LINEAR ALGEBRA II
SPRING 2024 - HKU

ARON HELEODORO

These notes will be updated as the semester progresses. Their goal is to present the material from the textbook and the class in a more concise form. I will often try to give slightly different phrasing (and/or proofs) than the one provided in the textbook. The intent is to make you think the concepts through and to work the concepts by yourself.

You are strongly encouraged to do the exercises as you read. They will help you parse the definitions, examples, and concepts used in the proofs of the theory. Some of these exercises will be assigned as Homework or will be discussed in class.

Points in red and blue are still being edited.

I would appreciate any comments. If you find mistakes, which are probably present, please let me know too. I normally revise part of the notes after the class in which we discussed the material, so please refer frequently to the website for the most up-to-date version.

1. Jan. 15, 2024

1.1. **Fields.** In your previous linear algebra class (Math 2101) you defined a vector space over the real numbers. The very same definition works in a slightly more general context, we start by introducing some terminology for that.

Definition 1. A *field* is a triple $(\mathbb{F}, +, \cdot)$, where $\mathbb{F}$ is a set, and we have operations (i.e. functions):

- addition $+ : \mathbb{F} \times \mathbb{F} \to \mathbb{F}$,
- multiplication $\cdot : \mathbb{F} \times \mathbb{F} \to \mathbb{F}$;

satisfying the following list of axioms:

(a) addition and multiplication are associative;

(b) addition and multiplication are commutative;

(c) there exists $0 \in \mathbb{F}$, such that $a + 0 = 0 + a = a$, for all $a \in \mathbb{F}$;

(d) there exists $1 \in \mathbb{F}$, such that $a \cdot 1 = 1 \cdot a = a$, for all $a \in \mathbb{F}$;

(e) $0 \neq 1$;

(f) every $a \in \mathbb{F}$ has an *additive inverse*, i.e. an element $b \in \mathbb{F}$ such that $a + b = b + a = 0$;

(g) every $a \in \mathbb{F}\backslash\{0\}$ has a multiplicative inverse;

(h) distributivity, i.e. for every $a, b, c \in \mathbb{F}$ one has: $a \cdot (b + c) = a \cdot b + a \cdot c$.

Notation 1. We will omit the $\cdot$ when writing the multiplication operation, i.e. for any $a, b \in \mathbb{F}$ we will write ab for $a \cdot b$.

Example 1. (i) The real numbers $\mathbb{R}$ form a field with usual addition and multiplication.

(ii) The complex numbers $\mathbb{C}$ form a field with usual addition and multiplication.

(iii) The *rational numbers* $\mathbb{Q} := \{\frac{p}{q} \mid p \in \mathbb{Z},\ q \in \mathbb{Z}\backslash\{0\}\}$ are a field.

Exercise 1. Write out explicitly what conditions (a-b) and (g) above are and check them in one of the examples in Example 1.

Date: Last updated April 22, 2024.

Exercise 2. Let p be a prime number and consider $\mathbb{F}_p = \{0, 1 \ldots, p-1\}$, then for $a, b, c \in \mathbb{F}_p$ we define:

$$a + b := c \text{ if } (a + b - c) \text{ is a multiple of } p, \qquad a \cdot b := c \text{ if } (a \cdot b - c) \text{ is a multiple of } p.$$

(i) Check that $+ : \mathbb{F}_p \times \mathbb{F}_p \to \mathbb{F}_p$ and $\cdot : \mathbb{F}_p \times \mathbb{F}_p \to \mathbb{F}_p$ are well-defined.

(ii) Prove that $\mathbb{F}_p$ is a field.

Exercise 3. Can you come up with another example of a field?

1.2. **Vector spaces.** In a previous Linear Algebra class you probably approached vector spaces by concrete examples. The main point of this class is to develop the theory from an abstract point of view focused on proofs, mostly basis-free, and applicable to general fields of characteristic zero, until later results that might require $\mathbb{F}$ to be the real or complex numbers.

Let $\mathbb{F}$ be a field.

Definition 2. A *vector space* over $\mathbb{F}$ is the data of

(i) a set V;

(ii) an operation $+ : V \times V \to V$;

(iii) a *scalar multiplication* operation $\cdot : \mathbb{F} \times V \to V$.

These are subject to the following axioms:

(a) the operation $+$ is associative, commutative, it admits an identity $0_V \in V$ and inverse;

(b) the operation $\cdot$ is associative;

(c) for every $v \in V$ one has $1 \cdot v = v$;

(d) scalar multiplication distributes over vector addition (i.e. the operation $+$ on V) and vector addition distributes over scalar multiplication[1].

Example 2. (i) The set $\{0\}$ is a vector space over any field $\mathbb{F}$.

(ii) Given a set S consider $\mathbb{F}^S$ the set of functions $f : S \to \mathbb{F}$. The operations are defined by pointwise addition and multiplication, i.e. given $f, g \in \mathbb{F}^S$ and $a \in \mathbb{F}$ we let:

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

and $0_{\mathbb{F}^S}$ is the zero function.

(iii) For any $n \geq 1$, the set $\mathbb{F}^n$ is a vector space, where the operations are defined as follows. Let $v = (v_1, \ldots, v_n) \in \mathbb{F}^n$, $w = (w_1, \ldots, w_n) \in \mathbb{F}$, and $a \in \mathbb{F}$, then:

$$v + w := (v_1 + w_1, \ldots, v_n + w_n), \qquad a \cdot v := (av_1, \ldots, av_n),$$

and $0_{\mathbb{F}^n} := (0, \ldots, 0)$.

(iv) For any $n, m \geq 1$ the set $M_{m \times n}(\mathbb{F})$ of $m \times n$ matrices with coefficients in $\mathbb{F}$ equipped with matrix addition and scalar multiplication is a vector space over $\mathbb{F}$.

(v) The set $\mathbb{F}^{\mathbb{N}}$ of sequences with value in $\mathbb{F}$ is a vector space with termwise addition and scalar multiplication.

Remark 1. A set G equipped with an operation $+ : G \times G \to G$ satisfying condition (a) above is an *Abelian group.* These objects are very important in algebra and are studied in more detail in an abstract algebra course, e.g. Math3301 (Algebra I).

Lemma 1. *Let V be a vector space over $\mathbb{F}$.*

(1) Given $v \in V$ such that $v + w = w$ for all $w \in V$, then $v = 0_V$.

[1] See Exercise 1 in Worksheet 1 for an example where this fails.

(2) *The additive inverse is unique.*

(3) *For every $v \in V$, we have $0 \cdot v = 0_V$.*

(4) *For every $a \in \mathbb{F}$, we have $a \cdot 0_V = 0_V$.*

(5) *For every $v \in V$ we have $v + (-1) \cdot v = 0$, i.e. the additive inverse of v is given by $-v := (-1) \cdot v$.*

Proof. (1) We have $v = v + 0_V = 0_V$, where the first equality follows from Definition 2 (a) and the second from the assumption.

(2) Assume there exists $u_1, u_2 \in V$ such that $u_1 + v = 0_V = u_2 + v$. Then we have: $u_1 = u_1 + 0_V = u_1 + u_2 + v = u_2 + u_1 + v = u_2 + 0_V = u_2$.

(3) Notice $v + 0 \cdot v = (1 + 0) \cdot v = 1 \cdot v = v$. Thus by (1), we have $0 \cdot v = 0_V$.

(4) For any $a \in \mathbb{F}$, we have: $a \cdot 0_V = a \cdot (0_V + 0_V) = a \cdot 0_V + a \cdot 0_V$. By (1), we have $a \cdot 0_V = 0_V$.

(5) Notice $v + (-1) \cdot v = (1 + (-1)) \cdot v = 0 \cdot v = 0_V$, where in the last step we used (3).

□

Notation 2. (1) Notice that for $a, b \in \mathbb{F}$ and $v \in V$ we have:

$$(ab) \cdot v = a \cdot (b \cdot v)$$

by Defintition 2 (b). Thus, we can omit the $\cdot$ for the operation of scalar multiplication as we omitted it for multiplication in a field (see) without causing ambiguity.

(2) We will denote the additive inverse of v by $-v$.

(3) We will denote 0_V simply by 0. This should not be confused with $0 \in \mathbb{F}$ the identity of the operation $+$ in $\mathbb{F}$, as these live in different sets, except when $V = \mathbb{F}$, in which case the notation is consistent.

Remark 2. (i) The empty set $\emptyset$ is not a vector space. Namely, it fails condition (a) from Definition 2.

(ii) Condition (a) from Definition 2 can be substituted by

(a)' the operation $+$ is associative, commutative, it admits an identity $0_V \in V$ and (3) from Lemma 1 holds.

Indeed, assume (a)', then we have $0_V = 0 \cdot w = (1 + (-1)) \cdot w = w + (-1)w$ for every $w \in V$. Thus (a) holds.

Example 3. Let V be a vector space over $\mathbb{R}$. We can define a vector space over the complex numbers $V_{\mathbb{C}}$, called the *complexification* of V as follows:

- as a set we let $V_{\mathbb{C}} := V \times V$;
- $+ : V_{\mathbb{C}} \times V_{\mathbb{C}} \to V_{\mathbb{C}}$ is given by $(u_1, v_1) + (u_2, v_2) := (u_1 + u_2, v_1 + v_2)$;
- scalar multiplication is defined as $(a + bi) \cdot (u_1, v_1) = (au_1 - bv_1, bu_1 + av_1)$.

The reader should check that $V_{\mathbb{C}}$ is a vector space over $\mathbb{C}$.

Exercise 4. Universal property of complexification. Let V be a vector space over $\mathbb{R}$ and W a vector space over $\mathbb{C}$. Notice that W can be seen as a vector space over $\mathbb{R}$, where $a \cdot w := (a + i0) \cdot w$, i.e. using the natural inclusion of $\mathbb{R}$ into $\mathbb{C}$. Let $\mathrm{Hom}_{\mathbb{R}}(V, W)$ denote the set of linear operators between V and W, where W is seen as a vector space over $\mathbb{R}$ and let $\mathrm{Hom}_{\mathbb{C}}(V_{\mathbb{C}}, W)$ denote the set of linear operator between $V_{\mathbb{C}}$ and W as vector spaces over $\mathbb{C}$. Prove that there exists a bijection:

$$\mathrm{Hom}_{\mathbb{R}}(V, W) \xrightarrow{\sim} \mathrm{Hom}_{\mathbb{C}}(V_{\mathbb{C}}, W).$$

2. Jan. 18, 2024

2.1. **Subspaces.**

Definition 3. Let V be a vector space, a subset $U \subseteq V$ is said to be a *subspace* if:

(a) $0 \in U$;

(b) the restrictions $+_U : U \times U \to V$ and $\cdot_U : \mathbb{F} \times U \to V$ factors as:

$$\begin{tikzcd} U \times U \arrow[r, "+'_U"] \arrow[dr, "+_U"'] & U \arrow[d] \\ & V \end{tikzcd} \qquad \begin{tikzcd} \mathbb{F} \times U \arrow[r, "\cdot'_U"] \arrow[dr, "\cdot_U"'] & U \arrow[d] \\ & V \end{tikzcd} .$$

Given a subspace $U \subseteq V$ we will simply write $+ : U \times U \to U$ and $\cdot : \mathbb{F} \times U \to U$ for $+'_U$ and $\cdot'_U$, respectively.

Exercise 5. (i) Check that Definition 3 agrees with Definition (1.33) from the textbook.

(ii) Show that only requiring condition (b) in Definition 3 would not agree with the notion as defined in the textbook.

Example 4. (i) let $U \subset \mathbb{F}^n$ defined as $U := \{(v_1, \dots, v_n) \in \mathbb{F}^n \mid v_1 + 2v_2 + \cdots + nv_n = 0\}$;

(ii) let $p \in \mathbb{F}[x, y, z]$ be a polynomial of the form $p(x, y, z) = ax + by + cz$, for some constants $a, b, c \in \mathbb{F}$, then $U := \{(v_1, v_2, v_3) \in \mathbb{F}^3 \mid p(v_1, v_2, v_3) = 0\}$ is a subspace;

(iii) the subset of functions $f : [0, 1] \to \mathbb{R}$ which are continuous is a subspace of all the functions from $[0, 1]$ to $\mathbb{R}$;

(iv) let $U \subset \mathbb{F}[x]$ denote the subset of polynomials p such that $p(0) = p'(0) = \cdots = p^{(k)}(0) = 0$;

(v) the set of all sequences of complex numbers whose limit is 0 is a subspace of $\mathbb{C}^{\mathbb{N}}$.

Exercise 6. (i) Let $p \in \mathbb{R}[x, y, z]$ be a polynomial of degree 1 and define the subset:

$$U_p := \{(v_1, v_2, v_3) \in \mathbb{R}^3 \mid p(v_1, v_2, v_3) = 0\}.$$

Show that U_p is a subspace if and only if p is of the form taken in (ii) of Example 4.

(ii) With the notation as in (i), assume that $p_1, p_2 \in \mathbb{R}[x, y, z]$ are polynomials of degree 1 with no constant term, prove that

$$U_{p_1} \cup U_{p_2} = U_{p_1 p_2}.$$

Is $U_{p_1 p_2}$ a subspace of $\mathbb{R}^3$? What about $U_{p_1} \cap U_{p_2}$?

(iii) Can you guess which types of polynomials $p \in \mathbb{R}[x, y, z]$ have the property that U_p is a subspace of $\mathbb{R}^3$.

Exercise 7. Let $\mathbb{F}^{\mathbb{N}}$ be the vector space of sequences over $\mathbb{F}$. For an integer $p \geq 1$, we define the subset $S_p \subset \mathbb{F}^{\mathbb{N}}$ of sequences $(a_n)_{n \geq 1}$ satisfying:

$$\sum_{n=1}^{\infty} |a_n|^p < \infty.$$

Proof or disproof S_p is a subspace for every integer $p \geq 1$.

Definition 4. Given $U_1, U_2 \subseteq V$ two subspaces of V we define the *sum* $U_1 + U_2 \subseteq V$ as the subset of elements $v \in V$ such that there exist $u_1 \in U_1$ and U_2 such that $u_1 + u_2 = v$. For $U_1, \dots, U_k$ a collection of k subspaces of V, we inductively define:[2]

$$U_1 + \cdots + U_k := U_1 + (U_2(\cdots + U_k)).$$

Example 5. (i) let $U_i = \{(v_1, v_2, v_3, v_4) \in \mathbb{F}^4 \mid v_j = 0 \text{ for } j \neq i\}$ for $i = 1, 2, 3, 4$. Then

$$U_2 + U_3 + U_4 = \{(v_1, v_2, v_3, v_4) \in \mathbb{F}^4 \mid v_1 = 0\}, \qquad U_1 + U_2 + U_3 + U_4 = \mathbb{F}^4.$$

[2] In fact, this definition is independent of the choice of parenthesization, hence justifying the notation.

(ii) let $U_1 = \{(v_1, v_2, v_3, v_4) \in \mathbb{F}^4 \mid v_3 + v_4 = 0 \text{ and } v_1 + v_2 = 0\}$, let $U_2 = \{(v_1, v_2, v_3, v_4) \in \mathbb{F}^4 \mid v_1 = 0\}$, then $U_1 + U_2 = \mathbb{F}^4$.

Exercise 8. In the condition in Definition 4 are the vectors u_1 and u_2 uniquely determined? Compare (i) and (ii) in Example 4.

Definition 5. Given $U_1, U_2 \subseteq V$ two subspaces of V we say that $U_1 + U_2$ is a *direct sum* if u_1 and u_2 are uniquely determined. In this case, we use the notation $U_1 \oplus U_2$[3]. Similarly, given subspaces $U_1, \dots, U_k \subset V$ we say that $U_1 + \cdots + U_k$ is a direct sum if any vector $v \in U_1 + \cdots + U_k$ can be written in an unique way as $v = u_1 + \cdots u_k$, where $u_j \in U_j$ for $1 \leq j \leq k$. We denote the direct sum by $U_1 \oplus \cdots \oplus U_k$.

Exercise 9. Let $V = \mathbb{F}^4$. Provide three distinct subspaces $U_1, U_2, U_3 \subseteq V$ such that:

$$V_1 + V_2 = V_1 \oplus V_2, V_2 + V_3 = V_2 \oplus V_3, \text{ but } V_1 + V_2 + V_3 \neq V_1 \oplus V_2 \oplus V_3.$$

Remark 3. Let $U_1, \dots, U_k \subseteq V$ be a family of subspaces, then we have $U_1 + \cdots + U_k$ is a direct sum if and only if for every $j \in \{1, \dots k\}$ we have $U_j \cap (\sum_{i=1, i\neq j}^{n} U_i) = \{0\}$, where $\sum_{i=1, i\neq j}^{n} U_i$ denotes the sum of $U_1, \dots, U_k$ where we omit U_j.

2.2. **Span and linear dependence.**

Definition 6. Given a subset $S \subseteq V$ we define $\operatorname{Span} S$ the *span of* S to be the subset of V consisting of vectors $v \in V$ such that

$$v = a_1 v_1 + \ldots + a_k v_k$$

for some $k \in \mathbb{N}$, $a_1, \dots, a_k \in \mathbb{F}$, and $v_1, \dots, v_k \in S$. It is convenient to define $\operatorname{Span} \emptyset = \{0\}$. If $\operatorname{Span} S = V$ we say that S *spans* V.

Remark 4. It is clear that $\operatorname{Span} S$ is a vector space and that it contains S. We claim that $\operatorname{Span} S$ is the smallest subspace of V containing S. Consider a subspace $U \subseteq V$ such that $S \subseteq U$, we claim that $\operatorname{Span} S \subseteq U$. Indeed, given $v \in \operatorname{Span} S$ we have $v = v = a_1 u_1 + \ldots + a_k u_k$ for some $a_1, \dots, a_k \in \mathbb{F}$, and $v_1, \dots, v_k \in S$. Since $u_1, \dots, u_k \in U$ we have $v \in U$. Thus, it follows that $\operatorname{Span} S$ belongs to the intersection of all subspaces of V containing $\operatorname{Span} S$.

Example 6. (i) Consider $\{e_1, \dots, e_n\} \subseteq \mathbb{F}^n$, where $e_i = (0, \cdots, 0, 1, 0, \cdots, 0)$ where 1 is in the ith position. Then $\operatorname{Span} \{e_1, \dots, e_n\} = \mathbb{F}^n$.

(ii) Let $a, b, c \in \mathbb{F}$ and consider $S = \{(b, -a, 0), (0, c, -b)\}$, then we have $\operatorname{Span} S = U_p$, where U_p is defined as in Example 4 (ii).

Exercise 10. Let e_i be as in Example 6 (i) and consider the set $S = \{je_i - ie_j\}_{1 \leq i < j \leq n}$, then

$$\operatorname{Span} S = \{(v_1, \dots, v_n) \in \mathbb{F}^n \mid v_1 + 2v_2 + \cdots + nv_n = 0\}.$$

Also notice that S is not a basis in general. How could you change it to be a basis?

Definition 7. A vector space U is *finite-dimensional* if there exists a finite subset $S \subseteq U$ such that $\operatorname{Span} S = U$.

Example 7. (i) $\mathbb{F}^n$ is finite-dimensional;

(ii) the set $\mathcal{P}_n(\mathbb{F})$ of polynomials of degree at most n;

(iii) for any S a finite set $\mathbb{F}^S$ is a finite-dimensional vector space.

Exercise 11. Check which of the examples of vector spaces defined so far are finite-dimensional.

Definition 8. (1) A *polynomial* with coefficients in $\mathbb{F}$ is a function $p : \mathbb{F} \to \mathbb{F}$ such that

$$p(x) = a_0 + a_1 x + \cdots a_n x^n \tag{1}$$

for some $n \in \mathbb{N}$ and $a_i \in \mathbb{F}$.

(2) We let $\mathbb{F}[x]$ denote the set of polynomials in $\mathbb{F}$, notice that the textbook uses the notation $\mathcal{P}(\mathbb{F})$.

[3] At the moment this notation might seem unmotivated, but it will be clearer when we consider this operation on vector spaces.

(3) Given a polynomial $p \in \mathbb{F}[x]$ the *degree* of p is the smallest natural number $n \in \mathbb{N}$ such that p can be written as (1). By convention, we set the degree of the zero polynomial to be $-\infty$.

(4) Let $\mathcal{P}_n(\mathbb{F})$ denote the set of polynomials of degree at most n.

Exercise 12. Check that $\mathcal{P}_n(\mathbb{F})$ forms a vector space.

Exercise 13. Assume that $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. Let $p : \mathbb{F} \to \mathbb{F}$ be a function. Check that $p \in \mathcal{P}_n(\mathbb{F})$ if and only if $p^{(n+1)} = 0$.

Exercise 14. The set $\mathbb{F}[x] = \mathcal{P}(\mathbb{F})$ is a vector space. Think about how we can formally define this.

Definition 9. Let V be a vector space over $\mathbb{F}$. Given a finite subset $S = \{v_1, \ldots, v_n\} \subset V$ we say that S is *linearly independent* if

$$a_1 v_1 + \cdots + a_n v_n = 0 \implies a_1 = \cdots = a_n = 0,$$

where $a_1, \cdots, a_n \in \mathbb{F}$. By convention, we declare that $S = \emptyset$ is linearly independent. We say that a subset $S \subset V$ is linearly dependent if it is not linearly independent.

Example 8. (i) For every $k \in \{1, \ldots, n\}$, the set $S = \{e_1, \ldots, e_k\} \subset \mathbb{F}^n$, where e_i's are defined as in Example 6 (i), is linearly independent.

(ii) For any $k \geq 0$ the set $S_k := \{1, x, \ldots, x^k\} \subset \mathbb{F}[x]$ is linearly independent.

(iii) Given $\{v, w\} \subset V$, then $\{v, w\}$ is linearly independent if and only if $v \neq aw$ for every $a \in \mathbb{F}$ and $bv \neq w$ for every $b \in \mathbb{F}$.

Exercise 15. Given a finite subset $S \subset V$. Prove that S if $0 \in S$ then S is linearly dependent.

Example 9. (1) the subset $S = \{e_1 - e_2, e_2 - e_3, e_3 - e_1\} \subset \mathbb{F}^3$ is linearly dependent.

(2) the subset $S = \{x^2, x^2 - 2x, 3x\} \subset \mathbb{F}[x]$ is linearly dependent.

Exercise 16. Given $S = \{(2,3,1), (1,-1,2), (7,3,c)\} \subset \mathbb{F}^3$. Check that S is linearly dependent if and only if $c = 8$.

Exercise 17. Given $\{v_1, v_2, v_3, v_4\} \subset V$ a linearly independent set. Prove that $\{v_1, v_1 + v_2, v_1 + v_2 + v_3, v_1 + v_2 + v_3 + v_4\}$ is a linearly independent set.

The next result is extremely useful in many future proofs since it allows one to make a linearly dependent set smaller.

Lemma 2. *Let $\{v_1, \ldots, v_n\} \subset V$ be a linearly dependent subset of a vector space V. Then there exists $k \in \{1, \ldots, n\}$ such that*

$$v_k \in \operatorname{Span}\{v_1, \ldots, v_{k-1}\},$$

when $k = 1$ the right-hand side above should be interpreted as $\operatorname{Span}\emptyset$. Moreover, one has:

$$\operatorname{Span}\{v_1, \ldots, v_n\} = \operatorname{Span}\{v_1, \ldots, v_n\} \backslash \{v_k\}.$$

Proof. Since $\{v_1, \ldots, v_n\}$ is linearly dependent there exists $a_1, \ldots, a_n \in \mathbb{F}$ not all zero such that

$$a_1 v_1 + \cdots + a_n v_n = 0.$$

Thus, let $k \in \{1, \ldots, n\}$ such that $a_k \neq 0$, then we have:

$$v_k = -a_k^{-1}(a_1 v_1 + \cdots + a_{k-1} v_{k-1} + a_{k+1} v_{k+1} + \cdots + a_n v_n),$$

where the expression on the right above works for $k \in \{2, \ldots, n-1\}$, we leave it to the reader to write the correct expression for the edge cases. To prove the last assertion we notice that clearly $\operatorname{Span}\{v_1, \ldots, v_n\} \backslash \{v_k\} \subseteq \operatorname{Span}\{v_1, \ldots, v_n\}$. Now suppose that $w \in \operatorname{Span}\{v_1, \ldots, v_n\}$ and let $w = a_1 v_1 + \cdots + a_n v_n$. Since $v_k \in \operatorname{Span}\{v_1, \ldots, v_{k-1}\}$, there exists $b_1, \ldots, b_{k-1} \in \mathbb{F}$ such that $v_k = b_1 v_1 + \cdots + b_{k-1} v_{k-1}$. Then

$$w = (a_1 + b_1) v_1 + \cdots + (a_{k-1} + b_{k-1}) v_{k-1} + \sum_{i=k+1}^{n} a_i v_i,$$

so $w \in \operatorname{Span}\{v_1, \ldots, v_n\} \backslash \{v_k\}$. This finishes the proof. $\square$

Definition 10. Let $T \subseteq V$ be a subset of a vector space. We say that T is a *spanning set of* V if $\operatorname{Span} T = V$.

Lemma 3. *Let V be a finite-dimensional vector space. Consider $S, T \subseteq V$ subsets of a vector space V. Suppose that $\operatorname{Span} S = V$ and that T is a linearly independent subset. Then $|T| \leq |S|$.*

Proof. Let $v_1 \in T$ and consider $S \cup \{v_1\}$. Since $\operatorname{Span} S = V$ we have that $v_1 \in \operatorname{Span} S$, so $S \cup \{v_1\}$ is linearly dependent. By Lemma 2 there exists $u_1 \in S$ such that $\operatorname{Span} S \cup \{v_1\} = \operatorname{Span} S \cup \{v_1\} \backslash \{u_1\}$. Now let $T_1 := T \backslash \{v_1\}$, $S_1' := S \backslash \{u_1\}$, and $S_1 := S_1' \cup \{v_1\}$.

Let $v_2 \in T_1$ and consider $S_1 \cup \{v_2\}$, as argued in the previous paragraph we can find $u_2 \in S_1$ such that $\operatorname{Span} S_1 = \operatorname{Span}(S_1 \cup \{v_2\} \backslash \{u_2\})$. Then we let $T_2 := T_1 \backslash \{v_2\}$, $S_2' := S_1' \backslash \{u_1, u_2\}$, and $S_2 := S_1 \cup \{v_2\}$. Notice that we can repeat this process k times, where $k = |T|$ to obtain two sequences:

$$\emptyset \subset T_k \subset \cdots \subset T_1 \subset T, \qquad \text{and} \qquad S_k' \subset \cdots \subset S_1' \subset S$$

where $\operatorname{Span} S_i = \operatorname{Span} S$, $|S_i'| = |S| - i$ and $|T_i| = |T| - i$ for every $i \in \{1, \ldots, k\}$. This implies that $|S| \geq k = |T|$.

This is the same argument as in (2.22) in the textbook. $\square$

Corollary 1. *Let $U \subseteq V$ be a subset of a finite-dimensional vector space V, then U is finite-dimensional.*

Proof. We do an induction on the number of vectors necessary to span U. The base case is $U = \operatorname{Span} \emptyset = \{0\}$, in which case U is finite-dimensional. Assume that $U \neq \{0\}$ and let $v_1 \in U$ be a non-zero vector. Then if $U = \operatorname{Span} v_1$ we are done, otherwise there exists $v_2 \in U$ such that $v_2 \notin \operatorname{Span} v_1$ and we can consider $\operatorname{Span}\{v_1, v_2\}$. We claim that repeating this step k times gives $\operatorname{Span}\{v_1, \ldots, v_k\} = U$ for some $k \in \mathbb{N}$. Indeed, let $S \subset V$ be a finite set such that $\operatorname{Span} S = V$. Such a set exists since V is finite-dimensional. Then consider $\{v_1, \ldots, v_k\} \subseteq \{v_1, \ldots, v_k\} \cup S$. Since $S \subseteq \{v_1, \ldots, v_k\} \cup S$ spans V, we have that $k \leq |S|$, thus k is finite. $\square$

3. Jan. 22, 2024

3.1. **Basis.** The following concept is extremely important in linear algebra. One could say that the main difference between this course and Math2101 is that in Math2101 one is choosing a basis for every vector space that is considered by default, whereas in Math2102 we are not.

Definition 11. A subset $S \subset V$ is a *basis* if it satisfies:

(a) $\operatorname{Span} S = V$;

(b) S is linearly independent.

Example 10. (i) The set $\{e_1, \ldots, e_n\}$ as defined in Example 6 (i) is a basis of $\mathbb{F}^n$.

(ii) The set $\{1, \ldots, x^4\}$ is a basis of $\mathcal{P}_4(\mathbb{F})$ the vector space of polynomials of degree at most 4.

(iii) The sets $\{(7, 5), (-4, 9)\}$ and $\{(1, 2), (3, 5)\}$ are both basis of $\mathbb{F}^2$.

Remark 5. A subset $S = \{v_1, \ldots, v_n\} \subset V$ is a basis of V if and only if every element $u \in V$ can be written as:

$$u = a_1 v_1 + \cdots + a_n v_n$$

for an unique choice of $a_1, \ldots, a_n \in \mathbb{F}$. Indeed, suppose that there are two n-uples $(a_1, \ldots, a_n), (b_1, \ldots, b_n) \in \mathbb{F}^n$ such that

$$u = a_1 v_1 + \cdots + a_n v_n, \qquad u = b_1 v_1 + \cdots + b_n v_n,$$

and $a_i \neq b_i$ for some $i \in \{1, \ldots, n\}$. Then we have:

$$\begin{aligned} 0 = u - u &= (a_1 v_1 + \cdots + a_n v_n) - (b_1 v_1 + \cdots + b_n v_n) \\ &= (a_1 - b_1) v_1 + \cdots + (a_n - b_n) v_n. \end{aligned}$$

Since S is linearly independent, we have that $a_i = b_i$ for all $i \in \{1, \ldots, n\}$.

One of the consequences of Lemma 2 is that any finite spanning set contains a subset which is a basis.

Lemma 4. *Let $T \subset V$ be a finite spanning subset of V. Then there exists $S \subseteq T$ such that S is a basis.*

Proof. We proceed by downward induction. If T is linearly independent we are done. If T is linearly dependent, by Lemma 2 there exists $v \in T$ such that $\operatorname{Span} T \backslash \{v\} = \operatorname{Span} T = V$ and $|T \backslash \{v\}| < |T|$. Since T is finite this process stops and we obtain a basis. □

We get two immediate consequences:

Corollary 2. *(1) Every finite-dimensional vector space V admits a basis.*

(2) Any linearly independent subset $S = \{v_1, \dots, v_k\} \subset V$ extends to a basis.

Proof. For (1) let T be a finite set such that $\operatorname{Span} T = V$. By Lemma 4 there exists $S \subseteq T$ such that S is a basis of V.

For (2) let $T = \{w_1, \dots, w_n\}$ be a finite set such that $\operatorname{Span} T = V$. Then $\operatorname{Span} S \cup T = V$. Order the set $T \cup S$ as follows $\{v_1 < v_2 < \cdots < v_k < w_1 < \cdots < w_n\}$, then running the argument in the proof of Lemma 4 we notice that we obtain a subset $R \subset V$ such that:

$$S \subseteq R \subset S \cup T \qquad \text{and} \qquad \operatorname{Span} R = V.$$

□

The following result is interesting because it uses that $\mathbb{F}$ is a field in a serious way. In other words, certain concepts so far would make sense for more general objects as (commutative) rings, i.e. the integers $\mathbb{Z}$, however, the following result is on of the first to fail.

Lemma 5. *Let V be a finite vector space and consider a subspace $U \subseteq V$. Then there exists a subspace $W \subseteq V$ such that $U \oplus W = V$.*

Proof. Notice that U is also finite-dimensional. Let T be a basis for U (it exists by Corollary 2 (1)). By Corollary 2 (2) we can find $T \subset R$ such that R is a basis of V. We claim that $W := \operatorname{Span} R \backslash T$ satisfies $U \oplus W = V$. Indeed, it is clear that $U + W = V$, by Remark 3 we need to check that $U \cap W = \{0\}$. We give names to the elements of $U = \{v_1, \dots, v_k\}$ and $W = \{v_{k+1}, \dots, v_n\}$. Assume by contradiction that there exists a non-zero vector $v \in U \cap W$, then we have

$$v = a_1 v_1 + \cdots a_k v_k = a_{k+1} v_{k+1} + \cdots + a_n v_n.$$

Thus, $a_1 v_1 + \cdots a_k v_k - (a_{k+1} v_{k+1} + \cdots + a_n v_n) = 0$, and since $\{v_1, \dots, v_k, v_{k+1}, \dots, v_n\}$ is linearly independent, we have that $a_i = 0$ for all $i \in \{1, \dots, n\}$. So we get a contradiction with $U \cap W \neq \{0\}$. This finishes the proof. □

3.2. **Dimension.** The notion of dimension is rather intuitive. The next result justifies that one can define it in a naïve way.

Lemma 6. *Given T, S two basis of a vector space V, we have $|S| = |T|$.*

Proof. Notice that S and T are both linearly independent sets and spanning sets for V. Thus, Lemma 3 implies that $|S| \leq |T|$ and $|T| \leq |S|$. □

Definition 12. The *dimension* of a vector space V, denoted by $\dim V$, is the size of any basis of V.

Exercise 18. Go through all the examples of finite-dimensional vector spaces we had so far and find out their dimension.

Here are a couple of easy consequences of the defintion.

Lemma 7. *Assume that V is finite-dimensional.*

(1) For any subspace $U \subseteq V$, we have $\dim U \leq \dim V$.

(2) Let $S \subseteq V$ be a linearly independent set, if $|S| = \dim V$, then S is a basis.

(3) Given a subspace $U \subseteq V$ such that $\dim U = \dim V$, then $U = V$.

(4) Let $S \subseteq V$ such that $\operatorname{Span} S = V$ and $|S| = \dim V$, then S is a basis.

Proof. (1) Let $S \subset U$ be a basis of U. Notice that $S \subset V$ is also linearly independent. Now Lemma 3 implies that $|S| \leq |T|$ for any basis T of V, i.e. $|S| \leq \dim V$.

(2) Assume that S is not a basis, then by Corollary 2 (2) there exists $S \subset S'$ such that S' is a basis. However, this would imply that $\dim V = |S| < |S'|$ where S' is a basis of V, which is a contradiction.

(3) Let $S \subset U$ be a basis of U. Since $S \subset V$ is linearly independent in V and $|S| = \dim U = \dim V$, by (2) we have S is a basis of V. Thus, $U = \operatorname{Span} S = V$.

(4) Assume that S is not a basis, i.e. S is linearly dependent, then by Lemma 2 there exist $v \in S$ such that $\operatorname{Span} S \backslash \{v\} = V$. This gives $\dim V \leq |S| - 1$, which is a contradiction. □

Now we investigate how the notion of dimension interacts with sums of subspaces.

Lemma 8. *Let $U_1, U_2 \subset V$ be subspaces of V. Then we have:*

$$\dim U_1 + U_2 = \dim U_1 + \dim U_2 - \dim U_1 \cap U_2.$$

Proof. Let S_{12} be a basis of $U_1 \cap U_2$. By Corollary 2 (2) there exists $S_{12} \subset S_1$ and $S_{12} \subset S_2$ such that S_i is a basis of U_i, for $i = 1, 2$. We claim that $S_1 \cup S_2$ is a basis of $U_1 + U_2$. Clearly, we have $S_1 \cup S_2 \subset U_1 + U_2$, this gives that $\operatorname{Span} S_1 \cup S_2 \subset U_1 + U_2$. Since $U_1 \subseteq \operatorname{Span} S_1 \cup S_2$ and $U_2 \subseteq \operatorname{Span} S_1 \cup S_2$, we obtain $\operatorname{Span} S_1 \cup S_2 = U_1 + U_2$.

Now, we need to check that $S_1 \cup S_2$ is linearly independent. For this we actually need to give names to the elements of S_{12}, S_1 and S_2. Let $S_{12} = \{u_1, \ldots, u_i\}$, $S_1 \backslash S_{12} = \{v_1, \ldots, v_j\}$ and $S_2 \backslash S_{12} = \{w_1, \ldots, w_k\}$. Suppose we have an equation:

$$a_1 u_1 + \cdots + a_i u_i + b_1 v_1 + \cdots + b_j v_j + c_1 w_1 + \cdots + c_k w_k = 0,$$

for some $a_1, \ldots, a_i, b_1, \ldots, b_j, c_1, \ldots, c_k \in \mathbb{F}$. Then solving for $w := c_1 w_1 + \cdots + c_k w_k$ we have that $w \in U_1$. But this also gives that $w \in U_2$. Thus, there exist some scalars $d_1, \ldots, d_i$ such that

$$c_1 w_1 + \cdots + c_k w_k = d_1 u_1 + \cdots d_i u_i.$$

Now, since $\{w_1, \ldots, w_k\} \cup \{u_1, \ldots, u_i\} = S_2$ is a linearly independent set, we get that all c_ℓ's vanish. Thus we have that $a_1 u_1 + \cdots + a_i u_i + b_1 v_1 + \cdots + b_j v_j = 0$. Since $S_1 = \{v_1, \ldots, v_j\} \cup \{u_1, \ldots, u_i\}$ is linearly independent, we get that all a_ℓ's and b_ℓ's also vanish. This finishes the proof. □

The previous result is an example of how questions about vector subspaces can be reduced to set-theoretic questions by using bases. We will return to this in later sections.

Exercise 19. Let V be a ten-dimensional vector space.

(1) Suppose that $U_1, U_2 \subset V$ are subspaces of dimension 6. Prove that there exists two vectors $u_1, u_2 \in U_1 \cap U_2$ such that neither is a scalar multiple of the other.

(2) Suppose that $U_1, U_2, U_3 \subset V$ are subspaces such that $\dim U_1 = \dim U_2 = \dim U_3 = 7$, prove that $U_1 \cap U_2 \cap U_3 \neq \{0\}$.

Exercise 20. Let $U_1, \ldots, U_m \subseteq V$ be a set of subspaces such that $U_1 \oplus \cdots \oplus U_m = V$.

(i) Prove that $\dim V = \sum_{i=1}^m \dim U_i$.

(ii) Let $B_i = \{v_1^i, \ldots, v_{k_i}^i\}$ be bases of U_i for $1 \leq i \leq m$, where $k_i = \dim U_i$. Prove that $B := \sqcup_{1 \leq i \leq m} B_i$ is a basis of V.

4. Jan. 25, 2024

4.1. **Linear Maps.** Most objects we encounter in mathematics only have "real" meaning when compared in an appropriate way to other objects of the same type. For instance, when study sets we are naturally lead to studying functions and comparing sets using them.

Vector spaces, a more structure kind of set, need to be compared with each other using a more structured kind of function. We introduce that now:

Definition 13. Let V and W be two vector spaces. A *linear map*, sometimes also called a *linear transformation*, is a function $T : V \to W$ satisfying:

(a) (additivity) $T(u + v) = T(u) + T(v)$ for every $u, v \in U$;

(b) (homogeneity) $T(au) = aT(u)$ for every $a \in \mathbb{F}$ and $u \in U$.

We will let $\mathcal{L}(V, W)$ denote the set of linear maps between V and W, and simply write $\mathcal{L}(V) := \mathcal{L}(V, V)$ for the set of linear maps from V to itself. Also notice that for any linear map T, one has $T(0) = 0$.

Example 11. Let V be a vector space.

(i) The zero map $0 : V \to V$, where $0(V) := 0_V = 0$ is a linear map.

(ii) The identity map $\mathrm{Id}_V : V \to V$ given by $\mathrm{Id}_V(v) = v$.

(iii) Given any $a \in \mathbb{F}$ then $T_a(v) := a \cdot v$ is a linear map.

(iv) Differentiation $D : \mathbb{R}[x] \to \mathbb{R}[x]$ is a linear map, i.e. $D(p) := p'$.

(v) Integration $T_{[0,1]} : \mathbb{R}[x] \to \mathbb{R}$ given by $T_{[0,1]}(p) := \int_0^1 p(x)dx$.

(vi) Let $q \in \mathbb{R}[x]$ then $T_q : \mathbb{R}[x] \to \mathbb{R}[x]$ given by $p(x) \mapsto p(q(x))$ is a linear map.

(vii) Let consider a collection of scalars $a_{i,j} \in \mathbb{F}$ for $1 \leq i \leq m$ and $1 \leq j \leq n$.
$a_{1,1}, \ldots, a_{n,1}, a_{1,2}, a_{2,2}, \ldots, a_{n,2}, \ldots, a_{1,m}, \ldots, a_{n,m} \in \mathbb{F}$, then

$$T(v_1, \ldots, v_n) := (a_{1,1}v_1 + \cdots + a_{1,n}v_n, \ldots, a_{m,1}v_1 + \cdots + a_{m,n}v_n)$$

is a linear map.

Exercise 21. Prove that any linear map $T : \mathbb{F}^n \to \mathbb{F}^m$ is of the form given in Example 11.(vii).

Lemma 9. *Let V and W be finite-dimensional vector spaces and $\{v_1, \ldots, v_n\} \subset V$ be a basis of V. Given any subset $\{w_1, \ldots, w_n\} \subset W$ there exists an unique linear map $T : V \to W$ such that*

$$T(v_i) = w_i \qquad \text{for } i \in \{1, \ldots, n\}. \tag{2}$$

Proof. We first define T. Given any $v \in V$ can be written as $v = a_1v_1 + \cdots + a_nv_n$ we let $Tv := c_1w_1 + \cdots + c_nw_n$. Notice that this is well-defined, since there is only one single way of written v as above and that it satisfies the conditions required. It is clear that it is a linear operator, we leave the details to be checked to the reader. Finally, assume that there exists $T' \in \mathcal{L}(V, W)$ satisfying equations (2). Then for any $a_i \in \mathbb{F}$ we have $T'(a_iv_i) = a_iw_i$, thus for any $v \in V$, which can be written uniquely as $v = a_1v_1 + \cdots + a_nv_n$ we obtain:

$$T'(v = a_1v_1 + \cdots + a_nv_n) = a_1w_1 + \cdots + a_nw_n = T(v).$$

This finishes the proof of uniqueness.

We leave the details to check that T is a linear map to the reader. □

We now observe that the set $\mathcal{L}(V, W)$ can be naturally endowed with the structure of a vector space.

Lemma 10. *For any two vector spaces V, W the set $\mathcal{L}(V, W)$ is a vector space with addition and scalar multiplication defined as:*

$$(T_1 + T_2)(v) := T_1(v) + T_2(v) \qquad \textit{and} \qquad (a \cdot T)(v) := a \cdot T(v).$$

Proof. The details are left to the reader. □

Exercise 22. Given U, V, and W vector spaces we can consider the composition operation:

$$\begin{aligned}(-) \circ (-) : \mathcal{L}(V, W) \times \mathcal{L}(U, V) &\to \mathcal{L}(U, W)\\ (S, T) &\mapsto S \circ T(v) := S(T(v)).\end{aligned}$$

(i) Check that $\circ$ defined above is a linear map.

(ii) Check that the operation $\circ$ is associative and that it has identity elements. Part of the exercise is making sense of what that means.

Remark 6. This is an abstract remark and can be skipped as we will not use this concept in this course.. A mathematical concept that is really helpful in organizing certain mathematical objects is that of a category. You can look up its definition here. We essentially just showed that vector spaces together with linear maps form a category Vect. In fact, the category Vect has many nice properties.

Exercise 23. Let V be a vector space, such that $\dim V > 1$. Prove that there exists $T, S \in \mathcal{L}(V)$ such that $ST \neq TS$.

4.2. **Null spaces and ranges.** In this subsection we define subspaces that are naturally associated to a linear operator.

Definition 14. Let $T : V \to W$ be a linear map.

(1) the *null space* of T:
$$\operatorname{null} T := \{v \in V \mid Tv = 0\};$$

(2) the *range* of T:
$$\operatorname{range} T := \{w \in W \mid Tv = w, \text{ for some } v \in V\}.$$

We check that $\operatorname{null} T$ and $\operatorname{range} T$ are in fact subspaces of V and W, respectively. Let $u, v \in \operatorname{null} T$ we have
$$T(u+v) = T(u) + T(v) = 0, \qquad \text{and} \qquad T(a \cdot u) = a \cdot T(u) = a \cdot 0 = 0.$$

Assume that $w_1, w_2 \in \operatorname{range} T$, then there exist $u_1, u_2 \in U$ such that $T(u_i) = w_i$, for $i = 1, 2$. Then we have
$$T(u_1 + u_2) = w_1 + w_2, \qquad \text{and} \qquad T(a \cdot u_1) = a \cdot w_1,$$
thus $\operatorname{range} T$ is a vector space.

Remark 7. The null space is sometimes also called *kernel* of T. The range is sometimes called *image* of T.

The next results shows how the kernel and range related to the notions of injective and surjective.

Lemma 11. *Let $T : V \to W$ be a linear operator.*

(i) T is injective if and only if $\operatorname{null} T = \{0\}$*;*

(ii) T is surjective if and only if $\operatorname{range} T = W$*.*

Proof. For (i), first assume that T is injective. Assume that there exists a non-zero vector $v \in \operatorname{null} T$. However, this is a contradiction with $Tv \neq 0$, since $v \neq 0$. Now assume that T is injective, and suppose that $Tv = 0$. Since $Tv = T(0) = 0$ and T is injective, we get $v = 0$. For (ii) there is nothing to check. □

Theorem 1 (Fundamental theorem of linear maps)**.** *Let V be a finite-dimensional vector space and $T \in \mathcal{L}(V, W)$. Then*
$$\dim V = \dim \operatorname{null} T + \dim \operatorname{range} T.$$

Proof. Let $\{u_1, \dots, u_n\}$ be a basis of $\operatorname{null} T$. By Lemma 2 (2) we can extend it to $\{u_1, \dots, u_n, u_{n+1}, \dots, u_{n+m}\}$ a basis of V. We claim that $\{Tu_{n+1}, \dots, Tu_{n+m}\}$ is a basis of $\operatorname{range} T$. Indeed, let $w \in W$, then there exist $v \in V$ such that $Tv = w$. Since $\{u_1, \dots, u_n, u_{n+1}, \dots, u_{n+m}\}$ is a basis of V, there exists scalars a_i's such that $v = a_1 u_1 + \dots + a_{n+m} u_{n+m}$ and we have:
$$w = T(v) = T(a_1 u_1 + \dots + a_{n+m} u_{n+m}) = a_1 T(u_1) + \cdots a_n T(u_n) + a_{n+1} T(u_{n+1}) + \dots + a_{n+m} T(u_{n+m}).$$
Since the first n terms vanish, we get $w = a_{n+1} T(u_{n+1}) + \dots + a_{n+m} T(u_{n+m})$. This shows that $\{Tu_{n+1}, \dots, Tu_{n+m}\}$ is a spanning set. We now check that it is also linearly independent. Assume that there exists scalars $b_1, \dots, b_m \in \mathbb{F}$, not all zero, such that we have
$$b_1 T(u_{n+1}) + \dots + b_m T(u_{n+m}) = 0$$
this implies that $b_1 u_{n+1} + \dots + b_m u_{n+m} \in \operatorname{null} T$. However, since $\{u_1, \dots, u_n\}$ is a basis of $\operatorname{null} T$, it means there are scalars $c_1, \cdots, c_n \in \mathbb{F}$, not all zero, such that:
$$c_1 u_1 + \dots + c_n u_n = b_1 u_{n+1} + \dots + b_m u_{n+m}.$$
However, this is a contradiction with $\{u_1, \dots, u_{n+m}\}$ being a basis of V.

This finishes the proof. □

Here are a couple of easy consequences of the previous result.

Corollary 3. *Let V and W be finite-dimensional vector spaces.*

(1) Assume that $\dim V > \dim W$*, then any linear map $T : V \to W$ is not injective.*

(2) *Assume that* $\dim V < \dim W$*, then any linear map* $T : V \to W$ *is not surjective.*

Proof. (1) Assume by contradiction that there exist $T : V \to W$ an injective linear map. By Lemma 11.(1) we have $\operatorname{null} T = 0$ and Theorem 1 implies $\dim V = \dim rangeT \leq \dim W$, which is a contradiction. (2) Assume by contradiction that there exist $T : V \to W$ a surjective linear map. Then Lemma 11.(2) and Theorem 1 implies that $\dim W = \dim \operatorname{range} T \leq \dim T$, since $\dim \operatorname{null} T \geq 0$, which is also a contradiction. □

In fact, we can deduce some statements you will be familiar with from Math 2101 from Corollary 3.

Remark 8. (1) Any homogeneous system of linear equations with more variables than equations has a nonzero solution. Consider $V = \mathbb{F}^n$ and $W = \mathbb{F}^m$, a system of linear equations is given by an $m \times n$ matrix A, which gives a linear operator $T_A : \mathbb{F}^n \to \mathbb{F}^m$, by Exercise 21. By Corollary 3(1) we get that there exists $v \in \mathbb{F}^n$ such that $Tv = 0$.

(2) Any inhomogeneous system of linear equations with more equations than variables has no solution for some choice of the constant terms. Again consider $V = \mathbb{F}^n$ and $W = \mathbb{F}^m$ and a linear transformation $T : V \to W$ encoding the system of linear equations. If $m > n$, then by Corollary 3.(2) there exists $w \in W$ such that $Tv = w$ has no solution.

4.3. **Matrices.** Given two positive natural numbers $n, m \geq 1$ and a field $\mathbb{F}$ an *m-by-n matrix* with coefficients in $\mathbb{F}$ is a list $(a_{i,j})_{1 \leq i \leq m,\, 1 \leq j \leq n}$ sometimes denoted by:

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

We let $\mathbb{F}^{m,n}$ denote the vector space of m by n matrices. You should make sure you understand why this is a vector space.

Definition 15. Given $T \in \mathcal{L}(V, W)$ a linear map between two vector spaces V, W. Let $B_V = \{v_1, \ldots, v_n\}$ be a basis of V and $B_W = \{w_1, \ldots, w_m\}$ be a basis of W. The *matrix* A *associated to* T *and these basis* is defined by:

$$Tv_j = a_{1,j} w_1 + \cdots + a_{m,j} w_m \qquad \text{for } 1 \leq j \leq n.$$

Sometimes we emphasize the dependence on the basis by denoting $A := \mathcal{M}(T, B_V, B_W)$.

Example 12. (i) Let $T : \mathcal{P}_3(\mathbb{F}) \to \mathcal{P}_2(\mathbb{F})$ be the linear map associated to the differentiation. Consider the bases $B_3 = \{1, x, x^2, x^3\}$ and $B_2 = \{1, x, x^2\}$ of $\mathcal{P}_3(\mathbb{F})$ and $\mathcal{P}_2(\mathbb{F})$, respectively. Then we have:

$$\mathcal{M}(T, B_3, B_2) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}.$$

(ii) For T as in (i) if we take the basis $B_3 = \{1, 1 + x, 1 + x + x^2, 1 + x + x^2 + x^3\}$ and $B_2 = \{1, x, x^2\}$ for $\mathcal{P}_3(\mathbb{F})$ and $\mathcal{P}_2(\mathbb{F})$, respectively. Then we obtain a *different* matrix:

$$\mathcal{M}(T, B_3, B_2) = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 0 & 0 & 2 & 2 \\ 0 & 0 & 0 & 3 \end{pmatrix}.$$

(iii) Consider $T : \mathbb{R}^4 \to \mathbb{R}^4$ the linear map given by $T(v_1, v_2, v_3, v_4) = (v_1 + v_2, v_3, 0, v_2)$. On the basis $B_V = \{e_1, e_2, e_3, e_4\}$ for the source $V = \mathbb{R}^4$ and $B_w = \{e_1, e_2, e_3, e_4\}$ for the target, we have:

$$\mathcal{M}(T, B_V, B_W) = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}.$$

However, if we take the basis $B_W^{(2)} = \{e_1 + e_2, e_2 + e_3, e_3 + e_4, e_4\}$ then we have:

$$\mathcal{M}(T, B_V, B_W^{(2)}) = \begin{pmatrix} 1 & 1 & 0 & 0 \\ -1 & -1 & 1 & 0 \\ 1 & 1 & -1 & 0 \\ -1 & 0 & 1 & 0 \end{pmatrix}.$$

Exercise 24. (i) Make sure that you understand how to obtain the matrices in the Example above.

(ii) Can you come up with a basis of $\mathbb{R}^4$ in Example 12.(iii) where all of the columns are non-zero?

(iii) Can you come you with a basis B_W such that the set of column vectors of the matrix in Example 12.(iii) are linearly independent?

Remark 9. The construction from Definition 15 sends the operations of addition and scalar multiplication of linear maps to the corresponding operations between matrices as defined in Math 2101. Indeed, given two linear maps $T, S \in \mathcal{L}(V, W)$, a scalar $a \in \mathbb{F}$, and two bases B_V of V and B_W of W. Then we have:

$$\mathcal{M}(T+S, B_V, B_W) = \mathcal{M}(T, B_V, B_W) + \mathcal{M}(T, B_V, B_W) \qquad \text{and} \qquad \mathcal{M}(aT, B_V, B_W) = a\mathcal{M}(T, B_V, B_W).$$

Question: When did you learn matrix multiplication? Have you every thought what was the *meaning* behind the rule of how to multiply matrices?

Lemma 12. *Consider U, V, and W three vector spaces and suppose that we picked bases B_U, B_V and B_W, respectively. Consider $T : U \to V$ and $S : V \to W$ two linear maps. Then*

$$\mathcal{M}(S \circ T, B_U, B_W) = \mathcal{M}(S, B_V, B_W)\mathcal{M}(T, B_U, B_V), \tag{3}$$

where on the righthand side above we consider the multiplication of matrices.

Proof. This is done on page 73 of the textbook. We leave the details to the reader. □

Problem 1. Does the formula fail if we take different bases for V, i.e. do we have

$$\mathcal{M}(S \circ T, B_U, B_W) = \mathcal{M}(S, B'_V, B_W)\mathcal{M}(T, B_U, B_V)$$

for B'_V different than B_V?

Let's recall a couple of concepts from Math 2101.

Definition 16. Given a matrix A represented as follow:

$$A = \begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & & \vdots \\ a_{m,1} & \cdots & a_{m,n} \end{pmatrix}. \tag{4}$$

(1) The *column space of* A, denoted $\mathrm{col}(A)$, is defined as the span of $\{v_1, \ldots, v_n\}$ in $\mathbb{R}^m$, where $v_i = (a_{1,i}, \ldots, a_{m,i})$ for $1 \le i \le n$.

(2) The *row space of* A is the span of $\{w_1, \ldots, w_m\}$ in $\mathbb{R}^n$ spanned by $w_j = (a_{j,1}, \ldots, a_{j,n})$ for $1 \le j \le m$.

Lemma 13. *For any $m \times n$ matrix A we have*

$$\dim \mathrm{col}(A) = \dim \mathrm{row}(A).$$

Proof. **TODO:** Write this. □

5. Jan. 29, 2024

5.1. **Isomorphisms.** The following notion is going to allow us to compare vector spaces and identify when there are "essentially" the same for all purposes of linear algebra.

Definition 17. A linear map $T : V \to W$ is *invertible* if there exists a map $S : W \to V$ such that

$$S \circ T = \mathrm{Id}_V \qquad \text{and} \qquad T \circ S = \mathrm{Id}_W.$$

Remark 10. Notice that if an inverse exists it is unique. Indeed, assume that S_1 and S_2 are inverses of T. Then we have $S_1 = S_1 \circ T \circ S_2 = S_2$. Thus, we will denote by T^{-1} the uniquely determied inverse, if it exists.

Example 13. Consider $T : \mathbb{R}^3 \to \mathbb{R}^3$ given by $T(v_1, v_2, v_3) = (v_1 + v_2 + v_3, v_2, v_3)$. Then $T^{-1}(u_1, u_2, u_3) = (u_1 - u_2 - u_3, u_2, u_3)$.

Exercise 25. Determine which ones of the linear maps that we consider so far are invertible and determine their inverses.

Lemma 14. *Let $T : V \to W$ be a linear map. The following are equivalent:*

(1) T is invertible;

(2) T is injective and surjective;

(3) $\operatorname{null} T = \{0\}$ and $\operatorname{range} T = W$.

Proof. The equivalence (2) $\Leftrightarrow$ (3) is Lemma 11. Assume (1). Let $v, u \in V$ such that $Tv = Tu$ then we have $v = T^{-1}Tv = T^{-1}Tu = u$, thus T is injective. Now let $w \in W$ be any vector, then $T^{-1}w \in V$ satisfies $TT^{-1}w = w$, so T is surjective.

Assume (2). For any $w \in W$ we let $S(w) = v$ for any $v \in V$ such that $Tv = w$. This is well-defined, since by the injectivity of T, there is only one v satisfying $Tv = w$. We claim that S is an inverse of T. Indeed, we have $TSw = w$ and $STv = v$ for every $w \in W$ and $v \in V$. $\square$

In the case where our vector space is finite-dimensional, then we have a stronger version of Lemma 14.

Corollary 4. *Assume that V and W are finite-dimensional vector space and that $\dim V = \dim W$ and consider $T \in \mathcal{L}(V, W)$. The following are equivalent:*

(1) T is invertible;

(2) T is injective;

(2)' $\operatorname{null} T = \{0\}$;

(3) T is surjective;

(3)' $\operatorname{range} T = W$.

Proof. The implications (1) $\Rightarrow$ (2)/(2)$'$ and (1) $\Rightarrow$ (3)/(3)$'$ are clear.

The equivalences (2) $\Leftrightarrow$ (2)$'$ and (3) $\Leftrightarrow$ (3)$'$ were establishes in Lemma 11.

Assume (2)', then by Theorem 1 we have $\dim V = \dim \operatorname{range} T = \dim W$, so we have (3)'. Now (2)' and (3)' imply (1) by Lemma 14.

Assume (3)', then by Theorem 1 we have $\dim V = \dim \operatorname{range} T - \dim \operatorname{null} T = \dim W - \dim \operatorname{null} T$. Thus, $\dim \operatorname{null} T = 0$, so we have (2)'. Now (2)' and (3)' imply (1) by Lemma 14. $\square$

Remark 11. It is important to notice that the assumption that V and W are finite-dimensional in Corollary 4 is crucial. Indeed, consider the linear map $T : \mathbb{R}^{\mathbb{N}} \to \mathbb{R}^{\mathbb{N}}$ given by

$$T(f)(n) = \begin{cases} 0 \text{ if } n = 0, \\ f(n+1) \text{ else.} \end{cases}$$

This is injective but not an isomorphism.

Exercise 26. Write an example of a linear map which is surjective but not an isomorphism.

Exercise 27. Prove that there exists a polynomials $p \in \mathbb{R}[x]$ such that $((x^2 + 5x + 7)p)'' = q$ for any $q \in \mathbb{R}[x]$.

Exercise 28. Consider $S \in \mathcal{L}(V, W)$ and $T \in \mathcal{L}(W, V)$ two linear maps. Assume that $\dim V = \dim W < \infty$. Prove that $ST = \mathrm{Id}_W$ if and only if $TS = \mathrm{Id}_V$.

The next concept plays the role for vector spaces of what bijections are for sets.

Definition 18. Two vector spaces V and W are said to be *isomorphic* if there exists an invertible linear map $T : V \to W$, equivalently if there exists an invertible linear map $S : W \to T$. In this case either morphism $T : V \to W$ or $S : W \to V$ are called *isomorphisms*.

Notation 3. We will sometimes simply write $V \simeq W$ to say that V and W are isomorphic.

It turns out that it is rather easy to determine if two finite-dimensional vector spaces are isomorphic or not as the next result shows.

Lemma 15. *Let V and W be two finite-dimensional vector spaces. The following are equivalent:*

(1) $V \simeq W$.

(2) $\dim V = \dim W$.

Proof. Assume (1) and let $T : V \to W$ be an invertible linear map. By Theorem 1 we have:

$$\dim T = \dim \operatorname{null} T + \dim \operatorname{range} T = \dim \operatorname{range} T = \dim W,$$

where the second and third equalities above follow from Corollary 4.

Assume (2) and let $\{v_1, \dots, v_n\}$ be a basis of V and $\{w_1, \dots, w_n\}$ be a basis of W. Let $T(c_1v_1 + \cdots + c_nv_n) := c_1w_1 + \cdots + c_nw_n$, which is well-defined as argued in the proof of Lemma 9. We claim T is injective. Indeed, assume that $T(c_1v_1 + \cdots + c_nv_n) = c_1w_1 + \cdots + c_nw_n = 0$ for some non-zero combination of c_i's, then linear independence of w_i's imply that all c_i's are zero which gives that $c_1v_1 + \cdots + c_nv_n = 0$. Thus, by Corollary 4 we are done. □

The next result justify the idea that on can always think of a linear map between finite-dimensional vector spaces as a matrix. Notice however that this depends on the choice of bases.

Lemma 16. *Let V and W be finite-dimensional vector spaces of dimensions $n = \dim V$ and $m = \dim W$. Then $\mathcal{L}(V, W) \simeq \mathbb{F}^{m,n}$.*

Proof. Let $B_V = \{v_1, \dots, v_n\}$ be a basis of V and $B_W = \{w_1, \dots, w_m\}$ be a basis of W. We claim that

$$\mathcal{M} := \mathcal{M}(-, B_V, B_W) : \mathcal{L}(V, W) \to \mathbb{F}^{m,n}$$

as defined in Definition 15 is an isomorphism. First notice that $\mathcal{M}$ is a linear map by Remark 9. First we prove that $\mathcal{M}$ is injective. Indeed, assume that $\mathcal{M}(T) = 0$, then $Tv = 0$ for every $v \in B_V$ since B_V is a basis we get $Tu = 0$ for every $u \in V$, thus $T = 0$. Now we prove that $\mathcal{M}$ is surjective. Let $A \in \mathbb{F}^{m,n}$, represented as equation (4). Let $Tv_i := a_{1,i}w_1 + \cdots + a_{m,i}w_m$ for $1 \le i \le n$. This is a well-defined linear operator and it is clear that $\mathcal{M}(T) = A$. □

Here is a nice consequence of the discussion so far:

Corollary 5. *For any finite-dimensional vector spaces V and W, we have $\dim \mathcal{L}(V, W) = \dim V \cdot \dim W$.*

Proof. By Lemma 16 and Lemma 15 we have that

$$\dim \mathcal{L}(V, W) = \mathbb{F}^{\dim W, \dim V}.$$

The result now follows from calculating the dimension of the space of $\dim V$ by $\dim W$ matrices. □

The following is variation on Definition 15.

Definition 19. Let V be a vector space of dimension n, given a basis $B_V = \{v_1, \dots, v_n\}$ we let $\mathcal{M}(V, B_V) : V \to \mathbb{F}^n$ denote the linear map determined as follows

$$\mathcal{M}(V, B_V)(v) = (a_1, \dots, a_n) \text{ if } a_1v_1 + \cdots + a_nv_n = v.$$

You should check this is well-defined and indeed a linear map.

Remark 12. Notice that $\mathcal{M}(V, B_V)$ is always an isomorphism. Indeed, the same argument as in the proof of Lemma 16 works, we simply take W to be $\mathbb{F}$.

Remark 13. This remark is a bit abstract, but the reason we need it is to obtain certain matrix and vector multiplication compatibility without doing many calculations. Given a vector $v \in V$ in a finite-dimensional vector space V, we can think of v as a linear operator $L_v : \mathbb{F} \to V$ given as $L_v(1) := v$. Notice that there was a choice in determining this linear operator, namely a basis of $\mathbb{F}$, in this case $\{1\} \subset \mathbb{F}$[4]. Now we claim that:

$$\mathcal{M}(L_v, \{1\}, B_V) = \mathcal{M}(V, B_V)(v).$$

[4]So there are other linear operators associated to v, but in a sense picking $\{1\} \subset \mathbb{F}$ as a basis of $\mathbb{F}$ over itself is rather natural, i.e. it makes sense for any field.

Indeed, let $\{v_1, \dots, v_n\} = B_V$, then $\mathcal{M}(L_v, \{1\}, B_V) = (a_{i,1})_{1 \le i \le n}$ as defined in Definition 15 is given by:

$$L_v(1) = v = a_{1,1}v_1 + \cdots + a_{n,1}v_n,$$

whereas $\mathcal{M}(V, B_V)(v) = (b_{i,1})_{1 \le i \le n}$, as defined in Defintion 19, is the set $(b_{1,1}, \dots, b_{n,1})$ such that

$$v = b_{1,1}v_1 + \cdots + b_{n,1}v_n.$$

The following is a consequence of Remark 13.

Corollary 6. *Let $T : V \to W$ be a linear map between finite-dimensional vector spaces and B_V and B_W bases, respectively. Then for any $v \in V$ we have:*

$$\mathcal{M}(T, B_V, B_W)\mathcal{M}(V, B_V)(v) = \mathcal{M}(W, B_W)(Tv).$$

Proof. Indeed, we compute:

$$\begin{aligned}\mathcal{M}(T, B_V, B_W)\mathcal{M}(V, B_V)(v) &= \mathcal{M}(T, B_V, B_W)\mathcal{M}(L_v, \{1\}, B_V) \\ &= \mathcal{M}(T \circ L_v, \{1\}, B_W) \\ &= \mathcal{M}(Tv, \{1\}, B_W) \\ &= \mathcal{M}(W, B_W)(Tv).\end{aligned}$$

□

The next result relates a notion from Math 2101 which was defined using a basis, namely the column space of a matrix, with the range of a linear map which didn't depend on a basis to be defined.

Proposition 1. *Let $T : V \to W$ be a linear map between finite-dimensional vector spaces and let $B_V = \{v_1, \dots, v_n\}$ and $B_W = \{w_1, \dots, w_n\}$ be basis of V and W, respectively. The restriction of $\mathcal{M}(W, B_W)$ to* $\operatorname{range} T$ *has the following factorization:*

$$\begin{tikzcd}
\operatorname{range} T \arrow[r, "\varphi"] \arrow[d, "\subset"'] & \operatorname{col}(\mathcal{M}(T, B_V, B_W)) \arrow[d, "\subset"] \\
W \arrow[r, "\mathcal{M}(W,B_W)"'] & \mathbb{F}^{\dim W}
\end{tikzcd}$$

and φ is an isomorphism.

Proof. TODO: maybe this proof becomes simpler after the previous Remark.

Let $(a_{i,j})_{1 \le i \le m; 1 \le j \le n}$ denote the entries of the matrix $\mathcal{M}(T, B_V, B_W)$. Given $w \in \operatorname{range} T$ there exists $b_1, \dots, b_n \in \mathbb{F}$ such that $w = T(b_1v_1 + \cdots + b_nv_n)$, we have

$$w = b_1T(v_1) + \cdots + b_nT(v_n).$$

Now we notice that $\mathcal{M}(W, B_V)(T(v_j)) = (a_{j,1}, \dots, a_{j,m}) \in \mathbb{F}^m$ for $j \in \{1, \dots, n\}$. (Check this!). By definition we have

$$\operatorname{col}(\mathcal{M}(T, B_V, B_W)) = \operatorname{Span}\left(\{(a_{1,1}, \dots, a_{1,m}), \dots, (a_{n,1}, \dots, a_{n,m})\}\right).$$

Thus, linearity of $\mathcal{M}(W, B_W)$ implies

$$\mathcal{M}(W, B_W)(w) \in \operatorname{Span}\{\mathcal{M}(W, B_W)(Tv_1), \dots, \mathcal{M}(W, B_W)(Tv_n)\} = \operatorname{col}(\mathcal{M}(T, B_V, B_W))$$

as required. This shows that φ factors as claimed.

We now check that φ is an isomorphism. We first notice that it is injective since it is the restriction of an injective linear map. Finally, let $w \in \operatorname{col}(\mathcal{M}(T, B_V, B_W))$, then it can be written as $w = d_1\mathcal{M}(W, B_W)(Tv_1) + \cdots + d_n\mathcal{M}(W, B_W)(Tv_n)$ for some $d_1, \dots, d_n \in \mathbb{F}$. By the linearity of $\mathcal{M}(W, B_W)$ we have

$$w = \mathcal{M}(W, B_W)(d_1Tv_1 + \cdots d_nTv_n)$$

and it is clear that $d_1Tv_1 + \cdots d_nTv_n \in \operatorname{range} T$. This finishes the proof. □

The proposition immediately imply:

Corollary 7. *Let V, W be finite-dimensional vector spaces, $T : V \to W$ a linear map and B_V and B_W bases, then*

$$\dim \operatorname{range} T = \dim \operatorname{col}(\mathcal{M}(T, B_V, B_W)) = \dim \operatorname{row}(\mathcal{M}(T, B_V, B_W)).$$

We emphasize that in Corollary 7 the quantity $\dim \operatorname{range} T$ is defined without any appeal to a basis either of V or W.

We end this subsection stating a useful formula for change of bases. Before it we introduce some notation to simplify the formula.

Notation 4. Let V be a finite-dimensional vector space and consider $T : V \to V$ a linear map and B_V a basis of V, then we simply write:

$$\mathcal{M}(T, B_V) := \mathcal{M}(T, B_V, B_V).$$

Lemma 17. *Let V be a finite-dimensional vector space and consider $T : V \to V$ a linear map and B_V^1 and B_V^2 two bases of V. Then we have:*

$$\mathcal{M}(T, B_V^1) = \mathcal{M}(\mathrm{Id}_V, B_V^1, B_V^2)^{-1}\mathcal{M}(T, B_V^2)\mathcal{M}(\mathrm{Id}_V, B_V^1, B_V^2). \tag{5}$$

Proof. This is (3.84) from the textbook. Write details. □

6. Feb. 1, 2024

6.1. **Products and Quotients of Vector spaces.** In this section we introduce some constructions that define a new (abstract) vector space from a number of other vector spaces.

Definition 20. Let $n \geq 1$ be a natural number. Given a collection of vector spaces $\{V_1, \dots, V_n\}$ the *product* $V_1 \times \cdots \times V_n$ of $\{V_1, \dots, V_n\}$ is the vector space whose:

- set is the Cartesian product $V_1 \times \cdots \times V_n$;
- $+ : (V_1 \times \cdots \times V_n) \times (V_1 \times \cdots \times V_n) \to V_1 \times \cdots \times V_n$ is defined coordinate-wise, i.e. $(v_i) + I + (w_i)_I := (v_i + w_i)_I$, where $I = \{1, \dots, n\}$;
- $\cdot : \mathbb{F} \times (V_1 \times \cdots \times V_n) \to V_1 \times \cdots \times V_n$ is also given coordinate-wise.

Remark 14. The following is the universal property of the product. Given another vector space W and a collection of linear maps $f_i : W \to V_i$, then there exists an *unique* linear map $\varphi : W \to V_1 \times \cdots \times V_n$ such that for every $i \in I$ the following diagrams commute:

$$\begin{tikzcd} W \arrow[r, dashed, "\varphi"] \arrow[dr, "f_i"'] & V_1 \times \cdots \times V_n \arrow[d, "p_i"] \\ & V_i \end{tikzcd},$$

where $p_i : V_1 \times \cdots \times V_n \to V_i$ sends $(v_j)_{j \in I}$ to v_i. You will prove this for $I = \{1, 2\}$ in Homework 1.

Example 14. (i) Let $V_1 = \mathbb{C}^2$ and $V_2 = \mathbb{C}^3$, then we have $V_1 \times V_2 = \{((v_1, v_2), (v_3, v_4, v_5)) \in \mathbb{C}^2 \times \mathbb{C}^3\}$. Notice that $\mathbb{C}^2 \times \mathbb{C}^3 \neq \mathbb{C}^5$, because of the parenthesization. We however normally simplify this to $\mathbb{C}^5$ because $\mathbb{C}^2 \times \mathbb{C}^3$ is isomorphic to $\mathbb{C}^5$.

(ii) Let $V_1 = \mathcal{P}(\mathbb{R})$ and $V_2 = \mathbb{R}^{m,n}$, then we have $V_1 \times V_2 = \mathcal{P}(\mathbb{R}) \times \mathbb{R}^{m,n}$.

Remark 15. In fact, it is more natural, and sometimes convenient, to take Remark 14 as the definition of product. More precisely, we say that the product U of vector spaces $\{V_1, \dots, V_n\}$, if it exists, is a vector space U satisfying the condition of Remark 14. Then we check two things:

(i) the construction given in Definition 20 solves the question asked in Remark 14;

(ii) any solution to Remark 14 is isomorphic to $V_1 \times \cdots \times V_n$.

Actually, one can do even better than (ii), we can prove that for any solution U, there exists an isomorphism $\psi : U \to V_1 \times \cdots \times V_n$ and that ψ is unique, if we require it to be compatible with the morphisms p_i for $V_1 \times \cdots \times V_n$ and p'_i for U.

Remark 16. In fact, $V_1 \times \cdots \times V_n$ also satisfies a dual universal property. Namely, let W be a vector space and $g_i : V_i \to W$ be a collection of linear maps. Then there exists an unique linear $\psi : V_1 \times \cdots \times V_n \to W$ such that for every $i \in I$ the following diagrams commute:

$$\begin{tikzcd} V_i \arrow[r, "\imath_i"] \arrow[dr, "f_i"'] & V_1 \times \cdots \times V_n \arrow[d, dashed, "\psi"] \\ & W \end{tikzcd},$$

where $\imath_i : V_i \to V_1 \times \cdots \times V_n$ sends v_i to $(0, \dots, 0, v_i, 0, \dots, 0)$. You will also prove this for $I = \{1, 2\}$ in Homework 1. A similar comment as in Remark 15 applies to the object that would satisfy this condition, it is called the coproduct of $\{V_1, \dots, V_n\}$.

Proposition 2. *Let S be a (not necessarily) finite set, and $\{V_s\}_S$ a collection of vector spaces. Consider:*

(1) $\prod_{s\in S} V_s$ the subset of functions $f : S \to \bigcup_{s\in S} V_s$ satisfying $f(s) \in V_s$ for every $s \in S$.

(2) $\oplus^{\mathrm{ext}}_{s\in S} V_s$ the subset of $\prod_{s\in S} V_s$ such that $f(s) \neq 0$ only for finitely many $s \in S$.

Then:

(i) $\prod_S V_s$ and $\oplus^{\mathrm{ext}}_S V_s$ are vector spaces.

(ii) $\prod_S V_s$ together with the morphisms $p_s : \prod_S V_s \to V_s$ given by $p_s(f) := f(s)$ satisfies the conditions of Remark 14.

(iii) $\oplus^{\mathrm{ext}}_S V_s$ together with the morphisms $\imath_s : V_s \to \oplus^{\mathrm{ext}}_S V_s$ given by $\imath(v)(t) = v$ for $t = s$ and 0 otherwise, satisfies the conditions of Remark 16.

Proof. We define $(a \cdot f + g)(s) := a \cdot f(s) + g(s)$ for $a \in \mathbb{F}$ and $f, g \in \prod_S V_s$. This clearly makes $\prod_S V_s$ into a vector space.

Let's specify the morphisms $p_s : \prod_S V_s \to V_s$. Given $f \in \prod_S V_s$ we let $p_s(f) := f(s)$.

We prove (ii). Let $\{g_s : W \to V_s\}$ be a collection of linear maps, then we define $h : W \to \prod_S V_s$ by

$$h(w)(s) := g_s(w).$$

We notice that $p_s \circ h(w) = h(w)(s) = g_s(w)$ as required. Assume we are given $h' : W \to \prod_S V_s$ another linear map satisfying

$$p_s \circ h' = g_s \text{ for all } s \in S.$$

Then we have that for any $w \in W$ and $s \in S$:

$$h'(w)(s) = p_s \circ h'(w) = g_s(w) = p_s \circ h(w) = h(w)(s),$$

so $h' = h$.

The argument for (iii) is analogous to (ii) and we leave the details for the reader. □

Exercise 29. Let $\{V_s\}_{s\in S}$ be a collection of vector spaces and assume that $\{B_s\}_{s\in S}$ is a collection of subsets $B_s \subseteq V_s$, such that each B_s is a basis of V_s. Let $B :=:= \cup_{s\in S} B_s$:

(i) prove that B is a basis of $\oplus_{s\in S} V_s$;

(ii) give an example to show that B is not a basis of $\prod_{s\in S} V_s$ in general.

Here are some immediate properties of the product of vector spaces.

Lemma 18. *(1) Let $V_1, \dots, V_n$ be finite-dimensional vector spaces, then we have $\dim(V_1 \times \cdots \times V_n) = (\dim V_1) + \cdots + (\dim V_n)$.*

(2) Suppose that $V_1, \dots, V_n$ are subspaces of an ambient vector space W. Then we have a linear map:

$$\Gamma : V_1 \times \cdots \times V_n \to V_1 + \cdots + V_n$$

given by $\Gamma(v_1, \dots, v_n) := v_1 + \cdots + v_n$. Then $V_1 + \cdots + V_n$ is a direct sum if and only if Γ is injective.

(3) *Given subspaces $V_1, \dots, V_n$ of a finite-dimensional vector space W, then $V_1 + \dots + V_n$ is a direct sum if and only if* $\dim(V_1 + \dots + V_n) = \sum_{i=1}^n \dim V_i$.

Proof. (1) can be proved by induction. We notice that given $\{v_1, \dots, v_n\}$ a basis of V_1 and $\{u_1, \dots, u_m\}$ a basis of V_2, then $\{(v_1, 0), \dots, (v_n, 0), (0, u_1), \dots, (0, u_m)\}$ is a basis of $V_1 \times V_2$. Indeed, assume we are given $(v, u) \in V_1 \times V_2$, then there exist unique $a_1, \dots, a_n$ in $\mathbb{F}$ such that $v = \sum_{i=1}^n a_i v_i$ and there are unique $b_1, \dots, b_m$ in $\mathbb{F}$ such that $u = \sum_{j=1}^m b_j u_j$. Thus we get:

$$(u, v) = \sum_{i=1}^n a_i(v_i, 0) + \sum_{j=1}^m b_j(0, u_j).$$

that is $a_1, \dots, a_n, b_1, \dots, b_m$ are unique coefficients writing (v, u) as a linear combination of the basis vectors.

(2) Again by induction it is enough to consider the case $n = 2$. Now recall that $\dim V_1 + V_2 = \dim V_1 + \dim V_2 - \dim V_1 \cap V_2$ and that $V_1 + V_2$ is a direct sum if and only if $\dim V_1 \cap V_2 = 0$. If we assume Γ is injective then Corollary 4 implies Γ is an isomorphism, thus $\dim V_1 + V_2 = \dim(V_1 \times V_2)$, which gives that $V_1 + V_2$ is a direct sum. Conversely, if $\dim V_1 \cap V_2 = 0$, then $\dim \operatorname{null} \Gamma = \dim(V_1 \times V_2) - \dim \operatorname{range} T$. But Γ is surjective, thus $\dim \operatorname{range} T = \dim V_1 + V_2 = \dim V_1 + \dim V_2$, which gives that $\operatorname{null} \Gamma = 0$.

(3) is essentially a simple restatement of (2). □

Notation 5. Given a set of vector spaces $\{V_s\}_{s \in S}$, we will simply denote by $\oplus_{s \in S} V_s$ their direction sum as defined in Proposition 2. Moreover, given linear maps $f_s : V_s \to U$ we will let $\oplus_{s \in S} f_S : \oplus_{s \in S} V_s \to U$ denote the map determined by Remark 16.

Assume that S is finite, i.e. $S = \{1, \dots, n\}$ for some $n \in \mathbb{N}$. Notice that if each $V_n \subseteq U$ is a subspace of a vector space U. Then Lemma 18 gives a linear map:

$$V_1 \oplus \dots \oplus V_n \to V_1 + \cdots V_n,$$

which is an isomorphism if and only if $V_i \cap V_j = \{0\}$ for every $i \neq j$. In particular, this shows that the notation $V_1 \oplus \dots \oplus V_n$ is unambiguous.

7. Feb. 5, 2024

7.1. **Quotient spaces.** The motivation for this session is the following. Suppose we are given a subspace $U \subseteq V$ and linear map $L : V \to W$ to another vector space such that $L|_U : U \to W$ is the identically 0 linear map. Can we describe the data of L in "smaller" terms, i.e. is there some vector space Q and a map $T : Q \to W$ such that we recover L from T. Notice that if V is finite-dimensional, one answer would be to pick a subspace $U' \subset V$ such that $U' \oplus U = V$, which exists by Lemma 5 and consider $T := L|_{U'} : U' \to W$. There are however two problems with this solution:

(1) we had to assume that V is finite-dimensional;

(2) we had to make a *choice* of U', which we know, by Exercise 5 from Worksheet 1, is *not* unique–so T is not unique either.

To solve these two problems we need some preparatory discussion.

Let $U \subset V$ be a subspace, the consider the relation $R_U \subseteq V \times V$ on V defined by

$$(v_1, v_2) \in R_U \text{ if } v_1 - v_2 \in U.$$

Exercise 30. For any subspace $U \subseteq V$ the relation R_U is an equivalence relation[5].

Let V/U denote the set of equivalence classes of R_U.

Lemma 19. *There exists an unique structure of vector space on V/U such the natural map:*

$$\begin{aligned} \pi : V &\to V/U \\ v &\mapsto \pi(v) \end{aligned}$$

is a linear map.

[5]If you don't remember this definition from set theory, see here.

Proof. First we give a concrete description of $\pi(v)$. Notice that $w \in \pi(v)$ if and only if $w - v \in U$, i.e. $w \in v + U := \{w \in V \mid w - v \in U\}$. Now we define $\cdot' : \mathbb{F} \times V/U \to V/U$ as follows:

$$a \cdot' \bar{v} := \pi(a \cdot v'), \text{ for any } v' \in V \text{ such that } \pi(v') = \bar{v}.$$

Similarly, we define $+' : V/U \times V/U \to V/U$ as follows:

$$\bar{v}_1 +' \bar{v}_2 := \pi(v_1' + v_2'), \text{ for any } v_1', v_2' \in V \text{ such that } \pi(v_1') = \bar{v}_1 \text{ and } \pi(v_2') = \bar{v}_2.$$

Let's check that these are well-defined. Let v' and v'' be such that $\pi(v') = \pi(v'')$, i.e. there exist $u \in U$ such that $v' + u = v''$. Then we have

$$a \cdot \pi(v') = \pi(a \cdot v') = \pi(a \cdot v' + a \cdot u)) = \pi(a \cdot (v' + u)) = \pi(a \cdot v'') = a \cdot \pi(v'').$$

The check that $+'$ is well-defined is similar. We leave it to the reader to check that π is a linear map. □

Definition 21. Given a subspace $U \subset V$ the quotient set V/U with the structure from Lemma 19 is called the *quotient space of V by U.*

Remark 17. Notice that $\operatorname{null} \pi = U$. Indeed, for any $u \in U$ we have $\pi(u) = \pi(u - u) = \pi(0) = 0$. Thus, Theorem 1 imply that $\dim V/U = \dim V - \dim U$, whenever V is finite-dimensional.

We now show that this construction solves the problem that motivated us in the begining of this subsection.

Lemma 20. *Given a subspace $U \subseteq V$ and a linear map $L : V \to W$ such that $L|_U = 0$. Then there exist an unique $T : V/U \to W$ such that the following diagram commutes*

$$\begin{tikzcd} V \arrow[r, "\pi"] \arrow[dr, "L"'] & V/U \arrow[d, dashed, "T"] \\ & W \end{tikzcd},$$

i.e. $T \circ \pi = L$.

Proof. We let $T(\bar{v}) := L(v')$ for any $v' \in V$ such that $\pi(v') = \bar{v}$. We check this is well-defined. Let v', v'' such that $\pi(v') = \pi(v'')$ then $v' = v'' + u$, which gives

$$L(v') = L(v'' + u) = L(v'').$$

We leave it to the reader to check that T is linear and uniquely determined. □

Remark 18. Notice that given a subspace $U \subseteq V$ of a finite-dimensional vector and a linear map $L : V \to W$ we can always modify L on U to make it satisfy the conditions of Lemma 20. Indeed, consider $L - \widetilde{L|_U}$, where $\widetilde{L|_U} : V \to W$ is the extension of $L|_U$ defined as follows: let $W \subseteq V$ be any subspace such that $W \oplus U = V$, then $\widetilde{L|_U} := 0|_W \oplus \widetilde{L|_U}$. Notice that $\widetilde{L|_U}$ does not depend on W. Explain the notation $0|_W \oplus \widetilde{L|_U}$ and why it doesn't depend on U.

Corollary 8. *Let $L : V \to W$ be a linear map, then one has a canonical map:*

$$\bar{L} : V/\operatorname{null} T \to W$$

satisfying the following:

(i) $\bar{L} \circ \pi = L$;

(ii) $\bar{L}$ *is injective;*

(iii) $\operatorname{range} \bar{L} = \operatorname{range} L$;

(iv) $V/\operatorname{null} T \simeq \operatorname{range} \bar{L}$.

Proof. The existence of $\bar{L}$ and (i) follow directly from Lemma 20, by noticing that $L|_{\operatorname{null} L} = 0$

For (ii), suppose that there exist $\bar{v} \in V/\operatorname{null} L$ such that $\bar{L}(\bar{v}) = 0$, but $\bar{v} \neq \pi(0)$. This implies that there exist $v' \in V$ such that $\pi(v') = \pi(0)$, i.e. $v' \in \operatorname{null} L$ and $\bar{L}(\pi(v')) = L(v') = 0$, which is a contradiction.

For (iii), we notice that $\operatorname{range} \bar{L} \subseteq \operatorname{range} L$ by set-theoretic consideration. Now assume that $w \in \operatorname{range} L$ and let $v \in V$ such that $L(v) = w$, then $\bar{L}(\pi(v)) = w$.

For (iv), we simply notice that $\bar{L}$ factors as:

$$\begin{tikzcd} V/\operatorname{null} T \arrow[r, "\varphi"] \arrow[dr, "\bar{L}"'] & \operatorname{range} \bar{L} \arrow[d, dashed, "\imath"] \\ & W \end{tikzcd},$$

where by (ii) and (iii) φ is injective and surjective, hence an isomorphism by Lemma 14. □

8. Feb. 8, 2024

8.1. **Duality.** The next definition associates a vector space to given vector space that is very useful.

Definition 22. Given a vector space V the vector space $V^* := \mathcal{L}(V, \mathbb{F})$ is called its *dual space*[6]. Elements $\lambda \in V^*$ are called *linear functionals* on V.

Example 15.
(1) Let $\mathcal{P}_3(\mathbb{R})$, then $\varphi : \mathcal{P}_3(\mathbb{R}) \to \mathbb{R}$ given by $\varphi(p) := \int_0^1 p(x)dx$ is a linear functional, i.e. $\varphi \in \mathcal{P}_3(\mathbb{R})^*$.
(2) For any $a := (a_1, \dots, a_n) \in \mathbb{F}^n$ we have $\lambda_a : \mathbb{F}^n \to \mathbb{F}$ given by $\lambda_a(v_1, \dots, v_n) := \sum_{i=1}^n a_i v_i$.
(3) Let $V = \mathbb{R}[[x]]$ denote the vector space of Taylor series, i.e. $f : \mathbb{R} \to \mathbb{R}$ given by a series $f(x) = \sum_{i\geq 0} a_i x^i$. Then $V^* \simeq \mathbb{R}[x]$. Give details.

Here are a couple of properties on the finite-dimensional case.

Lemma 21. *Given a finite-dimensional vector space V.*

(1) One has $\dim V = \dim V^*$.

(2) A choice of basis B_V for V determines a basis B_{V^*} for V^*.

Proof. By Corollary 5 we know that $\dim \mathcal{L}(V, \mathbb{F}) = \dim V \cdot \dim \mathbb{F} = \dim V$.

Let $B_v = \{v_1, \dots, v_n\}$ be a basis of V, then we define $B_{V^*} := \{\lambda_1, \dots, \lambda_n\} \subset V^*$ as follows:

$$\lambda_j(v_i) = \begin{cases} 1 \text{ if } j = i \\ 0 \text{ else} \end{cases}.$$

Notice that if there exists a collection of scalars $a_1, \dots, a_n \in \mathbb{F}$ not all zero such that $\lambda := \sum_i^n a_i \lambda_i = 0$, then $\lambda(v_1 + \cdots v_n) = 0$ would contradict $\{v_1, \dots, v_n\}$ being linearly independent. Thus, B_{V^*} is linearly independent. By Lemma 7 (2) B_{V^*} is a basis. □

Remark 19. The set B_{V^*} is called the *dual basis.* It has the following property:

$$v = \sum_{i=1}^n \lambda_i(v) v_i,$$

with the notation as in the proof of Lemma 21.

Definition 23. Given a linear map $T \in \mathcal{L}(V, W)$ its dual is the linear map $T^* \in \mathcal{L}(W^*, V^*)$ defined as follows:

$$T^*(\lambda) := \lambda \circ T \text{ for every } \lambda \in W^*.$$

Example 16. Let $D : \mathbb{R}[x] \to \mathbb{R}[x]$ be the linear map given by differentiation. Then $D^* : \mathbb{R}[x]^* \to \mathbb{R}[x]^*$ is given by $D^*(\varphi) = \varphi \circ D$. For example, if $\varphi \in \mathbb{R}[x]^*$ is given by $\varphi(p) = \int_0^1 p(x)dx$, then $D^* \circ \int_0^1 p(x)dx = p(1) - p(0)$ by the fundamental theorem of Calculus.

[6]The textbook uses the notation V' for the dual of V and similarly to other dual concepts. We adopt a notation that is consistent with Wikipedia.

Here are some properties of the construction from Definition 23.

Lemma 22. *Let V and W be two vector spaces. Then $(-)^* : \mathcal{L}(V,W) \to \mathcal{L}(W^*,V^*)$ is a linear operation. Moreover, we have $(T \circ S)^* = S^* \circ T^*$.*

Proof. Let $T, T' \in \mathcal{L}(V,W)$ then $(aT+T')^* : W^* \to V^*$ is given on an element $\lambda \in W^*$ by

$$(aT+T')^*(\lambda) = \lambda \circ (aT+T') = \lambda \circ aT + \lambda \circ T' = a\lambda \circ T + \lambda \circ T'$$

by Exercise 22. The second equation is left as an exercise. □

Exercise 31. Is the operation linear map that sends $T \in \mathcal{L}(V,W)$ to $T^* \in \mathcal{L}(W^*,V^*)$ an isomorphism? Prove or give a counter-example.

Exercise 32. Let V be a finite-dimensional vector space, prove that the map:

$$\Phi_V : V \to (V^\vee)^\vee \qquad v \mapsto \Phi_V(v)(\lambda) := \lambda(v), \text{ for every } \lambda \in V^\vee$$

is an isomorphism of vector spaces. Notice how we didn't need to pick a basis of V to write Φ_V.

We now study how the notion of dual vector space interacts with the concepts of null space and range. First we introduce the following:

Definition 24. Let $U \subset V$ be a subspace, the *annihilator of U* is:

$$U^0 := \{\lambda \in V^* \mid \lambda(u) = 0 \text{ for all } u \in U\}.$$

We notice that $U^0 \subset V^*$ is a subspace. Indeed, let $\lambda, \mu \in U^0$, then we have

$$(a \cdot \lambda + \mu)(v) = a \cdot \lambda(v) + \mu(v) = \lambda(a \cdot v) + \mu(v) = 0 + 0,$$

if $v \in U$, since this implies $a \cdot v \in U$ as well.

Example 17. (1) Let $\mathcal{P}_3(\mathbb{R}) \subset \mathcal{P}(\mathbb{R})$ be the subspace of degree at most 3 polynomials.

(2) Let $\mathbb{R}[x] \subset \mathbb{R}[[x]]$ be the subspace of polynomials. Then $\mathbb{R}[x]^0$ What is this?

Lemma 23. *Let $T : V \to W$ be a linear map, then:*

(i) $\operatorname{null} T^* = (\operatorname{range} T)^0$*;*

(ii) $\operatorname{range} T^* \subseteq (\operatorname{null} T)^0$ *is a subspace.*

Proof. For (i) let $\lambda \in \operatorname{null} T^*$ this implies that $\lambda \circ T : V \to \mathbb{F}$ vanishes. Consider $w \in \operatorname{range} T$, i.e. $w = Tv$ for some $v \in V$, then

$$\lambda(w) = \lambda(Tv) = 0.$$

So $\operatorname{null} T^* \subseteq (\operatorname{range} T)^0$. Now let $\lambda \in (\operatorname{range} T)^0$ then for any $w = Tv \in W$ we have

$$\lambda(w) = \lambda(Tv) = \lambda \circ T(v) = T^*(\lambda)(v) = 0.$$

Notice that the above equation holds for any v, thus $T^*(\lambda) = 0$, which gives $(\operatorname{range} T)^0 \subseteq \operatorname{null} T^*$.

For (ii), consider $\lambda \in \operatorname{range} T^*$, i.e. $\lambda = \mu \circ T$ for some $\mu \in V^*$. Let $v \in (\operatorname{null} T)^0$, then we have

$$\lambda(v) = \mu \circ T(v) = 0.$$

This implies that $\operatorname{range} T^* \subseteq (\operatorname{null} T)^0$ as sets. We leave it to the reader to check that one is a subspace of the other. □

Warning 1. Notice that in Lemma 23 (ii) we don't always have that $(\operatorname{null} T)^0 \subseteq \operatorname{range} T^*$. Can you give an example where this does *not* happen?

Remark 20. Let $U \subseteq V$ be a subspace of a finite-dimensional vector space V. Let $T : U \to W$ be a linear map, then there exist an *extension of T to V*, i.e. a linear map $\tilde{T} : V \to W$ such that $\tilde{T}\big|_U = T$. Indeed, let $\{u_1, \ldots, u_n\}$ be a basis of U, which can be extended to $\{u_1, \ldots, u_n, u_{n+1}, \ldots, w_m\}$ a basis of V. Then we define:

$$\tilde{T}(u_i) := \begin{cases} T(u_i) & \text{for } i \leq n; \\ 0 & \text{for } i \geq n+1. \end{cases}$$

Notice that this is well-defined and satisfy the required condition to be an extension.

We collect some properties of the annihilator when the ambient vector space is finite-dimensional.

Lemma 24. *Let $U \subset V$ be a subspace of a finite-dimensional vector space V.*

(i) $\dim U^0 = \dim V - \dim U$*;*

(ii) $U^0 = V^*$ *if and only if* $U = \{0\}$*;*

(iii) $U^0 = \{0\}$ *if and only if* $U = V$*.*

Proof. For (i) let $\imath : U \to V$ denote the inclusion linear map. Then we get a linear map $\imath^* : V^* \to U^*$ and Theorem 1 implies that:

$$\dim \operatorname{range} \imath^* + \dim \operatorname{null} \imath^* = \dim V^*.$$

Notice that $\operatorname{null} \imath^* = U^0$, by Lemma 23(i), since $\operatorname{range} \imath = U$.

Now we claim that $\operatorname{range} \imath^* = U^*$. Indeed, let $\lambda : U \to \mathbb{F}$ be a linear map, by Remark 20 there exist $\tilde{\lambda} : V \to \mathbb{F}$ such that $\tilde{\lambda} \circ \imath = \lambda$. Thus, $U^* \subseteq \operatorname{range} \imath^*$, which implies the equality, since we have $\operatorname{range} \imath^* \subseteq U^*$ by definition. Thus, we obtain:

$$\begin{aligned}
\dim V &= \dim V^* \\
&= \dim \operatorname{range} \imath^* + \dim \operatorname{null} \imath^* \\
&= \dim U^* + \dim U^0 \\
&= \dim U + \dim U^0,
\end{aligned}$$

where in the first and last equalities we used that $V \simeq V^*$ and $U \simeq U^*$ for finite-dimensional vector spaces.

For (ii) we notice that $U^0 \subseteq V^*$ being a subspace, we have that $U^0 = V^*$ if and only if $\dim U^0 = \dim V - \dim U = \dim V = \dim V^*$, that is $U = 0$.

Since (iii) is proved similarly, we leave the details to the reader.

□

Similar to Lemma 24 we collect some properties of the dual of a linear map when the vector spaces involved are finite-dimensional.

Lemma 25. *Suppose that V and W are finite-dimensional and let $T : V \to W$ be a linear map. Then:*

(i) $\dim \operatorname{null} T^* = \dim \operatorname{null} T + \dim W - \dim V$*;*

(ii) $\dim \operatorname{range} T^* = \dim \operatorname{range} T$*;*

(iii) $\operatorname{range} T^* = (\operatorname{null} T)^0$*;*

(iv) T^* *is injective if and only if* T *is surjective;*

(v) T^* *is surjective if and only if* T *is injective.*

Proof. For (i) notice that by Lemma 23 we have $\dim \operatorname{null} T^* = \dim(\operatorname{range} T)^0$. By the fundamental theorem of linear maps we have:

$$\dim V = \dim \operatorname{null} T + \dim \operatorname{range} T.$$

Notice that by Lemma 24 (1) we have $\dim(\operatorname{range} T)^0 = \dim W - \dim \operatorname{range} T$. Thus, substituting back we obtain:

$$\dim V = \dim \operatorname{null} T + \dim W - \dim(\operatorname{range} T)^0 \qquad = \dim \operatorname{null} T + \dim W - \dim \operatorname{null} T^*.$$

For (ii) we compute:

$$\begin{aligned}
\dim \operatorname{range} T^* &= \dim W^* - \dim \operatorname{null} T^* \\
&= \dim W - \dim(\operatorname{range} T)^0 \\
&= \dim \operatorname{range} T,
\end{aligned}$$

where the first equality is the fundamental theorem of linear maps applied to T^*, the second uses that W is finite-dimensional, so $\dim W = \dim W^*$ and Lemma 23(i). The last equality follows from Lemma 24(i).

For (iii), by Lemma 23(ii) we have that $\operatorname{range} T^* \subseteq (\operatorname{null} T)^0$ is a subspace. We will be done if we prove that they have the same dimension. Notice that:

$$\begin{aligned}\dim \operatorname{range} T^* &= \dim \operatorname{range} T\\ &= \dim V - \dim \operatorname{null} T\\ &= \dim (\operatorname{null} T)^0,\end{aligned}$$

where the first equality comes from (ii), the second from the fundamental theorem of linear maps, and the third from Lemma 24(i).

(iv) We have the following sequence of logical equivalences:

$$\begin{aligned}T^* \text{ is injective } &\Leftrightarrow \operatorname{null} T^* = 0\\ &\Leftrightarrow (\operatorname{range} T)^0 = 0\\ &\Leftrightarrow \operatorname{range} T = W\\ &\Leftrightarrow T \text{ is surjective }.\end{aligned}$$

Here the first equality is Lemma 11(i), the second is Lemma 23(i), the third is Lemma 24(ii) and the last is Lemma 11(ii).

The argument for (v) is similar to that for (iv) and we leave it as an exercise. □

Finally, we remark on how passing to the dual linear map interacts with associating a matrix to a linear map. Consider the function:

$$\begin{aligned}(-)^{\mathrm{t}} : \mathbb{F}^{m,n} &\to \mathbb{F}^{n,m}\\ (a_{ij}) &\mapsto (a_{ji})\end{aligned}$$

that sends a matrix to its transpose, i.e. we swap the indices of its terms.

Lemma 26. *Let V and W be two finite-dimensional vector spaces with bases B_V and B_W. These determine bases B_{V^*} and B_{W^*} of V^* and W^*, respectively, as explained in Remark 19. The following diagram commutes:*

$$\begin{tikzcd} \mathcal{L}(V,W) \arrow[r, "\mathcal{M}(-,B_V,B_W)"] \arrow[d, "(-)^*"'] & \mathbb{F}^{m,n} \arrow[d, "(-)^{\mathrm{t}}"] \\ \mathcal{L}(W^*,V^*) \arrow[r, "\mathcal{M}(-,B_{W^*},B_{V^*})"'] & \mathbb{F}^{n,m} \end{tikzcd},$$

where $n = \dim V$ and $m = \dim W$.

Proof. Write this! □

Here is a nice consequence of the previous result.

Corollary 9. *Let $A \in \mathbb{F}^{m,n}$ be an m by n matrix. Then $\dim \operatorname{col}(A) = \dim \operatorname{row}(A)$.*

Proof. Let $T : \mathbb{F}^n \to \mathbb{F}^m$ be the linear operator corresponding to A for B_n and B_m the standard bases of $\mathbb{F}^n$ and $\mathbb{F}^m$, respectively. That is we have $\mathcal{M}(T, B_n, B_m) = A$. Then we have:

$$\begin{aligned}\dim \operatorname{col}(A) &= \dim \operatorname{range} T\\ &= \dim \mathit{range} T^*\\ &= \dim \operatorname{col}(\mathcal{M}(T^*, B_m^*, B_n^*))\\ &= \dim \operatorname{col}(A^{\mathrm{t}})\\ &= \dim \operatorname{row}(A).\end{aligned}$$

□

8.2. **Polynomials.** This subsection is a digression that collects some facts about polynomials that we will need in the next topic.

Recall that $\mathbb{C}$ is the field of complex numbers. As a set one has $\mathbb{C} \simeq \mathbb{R}\times\mathbb{R}$ the addition and multiplication are given by:

$$(a,b)+(a',b')=(a+a',b+b') \quad \text{and} \quad (a,b)\cdot(a',b')=(aa'-bb',ba'+ab').$$

These are normally denoted by $a+ib:=(a,b)$. We have functions:

- $\overline{(-)}:\mathbb{C}\to\mathbb{C}$ given by $\overline{(a+ib)}=a-ib$ called *complex conjugation*;
- $\operatorname{Re}:\mathbb{C}\to\mathbb{R}$ given by $\operatorname{Re}(a+ib)=a$ called the *real part*;
- $\operatorname{Im}:\mathbb{C}\to\mathbb{R}$ given by $\operatorname{Im}(a+ib)=b$ called the *imaginary part*;
- $|-|:\mathbb{C}\to\mathbb{R}$ given by $|a+ib|=\sqrt{a^2+b^2}$ called the *absolute value.*

Here is a list of properties they satisfy:

Lemma 27. *For any two complex numbers $z,w\in\mathbb{C}$ we have:*

(i) $z+\overline{z}=2\operatorname{Re}z$;

(ii) $z-\overline{z}=2\operatorname{Im}zi$;

(iii) $z\overline{z}=|z|^2$;

(iv) $\overline{z+w}=\overline{z}+\overline{w}$ *and* $\overline{zw}=\overline{z}\overline{w}$;

(v) $\overline{\overline{z}}=z$;

(vi) $|\operatorname{Re}z|\le|z|$ *and* $|\operatorname{Im}z|\le z$;

(vii) $|\overline{z}|=|z|$;

(viii) $|zw|=|z||w|$;

(ix) $|z+w|\le|z|+|w|$.

Proof. Left as an exercise. □

Definition 25. Given a polynomial $p\in\mathbb{F}[x]$ a *zero of p* is an element $\alpha\in\mathbb{F}$ such that $p(\alpha)=0$.

Lemma 28. *Let $p\in\mathbb{F}[x]$ be a polynomial then the following are equivalent:*

(1) α is a zero of p;

(2) there exists a polynomial $q\in\mathbb{F}[x]$ such that $p(x)=(x-\alpha)q(x)$.

Proof. Assume (1) and let $p(z)=\sum_{i=0}^m a_iz^i$ for some coefficients $a_i\in\mathbb{F}$. Then we have:

$$p(z)-p(\alpha)=\sum_{i=1}^m a_i(z^i-\alpha^i).$$

Recall that each $z^i-\alpha^i$ can be factored as:

$$z^i-\alpha^i=(z-\alpha)\sum_{j=1}^i z^{i-j}\alpha^{j-1}.$$

Thus, we obtain:

$$\begin{aligned}\sum_{i=1}^m a_i(z^i-\alpha^i)&=\sum_{i=1}^m a_i(z-\alpha)\sum_{j=1}^i z^{i-j}\alpha^{j-1}\\&=(z-\alpha)\sum_{i=1}^m a_i\sum_{j=1}^i z^{i-j}\alpha^{j-1},\end{aligned}$$

that is we can take $q(z)=\sum_{i=1}^m a_i\sum_{j=1}^i z^{i-j}\alpha^{j-1}$. Notice that this is not unique.

Now assume (2), then clearly we get $p(\alpha)=(\alpha-\alpha)q(\alpha)=0$. □

Here is a basic result about polynomials that follows from the previous Lemma.

Corollary 10. *Let p be a polynomial of degree $m \geq 0$. Then p has at most m zeros.*

Proof. We proceed by induction on m. For $m = 0$ we have $p(z) = a_0$ where $a_0 \neq 0$[7]. Thus, $p(\alpha) = a_0$ for every $\alpha \in \mathbb{F}$, so p has no zeros.

Assume we proved the result for all polynomials of degree up to $m-1$. Let p have degree m. If it has no zeros, there is nothing to prove. Assume that p has a zero α, then $p(z) = (z-\alpha)q(z)$ by Lemma 28. Since $\deg p = \deg(z-\alpha)\deg(q)$ we obtain that q has degree $m-1$, so it has at most $m-1$ zeros by the inductive hypothesis. This finishes the proof. □

The following result is called the division algorithm for polynomials, it is analogous to the Euclidean algorithm that you learned to perform division of integers.

Lemma 29. *Let $p, s \in \mathbb{F}[x]$ be two polynomials such that $s \neq 0$. Then there exist unique polynomials $q, r \in \mathbb{F}[x]$ such that:*[8]

(i) $p = qs + r$;

(ii) $\deg r < \deg s$.

Proof. Exercise for the reader. □

The following result is a fact. In fact, it can be seen as the "raison d'être" of the complex numbers.

Theorem 2. *Every non-constant polynomial with complex coefficients has a root.*

Proof. See page 125 in the textbook. **TODO:** Give another argument. □

Remark 21. There are many proofs of this theorem. You should make sure you know at least one. Also, most places call the Theorem 2 the "fundamental theorem of algebra".

We also need a result on the factorization of real polynomials.

Proposition 3. *Let $p \in \mathbb{R}[x]$ be a nonconstant polynomial. Then p has an unique factorization:*

$$p(x) = a\prod_{i=1}^{m}(x-\lambda_i)\prod_{j=1}^{k}(x^2 + b_j x + c_j),$$

where $a, \lambda_1, \ldots, \lambda_m, b_1, \ldots, b_k, c_1, \ldots, c_k \in \mathbb{R}$ and for each $1 \leq j \leq k$ we have $b_j^2 < 4c_j$.

Proof. The proof is obtained by considering the factorization of Theorem 2, i.e. that $p(x) = \prod_{i=1}^{n}(x-\alpha_i)$ where $\alpha_i \in \mathbb{C}$. We then split this product into the factors which are already real and the factors where we have $(x-\alpha_i)(x-\alpha_j)$ where $\alpha = \alpha_i = \overline{\alpha_j}$. One can prove by induction that this most be the case for $p(x)$ to have real coefficients. Then the product:

$$x^2 - (\alpha + \overline{\alpha}) + \alpha\overline{\alpha}$$

can be easily seen to satisfy $(\alpha + \overline{\alpha})^2 < 4|\alpha|^2$ by Lemma 27 (i). □

9. Feb. 19, 2024

9.1. Invariant subspaces.

Definition 26. Let $T : V \to V$ be a linear map from V to itself. We will call T an *operator*. A subspace $U \subseteq V$ is *invariant (with respect to T)*[9] if the restriction $T|_U : U \to V$ factors as follows:

$$\begin{tikzcd} U \arrow[r, dashed, "T|_U"] \arrow[d, "\subset"'] & U \arrow[d, "\subset"] \\ V \arrow[r, "T"'] & V \end{tikzcd},$$

i.e. for every $u \in U$ we have $T(u) \in U$.

[7]Recall that we posed that $\deg 0 = -\infty$, which avoids us having to deal with the exception of the zero polynomial, which possibly has infinitely many zeros, e.g. if the field $\mathbb{F}$ is infinite.

[8]Notice that if we drop condition (ii) that there are no unique polynomials satisfying (i).

[9]When the operator T is clear from the context we will omit the part "with respect to T".

Example 18. (i) Given any operator $T : V \to V$, then $\{0\}, V, \operatorname{null} T$, and $\operatorname{range} T$ are invariant subspaces of V.

(ii) Let $D : \mathcal{P}(\mathbb{R}) \to \mathcal{P}(\mathbb{R})$ denote the differentiation operator. Then $\mathcal{P}_n(\mathbb{R}) \subset \mathcal{P}(\mathbb{R})$ the subspace of polynomials of degree at most n is invariant subspace for every $n \in \mathbb{N}$.

The following concepts are key definitions in this class.

Definition 27. Given an operator $T : V \to V$. A non-zero vector $v \in V$ such that $\operatorname{Span} v \subseteq V$ is an invariant subspace with respect to T is called an *eigenvector*.

Remark 22. Given an operator $T : V \to V$ and a non-zero vector $v \in V$ such that $\operatorname{Span} v \subseteq V$ is an invariant subspace with respect to T. Then we claim that $T(v) = \lambda \cdot v$ for an unique $\lambda \in \mathbb{F}$. Indeed, since $T(v) \in \operatorname{Span} v$ we have that $T(v) = \lambda v$ for some $\lambda \in \mathbb{F}$.

Definition 28. Let $T : V \to V$ be an operator.

(i) if $v \in V$ is an eigenvector of T, with $T(v) = \lambda v$ we say that λ is the eigenvalue of v;

(ii) we will say that $\lambda \in \mathbb{F}$ is an *eigenvalue of* T if there exists an eigenvector $v \in V$ such that λ is an eigenvector of v.

Remark 23. Suppose that $\lambda \in \mathbb{F}$ is an eigenvalue of T and $v \in V$ and eigenvector, then any non-zero multiple av is also an eigenvector of T with eigenvalue λ.

Lemma 30. *Given a vector space V and an operator $T : V \to V$ the following are equivalent:*

(1) λ is an eigenvalue of T;

(2) $T - \lambda \operatorname{Id}_V$ is not injective.

Assume that V is a finite-dimensional vector space, then the above are also equivalent to:

(3) $T - \lambda \operatorname{Id}_V$ is not surjective;

(4) $T - \lambda \operatorname{Id}_V$ is not invertible.

Proof. Left as an exercise. □

Example 19. (i) Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ denote the operator given by $T(x, y) = (-y, x)$. Assume that $T(x, y) = \lambda(-y, x)$ for some non-zero $(x, y) \in \mathbb{R}$. Then we have $x = -\lambda y = -\lambda^2 x$ and $y = \lambda x = -\lambda^2 y$, which gives $\lambda^2 = -1$, since either x or y is non-zero. However, there is no $\lambda \in \mathbb{R}$ such that $\lambda^2 = -1$. Thus, T has no eigenvectors and so no eigenvalues.

(ii) Now consider Let $T : \mathbb{C}^2 \to \mathbb{C}^2$ denote the operator given by $T(z, w) = (-w, z)$. The same argument as in (i) gives that $\lambda^2 = -1$ which gives $\lambda = \pm i$ in $\mathbb{C}$. By inspection we see that $T(z, -iz) = i(z, -iz)$ and $T(z, iz) = -i(z, iz)$, so for any $z \neq 0$ we have that $(z, -iz)$ is an eigenvector of i and (z, iz) is an eigenvector of $-i$.

Lemma 31. *Given an operator $T : V \to V$ and $\{v_1, \dots, v_n\}$ a list of eigenvectors whose eigenvalues $\{\lambda_i\}_{1 \leq i \leq n}$ satisfy $\lambda_i \neq \lambda_j$ for every $i \neq j$. Then $\{v_1, \dots, v_n\}$ are linearly independent.*

Proof. Let $\{v_1, \dots, v_n\}$ be the smallest set of vector such that we have a non-trivial linear combination

$$\sum_{i=1}^{n} a_i v_i = 0.$$

We apply $(T - \lambda_n)$ to the equation above to obtain:

$$\sum_{i=1}^{n} a_i(\lambda_i - \lambda_n) v_i = \sum_{i=1}^{n-1} a_i(\lambda_i - \lambda_n) v_i = 0,$$

where all of $a_i(\lambda_i - \lambda_n) \neq 0$ for $1 \leq i \leq n - 1$. This gives a non-trivial linear combination of smaller set of vectors, which is a contradiction with the initial assumption. □

Corollary 11. *Let $T : V \to V$ be an operator on a finite-dimensional vector space. Then T has at most $\dim V$ distinct eigenvalues.*

Proof. If T has m distinct eigenvalues then Lemma 31 gives that there are m distinct linearly independent vectors. Lemma 3 gives that $m \leq \dim V$. □

9.2. **Powers of operators.** Let $T : V \to V$ be a linear operator on a vector space V over a field $\mathbb{F}$. Then for any polynomial $p \in \mathbb{F}[x]$ we let $p(T) : V \to V$ be the linear operator given by:

$$p(T) := a_0 \operatorname{Id}_V + a_1 T + a_2 T^2 + \cdots + a_n T^n.$$

Notice that this assignment respect the following compatibilities, given $p, q \in \mathbb{F}[x]$ two polynomials and an operator $T : V \to V$ then we have:

$$(pq)(T) = p(T)q(T) \quad \text{and} \quad p(T)q(T) = q(T)p(T). \tag{6}$$

Exercise 33. Check the relations in (6).

It turns out that by applying polynomials to an operator T we can create, possibly new, invariant subspaces.

Remark 24. Given an operator $T : V \to V$ and a polynomial $p \in \mathbb{F}[x]$, then $\operatorname{null} p(T)$ and $\operatorname{range} p(T)$ are invariant under T. Indeed, assume that $v \in \operatorname{null} p(T)$, then $p(T) \circ T(v) = T \circ p(T)(v) = T(0) = 0$. Similarly, if for $w \in \operatorname{range} p(T)$ then $w = p(T)(v)$ for some $v \in V$ and we have $T(w) = T \circ p(T)(v) = p(T) \circ T(v) = p(T)(T(v))$, thus $T(w) \in \operatorname{range} p(T)$.

10. Feb. 22, 2024

10.1. **The minimal polynomial.** The following is one of the most important results of this course.

Proposition 4. *Let $T : V \to V$ be an operator on a finite-dimensional complex vector space V. Then T has an eigenvalue.*

Proof. Let $n = \dim V$ and consider $v \in V$ a non-zero vector. The set $\{v, Tv, \ldots, T^n v\}$ is not linearly independent, since it has size $n+1$ which is larger than the dimension of V. Thus, there exist $a_0, a_1, \cdots, a_n$ such that

$$\sum_{i=0}^{n} a_i T^i(v) = 0.$$

We let $p(T) = \sum_{i=0}^{n} a_i T^i$. Notice that we can assume that $p(T)$ has the smallest possible degree. By Theorem 2 there exist $\lambda \in \mathbb{C}$ such that $p(\lambda) = 0$. Let $q \in \mathbb{C}[z]$ such that $p(z) = (z - \lambda)q(z)$. Thus, we have:

$$0 = p(T)(v) = (T - \lambda \operatorname{Id}_V)(q(T)(v)). \tag{7}$$

Since $\deg q < \deg p$ we have $q(T)(v) \neq 0$. Thus, $q(T)(v)$ is an eigenvector of T with eigenvalue λ. □

Example 20. Let $T : \mathcal{P}(\mathbb{C}) \to \mathcal{P}(\mathbb{C})$ be given by $Tp(z) := zp(z)$. We claim that T has no eigenvalues. Indeed, assume by contradiction that $\lambda \in \mathbb{C}$ is an eigenvalue of T, i.e. there exist $q \in \mathcal{P}(\mathbb{C})$ such that

$$Tq(z) = zq(z) = \lambda q(z).$$

However, this is a contradiction, since $\deg Tq > \deg \lambda q$.

Proposition 5. *Given a finite-dimensional vector space V and an operator $T : V \to V$, there exist an unique monic[10] polynomial $p \in \mathbb{F}[x]$ such that $p(T) = 0$ and $\deg p \leq \dim V$.*

Proof. We induce on $\dim V$. For $\dim V = 0$ then $T = \operatorname{Id}_V$ and we can take $p = 1$ the constant polynomial with value 1.

Let $n = \dim V \geq 1$ and assume that the result holds for every vector space W over $\mathbb{F}$ of dimension (stricly) less than n. Consider $v \in V$ a non-zero vector, then $\{v, Tv, \ldots, T^n v\}\}$ are linearly dependent and Lemma 2 implies that there exist $m \leq n$ such that

$$T^m v = \sum_{i=0}^{m-1} a_i T^i v$$

[10] A polynomial $p(x) = \sum_{i=0}^{n} a_i x^i$ of degree n is said to be *monic* if $a_n = 1$.

for some scalars $a_0, \dots, a_{m-1}$. Thus, $p(T) := T^m - \sum_{i=}^{m-1} a_i T^i$ is such that $p(T)(v) = 0$. Notice that $p(T)(T^i v) = T^i \circ p(T)(v) = 0$ for every $i \geq 1$. The set $\{Tv, T^2 v, \dots, T^{m-1} v\}$ is linearly independent and $\operatorname{Span}\{Tv, T^2 v, \dots, T^{m-1} v\} \subseteq \operatorname{null} p(T)$, which implies that:

$$\dim \operatorname{range} p(T) \leq n - m.$$

By Remark 24 $\operatorname{range} p(T) \subseteq V$ is subspace invariant under T. Thus, we can consider:

$$T|_{\operatorname{range} p(T)} : \operatorname{range} p(T) \to \operatorname{range} p(T)$$

and we notice that $\dim \operatorname{range} p(T) < n$, since $m \geq 1$. Thus the inductive hypothesis implies that there exist an unique monic polynomial $r \in \mathbb{F}[x]$ such that

$$r(T|_{\operatorname{range} p(T)}) = 0, \quad \text{and} \quad \deg r < \dim \operatorname{range} p(T) \leq n - m.$$

Hence for all $v \in V$ we have:

$$rp(T)(v) = r(p(T)(v)) = 0,$$

i.e. $rp(T) = 0$. And $\deg rp \leq n - m + m = n$. This finishes the proof of the existence of a polynomial.

Assume that there exist two monic polynomials $p, q \in \mathbb{F}[x]$ such that $p(T) = q(T) = 0$ and they are both of minimal degree. Notice that $\deg p - q < \deg p$. Assume that $p - q \neq 0$, then we can divide $p - q$ by the coefficient of its highest power to make it a monic polynomial of degree less than p, which is a contradiction. Thus, $p = q$. This finishes the proof. $\square$

Definition 29. Given an operator $T : V \to V$ on a finite-dimensional vector space the unique monic polynomial p_T[11] determined by Proposition 5 is called the *minimal polynomial of T*.

Example 21. Consider $T : \mathbb{F}^n \to \mathbb{F}^n$ given by $T(x_1, \dots, x_n) = (x_1 + \cdots + x_n, \dots, x_1 + \cdots x_n)$. Notice that for any $i \in \{1, \dots, n\}$ we have $v_0 := T(e_i) = (1, \dots, 1)$. And that $T(v_0) = nv_0$, i.e. v_0 is an eigenvector with eigenvalue n.

Also notice that $T(e_i - e_{i+1}) = 0$ for $1 \leq i \leq n-1$, i.e. $e_i - e_{i+1}$ are all eigenvectors with eigenvalue 0. Thus, for any $v = \sum_{i=1}^n a_i e_i$ we have:

$$T(v) = (\sum_{i=1}^n a_i) v_0.$$

Since $T(v_0) = nv_0$, we obtain that

$$T^2(v) - nT(v) = 0$$

for every $v \in V$. As $T(v) \neq \lambda v$ for every $v \in V$, we have that $p(T) = T^2 - T$ is the minimal polynomial of T.

Remark 25. It can be sometimes tricky to compute the minimal polynomial of an operator, you should read the discussion on page 145-146 just after Definition 5.24. See also, Example 5.26 for a concrete example of how one can compute the minimal polynomial.

The following illustrate the relation between the minimal polynomial and eigenvalues of an operator.

Lemma 32. *Given $T : V \to V$ an operator on a finite-dimensional vector space over $\mathbb{F}$, let p_T be its minimal polynomial. For $\lambda \in \mathbb{F}$ the following are equivalent:*

(1) $p_T(\lambda) = 0$, i.e. λ is a root of p_T;

(2) λ is an eigenvalue of T.

Proof. (1) $\Rightarrow$ (2) by Lemma 28 we can write $p_T(x) = (x - \lambda) q(x)$ for some monic polynomial $q \in \mathbb{F}[x]$. Then we have $p(T) = (T - \lambda) q(T) = 0$, which gives

$$(T - \lambda \operatorname{Id}_V) q(T)(v) = 0,$$

for every $v \in V$. Now, since $\deg q < \deg p_T$, there exist at least one non-zero vector $v \in V$ such that $q(T)(v) \neq 0$, which implies that $(T - \lambda \operatorname{Id}_V)(v) = 0$, i.e. v is an eigenvector with eigenvalue λ.

(2) $\Rightarrow$ (1) let $v \in V$ be such that $T(v) = \lambda v$. By applying T repeatedly we obtain $T^k = \lambda^k v$, which implies:

$$p_T(T) v = p(\lambda) v.$$

Since $p_T(T)(v)$ and $v \neq 0$, we have $p_T(\lambda) = 0$, as required. $\square$

[11]We sometimes will use the notation p_T for the minimal polynomial associated to T.

Corollary 12. *Let $T : V \to V$ be an operator on a finite-dimensional vector space over $\mathbb{C}$ the minimal polynomial of T is given by:*

$$p(z) = \prod_{i=1}^{m}(z - \lambda_i),$$

where each λ_i is an eigenvalue of T.

Proof. This is a direct consequence of Theorem 2 and Lemma 32. □

Warning 2. Notice that in Corollary 12 the roots λ_i might appear multiple times. Here is an example. Consider $T : \mathbb{C}^2 \to \mathbb{C}^2$ given by $T(x_1, x_2) = (x_1, x_1+x_2)$. We know that T should have a minimal polynomial of degree 2, since T is not a multiple of the identity. We want to find $a, b \in \mathbb{C}$ such that

$$(T^2 + aT + b)(e_1) = 0, \text{ and } (T^2 + aT + b)(e_2) = 0.$$

The first equation gives $1 + a + b = 0$, while the second gives $1 + a + b = 0$ and $2 + a = 0$. We solve for $a = -2$ and $b = 1$. Thus, we have:

$$p(T) = T^2 - 2T + 1 = (T - 1)^2.$$

The following result is very useful for understanding how the minimal polynomial interacts with other polynomials we can write down from an operator.

Lemma 33. *Let $T : V \to V$ be an operator on a finite-dimensional vector space and consider a polynomial $q \in \mathbb{F}[x]$. Then $q(T) = 0$ if and only if p_T divides q.*

Proof. By Lemma 29 there exist $r, s \in \mathbb{F}[x]$ such that

$$q = p_T s + r$$

with $\deg r < \deg p$. Then we have:

$$q(T) = p_T(T)s(T) + r(T) = 0.$$

We claim that r is the 0 polynomial. Indeed, if that were not the case then r divided by the coefficient of its highest term would be a monic polynomial of degree smaller than $\deg p_T$ such that $r(T) = 0$, which is a contradiction with p_T being the minimal polynomial. Thus, q is a multiple of p_T as claimed.

Assume that $q(x) = p_T(x)s(x)$, then clearly we have $q(T) = p_T(T)s(T) = 0$. □

Corollary 13. *Let $T : V \to V$ be an operator on a finite-dimensional vector space. Then T is invertible if and only if p_T has a non-zero constant term.*

Proof. Recall that: by Lemma 30

$$0 \text{ is an eigenvalue } \Leftrightarrow\ T \text{ is not invertible},$$

by Lemma 30. And that:

$$0 \text{ is an eigenvalue } \Leftrightarrow\ 0 \text{ is a root of } p_T$$

by Lemma 32. It is clear that 0 is a root of p_T if and only if the constant term of p_T vanishes. □

The next result is the first general statement about the existence of eigenvalues for operators on real vector spaces.

Proposition 6. *Let $T : V \to V$ be an operator on a finite-dimensional real vector space V and assume that $\dim V$ is odd. Then T has an eigenvalue.*

Proof. Let $\dim V = 2n + 1$, we will use induction on n to prove the result.

The base case is clear, since when $V = \mathbb{R}$ any operator is simply given by scalar multiplication.

Let $n \geq 1$ and assume the result holds for all V such that $\dim V = 2k + 1$ and $k < n$. Let p_T be the minimal polynomial of T. If p_T is a multiple of $(x - \lambda)$ then by Lemma 32 λ is an eigenvalue and we are done. So we can assume that there exist $b, c \in \mathbb{R}$ and a monic polynomial $q \in \mathbb{R}[x]$ such that $p_T(x) = q(x)(x^2 + bx + c)$ and that $b^2 < 4c$. Then

$$p(T) = q(T)(T^2 + bT + c),$$

which implies that $q(T)|_{\text{range}(T^2+bT+c)} = 0$. Since $\deg q < \deg p$ we have that $\text{range}(T^2 + bT + c) \neq V$. By Theorem 1 we have:

$$\dim V = \dim \text{null}(T^2 + bT + c) + \dim \text{range}(T^2 + bT + c).$$

Now, $\dim \text{null}(T^2 + bT + c)$ is even by Lemma 34 below, which implies that $\dim \text{range}(T^2 + bT + c)$, since $\dim V$ is odd. Thus, $T|_{\text{range}(T^2+bT+c)}$ has an eigenvalue by the inductive hypothesis, which implies that T has an eigenvalue. □

Lemma 34. *Let $T : V \to V$ be an operator on a finite-dimensional real vector space. Consider $b, c \in \mathbb{R}$ such that $b^2 < 4c$. Then $\dim \text{null}(T^2 + bT + c)$ is even.*

Proof. By considering $T|_{\text{null}(T^2+bT+c)}$ we can assume that $\text{null}(T^2 + bT + c) = V$. Assume there are $\lambda \in \mathbb{R}$ and $v \in V$ such that $Tv = \lambda V$, then we have:

$$(T^2 + bT + c)v = (\lambda^2 + b\lambda + c)v = \left((\lambda + \frac{b}{2})^2 + c - \frac{b^2}{4} \right) v,$$

and since $(\lambda + \frac{b}{2})^2 + c - \frac{b^2}{4} > 0$, we obtain that $v = 0$, i.e. T has no eigenvectors.

Let $U \subseteq V$ be the largest even-dimensional subspace invariant under T. If $U = V$ we are done. So assume that there exist $w \in V$ such that $w \notin U$. Notice that $W := \text{Span}\{w, Tw\}$ is invariant under T, since $T^2 w = -bTw - cw$. Also $\dim W = 2$, otherwise w would be an eigenvector of T. Moreover, it is clear that $U \cap W = \{0\}$, otherwise it would be a one-dimensional subspace invariant under T. So we have:

$$\dim(U + W) = \dim U + \dim W - \dim U \cap W = \dim U + 2.$$

Thus, we produced a larger invariant subspace $U \subset U + W \subset V$, which contradicts the assumption that $U \neq V$. This finishes the proof. □

11. Feb. 26, 2024

11.1. **Upper-Triangular matrices.** In this section we discuss the relation between finding invariant subspaces of an operator and being able to write a matrix in upper-triangular form.

Lemma 35. *Let $T : V \to V$ be an operator on a finite-dimensional vector space and $B_V = \{v_1, \dots, v_n\}$ a basis of V. Then the following are equivalent:*

(1) the matrix $\mathcal{M}(T, B_V)$ associated to T (see Notation 4) is upper-triangular;

(2) $T(v_i) \in \text{Span}(v_1, \dots, v_i)$ for every $i \in \{1, \dots, n\}$;

(3) $\text{Span}(v_1, \dots, v_i)$ is invariant under T for every $i \in \{1, \dots, n\}$.

Proof. (1) ⇒ (2). By assumption for each $i \in \{1, \dots, n\}$ we have

$$T(v_i) = \sum_{j=1}^{i} a_j v_j,$$

for some constants a_j, i.e. $T(v_i) \in \text{Span}(v_1, \dots, v_i)$.

(2) ⇒ (3). For any $i \in \{1, \dots, n\}$, notice that $T(v_j)\text{Span}(v_1, \dots, v_i)$ for every $1 \leq j \leq i$. Thus $T(\text{Span}(v_1, \dots, v_i)) \subseteq \text{Span}(v_1, \dots, v_i)$.

(3) ⇒ (1). For each $i \in \{1, \dots, n\}$ we have

$$T(v_i) \in \text{Span}(v_1, \dots, v_n),$$

i.e. there exist $a_{i,j} \in \mathbb{F}$ such that

$$T(v_i) = \sum_{j=1}^{i} a_{i,j} v_j.$$

Since $\mathcal{M}(T, B_V) := (a_{i,j})_{1 \leq i,j \leq n}$, the claim follows. □

Exercise 34. Let $T : V \to V$ be an operator on a finite-dimensional vector space and $B_V = \{v_1, \ldots, v_n\}$ a basis of V. Assume that

$$\mathcal{M}(T, B_V) := \begin{pmatrix} \lambda_1 & * & \iddots & \cdots & * \\ 0 & \lambda_2 & \iddots & \ddots & * \\ \vdots & \ddots & \ddots & \ddots & * \\ 0 & \ddots & \ddots & \lambda_{n-1} & * \\ 0 & 0 & \cdots & 0 & \lambda_n \end{pmatrix} \tag{8}$$

for some scalars $\lambda_1, \ldots, \lambda_n \in \mathbb{F}$. Prove that

$$(T - \lambda_1 \cdot \mathrm{Id}_V) \cdot \ldots \cdot (T - \lambda_n \cdot \mathrm{Id}_V) = 0.$$

Corollary 14. *Let $T : V \to V$ be an operator on a finite-dimensional vector space. Assume that for some basis $B_V = \{v_1, \ldots, v_n\}$, the matrix $\mathcal{M}(T, B_V)$ has the form (8). Then the eigenvalues of T are precisely $\lambda_1, \ldots, \lambda_n$.*

Proof. Notice that $T(v_1) = \lambda_1 v_1$ and v_1 is non-zero, since it is part of a basis. Thus, λ_1 is one eigenvalue. For any $i \in \{2, \ldots, n\}$ notice that

$$(T - \lambda_i)v_i \in \mathrm{Span}\,(v_1, \ldots, v_{i-1}).$$

Thus, $T - \lambda_i|_{\mathrm{Span}\,(v_1,\ldots,v_i))}$ factors as follows:

$$T - \lambda_i|_{\mathrm{Span}\,(v_1,\ldots,v_i))} : \mathrm{Span}\,(v_1, \ldots, v_i) \to \mathrm{Span}\,(v_1, \ldots, v_{i-1}).$$

Since $i = \dim \mathrm{Span}\,(v_1, \ldots, v_i) > \mathrm{Span}\,(v_1, \ldots, v_{i-1}) = i - 1.$, the linear map $T - \lambda_i|_{\mathrm{Span}\,(v_1,\ldots,v_i))}$ is not injective. So there exists $v \in \mathrm{Span}\,(v_1, \ldots, v_i)$, $v \neq 0$ such that $(T - \lambda_i)v = 0$, so λ_i is an eigenvalue of T. This proves that $\{\lambda_1, \cdots, \lambda_n\}$ are all eigenvalues.

One needs to argue that there are none others. Let $q(z) := \prod_{i=1}^{n}(z - \lambda_i)$, then $q(T) = 0$; then Lemma 33 implies that q is a multiple of p_T, the minimal polynomial. Thus, Lemma 32 imply that all eigenvalues of T are of q, hence contained in the list $\{\lambda_1, \ldots, \lambda_n\}$. This finishes the proof. ☐

Let's now analyze an example that illustrates the relation between a property of the minimal polynomial and the existence of upper-triangular matrix representation.

Example 22. Consider $T : \mathbb{R}^4 \to \mathbb{R}^4$ given by $T(x_1, x_2, x_3, x_4) = (-x_2, x_1, 2x_1 + 3x_3, x_3 + 3x_4)$. For B the standard basis of $\mathbb{R}^4$ we have:

$$\mathcal{M}(B, T) = \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 2 & 0 & 3 & 0 \\ 0 & 0 & 1 & 3 \end{pmatrix}.$$

You can use some computation tool[12] to get that the minimal polynomial of T is:

$$p_T(x) = x^4 - 6x^3 + 10x^2 - 6x + 9.$$

Over the real numbers we have the factorization:

$$p_T(x) = (x^2 + 1)(x - 3)^2,$$

where the term $x^2 + 1$ can't be factored further. However, over the complex numbers we have:

$$p_T(z) = (z - i)(z + i)(z - 3)^2.$$

Now, with respect to the basis $B' = \{(4-3i, -3-4i, -3+i, 1), (4+3i, -3+4i, -3-i, 1), (0, 0, 1, 0), (0, 0, 0, 1)\}$ the operator T is given by:

$$\mathcal{M}(B', T) = \begin{pmatrix} i & 0 & 0 & 0 \\ 0 & -i & 0 & 0 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 3 \end{pmatrix}.$$

[12] I tried ChatGPT and the result was terrible. But WolframAlpha did, tough they called it the characteristic polynomial, so one needs to check that is actually the minimal polynomial.

Lemma 36. *Let $T : V \to V$ be an operator on a finite-dimensional vector space. Then the following are equivalent:*

(1) there exists a basis B of V, such that $\mathcal{M}(B, T)$ is upper-triangular;

(2) the minimal polynomial $p_T(x) = \prod_{i=1}^{m}(x - \lambda_i)$, for some $\lambda_1, \dots, \lambda_m \in \mathbb{F}$.

Proof. (1) $\Rightarrow$ (2) let $\{\alpha_1, \dots, \alpha_n\}$ be the diagonal entries of $\mathcal{M}(B, T)$ and define $q(x) := \prod_{i=1}^{n}(x - \alpha_i)$. Lemma 33 implies that p_T divides q, so $p_T = \prod_{i=1}^{m}(x - \lambda_i)$ for a subset $\{\lambda_1, \dots, \lambda_m\} \subseteq \{\alpha_1, \dots, \alpha_n\}$.

(1) $\Rightarrow$ (2) we proceed by induction on m. For $m = 1$, the claim is clear, since $T = \lambda_1 \operatorname{Id}_V$. Assume the result holds for all $k < m$. Let $U := \operatorname{range}(T - \lambda_m \operatorname{Id}_V)$, it is clear that U is invariant under T. Consider $T|_U$, by Exercise 34 and Lemma 33 we see that $\prod_{i=1}^{m-1}(x - \lambda_i)$ is a multiple of $p_{T|_U}$.

By the inductive hypothesis, there exists $B_U = \{u_1, \dots, u_k\}$ a basis of U, such that $\mathcal{M}(B_U, T|_U)$ is upper-triangular. Let $B_V = \{u_1, \dots, u_k, v_1, \dots, v_l\}$ be an extension of B_V to a basis of V. For each $i \in \{1, \dots l\}$ we have:

$$T(v_i) = (T - \lambda_m)v_i + \lambda_m v_i \in \operatorname{Span}\{u_1, \dots, u_k, v_i\} \subset \operatorname{Span}\{u_1, \dots, u_k, v_1, \dots, v_i\}.$$

Thus, Lemma 35 implies that $\mathcal{M}(B_V, T)$ is upper-triangular. $\square$

Corollary 15. *Let $T : V \to V$ be an operator on a finite-dimensional vector space over $\mathbb{C}$, then there exist a basis B_V of V, such that $\mathcal{M}(B_V, T)$ is upper-triangular.*

Proof. Left to the reader as an exercise. $\square$

Warning 3. Notice that in $B_V = \{v_1, \dots, v_n\}$ the basis given in Corollary 15, a priori, only the vector v_1 is an eigenvector with eigenvalue λ_1.

From Math2101, you learned how to transform a matrix into row echelon form (recall it here). The row echelon form of a matrix has *no relation* with a triangular form whose existence Corollary 15 guarantes. For instance, one can not read off the eigenvalues of a matrix from its row echelon form. However, in contrast with Gauss elimination, there is no method to compute the upper-triangular form from Corollary 15.

12. Mar. 11, 2024

12.1. **Diagonalizable Operators.** We start this session with a definiton.

Definition 30. Given a linear operator $T : V \to V$ and $\lambda \in \mathbb{F}$. The *eigenspace* corresponding to λ is

$$E(\lambda, T) := \operatorname{null}(T - \lambda \operatorname{Id}_V),$$

i.e. $E(\lambda, T) = \operatorname{Span}\{v_1, \dots, v_n\}$, where $v_1, \dots, v_n$ are all the eigenvectors of λ.

Lemma 37. *Let $T \in \mathcal{L}(V)$ and $\lambda_1, \dots, \lambda_m$ be distinct eigenvalues of T, then*

$$E(\lambda_1, T) + \cdots + E(\lambda_m, T) \subseteq V \tag{9}$$

is a direct sum.

Proof. This follows directly from Lemma 31. Make sure you understand why! $\square$

Proposition 7. *Assume that V is finite-dimensional. Let $T \in \mathcal{L}(V)$ and $\lambda_1, \dots, \lambda_m \in \mathbb{F}$ be a list of distinct eigenvalues. Then the following are equivalent:*

(1) there exist a basis B_V such that $\mathcal{M}(B_V, T)$ is a diagonal matrix;

(2) there exist a basis B_V consisting of eigenvectors of T;

(3) $V = E(\lambda_1, T) \oplus \cdots \oplus E(\lambda_m, T)$;

(4) $\dim V = \dim E(\lambda_1, T) + \cdots + \dim E(\lambda_m, T)$.

Proof. (1) $\Leftrightarrow$ (2) is clear.

(2) $\Rightarrow$ (3) let $\{\lambda_1, \ldots, \lambda_m\}$ be the set of eigenvalues of T and define $B_{V_i} := \{v \in B_V \mid T(v) = \lambda_i v\}$ for $1 \leq i \leq m$. It is clear that $E(\lambda_i, T) = \mathrm{Span}\,(B_{V_i})$. Moreover, since $B_V = \sqcup_{1\leq i\leq m} B_{V_i}$ is a disjoint union we obtain $V = \mathrm{Span}\, B_V = \oplus_{1\leq i\leq m}\mathrm{Span}\, B_{V_i}$.

(3) $\Rightarrow$ (4) follows from Exercise 20.

(4) $\Rightarrow$ (3) let B_{V_i} be a basis of $E(\lambda_i, T)$ for each $i \in \{1, \ldots, m\}$. We claim that $\{u_1, \ldots, u_n\} = B_V = \cup_{1\leq i\leq m} B_i$ is a basis of V. Indeed, since $|B_V| = \dim V$ it is enough to check that B_V is linearly independent. Assume that there exist $a_i \in \mathbb{F}$ such that

$$\sum_{i=1}^{n} a_i u_i = 0.$$

We can arrange such that the sum becomes:

$$\sum_{j=1}^{m} v_j = 0,$$

where $v_j \in E(\lambda_j, T)$. By Lemma 31 the set $\{v_1, \ldots, v_m\}$ is linearly independent, since they have different eigenvalues. Thus each $v_j = 0$, as each $v_j = \sum_{i\in I_j} a_i u_i$ with $\{u_i\}_{I_j}$ a basis of $E(\lambda_j, T)$ we have that each collection $\{a_i\}_{i\in I_j}$ consists of only zeros, which gives that $\{a_i\}_{1\leq i\leq n}$ all vanish.

(3) $\Rightarrow$ (2) for each $1 \leq i \leq m$ let $B_i := \{v^i_1, \ldots, v^i_{k_i}\}$ be a basis of $E(\lambda_i, T)$. We claim that $B_V = \cup_{1\leq i\leq m} B_i$ is a basis of V (see Exercise 20). $\square$

Definition 31. We say that an operator $T : V \to V$ is diagonalizable if it satisfies one of the equivalent conditions of Proposition 7.

Remark 26. Definition 31 gives a name to a collection of equivalent mathematical properties. Many definitions in Mathematics work the same way. Something is made into a definition to summarize many important facts about a mathematical object. An interesting example to read about is the definition derivative of a function. Notice that it is easier to envision how to generalize some of the conditions in Proposition 7 to infinite-dimensional vector space, but not others.

Corollary 16. *Let $T : V \to V$ be an operator on a finite-dimensional vector space. Assume that T has $\dim V$ distinct eigenvalues, then T is diagonalizable.*

Proof. Let $\{v_1, \ldots, v_n\}$ be a collection of eigenvectors for $\{\lambda_1, \ldots, \lambda_n\}$. By Lemma 31 we see that $\{v_1, \ldots, v_n\}$ is linearly independent. Also $\mathrm{Span}\,(v_i) \subseteq E(\lambda_i, T)$ for each i. Thus, by dimension reasons we obtain $\mathrm{Span}\,(v_i) = E(\lambda_i, T)$, which implies that $V = \oplus_{i=1}^m E(\lambda_i, T)$. That is T is diagonalizable by Proposition 7. $\square$

Exercise 35. Give an example of a diagonalizable operator $T : V \to V$ such that T does not have $\dim V$ distinct eigenvalues.

We record the following important relation between diagonalizable and the minimal polynomial.

Lemma 38. *Let $T : V \to V$ be a operator on a finite-dimensional vector space. Then T is diagonalizable if and only if the minimal polynomial $p_T = \prod_{i=1}^m (x - \lambda_i)$ where $\{\lambda_1, \ldots, \lambda_m\}$ are distinct eigenvalues.*

Proof. Left as an exercise in HW 3. $\square$

12.2. **Inner Products and Norms.** The following is a quick review of concepts discussed in Math2101. From now on $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$.

Definition 32. Let V be a vector space. An *inner product* is a function $\langle -, - \rangle : V \times V \to \mathbb{F}$ satisfying:

- (positivity) $\langle v, v\rangle \in \mathbb{R}_{>0}$ for every $v \in V$;
- (definiteness) $\langle v, v\rangle = 0$ if and only if $v = 0$;
- (additivity on the first slot) $\langle v_1 + v_2, u\rangle = \langle v_1, u\rangle + \langle v_2, u\rangle$ for every $v_1, v_2, u \in V$;
- (additivity on the first slot) $\langle av, u\rangle = a\,\langle v, u\rangle$ for every $v, u \in V$ and $a \in \mathbb{F}$;
- (conjugate symmetry) $\langle v, u\rangle = \overline{\langle u, v\rangle}$.

Notice that for $\mathbb{F} = \mathbb{R}$ the last equation becomes $\langle v, u\rangle = \langle u, v\rangle$. An inner produced space is the data of a pair $(V, \langle -, -\rangle_V)$, where V is a vector space over $\mathbb{F}$ and $\langle -, -\rangle_V : V \times V \to \mathbb{F}$ is an inner product.

Example 23. (i) Let $V = \mathbb{R}^n$, then $(-) \cdot (-) : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ given by

$$(x_1, \ldots, x_n) \cdot (y_1, \ldots, y_n) := \sum_{i=1}^{n} x_i y_i$$

is an inner product, sometimes refered to as the dot product.

(ii) Let $V = \mathbb{C}^n$, then $(-) \cdot (-) : \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}$ given by

$$(x_1, \ldots, x_n) \cdot (y_1, \ldots, y_n) := \sum_{i=1}^{n} x_i \overline{y_i},$$

is an inner product, also refered to as the dot product.

(iii) More generally, given $c_1, \ldots, c_n \in \mathbb{R}_{>0}$ a collection of positive numbers, then

$$\langle (x_1, \ldots, x_n), (y_1, \ldots, y_n)\rangle = \sum_{i=1}^{n} c_i x_i \overline{y_i}$$

defines an inner product on bR^n or $\mathbb{C}^n$.

(iv) $\langle -, -\rangle : \mathcal{P}(\mathbb{R}) \times \mathcal{P}(\mathbb{R}) \to \mathbb{R}$ given by

$$\langle p, q\rangle := p(0)q(0) + \int_{-1}^{1} p'q'.$$

We list some basic consequences of the definition.

Lemma 39. *Let V be an inner product space.*

(i) $\langle -, v\rangle \in V^*$ for every $v \in V$;

(ii) for $\mathbb{F} = \mathbb{R}$ we also have $\langle v, -\rangle \in V^*$ for every $v \in V$.

Proof. Left to the reader. □

Exercise 36. Convince yourself that (ii) in Lemma 39 does not hold for $\mathbb{F} = \mathbb{C}$.

Definition 33. A *norm* on a vector space is a function $\|-\| : V \to \mathbb{R}$ satisfying:

(i) $\|v\| \geq 0$ for every $v \in V$;

(ii) $\|v\| = 0$ if and only if $v = 0$;

(iii) $\|av\| = |a| \, \|v\|$ for every $v \in V$ and $a \in \mathbb{F}$;

(iv) (triangle inequality) for every $u, v \in V$ we have $\|u + v\| \leq \|u\| + \|v\|$.

A *normed vector space* is the data of a pair $(V, \|-\|)$ where V is a vector space and $\|-\|$ is a norm on V.

Example 24. (i) Given a vector space V with an inner product $\langle -, -\rangle : V \times V \to \mathbb{F}$. Then $\|v\| := \langle v, v\rangle^{1/2}$ is a norm on V.

(ii) Consider $V = \mathbb{R}^n$ and let $p \in \mathbb{Z}_{\geq 1}$ be an integer, then:

$$\|(x_1, \ldots, x_n)\| := (|x_1|^p + \cdots + |x_n|^p)^{1/p}$$

defines a norm.

The following are immediate properties:

Lemma 40. *Let $(V, \|-\|)$ be a normed vector space, then* *(i)* $v = 0$ *if and only if* $\|v\| = 0$*;*

(ii) $\|av\| = |a| \, \|v\|$ *for every* $a \in \mathbb{F}$ *and* $v \in V$*.*

Proof. Exercise for the reader. □

Definition 34. A pair of vectors $u, v \in V$ in an inner product space are said to be *orthogonal* if $\langle u, v \rangle = 0$.

Here are a couple of basic results relating to orthogonal vectors that we leave as an exercise as well.

Lemma 41. *Let* V *be an inner product space.*

(i) *let* $v \in V$ *such that* $\langle v, u \rangle = 0$ *for every* $u \in V$*, then* $v = 0$*;*

(ii) *(Pythagorean theorem)* $\|u + v\|^2 = \|u\|^2 + \|v\|^2$ *for every* $u, v \in V$*;*

(iii) *For every* $u, v \in V$ *such that* $v \neq 0$ *we have* $u = cv + w$*, where*

$$c = \frac{\langle u, v \rangle}{\|v\|^2}, \; w = u - \frac{\langle u, v \rangle}{\|v\|^2} v \text{ and } \langle w, v \rangle = 0,$$

i.e. u *can be written as a scalar multiple of* v *plus a scalar multiple of a vector* w*, which is orthogonal to* v*.*

The following is a rather important inequality:

Lemma 42 (Cauchy–Schwartz)**.** *For any* $u, v \in V$ *in an inner product space, we have:*

$$| \langle u, v \rangle | \leq \|u\| \, \|v\| .$$

Moreover, the above is an equality if and only if $v = au$ *for some* $a \in \mathbb{F}$*.*

The Cauchy–Schwartz inequality has the following two consequences which have a geometrical interpretation:

Corollary 17. *Let* $u, v \in V$ *be two vectors in an inner product space, then*

(i) *(triangle inequality)* $\|u + v\| \leq \|u\| + \|v\|$*;*

(ii) *(parallelogram equality)* $\|u + v\| + \|u - v\| = 2(\|u\|^2 + \|v\|^2)$*.*

Proof. Exercise! □

Example 24 poses the question if we can go back, that is, given $(V, \|-\|)$ a normed vector space, is there an inner product $\langle -, - \rangle : V \times V \to \mathbb{F}$ such that $\|v\| = \langle v, v \rangle$ for every $v \in V$?

Exercise 37. Given a normed vector space prove that there an inner product $\langle -, - \rangle : V \times V \to \mathbb{F}$ such that $\|v\| = \langle v, v \rangle$ for every $v \in V$ if and only if $\|-\|$ satisfies the parallelogram equality (see Corollary 17 (ii)).

12.3. **Orthonormal bases.**

Definition 35. Given V an inner product space, a subset of vectors $S \subset V$ is called *orthonormal* if for every $u, v \in S$ we have:

$$\langle u, v \rangle = \begin{cases} 1 & \text{if } u = v \\ 0 & \text{else} \end{cases} .$$

Exercise 38. (i) The standard basis of $\mathbb{C}^n$.

(ii) Let $V = \mathcal{C}([-\pi, \pi])$ denote the vector space of continuous real-valued functions with inner product given by $\langle f, g \rangle := \int_{-\pi}^{\pi} f(x) g(x) dx$. Then the set

$$\left\{ \frac{1}{\sqrt{2\pi}}, \frac{\cos x}{\sqrt{2\pi}}, \frac{\cos 2x}{\sqrt{2\pi}}, \dots, \frac{\cos nx}{\sqrt{2\pi}}, \frac{\sin x}{\sqrt{2\pi}}, \dots, \frac{\sin nx}{\sqrt{2\pi}} \right\}$$

is an orthonormal list of vectors. Check this!

Here are a couple of consequences of having a list of orthonormal vectors.

Lemma 43. *Let $S = \{e_1, \dots, e_n\} \subset V$ be an orthonormal set. Then*

(i) $\|a_1e_1 + \cdots + a_ne_n\|^2 = \sum_{i=1}^n |a_i|^2$;

(ii) *S is linearly independent;*

(iii) *(Bessel's inequality) for every $v \in V$ we have:*

$$\sum_{i=1}^n |\langle v, e_i\rangle|^2 \leq \|v\|^2 .$$

Proof. For (i) we proceed by induction. The case $n = 1$ is $\|a_1e_1\|^2 = |a_1|^2 \|e_1\|^2 = |a_1|^2$. Now assume the result holds for $n-1$, Lemma 41 (ii) gives:

$$\|a_1e_1 + \cdots + a_ne_n\|^2 = \|a_1e_1 + \cdots + a_{n-1}e_{n-1}\|^2 + \|a_ne_n\|^2 = \sum_{i=1}^{n-1} |a_i|^2 + |a_n|^2 = \sum_{i=1}^{n} |a_i|^2.$$

For (ii) assume that there are $a_1, \dots, a_n \in \mathbb{F}$ such that

$$a_1e_1 + \cdots + a_ne_n = 0.$$

Then we have $\|a_1e_1 + \cdots + a_ne_n\| = 0$, which implies $\sum_{i=1}^n |a_i|^2 = 0$ by (i). Since all $|a_i|^2 \geq 0$, we obtain that $a_1 = \cdots = a_n = 0$.

For (iii), let $u := \sum_{i=1}^n \langle v, e_i\rangle e_i$ and $w = v - u$. Notice that $\langle w, e_i\rangle = 0$ for all $e_i \in S$. Thus $\langle u, w\rangle = 0$. Now we compute:

$$\|v\|^2 = \|u\|^2 + \|w\|^2 \geq \|u\|^2 = \sum_{i=1}^n |\langle v, e_i\rangle|^2,$$

where the first equality is by Lemma 41 (ii) and the last follows from (i). □

Definition 36. Given V an inner product space an *orthonormal basis* is a basis S which is also an orthonormal set.

Remark 27. Notice that by Lemma 43 (ii) if one is given an orthonormal subset $S \subset V$ such that $|S| = \dim V$, then S is an orthonormal basis.

Example 25. The set $\{(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}), (\frac{1}{2}, \frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}), (\frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}, \frac{1}{2}), (-\frac{1}{2}, \frac{1}{2}, -\frac{1}{2}, \frac{1}{2})\} \subset \mathbb{F}^4$ is an orthonormal basis of $\mathbb{F}^4$.

Here is a nice consequence of using an orthonormal basis for an inner product space.

Lemma 44. *Let $\{e_1, \dots, e_n\} \subset V$ be an orthonormal basis of an inner product space V.*

(i) $v = \sum_{i=1}^n \langle v, e_i\rangle e_i$;

(ii) *(Parseval's identity)* $\|v\| = \sum_{i=1}^n |\langle v, e_i\rangle|^2$;

(iii) $\langle u, v\rangle = \sum_{i=1}^n \langle u, e_i\rangle \overline{\langle v, e_i\rangle}$.[13]

Proof. Left to the reader. □

The natural question to ask now is if given a basis B of an inner product space V can we modify B into an orthonormal basis. The following result says this is always possible in finite dimension.

Theorem 3 (Gram–Schmidt procedure)**.** *Let $\{v_1, \dots, v_n\}$ be a basis of an inner product space V. Inductively, define $e_1 := \frac{v_1}{\|v_1\|}$ and*

$$f_n := v_n - \sum_{i=1}^{n-1} \frac{\langle v_i, f_i\rangle}{\|f_i\|^2} f_i \quad \textit{and} \quad e_n := \frac{f_n}{\|f_n\|}.$$

Then for every $1 \leq k \leq n$:

$$\{e_1, \dots, e_k\} \textit{ is an orthonormal set and } \operatorname{Span}(e_1, \dots, e_k) = \operatorname{Span}(v_1, \dots, v_k). \tag{10}$$

In particular, $\{e_1, \dots, e_n\} \subset V$ is an orthonormal basis.

[13] In fact, I like to remember this formula as: $\langle u, v\rangle = \sum_{i=1}^n \langle u, e_i\rangle \langle e_i, v\rangle$.

Proof. We proceed by induction on (10). The case $k = 1$ is clear, since $\|e_1\| = 1$ and $v_1 = \|v_1\| e_1$.

Assume the result holds for every $1 \leq \ell \leq k-1$. Since $\{v_1, \ldots, v_k\}$ is linearly independent we have that $v_k \notin \mathrm{Span}\,(v_1, \ldots, v_{k-1}) = \mathrm{Span}\,(e_1, \ldots, e_{k-1}) = \mathrm{Span}\,(f_1, \ldots, f_{k-1})$. Thus, $f_k \neq 0$, so e_k is well-defined. It is clear that $\|e_k\| = 1$.

We check that $\langle e_k, e_j \rangle = 0$ for every $1 \leq j \leq k-1$.

$$\begin{aligned}
\langle e_k, e_j \rangle &= \frac{1}{\|f_k\| \, \|f_j\|} \langle f_k, f_j \rangle \\
&= \frac{1}{\|f_k\| \, \|f_j\|} \left(v_k - \sum_{i=1}^{k-1} \frac{\langle v_k, f_i \rangle}{\|f_i\|^2} f_i, f_j \right) \\
&= \frac{1}{\|f_k\| \, \|f_j\|} \left(\langle v_k, f_j \rangle - \sum_{i=1}^{k-1} \frac{\langle v_k, f_i \rangle}{\|f_i\|^2} \langle f_i, f_j \rangle \right) \\
&= \frac{1}{\|f_k\| \, \|f_j\|} \left(\langle v_k, f_j \rangle - \langle v_k, f_j \rangle \right) \\
&= 0.
\end{aligned}$$

This gives that $\{e_1, \ldots, e_k\}$ is orthonormal. Since by definition we have $v_k \in \mathrm{Span}\,(e_1, \ldots, e_k)$ by the inductive hypothesis we have:

$$\mathrm{Span}\,(v_1, \ldots, v_k) \subseteq \mathrm{Span}\,(e_1, \ldots, e_k).$$

Since both lists are linearly independent and have the same number of elements we have that their span are the same by Lemma 7 (4). This finishes the proof. □

Exercise 39. Work out an orthonormal basis for $\mathcal{P}_2(\mathbb{R})$.

These are direct consequence of Theorem 3.

Corollary 18. *(i) Every finite-dimensional inner product space V has an orthonormal basis.*

(ii) Every orthonormal subset $S \subset V$ extends to an orthonormal basis.

Proof. For (i) by Corollary 2 (1) a basis exists, then applying Theorem 3 we get the result.

For (ii), we know that S is linearly independent by 43 (ii), then Corollary 2 (2) allows us to extend this set to a basis, which then applying Theorem 3 can be made into an orthonormal basis. □

We also have a consequence of the Gram–Schmidt procedure to the question of whether there exist a basis such that a linear operator is upper-triangular with respect to such a basis.

Lemma 45. *Let V be a finite-dimensional inner product space and consider $T \in \mathcal{L}(V)$. The following are equivalent:*

(1) there exists a basis B_V of V such that $\mathcal{M}(T, B_V)$ is upper-triangular;

(2) there exists an orthonormal basis B'_V of V such that $\mathcal{M}(T, B'_V)$ is upper-triangular.

Proof. We only need to prove (1) $\Rightarrow$ (2). Let $B_V = \{v_1, \ldots, v_n\}$ be a basis for which T is upper-triangular. Notice that because of Lemma 35 it is enough to notice that applying Theorem 3 we have obtain a set $\{e_1, \ldots, e_n\}$ such that:

$$\begin{aligned}
T(\mathrm{Span}\,(e_1, \ldots, e_k)) &= T(\mathrm{Span}\,(v_1, \ldots, v_k)) \\
&\subseteq \mathrm{Span}\,(v_1, \ldots, v_k) \\
&= \mathrm{Span}\,(e_1, \ldots, e_k).
\end{aligned}$$

Thus, $\mathcal{M}(T, \{e_1, \ldots, e_n\})$ is upper-triangular. □

Corollary 19 (Schur's theorem)**.** *Let $T : V \to V$ be an operator on a complex inner product space. Then there exists an orthonormal basis B_V such that $\mathcal{M}(T, B_V)$ is upper-triangular.*

Proof. By Corollary 15 there exist a basis B'_V such that $\mathcal{M}(T, B'_V)$ is upper-triangular. By Lemma 45 we can find B_V as claimed. □

We end this subsection discussing the so-called Riesz representation theorem.

Proposition 8 (Riesz representation). *Let V be a finite-dimensional inner product space and $\lambda \in V^*$. Then there exists an unique $v \in V$, such that $\lambda = \langle -, v\rangle$.*

Proof. The proof comes from trying to find $v \in V$ such that the equation $\lambda(u) = \langle u, v\rangle$ holds for every u. Let $\{e_1, \dots, e_n\}$ be an orthonormal basis of V, we calculate:

$$\begin{aligned}\lambda(u) &= \lambda(\langle u, e_1\rangle e_1 + \cdots + \langle u, e_n\rangle e_n) \\ &= \langle u, e_1\rangle \lambda(e_1) + \cdot + \langle u, e_n\rangle \lambda(e_n) \\ &= \left\langle u, \overline{\lambda(e_1)}e_1 + \cdots + \overline{\lambda(e_n)}e_n \right\rangle.\end{aligned}$$

Thus, if we pose $v := \overline{\lambda(e_1)}e_1 + \cdots + \overline{\lambda(e_n)}e_n$, we obtain the desired equality.

Now assume that there exists $v' \in V$ a different vector such that $\langle u, v\rangle = \langle u, v'\rangle$ for every u. This implies that:

$$\langle u, v - v'\rangle = 0 \text{ for every } u \in V.$$

In particular, we obtain $\|v - v'\| = 0$, which gives that $v = v'$. □

13. Mar. 14, 2024

13.1. **Orthogonal Complements and Minimization Problems.** We begin this subsection with a definition.

Definition 37. Let $U \subseteq V$ be a subset of an inner product space V. The *orthogonal complement of U* is:

$$U^\perp := \{v \in V \mid \langle u, v\rangle = 0 \text{ for every } u \in U\}.$$

Here are a list of direct properties.

Lemma 46. *Given a subset $U \subseteq V$ we have:*

(i) $U^\perp$ is a subspace of V;

(ii) $\{0\}^\perp = V$;

(iii) $V^\perp = \{0\}$;

(iv) if U is a subspace, then $U \cap U^\perp = \{0\}$;

(v) if $W \subseteq U$ is a subset, then $U^\perp \subseteq W^\perp$.

Proof. We prove (i). Let $v_1, v_2 \in U^\perp$ and $a \in \mathbb{F}$, then by Lemma 39 we have:

$$\langle u, v_1 + av_2\rangle = \langle u, v_1\rangle + \overline{a}\langle u, v_2\rangle = 0 + \overline{a} \cdot 0 = 0,$$

that is $v_1 + av_2 \in U^\perp$.

For (ii), notice that for any $v \in V$ we have $\langle 0, v\rangle = \overline{\langle v, 0\rangle} = \overline{\langle v, u - u\rangle} = \overline{\langle v, u\rangle} - \overline{\langle v, u\rangle} = 0$.

For (iii) notice that $v \in V^\perp$ implies that $\langle v, v\rangle = 0$, which by (ii) from Definition 32 gives that $v = 0$.

For (iv), again notice that $u \in U \cap U^\perp$ implies $\langle u, u\rangle = 0$, which gives $u = 0$.

For (v), let $v \in U^\perp$ and consider $\langle v, w\rangle$. Since for every $w \in W \subseteq U$, this implies that $\langle v, w\rangle = 0$, we have $v \in W^\perp$. □

Example 26. (i) Let $V = \mathbb{R}^3$ and consider $U = \{(2, 3, 5)\}$, then

$$U^\perp = \{(x, y, z) \in \mathbb{R}^3 \mid 2x + 3y + 5z = 0\}.$$

(ii) Let $U = \operatorname{Span}\{(2, 3, 5)\}$, then $U^\perp = \{(x, y, z) \in \mathbb{R}^3 \mid 2x + 3y + 5z = 0\}$, as well.

(iii) Let $e_1, \dots, e_m, f_1, \dots, f_n$ be a orthonormal basis of V. Then

$$\operatorname{Span}(e_1, \dots, e_m)^\perp = \operatorname{Span}(f_1, \dots, f_n).$$

Corollary 20. *Given $U \subseteq V$ a subspace of a finite-dimensional inner product space, then $U \oplus U^\perp = V$.*

Proof. By Lemma 46 (iv) it is enough to prove that $U + U^\perp = V$. Let $v \in V$ if $v \in U^\perp$ then $v \in U + U^\perp$. Assume that $v \notin U^\perp$ and let $u \in U$ such that $\langle u, v\rangle \neq 0$. In particular, we have $u \neq 0$, thus Lemma 41 (iii) implies that we can write $v = c^{-1} \cdot (u - w)$ where $c = \frac{\langle u,v\rangle}{\|v\|^2} \neq 0$ and $\langle u, w\rangle = 0$, i.e. $w \in U^\perp$. This finishes the proof. ☐

We leave the next two statements as an exercise:

Lemma 47. *Let $U \subseteq V$ be a subspace of a finite-dimensional inner product space V. Then*

(i) $\dim U^\perp = \dim V - \dim U$*;*

(ii) $(U^\perp)^\perp = U$*.*

We end this section by giving another proof of Proposition 8.

Proposition 9. *Given V a finite-dimensional inner product space consider the map $\Phi : V \to V^*$ given by:*

$$\Phi(v)(u) := \langle u, v\rangle .$$

Then Φ is a bijection.

Proof. First we prove that Φ is surjective. Let $\varphi \in V^*$ be a linear functional on V. If $\varphi = 0$, then $\Phi(0) = \varphi$. So we assume that $\varphi \neq 0$. Notice that $\operatorname{null}\varphi \subset V$ is a subspace of dimension $\dim V - 1$. Indeed, by the fundamental theorem of linear algebra we have that $\dim V = \dim \operatorname{null}\varphi + \dim \operatorname{range}\varphi$ and $\operatorname{range}\varphi \neq 0$, so it is easy to see that $\operatorname{range}\varphi = \mathbb{F}$. Then Corollary 20 implies that $\dim(\operatorname{null}\varphi)^\perp = 1$. Let $0 \neq v_0 \in V$ be a basis of $(\operatorname{null}\varphi)^\perp$. We want to find v a scalar multiple of v_0 such that $\Phi(v)(u) = \varphi(u)$ for every $u \in V$. Let $v = \alpha v_0$ for some $\alpha \in \mathbb{F}$ and applying Lemma 41 (iii) we have that $u = cv_0 + w$ for some $c \in \mathbb{F}$ and $w \in W$ such that $\langle w, v_0\rangle = 0$. Then we have

$$\Phi(v)(u) = \langle u, v\rangle = \langle cv_0 + w, \alpha v_0\rangle = c\overline{\alpha} \left\|v_0\right\|^2 = \varphi(u) = \varphi(cv_0) = c\varphi(v_0),$$

since by definition of v_0, $w \in \operatorname{null}\varphi$. Thus, we notice that taking $\alpha = \overline{\frac{\varphi(v_0)}{\|v_0\|^2}}$ gives v with the desired property.

We leave it to the reader to check that Φ is injective. ☐

Warning 4. In Proposition 9 the case where V is a vector space over the real numbers, then Φ is an isomorphism, i.e. it is also linear. In particular, the proof of Proposition 9 becomes easier since we only need to check the map is injective. However, over the complex numbers Φ is *anti-linear*, i.e. $\Phi(av) = \overline{a}\Phi(v)$.

Definition 38. Let $U \subseteq V$ be a subspace of a finite-dimensional inner product space. The *orthogonal projection* $P_U : V \to V$ is given by

$$P_U(v) := u,$$

where $v = u + w$ is the unique decomposition of v in $U \oplus U^\perp$.

The following is a list of properties of the orthogonal projection:

Lemma 48. *Let $U \subseteq V$ be a subspace of a finite-dimensional inner product space. Then*

(i) P_U *is linear;*

(ii) $P_U(u) = u$ *for every* $u \in U$*;*

(iii) $P_U(w) = 0$ *for every* $w \in U^\perp$*;*

(iv) $\operatorname{range} P_U = U$*;*

(v) $\operatorname{null} P_U = U^\perp$*;*

(vi) $v - P_u(v) \in U^\perp$ *for every* $v \in V$*;*

(vii) $P_U^2 = P_U$*;*

(viii) $\|P_U(v)\| \leq \|v\|$ *for every* $v \in V$*;*

(ix) given $\{e_1, \dots, e_m\} \subset V$ *an orthonormal basis of* U*, we have:*

$$P_U(v) = \sum_{i=1}^{m} \langle v, e_i\rangle\, e_i.$$

Proof. Left as an exercise. ☐

13.2. **Self-Adjoint and Normal Operators.** **TODO:** Change all the dual notation from before here from V^* to $V^\vee$ and from T^* to $T^\vee$. Similarly for dual basis and all other related notions and constructions.

Warning 5. Change of notation. I will from this section on change notation so that we free the use of $(-)^*$ for adjoint operator. I will revisit the notation for dual space and morphisms to appropriately clear the ambiguity that is now being introduced.

Before introducing the definitions that follow it is interesting to try to motivate them. We saw in Lemma 45 that if an operator $T : V \to V$ is upper-triangular, then we can find an *orthonormal* basis B_V of V such that $\mathcal{M}(T, B_V)$ is upper-triangular. We can ask if the same happens for other conditions on T. For instance,

Question 1. Given $T : V \to V$ a diagonalizable operator with respect to some basis B_V, can we find an orthonormal basis B'_V such that $\mathcal{M}(T, B'_V)$ is a diagonal matrix?

The answer to the above question is given by the spectral theorem, which has a slightly different version over $\mathbb{R}$ and over $\mathbb{C}$. To formulate this theorem we need to introduce the following two concepts.

Notice that given $T : V \to W$ an operator between two inner product spaces, for every $w \in W$ we obtain a functional on V given by

$$\varphi_w : V \to \mathbb{F} \qquad \varphi_w(v) := \langle Tv, w\rangle_W .$$

By the Riesz representability theorem, there exist an unique $v' \in V$ such that $\langle v, v'\rangle_V = \varphi_w(v) = \langle Tv, w\rangle_W$. Let $T^*(w) := v'$, we claim that $T^* : W \to V$ is linear. Indeed, let $w_1, w_2 \in W$ and $a \in \mathbb{F}$ then we have that (Make sure you understand each step here!)

$$\begin{aligned}
\langle v, T^*(w_1 + aw_2)\rangle_V &= \langle Tv, w_1 + aw_2\rangle_W \\
&= \langle Tv, w_1\rangle_W + \overline{a}\,\langle Tv, w_2\rangle_W \\
&= \langle v, T^*w_1\rangle_V + \langle v, aT^*w_2\rangle_V \\
&= \langle v, T^*(w_1) + aT^*(w_2)\rangle_V
\end{aligned}$$

holds for every $v \in V$. By the Riesz representability theorem we obtain that $T^*(w_1 + aw_2) = T^*(w_1) + aT^*(w_2)$.

Definition 39. Let $T : V \to W$ be an operator between two inner product spaces. The *adjoint of* T is the unique linear map $T^* : W \to V$ such that

$$\langle Tv, w\rangle_V = \langle v, T^*w\rangle_W$$

for all $v \in V$ and $w \in W$.

Notation 6. From now one we will drop the subscript of V or W from the notaiton of the inner product.

Example 27. (i) Let $T : \mathbb{C}^3 \to \mathbb{C}^2$ be given by $T(x_1, x_2, x_3) = (2x_1, 3x_2 + ix_3)$. We calculate the adjoint of T:

$$\begin{aligned}
\langle T(x_1, x_2, x_3), (y_1, y_2)\rangle &= \langle (2x_1, 3x_2 + ix_3), (y_1, y_2)\rangle \\
&= 2x_1y_1 + 3x_2y_2 + ix_3y_2 \\
&= \langle (x_1, x_2, x_3), (2y_1, 3y_2, -iy_2)\rangle .
\end{aligned}$$

Thus, $T^* : \mathbb{C}^2 \to \mathbb{C}^3$ is given by $T^*(y_1, y_2) = (2y_1, 3y_2, -iy_2)$.

(ii) Let W be of dimension 1. Fix $u \in V$ and $x \in W$ define $T : V \to W$ by $T(v) := \langle v, u\rangle\, x$. We compute T^*:

$$\begin{aligned}
\langle Tv, w\rangle &= \langle \langle v, u\rangle\, x, w\rangle \\
&= \langle v, u\rangle\, \langle x, w\rangle \\
&= \left\langle v, \overline{\langle x, w\rangle} u \right\rangle \\
&= \langle v, \langle w, x\rangle\, u\rangle
\end{aligned}$$

Thus, $T^* : W \to V$ is given by $T(w) = \langle w, x\rangle\, u$.

Here is a list of properties of passing to the adjoint operator:

Lemma 49. *Let $T, S : V \to W$ be two operators between two finite-dimensional inner product spaces. Then:*

(i) $(S+T)^* = S^* + T^*$;

(ii) $(\lambda T)^* = \overline{\lambda} T^*$, *for every* $\lambda \in \mathbb{F}$;

(iii) $(T^*)^* = T$;

(iv) $(ST)^* = T^* S^*$;

(v) $(\mathrm{Id}_V)^* = \mathrm{Id}_V$;

(vi) T *is invertible if and only if* T^* *is invertible, in which case* $(T^*)^{-1} = (T^{-1})^*$.

Proof. Exercise. □

We also have properties similar to Lemma 50 describing the relation between the range and null spaces of the adjoint of an operator in terms of the range and null of the original operator.

Lemma 50. *Let $T, S : V \to W$ be two operators between two finite-dimensional inner product spaces. Then:*

(i) $(\operatorname{null} T^*) = (\operatorname{range} T)^\perp$;

(ii) $(\operatorname{range} T^*) = (\operatorname{null} T)^\perp$;

(iii) $(\operatorname{null} T) = (\operatorname{range} T^*)^\perp$;

(iv) $(\operatorname{range} T) = (\operatorname{null} T^*)^\perp$.

Proof. Write this! □

It is also very useful to have the following result which relates the adjoint of an operator T to the matrix representing T in an *orthonormal* basis.

Lemma 51. *Let B_V be an orthonormal basis of V and $T : V \to V$ an operator. Then*

$$\mathcal{M}(T, B_V) = \mathcal{M}(T^*, B_V)^\dagger,$$

where $\mathcal{M}(T^*, B_V)^\dagger$ is the conjugate transpose of the matrix $\mathcal{M}(T, B_V)$.

Proof. To do. □

Warning 6. In Lemma 51 it is necessary that B_V is an orthonormal basis. See Exercise 4 in Worksheet 7.

Finally we can pose the following definition:

Definition 40. Let $T : V \to V$ be an operator on an inner product space.

- we say that T is *self-adjoint* if $T = T^*$;
- we say that T is *normal* if it commutes with T^*, i.e. $TT^* = T^*T$.

Example 28. (i) Consider $T(x, y) = (3x + y, x + 3y)$ this operator is self-adjoint over $\mathbb{R}$.

(ii) Consider $T(x, y) = (2x + 2iy, -2ix + 2y)$. This operator is self-adjoint over $\mathbb{C}$.

(iii) Consider $T(x, y) = (2x - 3y, 3x + 2y)$. This operator is normal but not self-adjoint.

Here is a characterization of self-adjoint operators.

Lemma 52. *Let $T : V \to V$ be an operator on a finite-dimensional complex inner space. The following are equivalent:*

(1) T *is self-adjoint;*

(2) $\langle Tv, v\rangle \in \mathbb{R}$ *for every* $v \in V$.

Proof. Assume (1), then $T - T^* = 0$ and Remark 28 gives that

$$\langle (T - T^*)v, v\rangle = 0 \text{ for every } v \in V. \tag{11}$$

Since $\langle T^*v, v\rangle = \overline{\langle v, T^*v\rangle} = \overline{\langle Tv, v\rangle}$, we have

$$\langle Tv, v\rangle - \overline{\langle Tv, v\rangle} = 0 \text{ for every } v \in V. \tag{12}$$

Assume (2), then (12) holds, which implies (11) that by Remark 28 gives that $T = T^*$. □

We collect some properties of self-adjoint operators.

Lemma 53. *Let $T : V \to V$ be a self-adjoint operator, then:*

(i) all the eigenvalues of T are real;

(ii) $\langle Tv, v\rangle = 0$ for every $v \in V$ if and only if $T = 0$.

Remark 28. Before giving the proof we notice that if $T : V \to V$ is an operator on a *complex* vector space, then (ii) in Lemma 53 holds without the assumption that T is self-adjoint. Namely, $\langle Tv, v\rangle = 0$ for every $v \in V$ if and only if $T = 0$. Indeed, for every $u, w \in V$ consider:

$$\langle Tu, w\rangle = \frac{\langle T(u+w), u+w\rangle - \langle T(u-w), u-w\rangle}{4} + \frac{\langle T(u+iw), u+iw\rangle - \langle T(u-iw), u-iw\rangle}{4} i.$$

Thus, if $\langle Tv, v\rangle = 0$ for every $v \in V$, then we have $\langle Tu, w\rangle = 0$ for every $v, w \in V$, which gives that $T = 0$. The other direction is clear.

Proof of Lemma 53. For (i), let $T(v) = \lambda v$ for some non-zero $v \in V$. Then

$$\lambda \|v\|^2 = \langle \lambda v, v\rangle = \langle Tv, v\rangle = \langle v, Tv\rangle = \langle v, \lambda v\rangle = \overline{\lambda} \|v\|^2 .$$

Since $\|v\| \neq 0$ we obtain $\lambda = \overline{\lambda}$.

For (ii), one direction is clear. Since T is self-adjoint, we have $\langle Tv, w\rangle = \langle v, Tw\rangle$ for every $v, w \in V$. Thus, we obtain:

$$\langle Tu, w\rangle = \frac{\langle T(u+w), u+w\rangle - \langle T(u-w), u-w\rangle}{4},$$

which implies that $\langle Tu, w\rangle = 0$ for every $u, w \in V$, so $T = 0$. □

We collect some properties of normal operators.

Lemma 54. *Let $T : V \to V$ be an operator on a finite-dimensional complex inner space. The following are equivalent:*

(1) T is normal;

(2) $\|Tv\| = \|T^*v\|$ for every $v \in V$.

Proof. We have the following chain of equivalences:

$$\begin{aligned} T \text{ is normal } &\Leftrightarrow TT^* - T^*T = 0 \\ &\Leftrightarrow \langle (TT^* - T^*T)v, v\rangle = 0 \text{ for all } v \in V \\ &\Leftrightarrow \langle TT^*v, v\rangle = \langle T^*Tv, v\rangle \text{ for all } v \in V \\ &\Leftrightarrow \langle T^*v, T^*v\rangle = \langle Tv, Tv\rangle \text{ for all } v \in V \\ &\Leftrightarrow \|T^*v\|^2 = \|Tv\|^2 \text{ for all } v \in V \\ &\Leftrightarrow \|T^*v\| = \|Tv\| \text{ for all } v \in V, \end{aligned}$$

where the second equivalence follows from Lemma 53 (ii). □

We collect some further properties of normal operators that we won't need for now so we leave the proof as exercise.

Lemma 55. *Let $T : V \to V$ be a normal operator. Then* *(i)* $\operatorname{null} T = \operatorname{null} T^*$;

(ii) $\operatorname{range} T = \operatorname{range} T^*$;

(iii) $V = \operatorname{null} T \oplus \operatorname{range} T$;

(iv) $T - \lambda \operatorname{Id}_V$ *is normal for every* $\lambda \in \mathbb{F}$;

(v) *if* $v \in V$ *and* $\lambda \in \mathbb{F}$, *then* $Tv = \lambda v$ *if and only if* $T^* v = \overline{\lambda} v$.

Proof. Exercise. □

14. Mar. 21, 2024

14.1. **Spectral Theorem.** The goal of this section is to prove the following two results:

Theorem 4 (Real Spectral Theorem)**.** *Let* $\mathbb{F} = \mathbb{R}$ *and* $T \in \mathcal{L}(V)$ *an operator on a finite-dimensional vector space over* $\mathbb{R}$. *Then the following are equivalent:*

(1) T *is self-adjoint;*

(2) there exists an orthonormal basis B_V *such that* $\mathcal{M}(T, B_V)$ *is a diagonal matrix;*

(3) there exists an orthonormal basis B_V *consisting of eigenvectors of* T.

Theorem 5 (Complex Spectral Theorem)**.** *Let* $\mathbb{F} = \mathbb{C}$ *and* $T \in \mathcal{L}(V)$ *an operator on a finite-dimensional vector space over* $\mathbb{C}$. *Then the following are equivalent:*

(1) T *is normal;*

(2) there exists an orthonormal basis B_V *such that* $\mathcal{M}(T, B_V)$ *is a diagonal matrix;*

(3) there exists an orthonormal basis B_V *consisting of eigenvectors of* T.

Warning 7. We stress that condition (2) in Theorem 4 is *not* equivalent to T being diagonalizable.

Consider the basis $\{v_1, v_2\}$ of $V = \mathbb{R}^2$ given by $v_1 = e_1 + e_2$ and $v_2 = e_2$. Let $T : V \to V$ be defined by $T(v_1) = 2v_1$ and $T(v_2) = 3v_2$. By construction T is diagonalizable. However with respect to the standard basis $B = \{e_1, e_2\}$ we have that

$$\mathcal{M}(T, B) = \begin{pmatrix} 2 & 0 \\ -1 & 3 \end{pmatrix} \qquad \mathcal{M}(T^*, B) = \begin{pmatrix} 2 & -1 \\ 0 & 3 \end{pmatrix}.$$

Thus $T \neq T^*$. See this discussion as well.

Similarly, condition (2) in Theorem 5 is *not* equivalent to T being diagonalizable. See Exercise 7 in Worksheet 7. You can also check here for a counter-example.

Proof of Theorem 5. (1) $\Rightarrow$ (2). By Lemma 45 there exists $B_V = \{e_1, \ldots, e_n\}$ an orthonormal basis of V such that $\mathcal{M}(T, B_V)$ is upper-triangular. Let

$$\mathcal{M}(T, B_V) = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ 0 & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & a_{n,n} \end{pmatrix} \tag{13}$$

be the matrix representing T. We have

$$\|Te_1\|^2 = |a_{1,1}|^2 \quad \text{and} \quad \|T^* e_1\|^2 = \sum_{i=1}^{n} |a_{1,i}|^2.$$

Lemma 54 implies that $\|Te_1\| = \|T^* e_1\|$, thus $a_{1,j} = 0$ for $j \neq 1$. Substituting $a_{1,j} = 0$ for $j \in \{2, \ldots, n\}$ into (13) we obtain that:

$$|a_{2,2}|^2 = \|Te_2\|^2 = \|T^* e_1\|^2 = \sum_{i=2}^{n} |a_{2,i}|^2.$$

This implies that $a_{2,j} = 0$ for $j \neq 2$. We continue this process until we obtain that $\mathcal{M}(T, B_V)$ is in fact diagonal.

(2) $\Rightarrow$ (1) let $\mathcal{M}(T, B_V)$ be the diagonal matrix corresponding to T, by Lemma 51 we have that $\mathcal{M}(T^*, B_V) = \mathcal{M}(T, B_V)^\dagger$ so $\mathcal{M}(T^*, B_V)$ is also diagonal. In particular, we obtain that

$$\mathcal{M}(T^*T, B_V) = \mathcal{M}(T^*, B_V)\mathcal{M}(T, B_V) = \mathcal{M}(T, B_V)\mathcal{M}(T^*, B_V) = \mathcal{M}(TT^*, B_V),$$

since diagonal matrices commutes. Since $\mathcal{M}(-, B_V)$ is an isomorphism, we obtain $T^*T = TT^*$.

The equivalence (2) $\Leftrightarrow$ (3) is clear. □

Before proving Theorem 4 we need a couple of preliminaries. The first is the following:

Proposition 10. *Let $T : V \to V$ be a self-adjoint operator. Then the minimal polynomial is of the form $p_T = \prod_{i=1}^m (x - \lambda_i)$ for a collection of $\lambda_i \in \mathbb{R}$.*

Proof. First assume that $\mathbb{F} = \mathbb{C}$. In this case we have that $p_T = \prod_{i=1}^m (x - \lambda_i)$ by Corollary 12. By Lemma 53 i all $\lambda_i \in \mathbb{R}$.

For the case $\mathbb{F} = \mathbb{R}$, by the factorization of real polynomials Proposition 3 we have that

$$p_T(T) = a \prod_{i=1}^m (T - \lambda_i) \prod_{j=1}^k (T^2 + b_j T + c_j) = 0 \tag{14}$$

where $a, \lambda_1, \ldots, \lambda_m, b_1, \ldots, b_k, c_1, \ldots, c_k \in \mathbb{R}$ and for each $1 \leq j \leq k$ we have $b_j^2 < 4c_j$. We can assume that k is minimal. Now from Lemma 56 below we know that $(T^2 + b_j T + c_j)$ is invertible whenever $b_j^2 < 4c_j$. Thus, by multiplying (14) by the inverses of $(T^2 + b_j T + c_j)$ we obtain that

$$p_T(T) = a \prod_{i=1}^m (T - \lambda_i) = 0.$$

This finishes the proof. □

The following is a technical result needed in the proof of the above Proposition.

Lemma 56. *Let $T : V \to V$ be a self adjoint operator and consider $b, c \in \mathbb{R}$ such that $b^2 < 4c$. Then the operator $T^2 + bT + c\,\mathrm{Id}_V$ is invertible.*

Proof. Since $T^2 + bT + c\,\mathrm{Id}_V$ is a map from V to itself, it is enough to check that it is injective. Consider a non-zero $v \in V$ such that $(T^2 + bT + c\,\mathrm{Id}_V)v = 0$. Then we calculate:

$$\begin{aligned}
\left\langle (T^2 + bT + c\,\mathrm{Id}_V)v, v \right\rangle &= \left\langle T^2 v, v \right\rangle + b \langle Tv, v \rangle + c \langle v, v \rangle \\
&= \langle Tv, Tv \rangle + b \langle Tv, v \rangle + c \|v\|^2 \\
&\geq \|Tv\|^2 - |b| \|Tv\| \|v\| + c \|v\|^2 \\
&= \left(\|Tv\| - \frac{|b| \|v\|}{2} \right)^2 + \left(c - \frac{b^2}{4} \right)^2 \|v\| \\
&> 0,
\end{aligned}$$

where $-|b| \|Tv\| \|v\| \leq b \langle Tv, v \rangle$ follows from the Cauchy–Schwartz inequality, i.e. $| \langle Tv, v \rangle | \leq \|Tv\| \|v\|$. So we obtain that $0 = \langle 0, v \rangle > 0$, which is a contradiction. □

Proof of Theorem 4. (1) $\Rightarrow$ (2) By Proposition 10 we know that p_T is the product of linear factors. Then Lemma 36 implies that there exists a basis B_V such that $\mathcal{M}(T, B_V)$ is upper-triangular and Lemma 45 implies that we can assume that B_V is orthonormal.

Since B_V is orthonormal, Lemma 51 implies that $\mathcal{M}(T, B_V) = \mathcal{M}(T^*, B_V)^{\mathrm{t}}$, since the transpose is equal to the conjugate transpose. Since $\mathcal{M}(T, B_V)$ is upper-triangular, this implies that $\mathcal{M}(T, B_V)$ is diagonal, as we needed to prove.

(2) $\Rightarrow$ (1) there exists B_V such that $\mathcal{M}(T, B_V)$ is diagonal. Since B_V is orthonormal Lemma 51 implies that $\mathcal{M}(T^*, B_V) = \mathcal{M}(T, B_V)^{\mathrm{t}} = \mathcal{M}(T, B_V)$. Since $\mathcal{M}(-, B_V)$ is an isomorphism we obtain $T = T^*$.

The equivalence (2) $\Leftrightarrow$ (3) is clear. □

Example 29. (i) Let $T : V \to V$ be an operator on a finite-dimensional vector space over $\mathbb{C}$. Assume that T is normal and has a single eigenvalue. By the Spectral Theorem there exists an orthonormal basis B_V such that

$$\mathcal{M}(T, B_V) = \begin{pmatrix} \lambda & & 0 \\ & \ddots & \\ 0 & & \lambda \end{pmatrix}$$

for $\lambda \in \mathbb{C}$. Now consider any other basis B'_V, let $C = \mathcal{M}(\mathrm{Id}_V, B'_V, B_V)$, then

$$\mathcal{M}(T, B'_V) = C^{-1}\mathcal{M}(T, B_V)C = \mathcal{M}(T, B_V)$$

since diagonal matrices commute with all other matrices. Thus, we have that $\mathcal{M}(T, B'_V) = \mathcal{M}(\lambda \,\mathrm{Id}_V, B'_V)$ for any basis B'_V, which gives that $T = \lambda \,\mathrm{Id}_V$.

(ii) Let $T : V \to V$ be a self-adjoint operator on a real vector space. Assume that 2 and 3 are the only eigenvalues of T, then

$$T^2 - 5T + 6 = 0.$$

Indeed, there exists an orthonormal basis B_V such that $\mathcal{M}(T, B_V)$ is a diagonal matrix with only 2's and 3's in the diagonal. Then we have:

$$\mathcal{M}(T^2 - 5T + 6, B_V) = \mathcal{M}((T-2)(T-3), B_V) = \mathcal{M}((T-2), B_V)\mathcal{M}((T-3), B_V) = 0,$$

where the last equality follows because of the following. There exist a proper subset $S \subset \{1, \ldots, n\}$ such that:

$$\mathcal{M}((T-2), B_V)_{ij} = \begin{cases} 1 & \text{for } i = j \text{ and } i \in S \\ 0 & \text{else,} \end{cases}$$

and

$$\mathcal{M}((T-3), B_V)_{ij} = \begin{cases} -1 & \text{for } i = j \text{ and } i \in \{1, \ldots, n\}\backslash S \\ 0 & \text{else.} \end{cases}$$

Since $\mathcal{M}(-, B_V)$ is an isomorphism we obtain $T^2 - 5T + 6 = 0$.

14.2. **Positive operators.** Given an operator $T : V \to V$ a *square root* of T is an operator $S : V \to V$ such that $S^2 = T$.

Example 30. Let $T : \mathbb{F}^3 \to \mathbb{F}^3$ be given by $T(x_1, x_2, x_3) = (x_3, 0, 0)$. Then $S(x_1, x_2, x_3) = (x_2, x_3, 0)$ is a square root.

The motivation of this section is two-fold:

- we want to understand what happens if we impose a stronger condition than Lemma 52 (2);
- we want to understand when we can find square root of linear maps.

Proposition 11. *Let $T : V \to V$ be an operator on an inner product space. The following are equivalent:*

(i) $\langle Tv, v\rangle \geq 0$, i.e. $\langle Tv, v\rangle \in \mathbb{R}_{\geq 0}$, for every $v \in V$;

(ii) T is self-adjoint and all eigenvalues of T are non-negative[14]*;*

(iii) there exists an orthonormal basis B_V such that $\mathcal{M}(T, B_V)$ has only non-negative numbers in the diagonal;

(iv) T has a square root satisfying (i);

(v) T has a self-adjoint square root;

(vi) $T = R^*R$ for some $R \in \mathcal{L}(V)$.

[14]By non-negative we mean the eigenvalues belong to $\mathbb{R}_{\geq 0}$.

Proof. (i) $\Rightarrow$ (ii). By Lemma 52 we have that T is self-adjoint. Let λ be an eigenvalue for some non-zero vector $v \in V$. Then $\langle Tv, v\rangle = \lambda \|v\|^2 \geq 0$, since $\|v\| \neq 0$, we obtain $\lambda \geq 0$.

(ii) $\Rightarrow$ (iii). By Theorem 5 or Theorem 4 we know that there exists an orthonormal basis B_V such that $\mathcal{M}(T, B_V)$ is diagonal. Clearly the entries in the diagonal are eigenvalues, and hence we have the claim.

(iii) $\Rightarrow$ (iv). Let $\{\lambda_1, \ldots, \lambda_n\}$ denote the diagonal entries of T in the orthonormal basis $B_V = \{v_1, \ldots, v_n\}$ of V. Let $R(v_i) := \sqrt{\lambda_i} v_i$, where $\sqrt{\lambda_i}$ denotes the non-negative square root of λ_i. By expanding any $v \in V$ in the basis $\{v_1, \ldots, v_n\}$ we can check that $\langle Rv, v\rangle = \sum_{i=1}^n \sqrt{\lambda_i} |\langle v, v_i\rangle|^2$, which is non-negative.

(iv) $\Rightarrow$ (v). This follows from (i) $\Rightarrow$ (ii).

(v) $\Rightarrow$ (vi). Since $R = R^*$ and $R^2 = T$, we get $R^*R = T$.

(vi) $\Rightarrow$ (i). Given any $v \in V$ we compute $\langle Tv, v\rangle = \langle R^*Rv, v\rangle = \langle Rv, Rv\rangle = \|Rv\|^2 \geq 0$. ☐

Definition 41. We say an operator $T : V \to V$ is *positive* if it satisfies the equivalent conditions of Proposition 11.

Lemma 57. *Let T be positive operator, then there exists an unique positive operator R, such that $R^2 = T$.*

Proof. Exercise. ☐

Notation 7. Let $T : V \to V$ be a positive operator, we denote by $\sqrt{T}$ the unique positive operator which is a square root of T, whose existence is guaranteed by Proposition 11 and uniqueness by Lemma 57.

14.3. Singular Value Decomposition.

Definition 42. Let $T \in \mathcal{L}(V, W)$, the *singular values of T* are the non-negative square roots of the eigenvalues of T^*T, listed in decreasing order, each included as many times as the dimension of corresponding eigenspace of T^*T.

We need the following results for this section.

Lemma 58. *Let $T \in \mathcal{L}(V, W)$, then*

(i) $T^*T : V \to V$ *is a positive operator;*

(ii) $\operatorname{null} T^*T = \operatorname{null} T$;

(iii) $\operatorname{range} T^*T = \operatorname{range} T^*$;

(iv) $\dim \operatorname{range} T = \dim \operatorname{range} T^* = \dim \operatorname{range} T^*T$.

Proof. For (i), notice that $(T^*T)^* = T^*(T^*)^* = T^*T$ by Lemma 49 (iii) and (iv). Thus T^*T is self-adjoint. We now calculate:

$$\langle (T^*T)v, v\rangle = \langle Tv, Tv\rangle = \|Tv\|^2 \geq 0$$

for every $v \in V$, so T^*T is positive. Todo! ☐

Theorem 6. *Let $T : V \to W$ be a linear map and let $\{s_1, \ldots, s_m\}$ be the positive singular values of T. Then there exist orthonormal lists $\{e_1, \ldots, e_m\}$ in V and $\{f_1, \ldots, f_m\}$ in W such that*

$$Tv = s_1 \langle v, e_1\rangle f_1 + \cdots + s_m \langle v, e_m\rangle f_m$$

for every $v \in V$.

Proof. Let $\{s_1, \ldots, s_n\}$ be the singular values of T. By Lemma 58 the operator T^*T is positive, so Theorem 5 or Theorem 4 implies that there exists an orthonormal basis $B_V = \{e_1, \ldots, e_n\}$ such that

$$T^*Te_i = s_i^2 e_i \tag{15}$$

for $1 \leq i \leq n$. For each $j = 1, \ldots, m$ define $f_j := \frac{Te_i}{s_i}$. Then we have:

$$\langle f_j, f_k\rangle = \frac{1}{s_j s_k} \langle Te_i, Te_k\rangle = \frac{1}{s_j s_k} \langle e_i, T^*Te_k\rangle = \frac{s_k}{s_k} \langle e_i, e_k\rangle = \begin{cases} 1 & \text{if } j = k, \\ 0 & \text{if } j \neq k; \end{cases}$$

for every $j, k \in \{1, \ldots, m\}$. Thus the set $\{f_1, \ldots, f_m\}$ is orthonormal. Let $v \in V$, then we compute:

$$Tv = T(\sum_{i=1}^n \langle v, e_i\rangle e_i) = \sum_{i=1}^m \langle v, e_i\rangle Te_i = \sum_{i=1}^m s_i \langle v, e_i\rangle f_i,$$

where we drop Te_i, for $m + 1 \leq i \leq n$, since by assumption $\{s_{m+1}, \ldots, s_n\}$ are all zero. ☐

Example 31. Discussion Examples 7.67 and 7.79 from the textbook. Write the examples here.

15. Mar. 25, 2024

15.1. **Generalized Eigenvectors.** We start this section by understanding the null space of powers of an operator.

Lemma 59. *Let $T \in \mathcal{L}(V)$, then* $\operatorname{null} T^{\dim V} = \operatorname{null} T^{\dim V + k}$ *for any* $k \geq 0$.

Proof. Let $v \in \operatorname{null} T^k$ for some $k \geq 0$, then $T^{k+1}(v) = T(T^k v) = T(0) = 0$, thus $v \in \operatorname{null} T^{k+1}$.

Notice that if $\operatorname{null} T^k = \operatorname{null} T^{k+1}$ for some $k \geq 1$, then $\operatorname{null} T^k = \operatorname{null} T^{k+i}$ for every $i \geq 1$. Indeed, assume $v \in \operatorname{null} T^{k+i}$, then $T^{i-1}v \in \operatorname{null} T^{k+1}$, which implies that $T^{i-1}v \in \operatorname{null} T^k$, that is $v \in \operatorname{null} T^{k+i-1}$. Continuing in this way, we obtain that $v \in \operatorname{null} T^k$. The other inclusion is clear.

Finally, suppose that $\operatorname{null} T^{\dim V} \subset \operatorname{null}^{\dim V+1}$ is a proper subset. Then we have the sequence:

$$\{0\} = \operatorname{null} T^0 \subset \operatorname{null} T \subset \cdots \operatorname{null} T^{\dim V} \subset \operatorname{null} T^{\dim V}$$

of proper subsets. This implies that $\dim \operatorname{null} T^{\dim V+1} \geq \dim V + 1$, which is a contradiction since $\operatorname{null} T^{\dim V+1} \subseteq V$. ∎

Lemma 60. *For any $T \in \mathcal{L}(V)$ we have*

$$\operatorname{null} T^{\dim V} \oplus \operatorname{range} T^{\dim V} = V.$$

Proof. First we claim that $\operatorname{null} T^{\dim V} \cap \operatorname{range} T^{\dim V} = \{0\}$. Assume that $v \in \operatorname{null} T^{\dim V} \cap \operatorname{range} T^{\dim V}$ such that $v \neq 0$, then we have $v = T^n u$ for some $u \in V$ and $T^n v = 0$. This gives $T^{2n} u = 0$. Since $\operatorname{null} T^{2n} = \operatorname{null} T^n$, we get $v = T^n u = 0$.

Thus, $\operatorname{null} T^{\dim V} \oplus \operatorname{range} T^{\dim V} = \operatorname{null} T^{\dim V} + \operatorname{range} T^{\dim V}$. Moreover, by the fundamental theorem of linear algebra we have $\dim(\operatorname{null} T^{\dim V} \oplus \operatorname{range} T^{\dim V}) = \dim \operatorname{null} T^{\dim V} + \dim \operatorname{range} T^{\dim V} = \dim V$. Thus, we obtain that $\operatorname{null} T^{\dim V} \oplus \operatorname{range} T^{\dim V} = V$. ∎

Example 32. Let $T : \mathbb{C}^3 \to \mathbb{C}^3$ be given by $T(x_1, x_2, x_3) = (4x_2, 0, 5x_3)$. Notice that $\operatorname{null} T = \operatorname{Span} e_1$ and $\operatorname{range} T = \operatorname{Span}\{e_1, e_3\}$, so $\operatorname{null} T \cap \operatorname{range} T \neq \{0\}$ and $\operatorname{null} T + \operatorname{range} T \neq \mathbb{C}^3$. However, we have $\operatorname{null} T^3 = \operatorname{Span}\{e_1, e_2\}$ and $\operatorname{range} T^3 = \operatorname{Span} e_3$. Thus, we obtain $\operatorname{null} T^3 \oplus \operatorname{range} T^3 = \mathbb{C}^3$, as expected.

Definition 43. Let $\lambda \in \mathbb{F}$. A non-zero vector $v \in V$ is said to be a *generalized eigenvector* for λ if

$$(T - \lambda Id_v)^k v = 0, \quad \text{for some } k \geq 1.$$

Remark 29. Notice that by Lemma 59 v is a generalized eigenvector if and only if

$$(T - \lambda Id_v)^{\dim V} v = 0.$$

The following is a generalization of Remark 22.

Remark 30. Let v be generalized eigenvector for λ and λ', then $\lambda = \lambda'$. Indeed, let m be the smallest integer such that $(T - \lambda' \operatorname{Id}_V)^m v = 0$. Then we compute:

$$\begin{aligned} 0 &= (T - \lambda \operatorname{Id}_V)^{\dim V} v \\ &= ((T - \lambda' \operatorname{Id}_V) + (\lambda' - \lambda) \operatorname{Id}_V)^{\dim V} v \\ &= \sum_{k=0}^{\dim V} b_k (T - \lambda' \operatorname{Id}_V)^k (\lambda' - \lambda)^{\dim V - k} \operatorname{Id}_V v \end{aligned}$$

where $b_0 = 1$ and the other b_k are some binomial coefficients that we don't need to specfiy. By applying $(T - \lambda' \operatorname{Id}_V)^{m-1}$ we obtain:

$$0 = (\lambda' - \lambda)^{\dim V} (T - \lambda' \operatorname{Id}_V)^{m-1} v.$$

Since $(T - \lambda' \operatorname{Id}_V)^{m-1} v \neq 0$, then we get $(\lambda' - \lambda)^{\dim V}$ which implies that $\lambda' = \lambda$.

The next result shows that over the complex numbers for any operator $T : V \to V$ we can always decompose V as a directed sum of spaces which are generalized eigenvectors for distinct eigenvalues. This is in contrast to Theorem 5 and Lemma 38.

Proposition 12. *Let $T : V \to V$ be an operator on a complex vector space. There is a basis of V consisting of generalized eigenvectors.*

Proof. We proceed by induction on $n = \dim V$. The case $n = 1$ is clear, since then T is simply given by multiplication by a scalar, so any non-zero vector is an eigenvector, hence in particular a generalized eigenvector.

Assume that $n > 1$ and that the result holds for every vector space of dimension $k \leq n-1$. Let λ be an eigenvalue of T. Then by Lemma 60 applied to $T - \lambda \operatorname{Id}_V$ we have:

$$\operatorname{null}(T - \lambda \operatorname{Id}_V)^{\dim V} \oplus \operatorname{range}(T - \lambda \operatorname{Id}_V)^{\dim V} = V.$$

If $\operatorname{null}(T-\lambda \operatorname{Id}_V)^{\dim V} = V$, then every non-zero vector in V is a generalized eigenvector, hence the conclusion is clear. Assume that $\operatorname{null}(T - \lambda \operatorname{Id}_V)^{\dim V} \neq V$, then we have:

$$0 < \dim \operatorname{range}(T - \lambda \operatorname{Id}_V)^{\dim V} < \dim V = n.$$

Thus, let S be the restriction of T to $\operatorname{range}(T - \lambda \operatorname{Id}_V)^{\dim V}$, which is invariant under T by Remark 24. By the inductive hypothesis, there exists a basis of $\operatorname{range}(T - \lambda \operatorname{Id}_V)^{\dim V}$ consisting of generalized eigenvectors of S. It is clear that these are also generalized eigenvectors of T. Now we add to this basis a basis of $\operatorname{null}(T - \lambda \operatorname{Id}_V)^{\dim V}$. This finishes the proof. □

The following is a continuation of Example 32.

Example 33. Let $T : \mathbb{C}^3 \to \mathbb{C}^3$ be given by $T(x_1, x_2, x_3) = (4x_2, 0, 5x_3)$. We notice that $\lambda_1 = 0$ and $\lambda_2 = 5$ are eigenvalues of T. We can see that by considering the matrix representation of T in the standard basis and calculating its characteristic polynomial.

We notice that $(x_1, 0, 0)$ with $x_1 \neq 0$ are eigenvectors for $\lambda_1 = 0$ and that $(0, 0, x_3)$ with $x_3 \neq 0$ are eigenvectors for $\lambda_2 = 5$. These don't form a basis for V.

However, if we calculate $\operatorname{null} T^3 = \{(x_1, x_2, 0) \mid x_1, x_2 \in \mathbb{C}\}$, we notice that $(x_1, x_2, 0)$ where not both x_1 and x_2 are zero given generalized eigenvectors for λ_1. Whereas $\operatorname{null}(T - 5)^3 = \{(0, 0, x_3) \mid x_3 \in \mathbb{C}\}$, thus $(0, 0, x_3)$ with $x_3 \neq 0$ are generalized eigenvectors for $\lambda_2 = 5$. And we obtain a decomposition of $\mathbb{C}^3$ as claimed.

Here is a useful observation which generalizes Lemma 31.

Lemma 61. *Any finite list of generalized eigenvectors for distinct eigenvalues is linearly independent.*

Proof. Suppose by contradiction that there exists a list $\{v_1, \ldots, v_m\}$ of generalized eigenvectors for distinct eigenvalues $\lambda_1, \ldots, \lambda_m \in \mathbb{F}$ which is linearly depend. Notice that we can assume that m is minimal. Let $a_1, \ldots, a_m \in \mathbb{F}$ not all zero such that:

$$a_1 v_1 + \cdots + a_m v_m = 0.$$

By applying $(T - \lambda_m)^{\dim V}$ we obtain:

$$a_1 (T - \lambda_m)^{\dim V} v_1 + \cdots + a_{m-1} (T - \lambda_m)^{\dim V} v_{m-1} = 0. \tag{16}$$

Thus, Remark 30 imply that for every $i \in \{1, \ldots, m-1\}$ we have:

$$(T - \lambda_m)^{\dim V} v_i \neq 0.$$

Since $(T-\lambda_k)^{\dim V}(T-\lambda_m)^{\dim V} v_i = (T-\lambda_m)^{\dim V}(T-\lambda_k)^{\dim V} v_k = 0$. We have that each $(T-\lambda_m)^{\dim V} v_i$ is a generalized eigenvector of λ_i for $i \in \{1, \ldots, m-1\}$. Thus, the list $\{(T-\lambda_m)^{\dim V} v_1, \ldots, (T-\lambda_m)^{\dim V} v_{m-1}\}$ is list of generalized eigenvectors for distinct eigenvalues which by (16) is linearly depend. This contradicts the minimality of m. This finishes the proof. □

Corollary 21. *Let $T : V \to V$ be an operator on a real vector space. Then the following are equivalent:*

(1) there is a basis of V consisting of generalized eigenvectors;

(2) the minimal polynomial of T is of the form $p_T(x) = \prod_{i=1}^{m} (x - \lambda_i)$.

Proof. Exercise! □

We end this section with a brief discussion of nilpotent operators, which collects some properties that will be useful in the next section.

Definition 44. An operator $T : V \to V$ is said to be nilpotent if $T^k = 0$ for some $k \geq 1$.

Remark 31. Notice that if $T : V \to V$ is nilpotent, then $T^{\dim V} = 0$. Indeed, since $\operatorname{null} T^k = V$ for some $k \geq 1$ and $V = \operatorname{null} T^{k+i}$ for every $i \geq 1$ if $\operatorname{null} T^{\dim V} \neq V$ then we would have a sequence of proper subspaces $0 \subset \operatorname{null} T \subset \cdots \subset \operatorname{null} T^{\dim V} \subset \operatorname{null} T^{\dim V+1} \subseteq V$ of length $dimV + 1$, which is a contradiction.

Exercise 40. Let $T : V \to V$ be a nilpotent operator and let $m \geq 1$ be the smallest integer such that $T^m = 0$. Prove that there exists $v \in V$ such that $\{v, Tv, \ldots, T^{m-1}v\}$ is a linearly independent set.

Lemma 62. *Let $T : V \to V$ be a nilpotent operator. Then*

(i) the only eigenvalues of T are 0;

(ii) if $\mathbb{F} = \mathbb{C}$ and 0 is the only eigenvalue of T, then T is nilpotent.

Proof. For (i) let $v \neq 0$ in V such that $Tv = \lambda v$. Then we have $0 = T^{\dim V} v = \lambda^{\dim V} v$, which implies that $\lambda^{\dim V} = 0$, so $\lambda = 0$.

For (ii) by (i) and Corollary 12 the minimal polynomial of T has the form $p_T(x) = \prod_{i=0}^{m} x$, that is $p_T(T) = T^m = 0$ for some $m \geq 0$. ☐

Exercise 41. Give an example of $T : V \to V$ on a real vector space, such that 0 is the only eigenvalue of T but T is not nilpotent.

Lemma 63. *Let $T \in \mathcal{L}(V)$. Then the following are equivalent:*

(1) T is nilpotent;

(2) the minimal polynomial of T is z^m for some $m \geq 1$;

(3) there exists a basis B_V of V such that $\mathcal{M}(T, B_V)$ has zeros on the diagonal and all entries below it, i.e. it is strictly upper-triangular.

Proof. Exercise. ☐

Exercise 42. Let $T : V \to V$ be nilpotent, then $T^{\dim V} = 0$.

15.2. **Generalized Eigenspace Decomposition.** We start with a defintion:

Definition 45. Let $T \in \mathcal{L}(V)$ and $\lambda \in \mathbb{F}$. The *generalized eigenspace* of T corresponding to λ is:

$$G(\lambda, T) := \{v \in V \mid (T - \lambda)^k v = 0 \text{ for some } k \in \mathbb{Z}_{\geq 1}\}.$$

Notice that by Remark 29 one has $G(\lambda, T) = \operatorname{null}(T - \lambda)^{\dim V}$.

Example 34. Let $T : \mathbb{C}^3 \to \mathbb{C}^3$ be given by $T(x_1, x_2, x_3) = (4x_2, 0, 5x_3)$. Then

$$G(0, T) = Span\{e_1, e_2\} \text{ and } G(5, T) = \operatorname{Span} e_3,$$

which gives $V = G(0, T) \oplus G(5, T)$.

The following is a generalization of Proposition 7.

Theorem 7 (Generalized Eigenspace Decomposition)**.** *Assume that $\mathbb{F} = \mathbb{C}$. Let $T \in \mathcal{L}(V)$ be an operator and $\lambda_1, \ldots, \lambda_m$ its distinct eigenvalues. Then*

(i) $G(\lambda_i, T)$ is invariant under T for every $i \in \{1, \ldots, m\}$;

(ii) $T - \lambda_i \operatorname{Id}_V|_{G(\lambda_i,T)}$ is nilpotent for every $i \in \{1, \ldots, m\}$;

(iii) $V = G(\lambda_1, T) \oplus \cdots \oplus G(\lambda_m, T)$.

Proof. For (i), since $G(\lambda_i, T) = \operatorname{null}(T - \lambda_i \operatorname{Id}_V)^{\dim V}$ the result follows by applying Remark 24 to $q(T) := (T - \lambda_i \operatorname{Id}_V)^{\dim V}$.

For (ii), we notice that for any $v \in G(\lambda_i, T)$ we have $(T - \lambda_i \operatorname{Id}_V)^{\dim V} v = 0$, thus $(T - \lambda_i \operatorname{Id}_V|_{G(\lambda_i,T)})^{\dim V} = 0$.

For (iii), assume that we have

$$v_1 + \cdots + v_m = 0,$$

for $v_i \in G(\lambda_i, T)$. Since each v_i has a different eigenvalue Lemma 61 implies that $\{v_1, \ldots, v_m\}$ is linearly independent, which gives that each v_i is zero; that is $G(\lambda_1, T) + \cdots + G(\lambda_m, T)$ is a direct sum. Finally, Proposition 12 implies that each $v \in V$ can be written as a finite linear combination of generalized eigenvectors, i.e. $V \subseteq G(\lambda_1, T) + \cdots + G(\lambda_m, T)$. ☐

We know introduce the following:

Definition 46. Let $T \in \mathcal{L}(V)$ and λ be an eigenvalue of T.

- The *algebraic multiplicity* of λ is
$$\mu_T(\lambda_i) := \dim G(\lambda_i, T).$$
- The *geometric multiplicity* of λ is[15]
$$\gamma_T(\lambda_i) := \dim E(\lambda_i, T).$$

Notice that $\gamma_T(\lambda_i) \leq \mu_T(\lambda_i)$ for every λ_i.

Remark 32. Assume that $\mathbb{F} = \mathbb{C}$. It follows directly from Theorem 7 that
$$\sum_{i=1}^{m} \mu_T(\lambda_i) = \dim V,$$
i.e. the sum of the multiplicity of all eigenvalues of T equals the dimension of V.

Definition 47. Given $T \in \mathcal{L}(V)$ the *characteristic polynomial of* T is defined to be:[16]
$$c_T(x) := \prod_{i=1}^{m} (x - \lambda_i)^{\mu_T(\lambda_i)},$$
where $\lambda_1, \dots, \lambda_m$ are all the distict eigenvalues of T.

Remark 33. When $\mathbb{F} = \mathbb{C}$, the following are a direct consequence of Definition 47:

(i) $\deg c_T = \dim V$;

(ii) the zeros of c_T are precisely the eigenvalues of T.

Remark 34. Normally (for instance in Math 2101), the characteristic polynomial and algebraic multiplicity are defined using the determinant of $T - x\,\mathrm{Id}_V$. We will see in a couple of sections that Definition 46 and Definition 47 agree with this previous definition that you are probably familiar with.

The following result is harder to prove if one adopts the definition of characteristic polynomial via the determinant.

Theorem 8 (Cayley–Hamilton theorem)**.** *Assume that $\mathbb{F} = \mathbb{C}$ and let $T \in \mathcal{L}(V)$. The characteristic polynomial c_T is a multiple of the minimal polynomial p_T.*

Proof. By Lemma 33 it is enough to prove that $c_T(T) = 0$. Notice that since each $G(\lambda_i, T)$ is invariant under T, then given $q(T)$ any polynomial on T we have $q(T) = \oplus_{i=1}^{m} q(T)|_{G(\lambda_i,T)}$, where $\oplus_{i=1}^{m} q(T)|_{G(\lambda_i,T)}$ is the unique map from $\oplus_{i=1}^{m} G(\lambda_i, T)$ determined by the composites $G(\lambda_i, T) \overset{q(T)|_{G(\lambda_i,T)}}{\to} G(\lambda_i, T) \subseteq V$ via Proposition 2.

Thus, by Theorem 7 we have:
$$c_T(T) = \prod_{i=1}^{m} (T - \lambda_i)^{\mu_T(\lambda_i)} = \oplus_{i=1}^{m} c_T(T)|_{G(\lambda_i,T)},$$
but for each λ_k we have
$$c_T(T)|_{G(\lambda_k,T)} = \prod_{i=1,\, i\neq k}^{m} (T - \lambda_i)^{\mu_T(\lambda_i)}\Big|_{G(\lambda_k,T)} (T - \lambda_k)^{\mu_T(\lambda_k)}\Big|_{G(\lambda_k,T)} = 0,$$
where the last equality follows from $G(\lambda_k, T) = \mathrm{null}(T|_{G(\lambda_k,T)} - \lambda_k)^{\dim G(\lambda_k,T)}$. □

[15] I am not sure if there is a standard notation for the geometric multiplicity.

[16] I am also not sure what is the standard notation for characteristic polynomial. Some references use p_T, but we will reserve that for the minimal polynomial.

The last result gives us a way to compute $\mu_T(\lambda_i)$ from an upper-triangular matrix representing T.

Proposition 13. *Assume that $\mathbb{F} = \mathbb{C}$ and let $T \in \mathcal{L}(V)$. Assume that $B_V \subset V$ is a basis such that $\mathcal{M}(T, B_V)$ is upper-triangular. Then $\mu_T(\lambda_i)$ equals the number of times that λ_i shows up in the diagonal of $\mathcal{M}(T, B_V)$.*

Proof. Let $B_V = \{v_1, \ldots, v_n\}$ and $\lambda_1, \ldots, \lambda_n$ be the diagonal entries of $\mathcal{M}(T, B_V)$. Let $S \subset \{1, \ldots, n\}$ be the subset such that for $k \in S$ we have $\lambda_k \neq 0$. Thus, $T|_{\operatorname{Span} S}$ is injective and $\{Tv_i\}_{i \in S}$ is linearly independent. In particular, we obtain:

$$\dim \operatorname{range} T \geq n - d,$$

where $n = \dim V$ and $d := |\{1, \ldots, n\} \backslash S|$, i.e. d is the amount of 0's in the diagonal of $\mathcal{M}(T, B_V)$. The fundamental theorem of linear algebra gives that

$$\dim \operatorname{null} T \leq d. \tag{17}$$

Since $\mathcal{M}(T^n, B_V) = \mathcal{M}(T, B_V)^n$, it is clear that the diagonal entries of $\mathcal{M}(T^n, B_V)$ are $\lambda_1^n, \ldots, \lambda_n^n$. And we have 0's in the same positions in the diagonal of $\mathcal{M}(T^n, B_V)$ as in the diagonal of $\mathcal{M}(T, B_V)$. Thus (17) applied to T^n gives:

$$\dim \operatorname{null} T^n \leq d. \tag{18}$$

Let d_λ denote the number of times that λ appears in the diagonal of $\mathcal{M}(T, B_V)$ and let $\mu_T(\lambda)$ be the algebraic multiplicity of λ. Then by apply (18) to $(T - \lambda)$ we obtain:

$$\mu_T(\lambda) = \dim \operatorname{null}(T - \lambda)^n \leq d_\lambda. \tag{19}$$

However, now we notice that $\sum_{i=1}^m \mu_T(\lambda_i) = \dim V$ by Remark 32 and clearly $\sum_{i=1}^m d_{\lambda_i} = \dim V$, where $\lambda_1, \ldots, \lambda_m$ are a list of distinct eigenvalues of T. Thus (19) is an equality for every eigenvalue λ. This finishes the proof. $\square$

We end this section interpreting the decomposition of Theorem 7 in terms of matrices.

Definition 48. A *block diagonal* matrix is a square matrix of the form:

$$\begin{pmatrix} A_1 & & 0 \\ & \ddots & 0 \\ & & A_m \end{pmatrix}$$

where each A_i is a square matrix along the diagonal and all other entries are 0.

Lemma 64. *Assume that $\mathbb{F} = \mathbb{C}$ and let $T \in \mathcal{L}(V)$. Let $\lambda_1, \ldots, \lambda_m$ be distinct eigenvalues of T with (algebraic) multiplicities $d_1, \ldots, d_m$. Then there exists a basis B_V of V such that*

$$\mathcal{M}(T, B_V) = \begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_m \end{pmatrix} \tag{20}$$

where each A_i is an i-by-i upper-triangular matrix of the form:

$$A_i = \begin{pmatrix} \lambda_i & & * \\ & \ddots & \\ & & \lambda_i \end{pmatrix}.$$

Proof. Exercise. $\square$

16. Mar. 28, 2024

16.1. **Jordan normal form.** The goal of this section is to improve on the results of §15.2.

From the point of view of Theorem 7 we can ask: "Given an eigenvector $v_k \in E(\lambda, T)$ how can we extend this to a set that is a basis of $G(\lambda, T)$?"

From the point of view of Lemma 64 we can ask: "Can we do even better and make each block matrix A_i into a simpler form?".

Let's first consider a couple of examples for motivation. For this discussion the specific value of an eigenvalue λ is not relevant, so we will assume that the operator is nilpotent, hence its eigenvalue is 0.

Example 35. Let $T : \mathbb{C}^4 \to \mathbb{C}^4$ be given by $T(z_1, z_2, z_3, z_4) = (0, z_1, z_2, z_3)$. Notice that $T^4 = 0$. Moreover, for $v_1 = e_1$, we have $T^3v_1, T^2v_1, Tv_1, v_1$ is a basis of $\mathbb{C}^4$. Notice this case is particularly nice, since we have $v_1 \in E(0, T)$ and then $\{T^3v_1, T^2v_1, Tv_1, v_1\}$ becomes a basis of $G(0, T) = V$.

Example 36. Let $T : \mathbb{C}^6 \to \mathbb{C}^6$ be given by $T(z_1, z_2, z_3, z_4, z_5, z_6) = (0, z_1, z_2, 0, z_4, 0)$. Notice that $T^3 = 0$. However, there is *not* a vector $v \in \mathbb{C}^6$ such that $\{v, Tv, T^2v, T^3v, T^4v, T^5v\}$ is a basis of $\mathbb{C}^6$. However if we consider $v_1 = e_1$, $v_2 = e_4$ and $v_3 = e_6$, then we have that $B_V = \{T^2v_1, Tv_1, v_1, Tv_2, v_2, v_3\}$ form a basis of $\mathbb{C}^6$. In fact, we have

$$\mathcal{M}(T, B_V) = \begin{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} & \begin{matrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{matrix} & \begin{matrix} 0 \\ 0 \\ 0 \end{matrix} \\ \begin{matrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{matrix} & \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} & \begin{matrix} 0 \\ 0 \end{matrix} \\ \begin{matrix} 0 & 0 & 0 \end{matrix} & \begin{matrix} 0 & 0 \end{matrix} & (0) \end{pmatrix}. \tag{21}$$

In other words, we can think of $\mathcal{M}(T, B_V)$ as three block matrices A_1, A_2 and A_3, where A_1 is a 3 by 3 matrix, A_2 is a 2 by 2 matrix and A_3 is a 1 by 1 matrix. Moreover, notice that each of these block diagonal matrices only has non-zero entries on the first row above the diagonal.

It turns out that the behaviour of Example 36 is not special in two regards.

Firstly, for any nilpotent operator we can find a basis of V such that the matrix representing T in this basis has a form similar to that of (21).

Secondly, even for arbitrary operators (on a complex vector space) we are able to find a basis such that we have a block diagonal matrix as in (20), where each A_i only has non-zero entries at the diagonal and on the first row above the diagonal.

To prove the next results we introduce the following:

Definition 49. Given an operator $T : V \to V$ on a finite-dimensional space. A *Jordan basis* for T is a basis B_V such that

$$\mathcal{M}(T, B_V) = \begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_m \end{pmatrix}$$

where each A_i is an i-by-i upper-triangular matrix of the form:

$$A_i = \begin{pmatrix} \lambda_i & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda_i \end{pmatrix}.$$

Warning 8. The eigenvalues λ_i of A_i are not necessarily distinct. See Example 36.

Proposition 14. *Let $T : V \to V$ be a nilpotent operator. Then T has a Jordan basis B_V.*

Proof. We proceed by induction on $\dim V$. The case $\dim V = 1$ is clear, since any non-zero vector $v \in V$ would be a Jordan basis.

Let $n = \dim V$ and assume that the result holds for every vector space U and nilpotent operator on U such that $\dim U \leq n - 1$. Let m be the smallest positive integer such that $T^m = 0$. Consider $v \in V$ such that $T^{m-1}v \neq 0$, then by Exercise 40 $\{v, Tv, \ldots, T^{m-1}v\}$ is linearly independent. Let $U :=$

Span $\{v, Tv, \ldots, T^{m-1}v\}$, if $U = V$, then it is clear that $B_V = \{T^{m-1}v, \ldots, Tv, v\}$ gives a Jordan basis of T.

So we assume that $U \neq V$. Notice that U is invariant under T. Thus, the inductive hypothesis gives that $T|_U$ has a Jordan basis.

Let $\varphi \in V^\vee := \mathcal{L}(V, \mathbb{F})$, the dual space of V, be such that $\varphi(T^{m-1}v) \neq 0$. Define:

$$W := \{v \in V \mid \varphi(T^i v) = 0, \text{ for } 0 \leq i \leq m-1\}.$$

Since $\varphi(T^i v) = 0 \Rightarrow \varphi(T^i T v) = 0$ for $i \in \{0, \ldots, m-1\}$ we see that W is invariant under T. We claim that $V = U \oplus W$.

First, we show that $U \cap W = \{0\}$. Assume by contradiction that there is a non-zero $u \in U \cap W$, then we have

$$u = a_1 v + a_2 T v + \cdots + a_m T^{m-1} v, \tag{22}$$

for some constants $a_i \in \mathbb{F}$. Let $j \geq 1$ be the smallest integer such that $a_j \neq 0$, by applying T^{m-j} to both sides of (22) we obtain:

$$T^{m-j} u = a_j T^{m-1} v.$$

By applying φ we obtain

$$\varphi(T^{m-j} u) = a_j \varphi(T^{m-1} v) \neq 0,$$

which gives that $u \notin W$. Thus, $U \cap W = \{0\}$.

Now consider $S : V \to \mathbb{F}^m$ given by $S(v) := (\varphi(v), \varphi(Tv), \ldots, \varphi(T^{m-1}v))$. Clearly, we have $W = \operatorname{null} S$. By the fundamental theorem of linear algebra we obtain:

$$\dim W = \dim \operatorname{null} S = \dim V - \dim \operatorname{range} S \geq \dim V - m.$$

Thus, we calculate:

$$\dim(U \oplus W) = \dim U + \dim W \geq m + \dim V - m = \dim V,$$

which implies that $U \oplus W = V$.

Since $T|_W : W \to W$ is nilpotent and $\dim W < \dim V$, by the inductive hypothesis we have B_W a Jordan basis of W for $T|_W$. It is clear that $B = B_V \cup B_W$ is a Jordan basis of V for T. This finishes the proof. □

Finally, we have the following.

Theorem 9 (Jordan normal form). *Assume that $\mathbb{F} = \mathbb{C}$ and let $T : V \to V$ be an arbitrary operator. Then there exists a Jordan basis of V for T*

Proof. Let $\lambda_1, \ldots, \lambda_m$ be the distinct eigenvalues of T. By Theorem 7 we have $V = \oplus_{i=1}^m G(\lambda_i, T)$ and each $(T - \lambda_i \operatorname{Id}_V)|_{G(\lambda_i, T)}$ is nilpotent. By Proposition 14 each $(T - \lambda_i \operatorname{Id}_V)|_{G(\lambda_i, T)}$ has a Jordan basis B_i for $i \in \{1, \ldots, m\}$. Then it is clear that $B = B_1 \cup \cdots \cup B_m$ is a Jordan basis of T. □

Remark 35. Theorem 9 is sometimes also refered to as *Jordan canonical form.*

Remark 36. Let $T : V \to V$ be an operator on a complex vector space with minimal polynomial $p_T(x) = \prod_{i=1}^m (x - \lambda_i)^{d_i}$, where $\{\lambda_1, \ldots, \lambda_m\}$ are distinct eigenvalues and $d_i \geq 1$ for $i \in \{1, \ldots, m\}$. As a consequence of Exercise 11 in Worksheet 8 and Exercise 42 we know that the largest Jordan block for the eigenvalue λ_i is d_i.

Example 37. Let $T : \mathbb{C}^6 \to \mathbb{C}^6$ be an operator whose minimal polynomial is $p_T = (x-2)^2(x-3)(x-4)$. Then up to reordering of the blocks the possible Jordan forms of T are:

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

Exercise 43. Make sure you understand Example 37. Calculate the characteristic polynomial for each possibility.

17. Apr. 8, 2024

17.1. **Tensor Products.** There are many different ways to motivate the construction of tensor product. Here is one that I find particularly useful. Let V be a finite-dimensional vector space. Recall that $V^\vee := \mathcal{L}(V, \mathbb{F})$ is the dual vector space. Consider the function:

$$\begin{aligned} \mathrm{ev}_V : V^\vee \times V &\to \mathbb{F} \\ (\varphi, v) &\mapsto \varphi(v). \end{aligned}$$

We notice that this assignment satisfies the following:

$$\mathrm{ev}_V((a\varphi_1+\varphi_2, v)) = a\,\mathrm{ev}_V((\varphi_1, v)) + \mathrm{ev}_V((\varphi_2, v)) \quad \text{and} \quad \mathrm{ev}_V((\varphi, bv_1+v_2)) = \mathrm{ev}_V((\varphi, v_1)) + b\,\mathrm{ev}_V((\varphi, v_2)).$$

However, it is *not* linear. Namely, $V^\vee \times V$ is a vector space, where given $(\varphi, v) \in V^\vee \times V$ we have $a(\varphi, v) = (a\varphi, av)$[17], however:

$$\varphi(av) = a\varphi(v) = a\,\mathrm{ev}_V((\varphi, v)) \neq \mathrm{ev}_V((a\varphi, av)) = (a\varphi)(av) = a^2\varphi(v).$$

Here is a definition that abstracts what the function ev_V actually satisfies:

Definition 50. Given vector spaces U, V and W, a function $B : V \times W \to U$ is a *bilinear map* if

- $B(av_1 + v_2, w) = aB(v_1, w) + B(v_2, w)$ for every $a \in \mathbb{F}$, $v_1, v_2 \in V$ and $w \in W$; and
- $B(v, aw_1 + w_2) = aB(v, w_1) + B(v, w_2)$ for every $a \in \mathbb{F}$, $w_1, w_2 \in V$ and $v \in W$.

We let $\mathcal{B}(V \times W, U)$ denote the set of bilinear maps from $V \times W$ to U.

Remark 37. In fact, the set $\mathcal{B}(V \times W, U)$ is a vector space, with multiplication and addition point-wise, i.e. given B_1, B_2 and $a \in \mathbb{F}$ we let $(aB_1 + B_2)(v, w) := aB_1(v, w) + B_2(v, w)$ for every $v \in V$ and $w \in W$.

Example 38. (i) The function $\mathrm{ev}_V : V^\times \times V \to \mathbb{F}$ from the beginning of this section is a bilinear map;

(ii) Let $\varphi \in V^\vee$ and $\psi \in W^\vee$, then $B(v, w) := \varphi(v)\psi(w)$ is a bilinear map;

(iii) Let $B \in \mathcal{B}(V \times W, \mathbb{F})$, for any $T \in \mathcal{L}(V)$ we have $B'(v, w) := B(Tv, w)$ is another bilinear map.

(iv) The function $(-) \circ (-) : \mathcal{L}(V) \times \mathcal{L}(V) \to \mathcal{L}(V)$ is a bilinear map.

(v) The function $\mathrm{ev}(-, -) : V \times \mathcal{L}(V, W) \to W$ given by

$$\mathrm{ev}(v, T) := T(v)$$

is a bilinear map.

The following exercise is crucially important:

Exercise 44. Let $B : V \times W \to \mathbb{F}$ be a bilinear map. Suppose that B is also linear. Then $B(v, w) = 0$, for every $v \in V$ and $w \in W$.

The question that the tensor product tries to answer is the following:

Question 2. Can we formulate the data of a bilinear map $B : V \times W \to \mathbb{F}$ in terms of a *linear map*?

Notice that Exercise 44 tells us that $V \times W$ is not the correct vector space to consider.

18. Apr. 15, 2024

18.1. **Tensor Products (continued).**

Definition-Construction 1. Given two vector spaces V and W, the tensor product of V and W is a vector space T equipped with a bilinear map: $c : V \times W \to T$ such that for every bilinear map $B : V \times W \to U$ to

[17] Recall how we defined the structure of a vector space on a product of vector spaces.

an arbitrary vector space U, there exists an unique linear map $L : T \to U$ such that the following diagram commutes:

$$
\begin{tikzcd}
V \times W \arrow[d, "c"'] \arrow[dr, "B"] & \\
T \arrow[r, dashed, "L"'] & U
\end{tikzcd}
. \tag{23}
$$

The idea to construct T is to build a very large vector space out of $V \times W$ and then to take its quotient by the relations that we expect to hold for a linear map out of it to always exist.

Let $T' := \oplus_{(v,w)\in V\times W}\mathbb{F}_{(v,w)}$, where each $\mathbb{F}_{(v,w)}$ is the 1-dimensional vector $\mathbb{F}$. Let $\delta_{(v,w)} \in T'$ denote the vector that is 1 in $\mathbb{F}_{(v,w)}$ and 0 in every other copy $\mathbb{F}_{(v',w')}$, where $(v',w') \neq (v,w)$. Consider the set of vectors:[18]

(24)

$$S := \{\delta_{v+v',w} - \delta_{v,w} - \delta_{v',w}, \delta_{v,w+w'} - \delta_{v,w} - \delta_{v,w'}, \delta_{av,w} - a\delta_{v,w}, \delta_{v,aw} - a\delta_{v,w}, \delta_{av,w} - \delta_{v,aw}\}_{a\in\mathbb{F}\ v,v'\in V\ w,w'\in W}.$$

Let $R := \operatorname{Span} S \subseteq T'$ be the subspace spanned by all the vectors in (24). We claim that $T := T'/R$ satisfy the conditions to be a tensor product.

For any $v \in V$ and $w \in W$ we let $v \otimes w := \delta_{v,w} + R \in T$, i.e. the image of $\delta_{v,w} \in T'$ via the quotient map $\pi : T' \to T$.

Firstly, we notice that by construction we have the relations:

$$(v+v')\otimes w = v\otimes w + v'\otimes w,\ v\otimes(w+w') = v\otimes w + v\otimes w',\ (av)\otimes w = a(v\otimes w) = (v\otimes aw). \tag{25}$$

Secondly, we claim that the function $c : V \times W \to T$ given by $c(v,w) := v \otimes w$ is bilinear. We leave this for the reader to check!

Finaly, we prove the universal property. Let $L' : T' \to U$ be the linear map defined by $L'(\delta_{v,w}) := B(v,w)$, since $\{\delta_{v,w}\}_{(u,v)\in U\times V}$ is a basis of T' this is well-defined. We notice that L' vanishes on each element of (24). For instance, we have:

$$L'(\delta_{v+v',w} - \delta_{v,w} - \delta_{v',w}) = B(v+v',w) - B(v,w) - B(v',w) = 0.$$

One also needs to check the other relations, we leave that as an exercise.

Thus, $R \subseteq \operatorname{null} L'$ and Lemma 20 implies that we have map L such that the following diagram commutes:

$$
\begin{tikzcd}
T' \arrow[r, "\pi"] \arrow[dr, "L'"'] & T \arrow[d, "L"] \\
& U
\end{tikzcd}
.
$$

Now we notice that $L \circ c(v,w) = L(v \otimes w) = B(v,w)$, that is $c : V \times W \to T$ together with $L : T \to U$ make the diagram (23) commute.

We only need to argue that $L : T \to U$ is the unique linear map that makes the diagram (23) commute. Assume there exists $M : T \to U$ another linear map making the diagram (23) (with M in place of L) commute, by considering the image of (v,w) we would have $M(v\otimes w) = B(v,w) = L(v\otimes w)$. Notice that $\{\delta_{(v,w)}\}_{(v,w)\in V\times W}$ is a basis of T', by Exercise 29. Since $\pi : T' \to T$ is surjective, it is clear that the set of $v \otimes w = \pi(\delta_{v,w})$, for $(v,w) \in V \times W$ spans T. Thus, if $M(v\otimes w) = L(v\otimes w)$ for every $v\otimes w$ then $M = L$.

We leave it to the reader[19] to prove that T is unique up to isomorphism, i.e. any other vector space with the same property will be isomorphic to T. The idea is the same as in the proof that any direct sum or product of two vector spaces is isomorphic.

Notation 8. Given two vector spaces V and W we will simply denote by $V \otimes W$ their tensor product. Normally, the bilinear map from $V \times W$ to $V \otimes W$ is also denoted by $\otimes$. We will avoid this and denote it by:

$$\eta_{V,W} : V \times W \to V \otimes W.$$

[18]Strictly speaking, since we are working over a field of characteristic 0, these vector are not all linearly independent, so we could have considered a smaller set. But for the more general case we need all these relations.

[19]You can find a proof in the page 6 of the following notes. The set up of these notes is more general than our course, since they treat R-modules for a commutative ring R. However, they are a very good read for this material.

The description of Definition-Construction 1 is very abstract and not necessarily useful in practice. The thing that is useful is the universal property, as examplified in the proof of the following result.

Lemma 65. *Let V and W be vector spaces and $\{v_i\}_{i\in I}$ and $\{w_j\}_{j\in J}$ bases of V and W, respectively. Then $\{v_i \otimes w_j\}_{(i,j)\in I\times J}$ is basis of $V\otimes W$. In particular, if V and W are finite-dimensional, then we have $\dim V\otimes W = \dim V \cdot \dim W$.*

Proof. Since $\{v\otimes w\}_{(v,w)\in V\times W}$ is a spanning set of $V\otimes W$, it is easy to see that $\{v_i\otimes w_j\}_{(i,j)\in I\times J}$ is a spanning set (Make sure you understand why!).

We will prove that the set $\{v_i\otimes w_j\}_{(i,j)\in I\times J}$ is linearly independent. Assume that there are coefficients $a_{i,j}\in\mathbb{F}$ not all zero, such that $\sum_{i,j} a_{i,j} v_i\otimes w_j = 0$. Consider $(i',j')\in I\times J$ and define the following bilinear function $B_{i'j'} : V\times W\to\mathbb{F}$ by

$$B_{i'j'}(v,w) = c_{i'}d_{j'},$$

where $v=\sum_{i\in I} c_i v_i$ and $w = \sum_{j\in J} d_j w_j$ are expansions of v and w in the given bases. By the universal property, there exists an unique map $f_{i'j'}$ such that the following diagram commutes:

$$\begin{tikzcd} V\times W \arrow[r, "\eta_{V,W}"] \arrow[dr, "B_{i'j'}"'] & V\otimes W \arrow[d, "f_{i'j'}"] \\ & \mathbb{F} \end{tikzcd}.$$

In particular, we obtain that

$$f_{i'j'}(v_i\otimes w_j) = B_{i'j'}((v_i,w_j)) = \begin{cases} 1 & \text{if } (i,j)=(i',j') \\ 0 & \text{else.}\end{cases}$$

Since $f_{i'j'}$ is linear we obtain

$$f_{i'j'}(\sum_{i,j} a_{i,j} v_i\otimes w_j) = a_{i,j} f_{i'j'}(\sum_{i,j} v_i\otimes w_j) = a_{i',j'} = 0.$$

Since $(i',j')\in I\times J$ was arbitrary, we obtain that $a_{i,j}=0$ for every $(i,j)\in I\times J$. This finishes the proof. □

Example 39. (i) Let $V=\mathbb{C}^3$, then $V\otimes V\simeq\mathbb{C}^9$, with basis given by $e_i\otimes e_j$, where $i,j\in\{1,2,3\}$ and $\{e_1,e_2,e_3\}$ is the standard basis of $\mathbb{C}^3$.

(ii) Let V be a finite-dimensional vector space, then $V^\vee\otimes V\simeq\mathcal{L}(V)$. Indeed, let $\{e_1,\dots,e_n\}$ be a basis of V, Lemma 21 gives $\{\lambda_1,\dots,\lambda_n\}$ a basis of the dual vector space $V^\vee$. Consider the linear map $\Phi_V : V^\vee\otimes V\to\mathcal{L}(V)$ defined on the basis $\{\lambda_i\otimes e_j\}_{i,j\in I}$ by

$$\Phi_V(\lambda_i\otimes e_j)(v) = \lambda_i(v)e_j.$$

We claim that Φ_V is a isomorphism. Since the dimensions are the same it is enough to check that Φ_V is surjective. Let $T: V\to V$ be an operator, and let $T(e_k)=\sum_{j=1}^n a_{kj}e_j$ for some coefficients $(a_{kj})_{1\le k,j\le n}$. We claim that $T=\Phi_V(\sum_{i,j=1}^n a_{ij}\lambda_i\otimes e_j)$. Indeed, we compute:

$$\begin{aligned}\Phi_V(\sum_{i,j=1}^n a_{ij}\lambda_i\otimes e_j)(e_k) &= \sum_{i,j=1}^n a_{ij}\Phi_V(\lambda_i\otimes e_j)(e_k)\\ &= \sum_{i,j=1}^n a_{ij}\lambda_i(e_k)e_j\\ &= \sum_{i,j=1}^n a_{ij}\delta_{i,k}e_j\\ &= \sum_{j=1}^n a_{kj}e_j.\end{aligned}$$

Thus, $T(e_k)=\Phi_V(\sum_{i,j=1}^n a_{ij}\lambda_i\otimes e_j)(e_k)$ for every $k\in\{1,\dots,n\}$, which gives that $T=\Phi_V(\sum_{i,j=1}^n a_{ij}\lambda_i\otimes e_j)$ as desired.

(iii) Let $f : V_1 \to V_2$ and $g : W_1 \to W_2$ be linear maps between vector spaces. Then there exists an unique linear map $f \otimes g : V_1 \otimes W_1 \to V_2 \otimes W_2$. Indeed, consider the function:

$$B_{f,g} : V_1 \times W_1 \to V_2 \otimes W_2, \quad B_{f,g}((v_1, w_1)) := f(v_1) \otimes g(w_1).$$

We leave it to the reader to check that $B_{f,g}$ is bilinear, hence by the universal property there exist $f \otimes g$ as claimed. Moreover, for each $v_1 \in V_1$ and $w_1 \in W_1$, we have $(f \otimes g)(v_1 \otimes w_1) = f(v_1) \otimes g(w_1)$.

We go back to the motivating question in this section. We notice that $\mathrm{ev}_V : V^\vee \times V \to \mathbb{F}$ determines a linear map[20]:

$$\mathrm{ev}_V : V^\vee \otimes V \to \mathbb{F}.$$

Exercise 45. Let V be a vector space over $\mathbb{F}$ consider $B : \mathbb{F} \times V \to V$ given by $B(a, v) := av$.

(i) Check that B is bilinear and prove that the map $l_V : \mathbb{F} \otimes V \to V$ is an isomorphism.

(ii) Similarly prove that there exists an isomorphism $r_V : V \otimes \mathbb{F} \overset{\simeq}{\to} V$.

(iii) For every vector space W over $\mathbb{F}$. Consider the morphism $\sigma'_{V,W}$ defined as the composite:

$$V \times W \xrightarrow{S_{V,W}} W \times V \xrightarrow{\eta_{w,V}} W \otimes V.$$

Prove that $\sigma'_{V,W}$ is bilinear and that the induced linear map $\sigma_{V,W} : V \otimes W \xrightarrow{\sim} W \otimes V$ is an isomorphism.

(iv) What is the compatibility between l_V, r_V and $\sigma_{V,\mathbb{F}}$?

Notation 9. Given V a vector and $L : \mathbb{F} \to W$ a linear map between $\mathbb{F}$ and a vector space W. In view of Exercise 45 and Exercise 8 from Worksheet 9, we will simply write $\mathrm{Id}_V \otimes a : V \to V \otimes W$ for the composite:

$$V \xrightarrow[\simeq]{r_V^{-1}} V \otimes \mathbb{F} \xrightarrow{\mathrm{Id}_V \otimes a} V \otimes W,$$

and similarly $a \otimes \mathrm{Id}_V : V \to W \otimes V$ for the composite:

$$V \xrightarrow[\simeq]{l_V^{-1}} \mathbb{F} \otimes V \xrightarrow{a \otimes \mathrm{Id}_V} W \otimes V.$$

Given $b : W \to \mathbb{F}$ a linear map, we will also simply write $b \otimes \mathrm{Id}_V$ for the composite $W \otimes V \overset{b \otimes \mathrm{Id}_V}{\to} \mathbb{F} \otimes V \overset{\simeq}{\to} V$ and so on.

19. Apr. 18, 2024

19.1. **Further properties of tensor products.** Let U, V and W be three vector spaces. Given $T : V \to \mathcal{L}(V, W)$, we let $\varphi(T) : V \otimes W \to U$ be the linear map induced by the bilinear map:

$$\varphi'(T) : V \times W \to U \qquad \varphi'(T)(v, w) := T(v)(w).$$

Indeed, notice that $\varphi'(T)(av_1 + v_2, w) = T(av_1 + v_2)(w) = (aT(v_1) + T(v_2))(w) = aT(v_1)(w) + T(v_2)(w) = a\varphi'(T)(v_1, w) + \varphi'(T)(v_2, w)$, and similarly $\varphi'(T)(v, aw_1 + w_2) = a\varphi'(T)(v, w_1) + \varphi'(T)(v, w_2)$. So $\varphi'(T)$ is bilinear and induces $\varphi(T) : V \otimes W \to U$ which satisfies $\varphi(T)(v \otimes w) = T(v)(w)$ for every $v \otimes w \in V \otimes W$.

Lemma 66. *Let U, V and W be three vector spaces, the linear map constructed above:*

$$\varphi : \mathcal{L}(V, \mathcal{L}(W, U)) \overset{\simeq}{\to} \mathcal{L}(V \otimes W, U),$$
$$(V \overset{T}{\to} \mathcal{L}(V, W)) \mapsto \varphi(T),$$

is an isomorphism.

[20] We will still use the same notation for ev_V.

Proof. First we check that φ is injective. Let $T, T' \in \mathcal{L}(V, \mathcal{L}(W, U))$ such that $\varphi(T) = \varphi(T')$. That is for every $v \otimes w \in V \otimes W$ we have:

$$T(v)(w) = \varphi(T)(v \otimes w) = \varphi(T')(v \otimes w) = T'(v)(w).$$

Thus, $T(v) = T(v')$ for every $v \in V$, which implies that $T = T'$.

Now consider a linear map $S : V \otimes W \to U$. We let $\Phi(S) : V \to \mathcal{L}(W, U)$ be defined by:

$$\Phi(S)(v)(w) := S(v \otimes w).$$

We notice that $\Phi(S)(v)(aw_1 + w_2) = a\Phi(S)(v)(w_1) + \Phi(S)(v)(w_2)$ and $\Phi(S)(av_1 + v_2)(w) = a\Phi(S)(v_1)(w) + \Phi(S)(v_2)(w)$ so $\Phi(S)$ is well-defined. Finally, we leave it to the reader to check that

$$\varphi(\Phi(S)) = S,$$

which gives that φ is surjective. This finishes the proof. $\square$

Remark 38. This is an abstract remark and you can skip it for now. The statement of Lemma 66 is sometimes referred to as the Hom-$\otimes$-adjunction. It holds in my more general contexts and it is also *functorial*, e.g. given a linear map $T : V_1 \to V_2$ we obtain linear maps $(-) \circ T : \mathcal{L}(V_2, \mathcal{L}(W, U)) \to \mathcal{L}(V_1, \mathcal{L}(W, U))$ and $(-) \circ T \otimes \mathrm{Id}_W : \mathcal{L}(V_2 \otimes W, U) \to \mathcal{L}(V_1 \otimes W, U)$ such that the following diagram commutes:

$$\begin{tikzcd}
\mathcal{L}(V_2, \mathcal{L}(W, U)) \arrow[r, "\varphi_{V_2, W, U}"] \arrow[d, "(-)\circ T"'] & \mathcal{L}(V_2 \otimes W, U) \arrow[d, "(-)\circ T \otimes \mathrm{Id}_W"] \\
\mathcal{L}(V_1, \mathcal{L}(W, U)) \arrow[r, "\varphi_{V_1, W, U}"'] & \mathcal{L}(V_1 \otimes W, U)
\end{tikzcd}$$

The following is a result about the interaction between tensor products and dual spaces.

Lemma 67. *Let U and V be finite-dimensional vector spaces. Then one has an isomorphism*

$$U^\vee \otimes V^\vee \xrightarrow{\simeq} (U \otimes V)^\vee.$$

Proof. Let $\{u_1, \ldots, u_m\} \subset U$ and $\{v_1, \ldots, v_n\} \subset V$ be basis of U and V, respectively. We define: $B : U^\vee \times V^\vee \to (U \otimes V)^\vee$ as a linear map by:

$$B(\lambda, \mu)(u_i \otimes v_j) := \sum_{i=1}^{n} \lambda(u_i)\mu(v_j),$$

since $\{u_i \otimes v_j\}_{(i,j) \in [n] \times [m]}$[21] is a basis of $U \otimes V$. It is easy to see that B is a bilinear map, hence by the universal property of $U^\vee \otimes V^\vee$ (Definition-Construction 1) we obtain the linear map: $L : U^\vee \otimes V^\vee \to (U \otimes V)^\vee$. $\square$

Corollary 22. *Let V and U be finite-dimensional vector spaces, then we have*

$$\mathcal{L}(U, V) \xrightarrow{\sim} U^\vee \otimes V.$$

Proof. For (i), consider the chain of isomorphisms:

$$U^\vee \otimes V \xrightarrow{\simeq} (U \otimes V^\vee)^\vee \xrightarrow{=} \mathcal{L}(U \otimes V^\vee, \mathbb{F}) \xrightarrow[\varphi^{-1}_{U, (V^\vee)^\vee, \mathbb{F}}]{\simeq} \mathcal{L}(U, (V^\vee)^\vee) \xrightarrow{\simeq} \mathcal{L}(U, V),$$

where the first isomorphism is from Lemma 67, the second from Lemma 66, and the third from Exercise 32. $\square$

[21]Here we use $[n] = \{1, \ldots, n\}$ and $[m] = \{1, \ldots, m\}$.

19.2. **Trace of an operator.** The goal of this section is to define the trace of an operator on a finite-dimensional vector space V without refereeing to a basis of V. Given V a finite-dimensional vector space and let $B_V = \{v_1, \dots, v_n\}$ be a basis of V. We define the linear map:

$$\begin{aligned} c_V^{B_V} : \mathbb{F} &\to V^\vee \otimes V \\ a &\mapsto a \sum_{i=1}^n v_i^\vee \otimes v_i, \end{aligned}$$

where $\{v_1^\vee, \dots, v_n^\vee\}$ denotes the dual basis associated to B_V.

Let $B'_V = \{u_1, \dots, u_n\}$ be another basis of V. We claim that $c_V^{B'_V} = c_V^{B_V}$. Indeed, by using Example 39 (ii) we notice that $\Phi_V \circ \mathrm{coev}_V^{B_V}(a) : V \to V$ is simply

$$\Phi_V \circ c_V^{B_V}(a)(v) = av,$$

for any basis B_V. Since Φ_V is an isomorphism, we obtain that $c_V^{B_V}$ is independent of the basis. Henceforth we drop B_V from the notation and simply write $c_V : \mathbb{F} \to V^\vee \otimes V$.

The following result on the interaction between the coevaluation and evaluation maps is very useful:

Lemma 68. *Let $T : V \to W$ be a linear map between finite-dimensional vector spaces, then the following diagram commutes:*

$$\begin{tikzcd} V \arrow[r, "\mathrm{id}_V"] \arrow[ddd, "T"'] & V \arrow[d, "\mathrm{id}_V \otimes c_V"] \\ & V \otimes V^\vee \otimes V \arrow[d, "\mathrm{id}_{V \otimes V^\vee} \otimes T"] \\ & V \otimes V^\vee \otimes W \arrow[d, "\mathrm{ev}'_V"] \\ W \arrow[r, "\mathrm{id}_W"'] & W \end{tikzcd}$$

where $\mathrm{ev}'_V : V \otimes V^\vee \to \mathbb{F}$ is given by $\mathrm{ev}'_V(v \otimes \varphi) = \varphi(V)$.

Proof. Left as an exercise. □

Definition 51. Let $T \in \mathcal{L}(V)$ be an operator on a finite-dimensional vector space. We let $\mathrm{tr}(T)$ be the composite of the following linear maps:

$$\mathbb{F} \overset{c_V}{\to} V^\vee \otimes V \overset{\mathrm{Id}_{V^\vee} \otimes T}{\to} V^\vee \otimes V \overset{\mathrm{ev}_V}{\to} \mathbb{F},$$

i.e. any linear map from $\mathbb{F} \to \mathbb{F}$ is given by multiplication by a scalar, $\mathrm{tr}(T)$ is that scalar.

Lemma 69. *Let B_V be a basis of a finite-dimensional vector space V and $T : V \to V$ an operator. Then*

$$\mathrm{tr}(T) = \sum_{i=1}^n \mathcal{M}(T, B_V)_{i,i}.$$

Proof. Let $B_V = \{v_1, \dots, v_n\}$ and $\{v_1^\vee, \dots, v_n^\vee\}$ be the dual basis, we compute:

$$\begin{aligned}
\operatorname{tr}(V)(a) &= \operatorname{ev}_V \circ (\operatorname{Id}_{V^\vee} \otimes T)(a \sum_{i=1}^n v_i^\vee \otimes v_i) \\
&= \operatorname{ev}_T(a \sum_{i=1}^n v_i^\vee \otimes T(v_i)) \\
&= \operatorname{ev}_T(a \sum_{i=1}^n \sum_{j=1}^n v_i^\vee \otimes \mathcal{M}(T, B_V)_{j,i} v_j) \\
&= \sum_{i,j=1}^n \mathcal{M}(T, B_V)_{j,i} \operatorname{ev}_T(v_i^\vee \otimes v_j) \\
&= \sum_{i,j=1}^n \mathcal{M}(T, B_V)_{j,i} \delta_{i,j} \\
&= \sum_{i=1}^n \mathcal{M}(T, B_V)_{i,i}.
\end{aligned}$$

□

The following is an interesting way to think about the trace.

Lemma 70. *Suppose that for every finite-dimensional vector space V, we have a linear function $\tau_V : \mathcal{L}(V) \to \mathbb{F}$, which satisfies the following conditions:*

a) for every V and W and pair of linear maps $T : V \to W$ and $S : W \to V$ we have:

$$\tau_V(S \circ T) = \tau_W(T \circ S);$$

b) $\tau_V(\operatorname{Id}_V) = \dim V$ for every V.

Then

(i) for every decomposition $U_1 \oplus U_2 = V$ we have $\operatorname{tr}_V(T) = \operatorname{tr}_V(P_{U_1} T P_{U_1}) + \operatorname{tr}_V(P_{U_2} T P_{U_2})$, where $P_i : V \to V$ is the projection onto U_i, for $i = 1, 2$;

(ii) if such a data $\{\tau_V\}_V$ exist, it is unique.

Proof. We first prove (i). First we notice that $\operatorname{tr}_V(P_{U_i} T \circ P_{U_i}) = \operatorname{tr}_V(T \circ P_{U_i}^2)$ by a), and that $\operatorname{tr}_V(T \circ P_{U_i}^2) = \operatorname{tr}_V(T \circ P_{U_i})$, since any projection satisfy $P_{U_1}^2 = P_{U_i}$, for $i = 1, 2$. Moreover, one has $P_2 = \operatorname{Id}_V - P_1$. Since τ_V is linear, we obtain:

$$\tau_V(T \circ \operatorname{Id}_V) = \tau_V(T \circ P_{U_1}) + \tau_V(T \circ P_{V_2}) = \operatorname{tr}_V(P_{U_1} T P_{U_1}) + \operatorname{tr}_V(P_{U_2} T P_{U_2}).$$

We now prove (ii). Suppose that there are two τ_V, τ'_V for all vector spaces V satisfying both a) and b). We argue by induction on $n = \dim V$ that $\tau_V = \tau'_V$. For $n = 1$, since any linear operator $T \in \mathcal{L}(V)$ is $T = a \operatorname{Id}_V$ for some $a \in \mathbb{F}$ then we have:

$$\tau_V(T) = a\tau_V(\operatorname{Id}_V) = a \dim V = a\tau'_V(\operatorname{Id}_V) = \tau'_V(T).$$

Now we assume that $\tau_W = \tau'_W$ for all vector spaces W, such that $\dim W < n$. Let V be a subspace of dimension n and $U_1 \oplus U_2 = V$ and decomposition into non-zero subspaces U_1 and U_2. Let $V \overset{\overline{P_{U_i}}}{\to} U \overset{\imath_i}{\to} V$ denote the factorization of P_{U_i} via U_i, for $i = 1, 2$. We notice that $T \circ P_U = (T \circ \imath_i) \circ \overline{P_{U_i}}$ agree. Thus, we obtain:

$$\begin{aligned}
\tau_V(T) &= \tau_V(TP_{U_1}) + \tau_V(TP_{U_2}) \\
&= \tau_{U_1}(\overline{P_{U_1}} \circ (T \circ \imath_1)) + \tau_{U_2}(\overline{P_{U_2}} \circ (T \circ \imath_2)) \\
&= \tau'_{U_1}(\overline{P_{U_1}} \circ (T \circ \imath_1)) + \tau'_{U_2}(\overline{P_{U_2}} \circ (T \circ \imath_2)) \\
&= \tau'_V(TP_{U_1}) + \tau'_V(TP_{U_2}) \\
&= \tau'_V(T).
\end{aligned}$$

□

We end this section checking that the trace we constructed in Definition 51 is cyclic, namely it satisfies condition a) from Lemma 70.

Theorem 10 (Cyclicity)**.** *Let $T : V \to W$ and $S : W \to V$ be two linear maps between finite-dimensional vector spaces, then* $\mathrm{tr}(S \circ T) = \mathrm{tr}(T \circ S)$*.*

Proof. The proof is contained in the following two diagrams. For better readability we omit certain the horizontal arrows, they should all be identity arrows and the fact that the respective diagrams commute is tautological.

$$\begin{tikzcd}
\mathbb{F} \arrow[r, "\mathrm{id}"] \arrow[d, "c_V"'] & \mathbb{F} \arrow[r, "\mathrm{id}"] \arrow[d, "c_V"] & \mathbb{F} \arrow[r, "\mathrm{id}"] \arrow[d, "c_V"] & \mathbb{F} \arrow[d, "c_V"] \\
V^\vee \otimes V \arrow[d, "\mathrm{id}_{V^\vee} \otimes T"'] & V^\vee \otimes V \arrow[r, "\mathrm{id}"] \arrow[d, "\mathrm{id}_{V^\vee} \otimes T"] & V^\vee \otimes V \arrow[d, "\mathrm{id}_{V^\vee \otimes V} \otimes c_W"] & V^\vee \otimes V \arrow[d, "\mathrm{id}_{V^\vee \otimes V} \otimes c_W"] \\
V^\vee \otimes W \arrow[r, "\mathrm{id}"] \arrow[ddd, "\mathrm{id}_{V^\vee} \otimes S"'] & V^\vee \otimes W \arrow[d, "\mathrm{id}_{V^\vee} \otimes c_W"] & V^\vee \otimes V \otimes W^\vee \otimes W \arrow[r, "\mathrm{id}"] \arrow[d, "\mathrm{id}_{V^\vee} \otimes T \otimes \mathrm{id}_{\otimes W \otimes W^\vee}"] & V^\vee \otimes V \otimes W^\vee \otimes W \arrow[d, "\mathrm{id}_{V^\vee \otimes V \otimes W^\vee} \otimes S"] \\
 & V^\vee \otimes W \otimes W^\vee \otimes W \arrow[r, "\mathrm{id}"] \arrow[d, "\mathrm{id}_{V^\vee \otimes W \otimes W^\vee} \otimes S"] & V^\vee \otimes W \otimes W^\vee \otimes W \arrow[d, "\mathrm{id}_{V^\vee \otimes W \otimes W^\vee} \otimes S"] & V^\vee \otimes V \otimes W^\vee \otimes V \arrow[d, "\mathrm{id}_{V^\vee} \otimes T \otimes \mathrm{id}_{\otimes W \otimes W^\vee}"] \\
 & V^\vee \otimes W \otimes W^\vee \otimes V \arrow[d, "\mathrm{id}_{V^\vee} \otimes e'_W \otimes \mathrm{id}_V"] & V^\vee \otimes W \otimes W^\vee \otimes V \arrow[r, "\mathrm{id}"] \arrow[d, "\mathrm{id}_{V^\vee} \otimes e'_W \otimes \mathrm{id}_V"] & V^\vee \otimes W \otimes W^\vee \otimes V \arrow[d, "\mathrm{id}_{V^\vee} \otimes e'_W \otimes \mathrm{id}_V"] \\
V^\vee \otimes V \arrow[r, "\mathrm{id}"] \arrow[d, "\mathrm{ev}_V"'] & V^\vee \otimes V \arrow[d, "\mathrm{ev}_V"] & V^\vee \otimes V \arrow[d, "\mathrm{ev}_V"] & V^\vee \otimes V \arrow[d, "\mathrm{ev}_V"] \\
\mathbb{F} \arrow[r, "\mathrm{id}"] & \mathbb{F} \arrow[r, "\mathrm{id}"] & \mathbb{F} \arrow[r, "\mathrm{id}"] & \mathbb{F}
\end{tikzcd}$$

In the diagram bellow, on the third line, the horizontal map between the first and second column $\sigma : V^\vee \otimes V \otimes W^\vee \otimes W \to W^\vee \otimes W \otimes V^\vee \otimes V$ is given by $\sigma(v^\vee, v, w^\vee, w) \mapsto (w^\vee, w, v^\vee, v)$. To simplify the notation we use σ for the similarly defined map between the first and second column on the fourth and fifth lines.

$$\begin{tikzcd}
\mathbb{F} \arrow[r, "\mathrm{id}"] \arrow[d, "c_V"'] & \mathbb{F} \arrow[r, "\mathrm{id}"] \arrow[d, "c_W"] & \mathbb{F} \arrow[r, "\mathrm{id}"] \arrow[d, "c_W"] & \mathbb{F} \arrow[d, "c_W"] \\
V^\vee \otimes V \arrow[d, "\mathrm{id}_{V^\vee \otimes V} \otimes c_W"'] & W^\vee \otimes W \arrow[r, "\mathrm{id}"] \arrow[d, "\mathrm{id}_{W^\vee \otimes W} \otimes c_V"] & W^\vee \otimes W \arrow[d, "\mathrm{id}_{W^\vee} \otimes S"] & W^\vee \otimes W \arrow[d, "\mathrm{id}_{W^\vee} \otimes S"] \\
V^\vee \otimes V \otimes W^\vee \otimes W \arrow[r, "\sigma"] \arrow[d, "\mathrm{id}_{V^\vee \otimes V \otimes W^\vee} \otimes S"'] & W^\vee \otimes W \otimes V^\vee \otimes V \arrow[d, "\mathrm{id}_{W^\vee} \otimes S \otimes \mathrm{id}_{V^\vee \otimes V}"] & W^\vee \otimes V \arrow[r, "\mathrm{id}"] \arrow[d, "\mathrm{id}_{W^\vee \otimes V} \otimes c_V"] & W^\vee \otimes V \arrow[ddd, "\mathrm{id}_{W^\vee} \otimes T"] \\
V^\vee \otimes V \otimes W^\vee \otimes V \arrow[r, "\sigma"] \arrow[d, "\mathrm{id}_{V^\vee} \otimes T \otimes \mathrm{id}_{W^\vee \otimes V}"'] & W^\vee \otimes V \otimes V^\vee \otimes V \arrow[r, "\mathrm{id}"] \arrow[d, "\mathrm{id}_{W^\vee \otimes V \otimes V^\vee} \otimes T"] & W^\vee \otimes V \otimes V^\vee \otimes V \arrow[d, "\mathrm{id}_{W^\vee \otimes V \otimes V^\vee} \otimes T"] & \\
V^\vee \otimes W \otimes W^\vee \otimes V \arrow[r, "\sigma"] \arrow[d, "\mathrm{id}_{V^\vee} \otimes e'_W \otimes \mathrm{id}_V"'] & W^\vee \otimes V \otimes V^\vee \otimes W \arrow[d, "\mathrm{id}_{W^\vee} \otimes \mathrm{ev}'_V \otimes \mathrm{id}_W"] & W^\vee \otimes V \otimes V^\vee \otimes W \arrow[d, "\mathrm{id}_{W^\vee} \otimes \mathrm{ev}'_V \otimes \mathrm{id}_W"] & \\
V^\vee \otimes V \arrow[d, "\mathrm{ev}_V"'] & W^\vee \otimes W \arrow[d, "\mathrm{ev}_W"] & W^\vee \otimes W \arrow[r, "\mathrm{id}"] \arrow[d, "\mathrm{ev}_W"] & W^\vee \otimes W \arrow[d, "\mathrm{ev}_W"] \\
\mathbb{F} \arrow[r, "\mathrm{id}"] & \mathbb{F} \arrow[r, "\mathrm{id}"] & \mathbb{F} \arrow[r, "\mathrm{id"] & \mathbb{F}
\end{tikzcd}$$

We claim that all the above diagrams commute. We leave it to the reader to justify each one of them.

Notice that the composition of the left most column in the first diagram is $\mathrm{tr}(S \circ T)$, whereas the composition of the right most column in the second diagram is $\mathrm{tr}(T \circ S)$. $\square$

Exercise 46. Formulate which diagrams need to commute in the proof of Theorem 10 and prove that each of them commute. In fact, there are only three different type of abstract commuting diagrams that are needed to justify all the diagrams in the proof, one of them is in Lemma 68, so you only need to formulate and prove the other two.

Remark 39. Here are some comments about the proof above. If one traces what happens to an element $a \in \mathbb{F}$ via all of the above diagrams we will obtain equations proving $\operatorname{tr}(TS) = \operatorname{tr}(ST)$ in terms of matrix coefficients. The above proof has the advantage of never using matrices, however it is a bit abstract and hard to see what is happening. A nice perspective on this comes from thinking of the diagrams we wrote in terms of pictures. The following two blog posts explains this very nicely.

19.3. **Bilinear Forms.** The following is a special case of Definition 50:

Definition 52. Let V be a vector space over $\mathbb{F}$. A *bilinear form* on V is a bilinear map $B : V \times V \to \mathbb{F}$ in the sense of Definition 50.

Example 40.
 (i) An inner product on a finite-dimensional *real* vector space $\langle -, - \rangle : V \times V \to \mathbb{F}$ is a bilinear form.
 (ii) Let $\beta : \mathbb{F}^3 \times \mathbb{F}^3 \to \mathbb{F}$ be given by $\beta((x_1, x_2, x_3), (y_1, y_2, y_3)) = x_1y_2 - 5x_2y_3 + 2x_3y_1$.
 (iii) Let $\beta : \mathcal{P}_n(\mathbb{R}) \times \mathcal{P}_n(\mathbb{R}) \to \mathbb{R}$ be given by $\beta(p, q) = p(2)q'(3)$.
 (iv) Let $\varphi, \lambda \in V^\vee = \mathcal{L}(V, \mathbb{F})$ be two linear functions on V. Then $\beta(v, u) := \varphi(v)\lambda(u)$ is a bilinear form.

Exercise 47.
 (i) Check that all the examples in Example 40 are not linear.
 (ii) Let $\beta : V \times V \to \mathbb{F}$ be a bilinear form and linear functional on V. Prove that $\beta = 0$.

We can consider:

Definition 53. Given a vector space V, the set $\mathcal{L}^2(V)$[22] of bilinear forms on V is a vector space.

Warning 9. Make sure that you understand why the set $\mathcal{L}^2(V)$ is a vector space and that you understand why it is different from $\mathcal{L}(V \times V, \mathbb{F})$. This is extremely important!

Exercise 48.
 (i) Let $\beta \in \mathcal{L}^2(V)$, prove that, for every $v \in V$, $\beta(v, -)$ and $\beta(-, v)$ are linear maps.
 (ii) Consider the function:
$$\begin{aligned} \Phi_1 : \mathcal{L}^2(V) &\to \mathcal{L}(V, \mathcal{L}(V, \mathbb{F})) \\ \Phi_1(\beta)(v) &:= \beta(v, -). \end{aligned}$$
 Prove that Φ_1 is an isomorphism.

Remark 40. Let B_V be a basis of V and $\beta \in \mathcal{L}^2(V)$ a bilinear form. The *matrix of* β in the basis $B_V = \{v_1, \dots, v_n\}$ is defined by:
$$(\mathcal{M}(\beta, B_V))_{ij} := \beta(v_i, v_j).$$
In fact, the assignment $\mathcal{M}(-, B_V) : \mathcal{L}^2(V) \to \mathbb{F}^{\dim V, \dim V}$ is an isomorphisms of vector spaces. Indeed, it is clear that $\mathcal{M}(-, B_V)$ is a linear map. Assume that $\mathcal{M}(\beta, B_V) = \mathcal{M}(\beta', B_V)$, then for every $u_1, u_2 \in V$ we let $u_j = \sum_{i=1}^n a_i^j v_i$, be the expression of u_1 and u_2 in the basis B_V. Then we have:
$$\begin{aligned} \beta(u_1, u_2) &= \sum_{i,k=1}^n \mathcal{M}(\beta, B_V)_{ik} a_i^1 a_k^2 \\ &= \sum_{i,k=1}^n \mathcal{M}(\beta, B_V)_{ik} a_i^1 a_k^2 \\ &= \beta'(u_1, u_2). \end{aligned}$$
Thus, $\mathcal{M}(-, B_V)$ is injective. We can similarly prove that $\mathcal{M}(-, B_V)$ is surjective by defining β on a basis using $(b_{ik})_{1 \leq i,k \leq n} \in \mathbb{F}^{\dim V, \dim V}$.

In particular, we obtain that $\dim \mathcal{L}^2(V) = n^2$.

[22]The book uses the notation $V^{(2)}$, which I find confusing–but there does not seem to be a consistent notation in the literature.

The following are some things we can prove by using the matrix representation of a bilinear form introduced in Remark 40.

Lemma 71. *Let $\beta : V \times V \to \mathbb{F}$ be a bilinar form and $T : V \to V$ an operator. We define two linear forms $\beta_1 := \beta \circ T \times \mathrm{Id}_V$ and $\beta_2 := \beta \circ \mathrm{Id}_V \times T$. Consider a basis $B_V = \{v_1, \dots, v_n\}$, then we have:*

(i) $\mathcal{M}(\beta_1, B_V) = \mathcal{M}(T)^{\mathrm{t}}\mathcal{M}(\beta)$;

(ii) $\mathcal{M}(\beta_2, B_V) = \mathcal{M}(\beta)\mathcal{M}(T)$;

(iii) given another basis $B'_V = \{u_1, \dots, u_n\}$ we have:

$$\mathcal{M}(\beta, B_V) = \mathcal{M}(\mathrm{Id}_V, B_V, B'_V)^{\mathrm{t}}\mathcal{M}(\beta, B'_V)\mathcal{M}(\mathrm{Id}_V, B_V, B'_V). \tag{26}$$

Proof. Exercise. See textbook §9A. □

Warning 10. We stress how formula (26) is different than formula (5). The relation between the two will become clearer when we rewrite the vector spaces $\mathcal{L}^2(V)$ and $\mathcal{L}(V, V)$ in terms of tensor products involving V and its dual $V^\vee$.

Exercise 49. Consider $\beta : \mathcal{P}_2(\mathbb{R}) \times \mathcal{P}_2(\mathbb{R}) \to \mathbb{R}$ given by $\beta(p, q) = p(2)q'(3)$. Consider the basis $B_V = \{1, x, x^2\}$ and $B'_V = \{1, x-2, (x-3)^2\}$. Check the formula (26) in this case.

Definition 54. A bilinear form $\beta : V \times V \to \mathbb{F}$ is said to be

- *symmetric* if it satisfies: $\beta(v, u) = \beta(u, v)$ for every $u, v \in V$;
- *alternating* if it satisfies: $\beta(v, u) = -\beta(u, v)$ for every $u, v \in V$.

We will let $\mathcal{L}^2_{\mathrm{sym}}(V) \subseteq \mathcal{L}^2(V)$ and $\mathcal{L}^2_{\mathrm{alt}}(V) \subseteq \mathcal{L}^2(V)$ denote the subsets of symmetric and anti-symmetric bilinear forms, respectively.

Remark 41. Notice that $\mathcal{L}^2_{\mathrm{sym}}(V)$ is a subspace of $\mathcal{L}^2(V)$. Indeed, $\beta = 0$ belongs to both of them, so they are non-zero. Given $\beta_1, \beta_2 \in \mathcal{L}^2_{\mathrm{sym}}(V)$ we have:

$$(\beta_1 + a\beta_2)(u, v) = \beta_1(u, v) + a\beta_2(u, v) = \beta_1(v, u) + a\beta_2(v, u) = (\beta_1 + a\beta_2)(v, u),$$

for every $u, v \in V$ and $a \in \mathbb{F}$.

Similarly, one can check that $\mathcal{L}^2_{\mathrm{alt}}(V)$ is a subspace of $\mathcal{L}^2(V)$. This is left as an exercise.

Exercise 50. Let $\beta : V \times V \to \mathbb{F}$ be a bilinear form. Prove that the following are equivalent:

(1) $\beta(v, v) = 0$ for all $v \in V$;

(2) β is an alternating (bilinear) form.

Lemma 72. *We have a decomposition of $\mathcal{L}^2(V)$ as follows:*

$$\mathcal{L}^2(V) = \mathcal{L}^2_{\mathrm{sym}}(V) \oplus \mathcal{L}^2_{\mathrm{alt}}(V).$$

Proof. We leave this as an exercise. □

20. Apr. 22, 2024

20.1. **Multilinear maps.**

Definition 55. Let $n \geq 1$ be a positive integer, a function $\beta : V^n \to \mathbb{F}$ is an *n-linear map*, if β is linear on each factor, i.e. given $\{u_1, \dots, u_{n-1}\}$ fixed vectors in V the functions $f_i : V \to \mathbb{F}$ given by:

$$f_i(v) := \beta(u_1, \dots, u_{i-1}, v, u_i, \dots, u_{n-1})$$

is linear, for $1 \leq i \leq n$. We will denote the vector space of n-linear maps by $\mathcal{L}^n(V)$.

Remark 42. In the textbook they call such multilinear n-forms, however in certain contexts n-form is used for what we will call alternating n-linear map (see below). We will try to avoid this term so as not to cause too much confusion. Notice however that this is in contrast with what we called bilinear form in the previous section.

Example 41. (i) Let $\alpha, \beta \in \mathcal{L}^2(V)$, then $\gamma((v_1, v_2, v_3, v_4)) := \alpha(v_1, v_2)\beta(v_3, v_4)$ gives a 4-linear map γ.

(ii) Let $\beta : (\mathcal{L}(V))^m \to \mathbb{F}$ be given by

$$\beta(T_1, \ldots, T_m) = \operatorname{tr}(T_1 \circ \cdots \circ T_m).$$

Then β is an m-linear map on $\mathcal{L}(V)$.

Remark 43. Notice that for $n \geq 3$, there is no matrix representation of n-linear maps as we had for bilinear forms in Remark 40.

20.1.1. *Interlude on symmetric group.*

Definition 56. For every $n \geq 1$ we let S_n denote the set of bijections from $\{1, \ldots, n\}$ to itself. The set S_n forms a group[23] which we call the *symmetric group in n elements*.

Given $\sigma, \tau \in S_n$ we will write $\sigma\tau$ for the composition $\sigma \circ \tau$. An element $\sigma \in S_n$ is said to be a *transposition* if there are $i, j \in \{1, \ldots, n\}$ such that

$$\sigma(k) = \begin{cases} k & \text{if } k \neq i \text{ and } k \neq j, \\ j & \text{if } k = i; \\ i & \text{if } k = j. \end{cases}$$

In this case we might denote $\sigma = \tau_{ij}$.

We define a function $\operatorname{sign}(-) : S_n \to \{\pm 1\}$, via let $[n] := 1 < 2 < \cdots < n$ be a linear order on $\{1, \ldots, n\}$ and consider $\sigma([n]) := \sigma(1) < \sigma(2) < \cdots < \sigma(n)$ the new order on the set $\{1, \ldots, n\}$. We define $I(\sigma)$ the number of *inversions* on $\sigma([n])$ to be the number of pairs $(i, j) \in [n]$ such that $i < j$ and $\sigma(j) < \sigma(i)$. Then we let $\operatorname{sign}(\sigma) := (-1)^{I(\sigma)}$. Here are couple of things that one can check:

Lemma 73. *For every $\sigma, \tau \in S_n$ we have:*

$$\operatorname{sign}(\sigma \circ \tau) = \operatorname{sign}(\sigma) \operatorname{sign}(\tau).$$

Also $\operatorname{sign}(\tau) = (-1)$*, for any transposition* $\tau \in S_n$.

Proof. Exercise. □

For V a vector space and $\sigma \in S_n$ an element of the *symmetric group* we let:

$$\rho_\sigma : V^{\times n} \to V^{\times n}$$
$$(v_1, \ldots, v_n) \mapsto (v_{\sigma(1)}, \ldots, v_{\sigma(n)}),$$

i.e. the map from V^n to itself that sends the ith entry to the $\sigma(i)$th entry. Notice that ρ_σ is a linear map and when $\sigma = \tau_{ij}$, $\rho_{\tau_{ij}}(v_1, \ldots, v_n)$ simply swaps v_i with v_j.

The following generalizes Definition 54 to n-linear forms:

Definition 57. A n-linear map $\beta : V^{\times n} \to \mathbb{F}$ is said to be:

- *symmetric* if for every transposition $\tau_{ij} \in S_n$ we have $\beta = \beta \circ \tau_{ij}$;
- *alternating* if for every transposition $\tau_{ij} \in S_n$ we have $\beta = -\beta \circ \tau_{ij}$.

We let $\mathcal{L}^n_{\text{sym}}(V)$ and $\mathcal{L}^n_{\text{alt}}(V)$ denote the subspaces of symmetric and alternating n-linear maps.

Example 42. (i) Consider $\alpha : \mathbb{F}^n \times \mathbb{F}^n \times \mathbb{F}^n \to \mathbb{F}$ given by

$$\alpha((x_1, \ldots, x_n), (y_1, \ldots, y_n), (z_1, \ldots, z_n)) := x_1y_2z_3 - x_2y_1z_3 - x_3y_2z_1 - x_1y_3z_2 + x_3y_1z_2 + x_2y_3z_1.$$

Check that α is an alternating 3-linear map.

(ii) Let V be a real inner product space. Consider $\alpha : V \times V \times V \times V \to \mathbb{F}$ given by

$$\alpha(v_1, v_2, v_3, v_4) := \langle v_1, v_2 \rangle \langle v_3, v_4 \rangle + \langle v_1, v_3 \rangle \langle v_2, v_4 \rangle + \langle v_1, v_4 \rangle \langle v_2, v_3 \rangle.$$

We claim that α is a symmetric 4-linear map.

[23] See here for a definition.

Remark 44. Let $\beta \in \mathcal{L}^n(V)$, then β is an alternating n-linear map if and only if

$$\alpha(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_n) = 0, \text{ for every } \{v_1, \ldots, v_n\} \subset V, \text{ where } v_i = v_j, \text{ for some } i \neq j.$$

Here are a couple of simple properties of alternating n-linear maps.

Lemma 74. *(i) For every set $\{v_1, \ldots, v_m\} \subset V$ of linearly dependent vectors, we have*

$$\alpha(v_1, \ldots, v_m) = 0$$

for every $\alpha \in \mathcal{L}^m_{\text{alt}}(V)$.

(ii) Let $m > \dim V$, then $\mathcal{L}^m_{\text{alt}}(V) = \{0\}$.

Proof. For (i). Let $\sum_{i=1}^m a_i v_i = 0$ be a non-trivial linear combination of the given vectors. Let $j \in \{1, \ldots, m\}$ such that $a_j \neq 0$. Then we have:

$$\alpha(v_1, \ldots, v_j, \ldots, v_m) = -\sum_{i=1,\, i\neq j}^{m} a_i \alpha(v_1, \ldots, v_i, \ldots, v_m) = 0,$$

where each term in the sum vanishes by Remark 44.

For (ii), assume by contradiction that there exist $\{v_1, \ldots, v_m\} \subset V$ such that $\alpha(v_1, \ldots, v_m) \neq 0$. By (i), we have that $\{v_1, \ldots, v_m\}$ is linearly independent, which is a contradiction with $\dim V < m$. □

Here is a useful consequence of Lemma 74:

Corollary 23. *Let $\alpha \in \mathcal{L}^n_{\text{alt}}(V)$, where $n = \dim V$ and consider a set of vectors $\{e_1, \ldots, e_n\} \subset V$. Then the following are equivalent:*

(1) $\alpha(e_1, \ldots, e_n) \neq 0$;

(2) $\{e_1, \ldots, e_n\}$ is linearly independent.

Proof. (1) $\Rightarrow$ (2) is the converse of Lemma 74 (i).

For (2) $\Rightarrow$ (1) there exist $\{v_1, \ldots, v_n\} \subset V$ such that $\alpha(v_1, \ldots, v_n) \neq 0$. Since $\{e_1, \ldots, e_n\}$ are n linearly independent vectors, they form a basis. Let $v_i = \sum_{j=1}^n a_{i,j} e_j$ then we compute:

$$\begin{aligned}
\alpha(v_1, \ldots, v_n) &= \alpha\Big(\sum_{j_1=1}^n a_{i,j_1} e_{j_1}, \ldots, \sum_{j_n=1}^n a_{i,j_n} e_{j_n}\Big) \\
&= \sum_{j_1=1}^n \cdots \sum_{j_n=1}^n a_{1,j_1} \cdot \cdots \cdot a_{n,j_n} \alpha(e_{j_1}, \ldots, e_{j_n}) \\
&= \sum_{\sigma \in S_n} a_{1,\sigma(1)} \cdot \cdots \cdot a_{n,\sigma(n)} \alpha(e_{\sigma(1)}, \ldots, e_{\sigma(n)}) \\
&= \alpha(e_1, \ldots, e_n) \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) a_{1,\sigma(1)} \cdot \cdots \cdot a_{n,\sigma(n)},
\end{aligned}$$

where the second line is equal to the third line, because $\alpha(e_{j_1}, \ldots, e_{j_n})$ vanishes if $j_i = j_k$ for two $i, k \in \{1, \ldots, n\}$ such that $i \neq k$. Thus, $\alpha(e_1, \ldots, e_n) \neq 0$. □

Remark 45. Let $\{e_1, \ldots, e_n\}$ be a basis of V and consider a list of vectors $\{v_1, \ldots, v_n\}$. In the proof of Corollary 23 we obtained the following very useful formula. Let $v_i = \sum_{j=1}^n a_{i,j} e_j$, then

$$\alpha(v_1, \ldots, v_n) = \left(\sum_{\sigma \in S_n} \operatorname{sign}(\sigma) a_{1,\sigma(1)} \cdot \cdots \cdot a_{n,\sigma(n)}\right) \alpha(e_1, \ldots, e_n), \tag{27}$$

for any $\alpha \in \mathcal{L}^n_{\text{alt}}(V)$.

The following result is important to define the determinant.

Proposition 15. *Let V be a vector space of dimension n. Then $\dim \mathcal{L}^n_{\text{alt}}(V) = 1$.*

Proof. The proof is a bit lengthy, but the ideas are simple. See 9.37 in the textbook. □

20.2. **(Multi)linear maps in terms of tensor products.** The goal of this section is to reformulate the content of §19.3 and §20.1 in terms of tensor products.

Corollary 24. *Let V be a finite-dimensional vector space, then one has an equivalence:*

$$V^\vee \otimes V^\vee \xrightarrow{\sim} \mathcal{L}^2(V).$$

More generally, for every $n \geq 2$ we have $(V^\vee)^{\otimes n} \xrightarrow{\sim} \mathcal{L}^n(V)$.

Proof. Consider the following sequence of isomorphisms:

$$V^\vee \otimes V^\vee \xrightarrow{\simeq} (V \otimes V)^\vee \xrightarrow{=} \mathcal{L}(V \otimes V, \mathbb{F}) \xrightarrow[\varphi^{-1}_{V,V,\mathbb{F}}]{\simeq} \mathcal{L}(V, \mathcal{L}(V, \mathbb{F})) \xrightarrow[\Phi_1^{-1}]{\simeq} \mathcal{L}^2(V) \ ,$$

where the first isomorphism is Lemma 67, the third is Lemma 66, and the last isomorphism is from Exercise 48 (ii).

The case $n \geq 3$ can be proved by induction, we leave it to the reader to check the details. □

Finish this section in the future.

21. Apr. 22, 2024

21.1. **Determinant of an operator.** We notice that given $T : V \to V$ an operator, one has an induced map $T^* : \mathcal{L}^n_{\text{alt}}(V) \to \mathcal{L}^n_{\text{alt}}(V)$ defined as follows:

$$(T^*\alpha)(v_1, \ldots, v_n) := \alpha(Tv_1, \ldots, Tv_n).$$

Indeed, it is easy to check that the above is well-defined, i.e. $T^*\alpha$ is still an alternating n-linear map, and that the assignment $\alpha \mapsto T^*\alpha$ is linear.

Definition 58. Let $T : V \to V$ be an operator on a finite-dimensional vector space, then for any non-zero alternating form $\alpha \in \mathcal{L}^n_{\text{alt}}(V)$ we have:

$$T^*(\alpha) = (\det T)\alpha$$

for an unique $\det T \in \mathbb{F}$. We call $\det T$ the determinant of T.

Remark 46. It is clear, from Proposition 15 that $\det T$ does not depend on the choice of $\alpha \in \mathcal{L}^n_{\text{alt}}$, as long as it is non-zero.

Example 43. (i) Consider $\mathrm{Id}_V : V \to V$ then one has $(\mathrm{Id}_V)^*\alpha = \alpha$ for every $\alpha \in \mathcal{L}^{\dim V}_{\text{alt}}(V)$, thus $\det(\mathrm{Id}_V) = 1$.

(ii) More generally, for any $\lambda \in \mathbb{F}$ notice that $(\lambda\,\mathrm{Id}_V)^*\alpha = \lambda^{\dim V}\alpha$, so we have $\det(\lambda\,\mathrm{Id}_V) = \lambda^{\dim V}$.

(iii) Let $T : V \to V$ be diagonalizable and consider $\{v_1, \ldots, v_n\}$ a basis of eigenvectors with eigenvalues $\lambda_1, \ldots, \lambda_n$, respectively. Notice that for any $\alpha \in \mathcal{L}^n_{\text{alt}}(V)$ we have:

$$(\det T)\alpha(v_1, \ldots, v_n) = T^*\alpha(v_1, \ldots, v_n) = \alpha(Tv_1, \ldots, Tv_n) = \alpha(\lambda_1 v_1, \ldots, \lambda_n v_n) = (\prod_{i=1}^n \lambda_i)\alpha(v_1, \ldots, v_n).$$

Thus, if $\alpha(v_1, \ldots, v_n) \neq 0$, then $\det T = (\prod_{i=1}^n \lambda_i)$.

We notice how Definition 58 is very elegant and doesn't involve any choice of basis. We can actually use it to define determinant of matrices as follows:

Definition 59. Let $A \in \mathbb{F}^{n,n}$ be an n-by-n matrix, we let:

$$\det(A) := \det T_A,$$

where $T_A : \mathbb{F}^n \to \mathbb{F}^n$ is the operator on $\mathbb{F}$ such that $\mathcal{M}(T_A, B_V) = A$, where $B_V = \{e_1, \ldots, e_n\}$ is the standard basis of $\mathbb{F}^n$.

Lemma 75. *Consider $\alpha : \mathbb{F}^n \to \mathbb{F}$ defined by:*

$$\alpha(v_1, \ldots, v_n) = \det\left(v_1 \cdots v_n\right),$$

where $(v_1 \cdots v_n)$ is the matrix obtained by considering the n vectors $v_1, \ldots, v_n$ as its columns. Then α is an alternating n-linear map.

Proof. Let $\{e_1, \ldots, e_n\}$ be the standard basis of $\mathbb{F}^n$ and consider the operator $T : \mathbb{F}^n \to \mathbb{F}^n$ uniquely defined by $T(e_i) = v_i$. Then it is clear that $\det T = \det(v_1 \cdots v_n)$. Let $\beta : \mathbb{F}^n \to \mathbb{F}$ be an alternating form such that $\beta(e_1, \ldots, e_n) = 1$. Then

$$\begin{aligned}\det(v_1 \cdots v_n) &= \det T \\ &= (\det T)\beta(e_1, \ldots, e_n) \\ &= \beta(Te_1, \ldots, Te_n) \\ &= \beta(v_1, \ldots, v_n).\end{aligned}$$

□

Here is a consequence of the previous Lemma.

Lemma 76. *Let $A \in \mathbb{F}^{n,n}$ then we have:*

$$\det(A) = \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) A_{\sigma(1),1} \cdot \ldots \cdot A_{\sigma(n),n}. \tag{28}$$

Proof. Exercise. □

Lemma 77. *(i) Let $S, T \in \mathcal{L}(V)$, then $\det(S \circ T) = \det(S)\det(T)$.*

(ii) Let $A, B \in \mathbb{F}^{n,n}$ be two n by n matrices, then

$$\det(AB) = \det(A)\det(B).$$

Proof. For (i), let $\alpha \in \mathcal{L}^n_{\text{alt}}(V)$ be an alternating multilinear map on V, where $n = \dim V$. Then we have:

$$\begin{aligned}(ST)^*(\alpha)(v_1, \ldots, v_n) &= \alpha(STv_1, \ldots, STv_n) \\ &= \det(S)\alpha(Tv_1, \ldots, Tv_n) \\ &= \det(S)\det(T)\alpha(v_1, \ldots, v_n).\end{aligned}$$

Thus, by picking $v_1, \ldots, v_n$ such that $\alpha(v_1, \ldots, v_n) \neq 0$; we conclude that $\det(S)\det(T) = \det(ST)$.

(ii) follows from (i), we leave the details to the reader. □

Lemma 78. *Let $T \in \mathcal{L}(V)$, then T is invertible if and only if $\det T \neq 0$. Moreover, if T is invertible then $\det T^{-1} = (\det T)^{-1}$.*

Proof. Assume that T is invertible, then $1 = \det(\mathrm{Id}_V) = \det(TT^{-1}) = \det(T)\det(T^{-1})$, which implies that $\det(T) \neq 0$.

Now assume that $\det T \neq 0$. We will prove that T is injective. Let $v \in V$ be a non-zero vector. Consider $\{v, e_2, \ldots, e_n\}$ an extension to a basis of V. Let $\alpha \in \mathcal{L}^n_{\text{alt}}(V)$ be a non-zero alternating multilinear map, by Corollary 23 we have:

$$\alpha(Tv, Te_2, \ldots, Te_n) = (\det T)\alpha(v, e_2, \ldots, e_n) \neq 0,$$

which implies that $Tv \neq 0$ as we needed to prove. □

Here is a consequence for eigenvalues, which sometimes is taken as the definition of eigenvalues.

Corollary 25. *Let $T \in \mathcal{L}(V)$, then λ is an eigenvalue of T if and only if $\det(T - \lambda \mathrm{Id}_V) = 0$.*

Proof. By Lemma 78 we see that $T - \lambda \mathrm{Id}_V$ is not injective (definition of λ, see Definition 28) if and only if $\det(T - \lambda \mathrm{Id}_V) = 0$. □

Exercise 51. (i) Let $T \in \mathcal{L}(V)$ and $S : V \to W$ be an invertible linear map, i.e. an isomorphism. Prove that $\det(T) = \det(STS^{-1})$.

(ii) Let $T \in \mathcal{L}(V)$ and B_V a basis of V, then $\det T = \det(\mathcal{M}(T, B_V))$.

Corollary 26. *Let $T \in \mathcal{L}(V)$ where V is a finite-dimensional complex vector space.*

(i) $\det T$ is the product of all the eigenvalues of T;

(ii) *let $\lambda_1, \dots, \lambda_m$ be the list of eigenvalues of T and $d_1, \dots, d_m$ their (algebraic) multiplicities, respectively, then:*

$$\det(z \operatorname{Id}_V - T) = \prod_{i=1}^{m} (z - \lambda_i)^{d_i} = c_T(z). \tag{29}$$

Proof. There exists a basis B_V, such that $\mathcal{M}(T, B_V)$ is upper-triangular with the eigenvalues λ_i of T on the diagonal appearing exactly d_i times each. Insert reference for this. (i) and (ii) now follow directly from (28) and Exercise 51 (ii). □

Exercise 52. Let $T \in \mathcal{L}(V)$.

(i) Consider $T^\vee \in \mathcal{L}(V^\vee)$ the dual operator determined by T. Then $\det T = \det T^\vee$.

(ii) Assume that V is an inner product space. Then $\det T^* = \overline{\det T}$, where $T^* \in \mathcal{L}(V)$ is the adjoint operator.

Definition 60. Let V be a vector space either over $\mathbb{C}$ or $\mathbb{R}$ and $T \in \mathcal{L}(V)$, then the characteristic polynomial of T is defined as:

$$c_T(z) := \det(z \operatorname{Id}_V - T).$$

Notice that this agrees with Definition 47 in the case of $\mathbb{F} = \mathbb{C}$, by formula (29).

Theorem 11 (Cayley–Hamilton theorem). *Let V be a finite-dimensional vector space, then $c_T(T) = 0$. In particular, we obtain that the characteristic polynomial c_T is a multiple of the minimal polynomial p_T.*

Proof. The case of complex vector spaces is Theorem 8.

Assume that $\mathbb{F} = \mathbb{R}$ and let $A := \mathcal{M}(T, B_V)$, where B_V is the standard basis of V. Let S_A be the operator on $\mathbb{C}^{\dim V}$ whose matrix with respect to the standard basis of $\mathbb{C}^{\dim V}$ is A. Then for all $z \in \mathbb{R}$ we have:

$$q(z) = \det(z \operatorname{Id}_V - T) = \det(zI - A) = \det(z \operatorname{Id}_{\mathbb{C}^{\dim V}} - S).$$

Thus, $q(z)$ is the characteristic polynomial of S, and by the complex case, we obtain $q(T) = q(A) = q(S) = 0$. □

Lemma 79. *Let $T \in \mathcal{L}(V)$ and $n = \dim V$. Then the characteristic polynomial of T has the form:*

$$c_T(z) = z^n - (\operatorname{tr} T) z^{n-1} + \dots + (-1)^n (\det T). \tag{30}$$

Proof. The constant term is given by $\det(-T)$, which is given by $(-1)^n \det T$, since $((-1)^T)^* \alpha = (-1)^* T^* \alpha$ for any $\alpha \in \mathcal{L}^n_{\text{alt}}(V)$.

Now let B_V be a basis of V and $A = \mathcal{M}(T, B_V)$. Then the term coming from the identity permutation $e \in S_n$ in formula (28) applied to $\det(zI - A)$ gives:

$$(z - A_{1,1}) \cdot \ldots \cdot (z - A_{n,n}). \tag{31}$$

Notice that the terms coming from other permutations $\sigma \in S_n$ have at most $n-2$ factors of the form $(z - A_{k,k})$, thus they don't contribute to the coefficients of z^n and z^{n-1}. We can now read the parts of formula (30) from (31). □