|x\|\|y\|$$

The equivalence is possible since both sides are non-negative. Expanding:

$$LHS = \|x\|^2 + \|y\|^2 + \langle x, y \rangle + \langle y, x \rangle$$

We know from Cauchy-Schwarz that $\|x\|\|y\| \geq |\langle x, y \rangle|$, and therefore WTS that

$$\langle x, y \rangle + \langle y, x \rangle \leq 2|\langle x, y \rangle|$$

But this is borderline obvious, let $\langle x, y \rangle = m = a + bi$ then :

$$m + \overline{m} = 2a \leq 2\sqrt{a^2 + b^2} = 2|\langle x, y \rangle|$$

□

This exploration of the relationship between the norm and inner product gives us a more abstract idea of the norm!

16.3 Normed Space

Definition 16.2 (Normed Space)
A **normed** vector space is a vector space V over $\mathbb{K} = \mathbb{R}/\mathbb{C}$, equipped with a function:

$$\|\cdot\| : V \to \mathbb{R} \text{ s.t.}$$

1. $\|x\| \geq 0, \|x\| = 0 \Leftrightarrow x = 0$
2. $\forall\, \alpha \in \mathbb{K}, \|\alpha x\| = |\alpha|\|x\|$
3. $\|x+y\| \leq \|x\| + \|y\| \;\forall\, x, y \in V$

Remark
It is clear from our exploration above, that if $(V, \langle , \rangle)$ is an inner product space over $\mathbb{R}/\mathbb{C}$ then we can build $(V, \|\cdot\|)$ to be a normed space, by defining

$$\|x\| := \sqrt{\langle x, x \rangle}$$

Definition 16.3 (Perpendicular, Orthogonal)
Let $(V, \langle , \rangle)$ be an inner product space over $\mathbb{R}/\mathbb{C}$. Then $v, w \in V$ are said to be **perpendicular** to each other if

$$\langle v, w \rangle = 0$$

and we write

$$v \perp w$$

In general, if $E, F \subseteq V$ are 2 subspaces, then $E \perp F$ if $x \perp y \;\forall\, x \in E,\; \forall\, y \in F$
Similarly, define $x \perp E$ if $x \in V, x \perp w \;\forall\, w \in E$

Observe
If $x \perp y$ then

$$\begin{aligned}
\|x+y\|^2 &= \langle x+y, x+y \rangle \\
&= \langle x, x \rangle + \langle x, y \rangle + \langle y, x \rangle + \langle y, y \rangle \\
&= \|x\|^2 + \|y\|^2
\end{aligned}$$

Pythagoras!

16.4 Orthogonal System of Vectors

Definition 16.4 (Orthogonal System of Vectors)
Let $\{v_1, v_2, \ldots, v_r\}$ be vectors in V, then they are called an **orthogonal system of vectors** if $v_i \perp v_j \; \forall \, i \neq j$

Lemma 16.2
If $\{v_i\}$ is an orthogonal system then

$$\|\sum_{i=1}^{r} \alpha_i v_i\|^2 = \sum_{i=1}^{r} |\alpha_i|^2 \|v_i\|^2$$

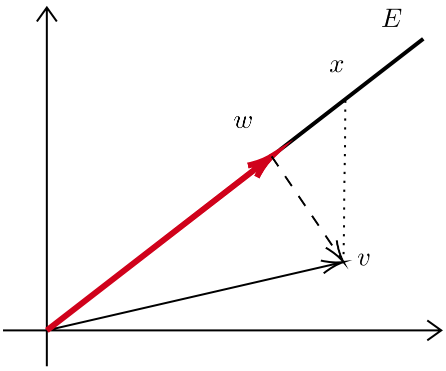
where $a_i \in \mathbb{K}$. The proof of which shall be left as an exercise.

Corollary 16.2
If $\{v_i\}, v_i \neq 0$ is an orthogonal system then they are linearly independent

Proof (Corollary)
Suppose $\sum \alpha_i v_i = 0 \Rightarrow \|\sum \alpha_i v_i\|^2 = 0 \Rightarrow \sum |\alpha_i|^2 \|v_i\|^2 = 0 \Rightarrow |\alpha_i| = 0 \Rightarrow \alpha_i = 0$ □

Definition 16.5 (Orthonormal)
$\{v_1, v_2, \ldots, v_r\}$ is **orthonormal** if it is orthogonal and $\|v_i\| = 1 \; \forall \, i$. It is easy to see that we can always transform an orthogonal system to an orthonormal system of vectors, simply by dividing each vector by their norm.

16.5 Orthogonal Projection



We want to find $w \in E$ s.t. $(x - w) \perp w$

Definition 16.6 (Projection)
Let V be an inner product space, $E \subseteq V$ be a subspace. Define

$$pr_E : V \to E$$

to be a projection s.t. $\forall \, x \in V, x - pr_E(x) \perp pr_E(x)$
So far, we do not know if it even exists, and if it does whether it's unique.

Lemma 16.3

Let $v \in V, w \in E$. Suppose $v - w \perp w$ then $\forall\, x \in E, \|v - w\| \leq \|v - x\|$. Essentially, the projection gives the minimum distance from $V \to E$, and it is unique.

Proof (Lemma)

We first enforce a stronger definition for projection, that is

$$x - pr_E(x) \perp E$$

The proof for the general case is similar to this.
We first prove that if there exists $w = pr_E(x) \in E$, then it is unique.
Write $\delta := w - x$ then $v - x = (v - w) + (w - x) = (v - w) + \delta$
But $\delta = w - x \in E$. Since $v - w \perp E \Rightarrow v - w \perp \delta$. Then,

$$\begin{aligned} \|v - x\|^2 &= \|(v - w) + \delta\|^2 \\ &= \|v - w\|^2 + \|\delta\|^2 (\Leftarrow (v - w) \perp \delta) \\ &\geq \|v - w\|^2 \\ \Rightarrow \|v - x\| &\geq \|v - w\| \end{aligned}$$

Equality holds iff $\delta = 0 \Leftrightarrow w = x$
Let us now prove the existence. We first assume that E has an orthogonal basis $\mathcal{B} = \{e_1, e_2, \ldots, e_n\}$. Then, we can construct

$$pr_E(v) := \sum_{i=1}^{m} \frac{\langle v, e_i \rangle}{\|e_i\|^2} e_i$$

and we can check that this construction indeed works!. For example, in the case of 1-dimensional E,

$$w = pr_E(v) = \frac{\langle v, e_1 \rangle}{\|e_1\|^2} e_1$$

then

$$\begin{aligned} \langle v - w, e_1 \rangle &= \langle v, e_1 \rangle - \langle w, e_1 \rangle \\ &= \langle v, e_1 \rangle - \frac{\langle v, e_1 \rangle}{\|e_1\|^2} \langle e_1, e_1 \rangle \\ &= \langle v, e_1 \rangle - \langle v, e_1 \rangle = 0 \end{aligned}$$

What remains is to show that E has an orthogonal basis, which can be shown using the **Gram-Schmidt process**, that is inductive as follows.
In the base case dim $= 1$, the basis is trivially orthogonal. When dim $= 2$, suppose $\{e_1, e_2\}$ forms the basis for E, then let E_1 be the subspace spanned by e_1. Take $pr_{E_1}(e_2) = w_1 \Rightarrow e_2' := e_2 - w_1 \perp e_1 \Rightarrow \{e_1, e_2'\}$ is an orthogonal system.
The inductive process is then trivial, as we take the projection $w_{n-1} = pr_{E_{n-1}}(e_n) \Rightarrow e_n' = e_n - w_{n-1} \perp E_{n-1} \Rightarrow \{e_1' = e_1, e_2', e_3', \cdots, e_n'\}$ forms an orthogonal system, since $e_n' \perp e_i'\ \forall\, i \leq n - 1 (\Leftarrow e_n' \perp E_{n-1})$
And V is finite-dimensional, so this inductive process will come to a stop. □