Math 110: Linear Algebra

Michael Pham

Fall 2023

CONTENTS

INTRODUCTION

1.1 Lecture - 8/24/2023

1.1.1 Complex Numbers

> **Definition 1.1** (Complex Numbers)**.** We can represent complex numbers in three different ways, for $a, b \in \mathbb{R}$:
>
> 1. We can represent them as pairs: (a, b).
>
> 2. We can write them as $a + bi$.
>
> 3. We can also represent them geometrically; we can imagine a plane where the x-axis is the real axis, and the y-axis is the imaginary axis.

1.1.2 Addition of Complex Numbers

For complex numbers $\alpha = a + bi$ and $\beta = c + di$, we have that $\alpha + \beta = (a + bi) + (c + di) := (a + c) + (b + d)i$.

Geometrically, we can think of the diagonal of the parallelogram constructed by the two vectors corresponding to α and β.

In terms of pairs, for $\alpha = (a, b)$ and $\beta = (c, d)$, we have $\alpha + \beta = (a + c, b + d)$.

We proceed similarly for subtraction.

> **Remark 1.2.** We note here that addition for complex numbers is commutative; that is, $\alpha + \beta = \beta + \alpha$.

We also note here that in the complex world, we have the additive identity as well: $0 + 0i$.

The additive inverse of α is simply $-\alpha$.

1.1.3 Multiplication of Complex Numbers

We observe that for $\alpha = a + bi$ and $\beta = c + di$, we can also define multiplication as: $\alpha\beta = (ac - bd) + (ad + bc)i$.

Similarly to addition, multiplication of complex numbers is also commutative.

For the multiplicative inverse, we can proceed as follow:

$$\frac{1}{a+bi} = \frac{a-bi}{(a+bi)(a-bi)}$$
$$= \frac{a-bi}{a^2+b^2}$$

And since we are working with some $\alpha \neq 0$, then we know that at least one of $a, b \neq 0$. Thus, we have that $a^2 + b^2 \neq 0$ as well.

Thus, we have the multiplicative inverse of $\alpha = a + bi$ as $\frac{a}{a^2+b^2} + \frac{b}{a^2+b^2}i$.

To divide complex numbers α and β, then we simply have to do $(a + bi) \cdot \frac{1}{c+di} = \frac{ac+bd}{c^2+d^2} + \frac{bc-ad}{c^2+d^2}i$.

We note here that both $\mathbb{R}$ and $\mathbb{C}$ are fields, with $\mathbb{R} \subsetneq \mathbb{C}$.

1.1.4 Complex Numbers, Polar Form, and Exponentials

We can picture a vector from the origin, at some angle θ from the real axis. We observe that $a = r\cos(\theta)$ and $b = r\sin(\theta)$, where r is the length of the vector (or the "modulus" of the complex number $a + bi$).

1.2 Discussion - 8/25/2023

> **Problem 1.1.** Let $a = 1 + i$, $b = 2 - i$. Find the following:
>
> 1. $a + b$
> 2. $2a - 3b$
> 3. ab
> 4. b/a

Solution.   1. $a + b = (1 + 2) + (1 + (-1))i = 3 + 0i = 3$.

2. $2a - 3b = 2(1 + i) - 3(2 - i) = (2 + 2i) - (6 + 3i) = -4 + 5i$.

3. $ab = (1(2) - (1)(-1)) + (1(-1) + 2)i = 3 + i$

4. To find b/a, can find the conjugate of the denominator, and proceed with $\frac{(2-i)(1-i)}{(1+i)(1-i)} = \frac{1-3i}{2}$.

∎

> **Problem 1.2.** Solve the equation $x^3 = 1$ in $\mathbb{C}$.

Solution. We can tackle this problem by working with the polar form of complex numbers instead. Define $\omega := r\cos(\theta) + ir\sin(\theta)$.

Now, we observe that since we want to find $\omega^3 = 1$, then we have that $r = 1$. Next, we can think of a circle (with radius $r = 1$), and we rotate by $120°$, yielding us three rotations before returning to the original position. Then solution then to our problem $\omega^3 = 1$ are the three points for these rotations.

We see then that the solutions are:

1. $\omega_1 = 1 + 0i = 1$

2. $\omega_2 = \cos(120) + i\sin(120) = -\frac{1}{2} + \frac{\sqrt{3}}{2}i$

3. $\omega_3 = \cos(240) + i\sin(240) = -\frac{1}{2} - \frac{\sqrt{3}}{2}i$

∎

> **Problem 1.3.** Find $x \in \mathbb{R}^4$ such that $(4, -3, 1, 7) + 2x = (5, 9, -6, 8)$.

Solution. Let $x = (0.5, 6, -3.5, 0.5)$ ∎

> **Problem 1.4.** Let $V = \mathbb{R}_+ = \{a \in \mathbb{R} : a > 0\}$. Define addition and scalar multiplication by $v \oplus w = vw$, $c \otimes v = v^c$. Convince yourself that $(V, \oplus, \otimes)$ is a vector space.
>
> Now, what is the zero vector of this vector space? What is the additive inverse of a vector?

Solution. We observe that the zero vector v is 1. We see that $v \oplus w = (1)w = w$. We also see that $v \otimes w = w^1 = w$.

The additive inverse of any vector $w \in V$ is $w' = \frac{1}{w}$, as $w \oplus w' = w(w') = \frac{w}{w} = 1$. ∎

> **Problem 1.5.** Prove that for all positive integers n, we have that $n(n+1)$ is even.

Solution. We proceed by cases.

If we have a positive integer which is even, we see then that we can rewrite $n = 2m$ for some $m \in \mathbb{Z}$. Then, we observe that $n(n+1) = 2m(2m+1) = 2(2m^2+m)$. By closure of integers, we have that $2m^2+m = j \in \mathbb{Z}$. Then, we have $n(n+1) = 2j$, which is even by definition.

If the integer n is odd, then we see that $n = 2m + 1$. Now, we have $n(n + 1) = (2m + 1)(2m + 2) = (2m + 1)(2(m + 1)) = 2(m + 1)(m + 1)$. Again, since $\mathbb{Z}$ is closed under multiplication and addition, we have then that $(m + 1)^2 = j \in \mathbb{Z}$, so $2j$ is also even. ∎

> **Problem 1.6.** Show that $\sqrt{2}$ is not a ratio of integers.

Solution. Suppose for the sake of contradiction that $\sqrt{2}$ can be expressed as a ratio of integers. Then, there exist $m, n \in \mathbb{Z}$ such that $\sqrt{2} = \frac{m}{n}$, where $\gcd(m, n) = 1$.

Now, observe that $2 = \frac{m^2}{n^2}$. Then, we see that $2n^2 = m^2$. But, we see then that this contradicts with the fact that $\gcd(m, n) = 1$, as this implies that m must have also been even. ∎

> **Problem 1.7.** Show that the opposite of a vector is unique.

Solution. We suppose for the sake of contradiction that for a vector v, there exists two different vectors which is the opposite of it.

Then, we see that $v + w = 0$ and $v + u = 0$. Then, we have that $v + w = 0 = v + u \implies w + u$. This thus contradicts with the fact that the two vectors are different; the opposite of a vector is unique. ∎

VECTOR SPACES

2.1 Lecture - 8/29/2023

2.1.1 Fields

Definition 2.1 (Fields)**.** We define a field to be a set $\mathbb{F}$ with two operations $(+, \cdot)$. We denote this by $\langle \mathbb{F}, +, \cdot \rangle$. These operations satisfying the following rules:

1. Closed under $+$ and $\cdot$

   - If we add or multiply two elements in $\mathbb{F}$, the result will also be in $\mathbb{F}$. In other words, we have $+ : \mathbb{F} \times \mathbb{F} \to \mathbb{F}$. The same for $\cdot$.

2. Commutativity

   - In other words, for all $a, b \in \mathbb{F}$, we have $a + b = b + a$. similarly, $a \cdot b = b \cdot a$.

3. Associativity

   - For all $a, b, c \in \mathbb{F}$, we have $a + (b + c) = (a + b) + c$, and $a \cdot (b \cdot c) = (a \cdot b) \cdot c$.

4. Identities

   - There exist elements $0 \neq 1 \in \mathbb{F}$, where we have $0 + a = a$, and $1 \cdot a = a$.

5. Inverses

   - For any $a \in \mathbb{F}$, there exists $-a \in \mathbb{F}$ such that $a + (-a) = 0$. Furthermore, for any $a \in \mathbb{F}$, where $a \neq 0$, we have $a^{-1} \in \mathbb{F}$ such that $aa^{-1} = 1$.

6. Distributivity

   - For $a, b, c \in \mathbb{F}$, we have $c \cdot (a + b) = c \cdot a + c \cdot b$.

Example 2.2 (Finite Field of Five Elements)**.** Suppose we have the integers mod 5. Now, we construct the following table for $+$:

$+$	0	1	2	3	4
0	0	1	2	3	4
1	1	2	3	4	0
2	2	3	4	0	1
3	3	4	0	1	2
4	4	0	1	2	3

We see that commutativity holds as the table is symmetric along the main diagonal; if we flip the columns and rows, we will still have the same table. We see that since there is a zero in every row and column, then an additive inverse exists for every number.

Now, for $\cdot$, we have the following table:

$\cdot$	0	1	2	3	4
0	0	0	0	0	0
1	0	1	2	3	4
2	0	2	4	1	3
3	0	3	1	4	2
4	0	4	3	2	1

Again, we see that it is commutative by the same argument. Same applies for finding an inverse, though we have to make sure that we are only looking at rows/columns that are non-zero, and that we check if every row/column has a 1.

Remark 2.3 ($0 \cdot a = 0$)**.** We will show that $0 \cdot a = 0$, for any $a \in \mathbb{F}$. We can proceed as follow:

$$0 \cdot a = (0 + 0) \cdot a$$
$$= 0 \cdot a + 0 \cdot a$$

So, we have that $0 \cdot a = 0 \cdot a + 0 \cdot a$.

From here, we know that there exists some $-a \in \mathbb{F}$ such that $0 \cdot a + (-a) = 0$.

So, we have $0 \cdot a + (-a) = 0 \cdot a + (0 \cdot a + (-a))$. So, we have $0 = 0 \cdot a + 0 \implies 0 \cdot a = 0$.

Remark 2.4 (Uniqueness of Additive Inverses)**.** We will prove that additive inverses are unique.

Suppose for contradiction, we have two elements $b, c \in \mathbb{F}$ such that $a + b = 0, a + c = 0$.

We see then that since $a + b = 0 = a + c$. Then, we have $a + b = a + c$. Then, we have $(a+b)+b = (a+c)+b$. From here, we see that $(a + b) + b = (a + b) + c$. Thus, we have $b = c$.

2.1.2   Vector Spaces

Definition 2.5 (Vector Spaces). A vector space V consists of two sets and two operations. We usually denote it by $\langle V, \mathbb{F}, +, \cdot \rangle$.

V is a set (of vectors).

$\mathbb{F}$ is a set (of numbers/scalars). We note here that $\mathbb{F}$ must be a field its own operations.

We define the operation $+ : V \times V \to V$, and $\cdot : \mathbb{F} \times V \to V$.

Example 2.6 (Canonical Example). The following are canonical examples: $\mathbb{R}^n, \mathbb{C}^n$ or, generally, $\mathbb{F}^n$.

Given an $n \in \mathbb{N}$, we define $\mathbb{R}^n := \{(x_1, x_2, \ldots, x_n) : x_j \in \mathbb{R}, \forall j = 1, 2, \ldots, n\}$.

We also have $\mathbb{R}^\infty := \{(x_1, x_2, \ldots) : x_j \in \mathbb{R}, \forall j \in \mathbb{N}\}$.

Definition 2.7 (Vector Addition). Given $(x_1, \ldots, x_n), (y_1, \ldots, y_n) \in \mathbb{F}^n$, we can define addition by adding the corresponding components together as follow:

$$(x_1, \ldots, x_n) + (y_1, \ldots, y_n) := (x_1 + y_1, \ldots, x_n + y_n).$$

We see that we have a neutral vector whose components are all equal to 0 in regards to addition. We also see $(-x_1, \ldots, x_n) + (x_1, \ldots, -x_n) = (0_1, \ldots, 0_n)$.

We see then that the additive structure of our vector space must follow the same conditions seen in fields as well; $\langle V, + \rangle$ must be an Abelian group.

Definition 2.8 ((Scalar) Multiplication). We define multiplication to be that as follow:

Suppose we have some $c \in \mathbb{F}$, and some $(x_1, \ldots, x_n) \in \mathbb{F}^n$, we define multiplication as follow:

$$c(x_1, \ldots, x_n) := (cx_1, \ldots, cx_n)$$

For rules, we observe that we have distributivity as follows, for scalars λ, μ and vectors u, v:

$$\lambda \cdot (u + v) = \lambda \cdot u + \lambda \cdot v$$
$$(\lambda + \mu) \cdot u = \lambda \cdot u + \mu \cdot u$$

We also have associativity for the scalars λ, μ:

$$\lambda \cdot (\mu \cdot v) = (\lambda \cdot \mu) \cdot v$$

We also have the following rules:

$$1 \cdot v = v$$

Remark 2.9. We can think of scalar multiplication geometrically as stretching/compressing(/flipping) the vector.

Remark 2.10. We also note here that, for the associativity of $\cdot$ for vector spaces, the $\cdot$ operator is different in $\lambda \cdot (\mu \cdot v)$ and $(\lambda \cdot \mu) \cdot v$; the latter occurs in the field between two scalars, whereas the former is between a scalar and a vector – the dot is being overloaded.

2.2 Lecture - 8/31/2023

Example 2.11. Let S be any given set which is non-empty, and $\mathbb{F}$ to be a field. We define $\mathbb{F}^S$ as follow:

$$\mathbb{F}^S := \{f : S \to \mathbb{F}\}.$$

Take $\mathbb{F}$ as our field. We define, for any functions $f, g \in \mathbb{F}^S$, addition to be $(f + g)(s) := f(s) + g(s)$, and for some $\lambda \in \mathbb{F}$, we define (scalar) multiplication to be $(\lambda \cdot f)(s) := \lambda \cdot f(s)$.

Remark 2.12. We note here that in $(f + g)(s) = f(s) + g(s)$, the first $+$ takes place in $\mathbb{F}^S$, while the second $+$ takes place in $\mathbb{F}$.

Similarly, the first $\cdot$ takes place in $\mathbb{F} \times \mathbb{F}^S$, whereas the second $\cdot$ takes place in $\mathbb{F}$.

Example 2.13. Define $\mathbb{F} := \mathbb{R}$, and $S := \{1, 2\}$. Any function $f \in \mathbb{R}^S$ is completely described by $(f(1), f(2))$.

We see that adding two functions will result in adding the pairs together:

$$(f(1), f(2)) + (g(1), g(2)) = ((f + g)(1), (f + g)(2))$$

Likewise, with scalar multiplication, we have:

$$\lambda \cdot (f(1), f(2)) = ((\lambda \cdot f)(1), (\lambda \cdot f)(2))$$

Remark 2.14. Notice that if $S = \mathbb{N}$ (or any other countable set), then $\mathbb{F}^s$ can be understood as $\mathbb{F}^\infty$, which is an infinite sequence.

2.2.1 Subspaces

The following are basic observations in V:

- The zero vector is unique.

- The additive inverses are unique.

- $0 \cdot v = 0$

- $(-1) \cdot v = -v$

Definition 2.15 (Subspaces)**.** Suppose we have a vector space V. Now, we have some $U \subseteq V$. We see that the following conditions must be met for U to be a subspace:

- Closure under addition and scalar multiplication.

- The zero vector is contained in U.

All the other rules for a vector space is inherited by virtue of the fact that $U \subseteq V$. We note here that U is also a vector space, contained within a vector space.

Subspaces can be added.

DIRECT SUMS AND LINEAR INDEPENDENCE

3.1   Lecture - 9/5/2023

3.1.1   Recap

Recall that a subspace $U \subseteq V$, where V is a vector space, must satisfy the following two requirements:

1. Closure under vector addition and scalar multiplication.

2. The zero vector is contained in U.

We note that subspaces can be added. For example, suppose we have subspaces U, W of V. Then, we define addition of subspaces as:

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

> **Remark 3.1.** We can add other things together as well. For example, we can add subsets together too.

> **Theorem 3.2.** The sum $U + W$ is also a subspace of V.

Proof. We can verify this as follows:

1. Since U, W are subspaces, they are both closed under addition and scalar multiplication. Then we see that $\lambda(u + w) = \lambda u + \lambda w$; $\lambda u \in U, \lambda w \in W$. Similarly, $(u_1 + w_1) + (u_2 + w_2) = (u_1 + u_2) + (w_1 + w_2)$. And we see that $(u_1 + u_2) \in U$ and $(w_1 + w_2) \in W$. So, we see that $U + W$ is closed under addition and scalar multiplication.

2. For the zero vector, since both U, W contains the zero vector, then we can take $0 \in U$ and $0 \in W$. Then, we see that $0 + 0 = 0$, so $U + W$ contains the zero vector as well.

$\blacksquare$

A Little Thought Experiment

Suppose we have the line $y = x$ in $\mathbb{R}^2$ and the point $(0, 2)$. We note that when adding them together, we have a line shifted up two units. However, while the line $y = x$ is a subspace of $\mathbb{R}^2$, the point $(0, 2)$ isn't; the result is no longer a subspace. We can verify this by the fact that we don't have the zero vector.

Now, suppose that we instead added the point $(2, 2)$; while the singleton isn't a subspace itself, the result of the addition yields a subspace still. This is because the point $(2, 2)$ is an element of the subspace formed by $y = x$.

> **Remark 3.3.** The key takeaway from this experiment is that, while adding two subspaces together guarantees a subspace, we can't guarantee the result if the sets aren't known subspaces.

3.1.2 Direct Sums

> **Definition 3.4** (Direct Sum of Subspaces)**.** A sum $U + W$ of 2 subspaces U, W of V is called a direct sum if every vector in this sum has a unique representation as $u + w$, where $u \in U, w \in W$.

> **Example 3.5.** Let us take two different lines going through the origin. We observe that the result is the entire plane $\mathbb{R}^2$.
>
> We observe from here that every vector has a unique vector representation because if we have two different vectors, then we have the following:
>
> $$v_1 + w_1 = v_2 + w_2$$
> $$v_1 - v_2 = w_2 - w_1$$
>
> Since $U \cap W = \{0\}$, we get $u_1 = u_2$ and $w_1 = w_2$.

> **Theorem 3.6.** $U + W$ is a direct sum (of subspaces) if and only if:
>
> - $U \cap W = \{0\}$, or
>
> - 0 has a unique representation $0_{\in U} + 0_{\in W}$.

Proof. We shall prove the equivalences by cycling through them.

1. Suppose we have a direct sum as per the definition. Then, every vector $u + w$ must have a unique representation, including the zero vector 0. Since $0 = 0 + 0$, we know that this must be the only representation.

2. We now want to show that if 0 has a unique representation, then $U \cap W = \{0\}$. We shall proceed by contraposition.

   Let us suppose that $U \cap W \neq \{0\}$. Then, we see that there is a non-zero vector $v \in U \cap W$. That means then that $v + (-v) = 0$, where $v \in U$ and $-v \in W$. So, we see that 0 does not have a unique representation.

3. The final implication of $U \cap W = \{0\}$ implying direct sum's proof is the same as what we saw in our example.

$\blacksquare$

Remark 3.7. Suppose we take two subspaces: is the intersection always at the origin? We see that this isn't the case, as if we take the intersection between some line in $\mathbb{R}^2$ and $\mathbb{R}^2$ itself (which is its own subspace!), we see that the intersection is actually the line itself.

Remark 3.8. If U and W are subspaces, then their intersection will also be a subspace. First, we observe that since $0 \in U$ and $0 \in W$, then $0 \in U \cap W$.

For closure, we see that since U, W are both closed, then suppose we have $v_1, v_2 \in U \cap W$. Since v_1, v_2 are in both U and W, then $v_1 + v_2 \in U$ and $v_1 + v_2 \in W$. Thus, we have that $v_1 + v_2 \in U \cap W$ as well.

Similarly, suppose we have some vector $v \in U \cap W$. Then, $v \in U \wedge v \in W$. This means then that $\lambda v \in U \wedge \lambda v \in W$, where $\lambda \in \mathbb{F}$; then we see that $\lambda v \in U \cap W$ too.

Thus, we see that $U \cap W$ is also a subspace.

3.2   Lecture - 9/7/2023

3.2.1   Recap

Last lecture, we discussed the sum and intersection of subspaces. Recall that the sum of subspaces will always yield a subspace, though in other scenarios all bets are off.

3.2.2   Extending to Multiple Subspaces

Intersections and Sum

Now, we note here that there isn't anything necessarily sacred about performing these operations with just two subspaces/subsets/etc.; we can extend this to multiple subspaces. We can intersect any number of subspaces of some vector space V and still get a subspace. We see that the more subspaces we intersect with each other, the smaller the resulting space will be.

Similarly, we can perform sums on any number of subspaces. Let us take three subspaces $U_1 + U_2 + U_2 \coloneqq \{u_1 + u_2 + u_3 : u_1 \in U_1, u_2 \in U_2, u_3 \in U_3\}$. Unlike the intersection, we see that the sum of subspaces will expand the size.

Direct Sum

Definition 3.9 (Direct Sum of k Subspaces)**.** $U_1 + U_2 + \ldots + U_k$ form a direct sum if every vector in $U_1 + U_2 + \ldots + U_k$ has a unique representation as $u_1 + u_2 + \ldots + u_k$.

For simplicity, let us consider the case where $n = 3$. The most important part to note about this is that if $k > 2$, then the intersection condition is not an adequate way to check whether or not the resulting sum is direct.

3.2.3   Linear Dependence and Independence

(Linear) Span(ning)

Definition 3.10 (Span). Given $v_1, v_2, \ldots, v_k \in V$, and some scalars $a_1, a_2, \ldots, a_k \in \mathbb{F}$, we can consider the following:

$$a_1 \cdot v_1 + a_2 \cdot v_2 + \ldots + a_k v_k.$$

This is what we refer to as a linear combination of $v_1, \ldots, v_k$, with coefficients $a_1, \ldots, a_k$.

Then, the span $\mathrm{Span}(v_1, \ldots, v_k) := \{a_1 v_1 + \ldots + a_k v_k : \forall j, a_j \in \mathbb{F}\}$

Example 3.11. Suppose we have $V = \mathbb{R}^2$, and vectors $v_1 = (1, 0)$, $v_2 = (\pi, 0)$.

Then, we see that $\mathrm{Span}(v_1, v_2) = \{(x, 0) : x \in \mathbb{R}\}$.

We say that $v_1, \ldots, v_k$ spans V if $\mathrm{Span}(v_1, \ldots, v_k) = V$.

Definition 3.12 (Finite Dimensional). V is called finite-dimensional if V is spanned by some finite list of vectors.

Example 3.13. We see that $\mathbb{R}^2$ is finite-dimensional: $\mathbb{R}^2 = \mathrm{Span}((1, 0), (0, 1))$. We can see similarly that for $\mathbb{R}^n$, we can construct a list of n vectors that spans the space.

Now, we observe that the space $\mathbb{R}^\infty$ is infinite dimensional. We shall prove this later.

Example 3.14. A polynomial $\mathscr{P}(\mathbb{F})$ is a function of the form $f(z) = a_0 + a_1 z + \ldots + a_m z^m$, where $a_0, \ldots, a_m \in \mathbb{F}$. Now, if $a_m \neq 0$ then we see that $\deg(f) = m$.

For this space, we see that the polynomial $f(z) = 0$, for all z, is the zero vector.

Now, let us fix m. We observe that the sum of two polynomials of degree $\leq= m$, their sum is degree of at most m. Similarly, multiplying by a scalar still yields a polynomial of degree m. Thus, we see that $\mathscr{P}(\mathbb{F})$ is a vector space. We see here that it is infinite-dimensional, as we can write it as being spanned by the vectors $\{1, z, \ldots, z^m\}$.

Now, suppose we don't restrict the degree and instead look at $\mathscr{P}(\mathbb{F})$. We see that it is a vector space, as the sum of polynomials will result in a polynomial, and multiplying by a scalar will still yield a polynomial. A quick note here is that we are working in non-finite fields; otherwise, we may run into weird situations.

we say that the vector space $\mathscr{P}(\mathbb{F})$ is infinite dimensional. We will prove this as follows:

Proof. Let us proceed by contradiction.

Let us suppose that $\mathscr{P}(\mathbb{F})$ is spanned by a finite list of n vectors $\{1, z, \ldots, z^{n-1}\}$. We see then that the maximum degree of the polynomials in the span of this set is $n - 1$; this means then that $z^n \notin \{1, z, \ldots, z^{n-1}\}$. This is a contradiction.

Thus, it must be that $\mathscr{P}(\mathbb{F})$ is infinite dimensional. ∎

3.2.4 Independence

Definition 3.15 (Linear Independence). Given $v_1, \ldots, v_k \in V$, this list is linearly independent if every vector in $\mathrm{Span}(v_1, \ldots, v_k)$ has a unique representation as $a_1 v_1 + \ldots + a_k v_k$.

We note here that $\mathrm{Span}(v_1, \ldots, v_k) = \mathrm{Span}(v_1) + \ldots + \mathrm{Span}(v_k)$.

Theorem 3.16 (Direct Sum and Linear Independence)**.** The subspaces $\mathrm{Span}(v_1), \ldots, \mathrm{Span}(v_2)$ are in direct sum $\iff$ the vectors $v_1, \ldots, v_k$ is linearly independent.

Theorem 3.17 (Linear Combination and Independence)**.** The list $v_1, \ldots, v_k$ is linearly independent $\iff$ The only way to obtain 0 as a linear combination of $v_1, \ldots, v_k$ is if $a_1, \ldots, a_k = 0$.

Example 3.18. Suppose we have a list of vectors which is linearly independent, then if we remove some vectors from the list, the resulting vector is linearly independent.

Similarly, if we have a dependent list, then adding most vectors will still result in a dependent list.

In other words, all sublists of an independent list is independent; all superlists of a dependent list is dependent.

Remark 3.19. From this example, we observe that the empty list $\{\}$ is linearly independent; $\mathrm{Span}() = \{0\}$.

Another Look at Dependence

Example 3.20. Suppose we have $V = \mathscr{P}_2(\mathbb{R})$. Let us have the vectors $f_1(x) = x^2 - 2$ and $f_2 = 2x^2 - 4$.

We see that the list containing only (f_1) is independent. However, the list (f_1, f_2) is dependent; $f_2 - 2f_1 = 0$.

Then, any linear dependency $a_1 v_1 + \ldots + a_m v_m = 0$, where not all $a_j = 0$ for $j = 1, \ldots, m$, can be rewritten as the following:

$$v_k \in \mathrm{Span}\left(v_1, \ldots, v_{k-1}\right) \qquad \text{(for some } k \leq m)$$

Then, there is in fact a smallest k with this property, 2.

Remark 3.21. We note that if we have list of only one element, and that elemetn is the zero vector, then the list is not linearly independent!

3.3   Discussion - 9/8/2023

Problem 3.1. Find a number t such that the vectors $(3, 1, 4), (2, -3, 5), (5, 9, t)$ is *not* linearly independent in $\mathbb{R}^3$.

Solution. We essentially want to find some number t such that $a_1 v_1 + a_2 v_2 + a_3 v_3 = 0$, but not all $a_j = 0$.

So, we observe the following:

$$3a_1 + 2a_2 + 5a_3 = 0$$
$$a_1 - 3a_2 + 9a_3 = 0$$
$$4a_1 + 5a_2 + ta_3 = 0$$

We can rewrite this as an augmented matrix, and find some value of t that makes it not diagonalisable.   ∎

> **Problem 3.2.** Let $U = \{(z_1, \ldots, z_5) : 6z_1 = z_2, z_3 + 2z_4 + 3z_5 = 0\} \subseteq \mathbb{R}^5$. Find a basis of U, and extend it to a basis of $\mathbb{R}^5$.

Solution. We observe that since we can write z_2 in terms of z_1, and we see that $z_3 + 2z_4 + 3z_5 = 0$

We observe that since $z_3 + 2z_4 + 3z_5 = 0$, then $z_3 = -2z_4 - 3z_5$. So, we can write z_3 as a linearly combination of z_4 and z_5. We also see that we can write z_2 as a linearly combination of z_1, since $z_2 = 6z_1$.

So, a basis we can write for U is $\mathrm{Span}(z_1, z_4, z_5)$. ∎

> **Problem 3.3.** Suppose that $v_1, \ldots, v_m$ are linearly independent in V, and $w \in V$. Show that $v_1, \ldots, v_m, w$ is linearly independent if and only if $w \notin \mathrm{Span}(v_1, \ldots, v_m)$.

Solution. Suppose that $v_1, \ldots, v_m$ are linearly independent and is in V, and that $w \in V$.

Forward Direction. Let us proceed with the forward direction. For the sake of contradiction, let us suppose that $w \in \mathrm{Span}(v_1, \ldots, v_m)$.

We see that for $v_1, \ldots, v_m, w$ is linearly independent, then by definition we have

$$a_1 v_1 + \ldots + a_m v_m + a_{m+1} w = 0$$

occurs only when all $a_1, \ldots, a_{m+1} = 0$. Now, if we suppose that $w \in \mathrm{Span}(a_1, \ldots, a_m)$, then w can be written as some linearly combination of $v_1, \ldots, v_m$.

Since this is the case however, then if we let $a_{m+1} = -1$, and we choose $a_1, \ldots, a_m$ such that $a_1 v_1 + \ldots + a_m v_m = w$, then we see that the result will be zero, and $a_1 v_1 + \ldots + a_m v_m + a_{m+1} w = 0$ but the coefficients are not all zero.

Therefore, we have a contradiction. ∎

Proof. Let us now prove the backwards direction instead. For this, we will proceed by contraposition.

Let us suppose that $v_1, \ldots, v_m, w$ is linearly dependent. Now, because $v_1, \ldots, v_m$ is linearly independent, then we know that $a_1 v_1 + \ldots + a_m v_m = 0$, and all coefficients are zero.

However, by adding in w, the list of vectors is now linearly dependent; then, $a_1 v_1 + \ldots + a_m v_m + a_{m+1} w = 0$, where at least one of $a_1, \ldots, a_m, a_{m+1} \neq 0$.

Then, we see that $a_1 v_1 + \ldots + a_m v_m + a_{m+1} w = 0$, where not all $a_1, \ldots, a_m, a_{m+1}$ are zero. Now, we have two cases:

1. If $w = 0$, then we observe that $a_1 v_1 + \ldots + a_m v_m = 0$, but at least one of $a_1, \ldots, a_m \neq 0$. However, this is a contradiction, as that would imply the vectors are linearly dependent.

2. Therefore, we only have $w \neq 0$. In this case, we see that w can be expressed as a linearly combination of $v_1, \ldots, v_m$:

$$\ldots$$

We see then that since w can be expressed as a linearly combination of vectors $v_1, \ldots, v_m$, then $w \in \mathrm{Span}(v_1, \ldots, v_m)$. ∎

Thus, we have proved both directions of the proof. Therefore, we can conclude that the statement is true. ∎

Problem 3.4. Let U, W be subspaces of V such that $U \oplus W = V$, and $u_1, \ldots, u_m$ and $w_1, \ldots, w_n$ be the basis of U and W respectively. Show that $u_1, \ldots, u_m, w_1, \ldots, w_n$ be the basis of V.

Solution. content... ∎

INDEPENDENCE AND BASES

4.1 Lecture - 9/12/2023

4.1.1 Removing Independent Vectors

Suppose we have a list of linearly independent vectors $v_1, \ldots, v_m$, and we look at its span $\mathrm{Span}(v_1, \ldots, v_m)$. Now, if we were to remove a vector from this list, the span will become strictly smaller.

The vector which we remove is independent from the rest, and thus doesn't live in the span of the other vectors. So, when we remove it, the span will be smaller.

4.1.2 Basis

> **Definition 4.1** (Basis)**.** If we have a list of linearly independent vectors, and they're spanning the entire vector space, then we have what is referred to as a "basis."

Recap

Recall from last time that if we have a list of dependent vectors, then there exists a nontrivial combination of these vectors which is equal to zero. This combination isn't necessarily unique. First, we can scale a combination by an arbitrary scalar, and there could also be multiple linear combinations as well. In other words,

$$a_1 v_1 + \ldots + a_m v_m = 0,$$

where not all $a_j = 0$. Now, we can express it as $v_k \in \mathrm{Span}(v_1, \ldots, v_{m-1})$.

We can start from the right side, and keep on working backwards until we see some vector which can be expressed as a linear combination of the other vectors.

Now, if the very first vectors is in the span of the previous vectors, then v_1 must be the zero vector.

Now, with this observation, it leads it to the following idea:

> **Theorem 4.2.** The length of any linearly independent list $\leq$ length of any spanning list.

Proof. Suppose the list $u_1, \ldots, u_k$ is linearly independent, and $v_1, \ldots, v_l$ is spanning for V.

Now, we proceed as follows:

1. We consider the list $u_1, \ldots, u_k, v_1, \ldots, v_l$. This list is linearly dependent. We note that $u_1 \notin \mathrm{Span}() = \{0\}$, since it comes from a linearly independent list.

2. So, we see that there's some $v_j \in \mathrm{Span}(u_1, v_1, \ldots, v_{j-1})$, where $j \leq l$. Now, let us remove this v_j from the list. Then, the remaining list still spans V.

3. Next, we consider the new list $u_1, u_2, v_1, \ldots, v_{j-1}, v_{j+1}, \ldots, v_l$. This list is still dependent, so we can apply the lemma and repeat the process.

4. Once we have used up all of the u vectors, we will have a list which has l vectors. We can repeat this process k times and bring the vectors $u_1, \ldots, u_k$ in. Thus, $k \leq l$.

∎

Now, the consequences of this theorem are:

> **Corollary 4.3.** Any spanning set for V can always be reduced to a basis.
>
> Likewise, any linearly independent list in a finite-dimensional space V can be transformed into a basis by adding vectors. We can append the spanning list for V after the linearly independent list, and then reduce down.

> **Corollary 4.4.** Next, any two bases of a finite-dimensional space have the same number of vectors.

Subspaces

Suppose U is a subspace of V and V is finite-dimensional.

Now, consider the two following conjectures:

1. $\dim V \overset{?}{<} \infty \implies \dim U < \infty$

2. $v_1, \ldots, v_n$ is a basis of $V \implies v_1, \ldots, v_n$ contains a basis for U.

For the second conjecture, we observe that it is false: we can consider the vector space $V = \mathbb{R}^2$ with basis $(1, 1)$ and $(1, -1)$. Then, the subspace $U = \{(x, 0) : x \in \mathbb{R}\}$ cannot be created using the basis vectors of V.

Then, if this is the case, how do we prove the first conjecture? In other words, how do we prove the following:

> **Theorem 4.5.** Every subspace U of a finite-dimensional space V must be finite-dimensional.

Proof. Well, to do this, let us first take some $u_1 \in U$. If $\mathrm{Span}(u_1) = U$, then we're done.

Otherwise, there is another vector $u_2 \notin \mathrm{Span}(u_1)$, where $u_2 \in U$.

If $\mathrm{Span}(u_1, u_2) = U$, we're done. If not, we repeat this process repeatedly.

We note here that the process must terminate after finitely many steps. This is because the length of an independent list $\leq$ the length of a spanning list.

Moreover, the number of vectors in any basis of U ($= \dim U$) is always $\leq$ the number of vectors in any basis of V ($= \dim V$). Thus, we can conclude that U must be finite-dimensional as well. ∎

From here, if U is a subspace of V, and V is finite-dimensional, then any basis of U can be extended to a basis of V. To do this, we can pick vectors $w_j \notin \operatorname{Span} U$ until the resulting list of vectors spans V, and is linearly independent.

> **Theorem 4.6.** Consider $W := \operatorname{Span}(w_1, \ldots, w_l)$.
>
> Then, $U \oplus W = V$.
>
> This sum is direct since the vectors in U, W are linearly independent form each other (as they form a basis for V). And by construction, we see that their sum will form the space V.

4.2   Lecture - 9/14/2023

4.2.1   Recap

We recall that last lecture we were discussing bases and dimension. The dimension of a vector space is the length of any basis.

Now, say $v_1, \ldots, v_k, v_{k+1}, \ldots, v_n$ is a basis. Then, $\operatorname{Span}(v_1, \ldots, v_k) + \operatorname{Span}(v_{k+1}, \ldots, v_n)$ spans the whole space. Furthermore, we know that the sum of these two spans is direct.

Thus, with these observations, we can have a subspace U with a basis, and then enlarge it to the basis of the whole space; then, the newly added vectors will form a basis for a new subspace W such that $U \oplus W = V$.

> **Example 4.7.** Suppose we have $V = \mathscr{P}_3(\mathbb{R})$. Then, say we have $U := \{f \in V : f'2(2) = 0\}$. Then, the basis of this subspace is $\{1, (x-2)^2, (x-2)^3\}$. We know that these vectors are all linearly independent of each other as they aren't scalar multiples of each other. We know its dimension then is at least 3. So, we have to rule out the possibility of it being of dimension 4.
>
> Now, if $\dim U = 4$, then U must be the whole space; if the dimensions were equal, there is no way to make the list larger, and thus we have a basis for the entire space already.
>
> In the case of U, we see that x fails the condition; we can verify by looking at $\{1, (x-2), (x-2)^2, (x-2)^3\}$.
>
> Now, we know that $\mathscr{P}_3$ is of dimension 4, since its basis is $\{1, x, x^2, x^3\}$. So, in order to enlarge the list we have, we can just add $x - 2$ to it.
>
> Thus, if we let $W := \operatorname{Span}\{(x-2)\}$ and we will have $U \oplus W = V$.

> **Remark 4.8.** Axler mentions an analogy between finite sets and finite-dimensional vector spaces. He builds a small table, where S represents finite sests, and V representing finite-dimensional vector spaces:
>
> | S | V |
> |---|---|
> | $\lvert S \rvert$ | $\dim V$ |
> | $T \subseteq S$ (subset) | $U \subseteq V$ (subspace) |
> | If $\lvert T \rvert = \lvert S \rvert$, then $T = S$ | If $\dim U = \dim V$, then $U = V$ |
> | $S_1 \cap S_2$ | $U_1 \cap U_2$ |
> | $S_1 \cup S_2$ | $U_1 + U_2$ |
> | $S_1 \dot\cup S_2$ | $U_1 \oplus U_2$ |
> | $\lvert S_1 \dot\cup S_2 \rvert = \lvert S_1 \rvert + \lvert S_2 \rvert$ | $\dim(U_1 \oplus U_2) = \dim U_1 + \dim U_2$ |
> | $\lvert S_1 \cup S_2 \rvert = \lvert S_1 \rvert + \lvert S_2 \rvert - \lvert S_1 \cap S_2 \rvert$ | $\dim(U_1 + U_2) = \dim U_1 + \dim U_2 - \dim(U_1 \cap U_2)$ |

Remark 4.9 (Proof of (sort-of) PIE for Dimensionality)**.** We will prove that $\dim(U_1 + U_2) = \dim U_1 + \dim U_2 - \dim(U_1 \cap U_2)$.

First, let us take a basis for $U_1 \cap U_2$. We will call it $u_1, \cap, u_k$. Now, we will extend it to a basis of U_1 — to do this, we get $u_1, \ldots, u_k, v_1, \ldots, v_l$.

We can also extend it so that it becomes a basis for U_2, yielding us $u_1, \ldots, u_k, w_1, \ldots, w_m$.

First, we see that $\mathrm{Span}(u_1, \ldots, u_k, v_1, \ldots, v_l, w_1, \ldots, w_m)$ contains U_1 and U_2. And since this spanning set contains the linear combination of all vectors from $U_1 + U_2$, then $U_1 + U_2$ is spanned by this set. In fact, it is exactly $U_1 + U_2$.

Now, to test for independence of these vectors, we imagine we have a linear combination of these vectors:
$$a_1 u_1 + \ldots + a_k u_k + b_1 v_1 + \ldots + b_l v_l = -c_1 w_1 - \ldots - c_m w_m.$$

Now, we see that the LHS lives in U_1, whereas the RHS lives in U_2. So, the RHS $-c_1 w_1 - \ldots - c_m w_m = d_1 u_1 + \ldots + d_k u_k$. Then, all of these coefficients have to be equal to zero, since this serves as a basis for U_2. Thus, they're linearly independent: $c_1 = \cdots = c_m = d_1 = \cdots = d_k = 0$.

Then, we have that the RHS of the original equation is equal to zero. Then, of course, the LHS coefficients must also be equal as it serves as a basis of U_1.

Thus, we observe that $\dim(U_1 + U_2) = k + l + m$. Meanwhile, $\dim U_1 = l + k$ and $\dim U_2 = l + m$. So, we see that $\dim U_1 + \dim U_2 = (k + l) + (k + m)$; this will yield $2k + l + m$. Now, we observe from here that in U_1 and U_2, they share the vectors $u_1, \ldots, u_k$; so, we see that $U_1 \cap U_2 = (u_1, \ldots, u_l)$. Then, $\dim(U_1 \cap U_2) = k$.

Thus, if we subtract this from $2k + l + m$, we get $(2k + l + m) - k = k + l + m$, as desired.

4.2.2 Linear Maps

We have the following set-up: we have two vectors spaces, V and W. Now, we start looking at functions, or maps, from V to W.

Now, given some $T : V \mapsto W$, we say that T is linear if it respects linear operations (or has linearity). In other words,

1. $T(v_1 + v_2) = T(v_1) + T(v_2)$.

2. $T(\alpha v_1) = \alpha T(v_1)$.

We note that this definition alone requires that V and W to be over the same field.

Example 4.10. Suppose we have $V = \mathbb{R}^2$ and $W = \mathbb{R}^3$.

Examples of linear maps includes:

1. $(x_1, x_2) \mapsto (0, 0, 0)$

2. $(x_1, x_2) \mapsto (x_1, x_2, 0)$

3. $(x_1, x_2) \mapsto (x_1, x_2, x_1 + x_2)$

We can also consider the polynomials P. Differentation is a linear map as well, with polynomials.

Example 4.11. Examples of non-linear maps are:

1. $(x_1, x_2) \mapsto (3x_1 + x_2, -x_2 - 25x_2, x_1 + 1)$

2. $(x_1, x_2) \mapsto (\sin(x_1), \cos(x_2), \tan(x_1 + x_2))$

4.3 Discussion - 9/15/2023

Remark 4.12. We will first discuss the difference beteen {} and ().

If we write the basis as a tuple, we can remember the order. Meanwhile, as a set, we have to remember the order ourselves. However, for the tuple, if we write something like $(1, i)$, the basis for $\mathbb{C}$ over $\mathbb{R}$, can be confused as a coordinate in $\mathbb{C}^2$.

Example 4.13. Suppose we have the following:

$$U := \{(a, a - b, b) : a, b \in \mathbb{R}\}.$$

We can find a basis by setting $a = 1, b = 0$, and vice versa. This gives us $\{(1, 1, 0), (0, -1, 1)\}$ as a basis.

Now, to show that this indeed spans, we observe that all vectors in U can be rewritten as a linear combination of these two vectors. Furthermore, we observe that they're linearly independent as neither of them are scalar multiples of each other.

Problem 4.1. Let $U = \{(z_1, \ldots, z_5) : 6z_1 = z_2, z_3 + 2z_4 + 3z_5 = 0\}$. Find a basis of U, and extend it to a basis of $\mathbb{R}^5$. What if in 'abstract' vector spaces not written as $\mathbb{F}^n$?

Solution. We observe that $6z_1 = z_2$. Next, we observe that $z_3 + 2z_4 + 3z_5 = 0$, it follows then that $-2z_4 - 3z_5 = z_3$.

So, $U = \{(z_1, 6z_1, -2z_4 - 3z_5, z_4, z_5) : z_1, z_4, z_5 \in \mathbb{R}\}$. Then, the following is a basis for U:

$$\{(1, 6, 0, 0, 0), (0, 0, -2, 1, 0), (0, 0, -3, 0, 1)\}.$$

We see that all vectors in U can be written as a linear combination of these basis vectors. Furthermore, to confirm that they are indeed linear independent, we observe the following:

Now, to extend it to a basis for $\mathbb{R}^5$, we can simply add the vectors $(0, 1, 0, 0, 0), (0, 0, 1, 0, 0)$ to our list. ∎

Problem 4.2. Suppose that $p_0, \ldots, p_m \in \mathscr{P}_m$ such that $p_k(2) = 0$. Show that $p_0, \ldots, p_m$ are not linearly independent.

Solution. We suppose that $p_0, \ldots, p_m$ are linearly independent. Then, this means that for

$$a_0 p_0 + \ldots + a_m p_m = 0$$

to be true, all the coefficients $a_0, \ldots, a_m = 0$.

However, because $p_k(2) = 0$, then we see that for

$$a_0 p_0(2) + \ldots + a_k p_k(2) + \ldots + a_m p_m(2) = 0,$$

when $a_1 = \ldots = a_{k-1} = a_{k+1} = \ldots = a_m = 0$, we have

$$a_k p_k(2) = 0.$$

If these vectors were linearly independent, a_k should be 0. However, because $p_k(2) = 0$, then any value of a_k will satisfy the equation above. But, this means then that these vectors are not linearly independent, as $a_k \neq 0$ still satisfies the equation.

Therefore, we can conclude that these vectors are in fact linearly dependent. ∎

> **Problem 4.3.** Show that $\dim V_1 + \ldots + \dim V_n \geq \dim(V_1 + \ldots + V_n)$.

Solution. We can proceed by induction.

Base Case: For the case of $n = 1$, we see that it is trivially true, as we get $\dim V_1 = \dim V_1$.

For the case of $n = 2$, we will use the following lemma:

> **Lemma 4.14.** $\dim(V_1 + V_2) = \dim V_1 + \dim V_2 - \dim(V_1 \cap V_2)$.

Then, with this lemma in mind, we observe that in the case of $n = 2$, we have

$$\dim V_1 + \dim V_2 = \dim(V_1 + V_2) + \dim(V_1 \cap V_2)$$
$$\dim V_1 + \dim V_2 \geq \dim(V_1 + V_2).$$

Induction Hypothesis: Let us suppose that for $n = k$, we have:

$$\dim V_1 + \ldots + \dim V_k \geq \dim(V_1 + \ldots + V_k).$$

Inductive Step: Now, we will look at the case of $n = k + 1$. We observe that we have:

$$\begin{aligned}
\dim(V_1 + \ldots + \dim V_k + \dim V_{k+1}) &= \dim((V_1 + \ldots + V_k) + V_{k+1}) \\
&= \dim(V_1 + \ldots + V_k) + \dim(V_{k+1}) - \dim((V_1 + \ldots + V_k) \cap V_{k+1}) \\
&\leq \dim(V_1 + \ldots + V_k) + \dim(V_{k+1}) \\
&\leq \dim V_1 + \ldots + \dim V_k + \dim V_{k+1}
\end{aligned}$$

Thus, our induction hypothesis holds for the case of $n = k + 1$. Therefore, by the principle of mathematical induction, we can conclude that our claim holds for all $n \in \mathbb{N}$. ∎

> **Problem 4.4.** Let V be the space of real polynomial functions, $D = \frac{\mathrm{d}}{\mathrm{d}x}$, and I be the identity map, both viewed as maps $V \to V$.
>
> Determine if I, D, D^2 are linearly independent or not.

Solution. We observe that for the vectors to be linearly independent, only the trivial solution of $a_1 = a_2 = a_3 = 0$ satisfies the following equation:

$$a_1 I + a_2 D + a_3 D^2 = T_0,$$

where T_0 is the linear map such that $T_0 v = 0$ for all $v \in V$.

Then, from here, we see that:

$$(a_1 I + a_2 D + a_3 D^2)(p) = T_0(p) \qquad\qquad = 0$$

So, let us define $p_1 = x^0 = 1$. Then, we have that

$$(a_1 I + a_2 D + a_3 D^2)(p_1) = T_0(p_1)$$
$$a_1 + 0 + 0 = 0$$
$$a_1 = 0$$

Next, define $p_2 = x$. Then, we have:

$$(a_2 D + a_3 D^2)(p_2) = T_0(p_2)$$
$$a_2 + 0 = 0$$
$$a_2 = 0$$

Finally, if we let $p_3 = x^2$, then we see that

$$(a_3 D^2)(p_3) = T_0(p_3)$$
$$2a_3 = 0$$
$$a_3 = 0$$

Since $a_1 = a_2 = a_3 = 0$, we conclude that they are indeed linearly independent. ∎

LINEAR MAPS

5.1 Lecture - 9/19/2023

5.1.1 Linear Maps

We recall that when discussing "linear map", we first need two vector spaces V, W over the same field $\mathbb{F}$.

Now, a "linear map" is defined as a map in which the following two conditions hold:

For all $v, w \in V$, along with some scalar $\lambda \in \mathbb{F}$, we have:

1. $T(v + w) = T(v) + T(w)$.

2. $T(\lambda v) = \lambda T(v)$.

In other words, T respects linearity.

A Little Thought Experiment (Save Me)

Suppose we have $V = \mathscr{P}(\mathbb{R}) = W$.

When we try to integrate some expression, such as $(x - 1)^2$, using an indefinite integral, we see that we are stuck with having a $+C$ at the end – what should this C be?

There is no canonical C, so in order to resolve this issue, we can instead take the definite integral $\int_0^x f(t)dt$. Of course, the lower bound isn't necessarily sacred; we could change it to be whatever; the difference between choosing, say, starting at 1 and 0 would be that constant C.

Then, let us take the following:

$$f(x) \mapsto \int_a^x f(t)\mathrm{d}t$$

Now, we want to show that this is indeed a linear map. Let us take $f, g \in \mathscr{P}(\mathbb{R})$ and $\alpha, \beta \in \mathbb{R}$. We then

observe the following:

$$(\alpha f + \beta g)(x) \mapsto \int_a^x (\alpha f + \beta g)(t)\mathrm{d}t$$

$$= \int_a^x (\alpha f)(t)\mathrm{d}t + \int_a^x (\beta g)(t)\mathrm{d}t$$

$$= \alpha \int_a^x (f)(t)\mathrm{d}t + \beta \int_a^x (g)(t)\mathrm{d}t$$

$$= \alpha(Tf)(x) + \beta(Tg)(x)$$

Remark 5.1. Now, we notice that if we start from integrating, then differentiating, then we get:

$$f(x) \mapsto \int_a^x f(t)\mathrm{d}t \mapsto f(x).$$

However, in the opposite direction, we get:

$$f(x) \mapsto f'(x) \mapsto \int_a^x f'(t)dt = f(x) - f(a).$$

Now, we note that if we're working in a finite-dimension then we are working in a more rigid structure, and thus one-sided inverses won't necessarily occur. But with infinite-dimensional vector spaces, it can occur.

Example 5.2. Suppose we have some mapping from $\mathscr{P}(\mathbb{R}) \to \mathscr{P}(\mathbb{R})$.

Let us consider $f(x) \mapsto (x^2 - 1)f(x)$. In this case, it is linear.

However, if we tried instead to do $\sin(x)$, then it wouldn't work, since we're violating the fact that we are working in the polynomials.

If we tried $f(x) \mapsto f(x^2 + x - 3)$, the resulting map will also be linear.

Example 5.3. Suppose we have $V = \mathscr{P}(\mathbb{R})$ and $W = \mathbb{R}$, both of which are vector spaces over $\mathbb{R}$.

Then, suppose we have the mapping $T : V \to W$. Then, if we have $f \mapsto f(\pi)$, then this will be a linear mapping. Similarly, $f \mapsto \int_0^5 f(t)\mathrm{d}t$ is also linear.

The Set of Linear Maps

Now, let us consider the set of all linear maps from V to W, denoted by $\mathcal{L}(V, W)$.

Likewise, the set of all linear maps from V to V is denoted by $\mathcal{L}(V, V)$, or also sometimes $\mathcal{L}(V)$.

It turns out that $\mathcal{L}(V, W)$ and $\mathcal{L}(V)$ also have some sort of linear structures.

First, we note that $\mathcal{L}(V, W)$ is essentially just a set of functions. Then, we know, provided the outputs allows us to, we can perform pointwise addition with functions. Then, let us proceed as follows:

Theorem 5.4 (Addition of Linear Maps)**.** Let us consider two linear maps $T \in \mathcal{L}(V, W), S \in \mathcal{L}(V, W)$. Then, we define $T + S$ as:
$$(T + S)(v) := T(v) + S(v), \qquad \forall v \in V.$$

We note that the result will also be linear as well.

Proof. We observe the following:

$$(T + S)(\alpha v + \beta w) = T(\alpha v + \beta w) + S(\alpha v + \beta w)$$
$$= \alpha T(v) + \beta T(w) + \alpha S(v) + \beta S(w)$$
$$= \alpha(T + S)(v) + \beta(T + S)(w)$$

Thus, we observe that $T + S$ is also a linear map. In other words, we have that $(T + S)(v) \in \mathcal{L}(V, W)$ as well. ■

Theorem 5.5 (Scalar Multiplication on Linear Maps)**.** We define $(c \cdot T)(v) = c \cdot T(v)$. The proof that $(c \cdot T)(v)$ is similar to the one above for addition.

To sum up, we observe that $\mathcal{L}(V, W)$ is a vector space over the same field as V and W.

Definition 5.6 (Map Composition)**.** Suppose we have a mapping T from V to W. Then, we have a mapping S from W to U.

In other words, suppose we have $T \in \mathcal{L}(V, W)$, $S \in \mathcal{L}(W, U)$. Then, we define this as $(S \circ T)(v) = S(T(v))$.

Evidently, this composition is also linear.

Special Spaces (actually, subspaces) Associated to a Linear Map

Consider $T \in (V, W)$. Then,

1. $T(0) = 0$. A small note is that the first zero is in V, whereas the second is in W.

   Generally, other vectors in V may be sent to 0_W by a linear map T. Suppose we have two vectors v_1, v_2 such that $T(v_1) = T(v_2)$. Then, we have $T(v_1) - T(v_2) = 0$.

   So, $T \iff (T(v) = 0 \implies v = 0)$.

Definition 5.7. We define the null space, denoted as $\mathrm{Null}(T) : \{v \in V : T(v) = 0_w\}$.

We note that T is injective $\iff \mathrm{Null}(T) = \{0\}$. Also, notice that $\mathrm{Null}(T)$ is a subspace of V.

Proof. We will prove some of the facts above.

First, for $T(0) = 0$, we observe the following:

$$T(0) = T(0 + 0)$$
$$= T(0) + T(0)T(0) + T(-1(0)) \qquad = T(0) + T(0) + T(-1(0))$$
$$T(0) - T(0) = T(0) + T(0) - T(0)$$
$$0 = T(0)$$

■

5.2 Lecture - 9/20/2023

5.2.1 True or False?

1. Linear Maps are injective – False.

2. $\operatorname{null} T = \{0\} \implies T$ is injective – True. However, we note that this holds only when T is linear.

3. T is surjective $\implies T$ is linear – False.

4. The map $T : \mathscr{P}(\mathbb{R}) \to \mathscr{P}(\mathbb{R}), f(x) \mapsto (x^2 + x - 1)f(x+1) - f'(x)$ is linear – True; we observe that this is a difference of two linear maps. The first part is a composition of two linear maps, with the first being a shift, then multiplication by a fixed polynomial. Meanwhile, the second part is simply the derivative, which is also linear.

5.2.2 Range

> **Definition 5.8** (Range)**.** We define the range of some linear map T as:
>
> $$\operatorname{range} T := \{T(v) : v \in V\}.$$
>
> Here, $T \in \mathcal{L}(V, W)$.

First, we observe that $\operatorname{range} T \subseteq W$. If T is surjective, then it means that $\operatorname{range} T = W$.

> **Theorem 5.9.** $\operatorname{range} T$ is a subspace of W.

Proof. First, we observe that since $T(0_v) = 0_w$, it follows then that $0_w \in \operatorname{range} T$. Thus, the zero vector is contained in $\operatorname{range} T$.

We now want to check whether closure holds for $\operatorname{range} T$.

To do this, let us look at vectors $w_1, w_2 \in \operatorname{range} T$ and some scalars $\alpha, \beta \in \mathbb{F}$.

Now, since $w_1, w_2 \in \operatorname{range} T$, it follows then that $Tv_1 = w_1$, and $Tv_2 = w_2$. It follows then that:

$$
\begin{aligned}
T(\alpha v_1 + \beta v_2) &= \alpha T(v_1) + \beta T(v_2) \\
&= \alpha w_1 + \beta w_2
\end{aligned}
$$

Thus, we observe that $\alpha w_1 + \beta w_2 \in \operatorname{range} T$.

Therefore, we have confirmed that $\operatorname{range} T$ is indeed a subspace of W. ∎

5.2.3 The Fundamental Theorem of Linear Maps

Suppose we have $T \in \mathcal{L}(V, W)$. Let us assume that V is finitely dimensional. Then, we have the following:

> **Theorem 5.10** (Fundamental Theorem of Linear Maps)**.** The following is one of the most important theorems of this course:
>
> $$\dim V = \dim \operatorname{null} T + \dim \operatorname{range} T$$

Proof. Let us suppose that $\dim V = m$.

We observe that since $\dim V$ is finite, it follows then that $\dim \operatorname{null} T$ must also be finite.

Now, we take some basis $v_1, \ldots, v_n$ of $\operatorname{null} T$.

Now, let us extend the basis of $\operatorname{null} T$ to the basis of V. Then, we have a list of vectors $v_1, \ldots, v_n, v_{n+1}, \ldots, v_m$ which forms a basis for V.

Now, we observe that for $v_1, \ldots, v_n$, these vectors get sent by T to 0. We see next that the vectors $v_{n+1}, \ldots, v_m$ must be sent to some non-zero vector in W by T.

Then, we see that $\operatorname{span} T = \{T(v_i) : i = n+1, \ldots, m\}$.

Now, we want to show that all of these vectors form a basis for $\operatorname{range} T$. Since we already have that they span $\operatorname{range} T$, then we want to show linear independence. To do this, we observe that since T is a linear map, we have:

$$a_{n+1}T(v_{n+1}) + \ldots + a_m T(v_m) = 0$$
$$T(a_{n+1}v_{n+1} + \ldots + a_m v_m) = T(0)$$
$$a_{n+1}v_{n+1} + \ldots + a_m v_m = 0$$

And since we know that the vectors $v_{n+1}, \ldots, v_m$ are part of a basis for V, it follows then that the coefficients must thus be all equal to zero as these vectors are linearly independent.

Therefore, we observe that these vectors $v_{n+1}, \ldots, v_m$ are a basis for $\operatorname{range} T$. Then, we have the following information:

$$\dim V = m$$
$$\dim \operatorname{null} T = n$$
$$\dim \operatorname{range} T = m - n$$

Therefore, we see that:

$$\dim V = m$$
$$= n + (m - n)$$
$$= \dim \operatorname{null} T + \dim \operatorname{range} T.$$

$\blacksquare$

Example 5.11. Suppose we have $V = \mathscr{P}(\mathbb{R})_k$. Then, we see that $\dim V = k + 1$.

Now, suppose we have the linear map T which differentiates all $p \in V$. Then, we see that $\dim \operatorname{range} T = (k+1) - 1 = k$. Similarly, we observe that $\dim \operatorname{null} T = 1$.

Thus, we see that $\dim \operatorname{null} T + \dim \operatorname{range} T = 1 + k = k + 1$.

Remark 5.12. We observe that this theorem is actually assymetric; the right-hand side is focused entirely on V. Meanwhile, the left-hand side is composed of a subspace of V and a subspace of W.

We note that this is due to the fact that there is nothing exactly "sacred" about the codomain; while we do need $\operatorname{range} T \subseteq W$, we can have W to be as big as we want.

Thus, it is unreasonable to expect that we have the dimension of W to appear in this formula.

Example 5.13. This formula can be used to answer questions of injectivity and surjectivity.

For example, let us consider $\dim V = \dim W = n$, $\dim \operatorname{null} T = 2$, and $\dim \operatorname{range} T = n - 2$.

From this, we observe that since $\dim \operatorname{null} T = 2$, it can't be injective. Furthermore, since $\dim \operatorname{range} T \neq n$, it follows that it can't be surjective either.

Lemma 5.14. If we have $\dim V = \dim W < \infty$, then we observe the following:

$$\text{Injectivity} \iff \text{Surjectivity}$$

Remark 5.15. We observe that if our assumption of $\dim V = \dim W$ isn't made, then this lemma fails.

For example, suppose we have some linear map T which takes polynomials in $\mathscr{P}(\mathbb{R})_k \to \mathscr{P}(\mathbb{R})_{k-1}$. While it may be surjective, we see that it isn't injective, because all the constants are sent to the zero vector.

Lemma 5.16. We cannot have an injective, linear mapping from a higher-dimensional vector space to a lower-dimensional one. This is because $\dim \operatorname{range} T$ must be less than the dimension of V, meaning that $\dim \operatorname{null} T \neq 0$. Thus, we can't have injectivity.

Similarly, we cannot send a lower-dimensional vector to a higher-dimensional vector with a linear map that is surjective. This is because ...

5.2.4   Finite-Dimensional Domains

Firstly, if $\dim V < \infty$, then we always have some basis $v_1, \ldots, v_n$ of V.

If $T \in \mathcal{L}(V, W)$, then the vectors $T(v_j) : j = 1, \ldots, n$ determine the action of T on any vector $v \in V$.

We can think of this as rewriting any vector in V as some linear combination of the basis vectors. Then, using linearity of the map, we can break it down.

Theorem 5.17. For $v_1, \ldots, v_n$, we can always form a linear map T which sends them to w_1

5.2.5   Matrices

Suppose we have $T \in \mathcal{L}(V, W)$ such that $\dim V, \dim W < \infty$.

Then, given a basis $v_1, \ldots, v_n$ and $w_1, \ldots, w_m$ of V and W respectively, we have:

$$T(v_j) = a_{1,j} w_1 + a_{2,j} w_2 + \ldots + a_{m,j} w_m.$$

5.3   Discussion - 9/22/2023

Problem 5.1. Given vectors $v_1, \ldots, v_m \in V$, construct a linear map $T : \mathbb{F}^n \to V$ by $T(x_1, \ldots, x_m) = \sum x_i v_i$. Show that T is injective if and only if $v_1, \ldots, v_m$ are linearly independent.

Solution. To begin with, we will prove the forward direction. Let us suppose that T is injective. Then, by definition, we observe that $T(v) = T(w) \implies v = w$.

Now, suppose we have the following:

$$T(x_1, \ldots, x_m) = T(y_1, \ldots, y_m)$$
$$\sum x_i v_i = \sum y_i w_i, \quad v_i, v_i \in V$$
$$x_1 v_1 + \ldots x_m v_m = y_1 w_1 + \ldots + y_m v_m$$
$$x_1 v_1 + \ldots + x_m v_m - y_1 w_1 - \ldots - y_m v_m = 0$$
$$(x_1 - y_1) v_1 + \ldots + (x_m - y_m) v_m = 0$$

Furthermore, we know that, by injectivity, we have that $x_i = y_i$ for each x_i, y_i. So, each $x_i - y_i = 0$. Thus, we have that every coefficient in the equation above is equal to zero, meaning that $v_1, \ldots, v_m$ is linearly independent.

Now, let us suppose that $v_1, \ldots, v_m$ is linearly independent.

We observe then that for $x_1, \ldots, x_m$ and $y_1, \ldots, y_m$, we have:

$$T(x_1, \ldots, x_m) = x_1 v_1 + \ldots + x_m v_m$$
$$T(y_1, \ldots, y_m) = y_1 v_1 + \ldots + y_m v_m$$

Setting each of these equations equal to zero yields us:

$$x_1 v_1 + \ldots + x_m v_m = 0$$
$$y_1 v_1 + \ldots + y_m v_m = 0$$

And because $v_1, \ldots, v_m$ is linearly independent, then we see that each $x_1 = \cdots = x_m = y_1 = \cdots = y_m = 0$. Then, it follows that $x_1, \ldots, x_m = y_1, \ldots, y_m$. So, we have shown that if $T(x_1, \ldots, x_m) = T(y_1, \ldots, y_m)$, then $x_1, \ldots, x_m = y_1, \ldots, y_m$ as well. ∎

> **Problem 5.2.** Now, show that T is injective if and only if $v_1, \ldots, v_m$ spans V.

Solution. We will proceed with the forward direction. Let us suppose that T is injective. Then, by definition, we see that $\dim \operatorname{null} T = 0$. Furthermore, we observe that because $T(x_1, \ldots, x_m) = x_1 v_1 + \ldots + x_m v_m$, it follows then that any vector $t \in T$ can be expressed as a linear combination of m vectors. Since T is injective, we note that it must be linearly independent. Thus, we have that $\operatorname{range} T = m$.

Then, we see the following:

$$\dim \operatorname{null} T + \dim \operatorname{range} T = \dim V$$
$$0 + m = \dim V$$
$$\dim V = m.$$

Thus, we see that $v_1, \ldots, v_m$ must be a spanning set of V.

Now, let us suppose that $v_1, \ldots, v_m$ spans V. Then, we know that all vectors $v \in V$ can be written as a linear combination of $x_1 v_1 + \ldots + x_m v_m$. So, $\dim V \leq m$.

Next, we note here that this linear combination is exactly $T(x_1, \ldots, x_m)$, and we further note that $\dim \operatorname{range} T = \dim V$.

And because we have that $\dim V = \dim \operatorname{range} T$, then it must follow that $\dim \operatorname{null} T = 0$, meaning that T is injective. ∎

Problem 5.3. Let $S : V \to W$, $T : U \to V$ be linear maps. If ST is injective, show that T is injective.

Solution. We will proceed by contrapositive.

We suppose that T is not injective. Then it follows that there's some non-zero $v \in V$ such that $Tv = 0$.

From here, we observe that $S(Tv) = S(0) = 0$. Thus, we see that STv is also not injective. ∎

Problem 5.4. Define $T : \mathscr{P}_2(\mathbb{R}) \to \mathbb{R}^2$ by

$$Tp = (p(1) + p'(2), p(0)).$$

What is $T1$ and Tx?

Solution. We observe that $T1$ and Tx are:

$$T1 = (1, 1)$$
$$Tx = (2, 0)$$

∎

Problem 5.5. What is the null space of T?

Solution. We observe that the null space of T is the list of polynomials p such that $Tp = (0, 0)$.

Then, we must have that $p(1) + p'(2) = 0$, and $p(0) = 0$.

In other words, we have:

$$c = 0$$
$$5a + 2b + c = 0$$
$$5a + 2b = 0$$

So, any polynomial of the form $ax^2 - \frac{5}{2}ax$ will be within the null space of T, for $a \in \mathbb{R}$. Thus, we have:

$$\operatorname{null} T = \left\{ ax^2 - \frac{5}{2}ax : a \in \mathbb{R} \right\}$$

∎

Problem 5.6. Define $T_1, T_2, T_3 : \mathscr{P}_2(\mathbb{R}) \to \mathbb{R}$ as:

$$T_1(p) = p(1)$$
$$T_2(p) = p(2)$$
$$T_3(p) = p(3).$$

Are T_1, T_2, T_3 linearly independent?

Solution. We observe that for T_1, T_2, T_3 to be linearly independent, it must follow that only the trivial solution satisfies the following equation:

$$a_1 T_1 + a_2 T_2 + a_3 T_3 = T_0$$
$$(a_1 T_1 + a_2 T_2 + a_3 T_3)(p) = (T_0)(p)$$

where T_0 is the linear map which maps all polynomial p to 0.

Now, we observe that for $p = ax^2 + bx + c$, we have:

$$a_1(a + b + c) + a_2(4a + 2b + c) + a_3(9a + 3b + c) = 0$$
$$(a_1 + 4a_2 + 9a_3)a + (a_1 + 2a_2 + 3a_3)b + (a_1 + a_2 + a_3)c = 0$$

Then, we can write this as a system of equations as follow:

$$a_1 + 4a_2 + 9a_3 = 0$$
$$a_1 + 2a_2 + 3a_3 = 0$$
$$a_1 + a_2 + a_3 = 0$$

From here, after some computation, we observe that $a_1 = a_2 = a_3 = 0$. ∎

> **Problem 5.7.** Show that there is no linear map $T : \mathbb{R}^3 \to \mathbb{R}^3$ with $\operatorname{null} T = \operatorname{range} T$.

Solution. We observe that the linear map T takes vectors from a vector space $V = \mathbb{R}^3$ to $W = \mathbb{R}^3$.

Then, we observe that $T \in \mathcal{L}(\mathbb{R}^3, \mathbb{R}^3)$ ∎

> **Problem 5.8.** We observe that for $T(v + v)$, we can rewrite it as $T(2v)$, and thus if we keep on going we can get $T(kv)$. But this isn't good, because $k \in \mathbb{Z}_+$, so we can't have something like $T(0.5v)$.
>
> The problem now is, let's construct a map where $T(v + v) = T(v) + T(v)$, but not that $T(\lambda v) = \lambda T(v)$.
>
> Similarly, construct a map such that $T(\lambda v) = \lambda T(v)$, but $T(v + v) \neq T(v) + T(v)$.

Solution. Hint 1: 2 but not 1. Map from $\mathbb{R}^2 to \mathbb{R}$. $\sqrt[3]{a^3 + b^3}$ 1 but not 1: $T : \mathbb{C} \to \mathbb{C}$, then we have $T(a + bi) = b$. Consider something where we have $\sqrt{a^2 + b^2}$. ∎

MATRIX REPRESENTATION, INVERTIBILITY, AND ISOMORPHISMS

6.1 Lecture - 9/26/2023

6.1.1 Matrix Representations

Example 6.1. Let $V = \text{span}_{\mathbb{C}}(1, \cos x, \sin x, \cos 2x, \sin 2x)$.

We define $T : f(x) \mapsto f''(x) + 4f(x)$. Now, consider the following:

1. Is T in $\mathcal{L}(V, V)$?

2. If yes, what is the representation of T in the above basis?

We note that since linearity and closure checks out, we see that T is indeed in $\mathcal{L}(V, V)$.

Now, we observe that the list $1, \cos x, \sin x, \cos 2x, \sin 2x$ are linearly independent, and thus forms a basis for V. Then, in order to determine what the representation of T, we simply see what it maps each of the basis to.

As it turns out, we get the following matrix:

$$\begin{bmatrix} 4 & 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Then, from here, we observe the following: $\text{Range}\, T = \text{span}_{\mathbb{C}}(1, \sin x, \cos x)$, and $\text{Null}\, T = \text{span}_{\mathbb{C}}(\cos 2x, \sin 2x)$.

Now, we consider: is it

6.2   Lecture - 9/28/2023

6.2.1   Invertibility

We recall that for a map to be invertible, it must be that it is both injective and surjective (in other words, it's bijective).

Now, suppose that $T : V \to W$ bijectively. Now, we want to find some inverse T^{-1}.

We define an inverse $S : W \to V$ by taking $S(w) = v$, where $\{v\} \in T^{-1}(\{w\})$. We note that this is guaranteed by the bijectivity of T.

Observe then that

$$
\begin{aligned}
(TS)(w) &= Tv \\
&= v \\
(ST)(v) &= S(T(v)) \\
&= v
\end{aligned}
$$

Thus, we observe that $ST = I_V$, and $ST = I_W$, where I is the identity map.

> **Remark 6.2.** We note here that this is indepedent of the linearity of the map T.

Linearity of Inverse

Consider the following question: if T happens to be linear and invertible, what about its inverse S?

> **Theorem 6.3.** If T is linear and invertible, then its inverse S must be linear as well.

We note that since T is linear, it follows then that

$$
\begin{aligned}
T(v + w) &= T(v) + T(w) \\
T(\lambda v) &= \lambda T(v).
\end{aligned}
$$

Now, we observe the following:

$$
\begin{aligned}
S(w_1 + w_2) &= v_1 + v_2 \\
&= S(w_1) + S(w_2)
\end{aligned}
$$

A similar check is done for scalar multiplication.

Dimensionality of W

Now, suppose we have some $T \in \mathcal{L}(V, W)$ which is invertible. Moreover, we know $\dim V = n$. Then, what about $\dim W$?

> **Theorem 6.4.** If $T \in \mathcal{L}(V, W)$ is invertible, and $\dim V = n$, then $\dim W = n$ as well.

Proof.   We observe here that $\dim W$ must be equal to m. We know that since T is invertible, it means that it is bijective; in other words, we have that T is both injective and surjective.

Now, we recall that since T is injective, it follows that $\dim \operatorname{null} T = 0$. Then, we see that $\dim \operatorname{range} T = n$.

However, since T is surjective, we see then that $\operatorname{range} T = W$. Thus, it must be that $\dim \operatorname{range} T = n = \dim W$.

Therefore, we see then that $\dim V = \dim W$. ∎

Alternatively, we can proceed as follows:

Proof. Let us define T linearly by

$$T(v_j) := w_j \qquad \text{for j = 1, ..., n.}$$

Now, we observe that this map is indeed surjective, as we note that $\operatorname{range} T$ has to contain the span of $w_1, \ldots, w_n$ which is precisely W.

We note as well that T is injective, since the condition $T(\sum a_j v_j) = 0$ means that $\sum a_j w_j = 0$. The latter implies that $a_j = 0$, for all j's as $w_1, \ldots, w_n$ is linearly independent. Thus, it must be that T is injective as well.

From here, we observe that $\dim W = n$. ∎

6.2.2 Isomorphisms

Theorem 6.5. We say that a linear map invertible map from V to W is an isomorphism. And in this case, V and W are isomorphic.

In other words, we see that if $V \cong W$ means that there exists some $T : V \to W$ such that T is invertible.

We note here that it must be that $\dim V = \dim W$ (as proven earlier).

Furthermore, we see that if T maps the basis of V to the basis of W, then we have that T is thus invertible.

We also note that if $\dim \operatorname{null} T = 0$, and we know that $\dim V = \dim W$, then of course it must be that T is invertible.

Now, let us revisit the map

$$\mathcal{M} : \mathcal{L}(V, W) \to \mathbb{F}^{m \times n},$$

where $\dim V = n$ and $\dim W = m$. We note here that this map actually maps from one vector to another vector space. Also, note that $\mathbb{F}^{m \times n}$ consists of the $m \times n$ matrices.

Now, we posit the following question: is it true that $\mathcal{M}(\alpha T_1 + \beta T_2) = \alpha \mathcal{M}(T_1) + \beta \mathcal{M}(T_2)$.

It follows that this is indeed true by construction.

We further note now that this map $\mathcal{M}$ is indeed invertible.

Example 6.6. To see that $\mathcal{M}$ is invertible, let's proceed with the following example:

$$T := \begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & 2 \\ 0 & 0 & 3 \end{bmatrix}$$

Next, we suppose that $V = \mathbb{P}_2$ and $W = \mathbb{R}^3$.

Now, say the basis of V is $7, x + 5, x^2 - 8$. And the basis of W is $(3, 0, 0), (0, 5, 0), (1, 1, 1)$.

Now, we observe that:

$$T : a7 + b(x + 5) + c(x^2 - 8) \to a(3, 0, 0) + 2b(0, 5, 0) + c(3, 0, 0) + 2c(0, 5, 0) + 3c(1, 1, 1)$$
$$= (3a + 6c, 10b + 13c, 3c)$$

Now, we write $a + bx + cx^2 = \tilde{a}7 + \tilde{b}(x + 5) + \tilde{c}(x^2 - 8)$.

Then, we see that $\tilde{c} = c, \tilde{b} = b, \tilde{a} = \frac{1}{7}(a - 5b + 8c)$.

Thus, we see that T maps $a + bx + cx^2$ to $(\frac{3}{7}(a - 5b + 8c) + 6c, 10b + 13c, 3c)$.

From here, we see that we can always reverse the steps and thus find an inverse for T.

Now, we see that because $\mathcal{M}$ is an isomorphism, it must be then that $\dim \mathcal{L}(V, W) = F^{m,n}$.

Theorem 6.7. For finite-dimensional vector spaces V, W, we observe that we have

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

6.2.3 Duality

Duality comes from a very non-scary subset.

Suppose that V is some finite-dimensional vector space, over some field $\mathbb{F}$ (for example, $\mathbb{R}$ or $\mathbb{C}$).

Now, we define the dual space V' to be:
$$V' := \mathcal{L}(V, \mathbb{F}).$$

Now, we already know that $\dim V' = \dim V$. The elements of V' are called linear functionals.

Example 6.8. Suppose $V = \mathscr{P}_2(\mathbb{R})$. Now, can we find a basis for V'?

First, $f \mapsto f(0)$ is a linear functional. Similarly, $f \mapsto f'(0)$ and $f \mapsto f''(0)$.

We denote these are $\delta_0, \delta_0', \delta_0''$. Now, we want to check whether

$$a\delta_0 + b\delta_0' + c\delta_0'' = 0.$$

So, let us proceed as follows:

$$(a\delta_0 + b\delta_0' + c\delta_0'')(1) = a$$
$$(a\delta_0 + b\delta_0' + c\delta_0'')(x) = b$$
$$(a\delta_0 + b\delta_0' + c\delta_0'')(x^2) = 2c$$

Thus, they form the basis for V', the dual space of V.

6.3   Discussion - 9/29/2023

6.3.1   Coordinates

Say we have a vector space V and a basis $\mathcal{B} = \{v_1, \ldots, v_n\}$. So, for any $v \in V$, we can write it as a linear combination of the v_i's.

Furthermore, we note that each vector v has a unique linear combination of the v_i's.

Then, each of these coefficients form the coordinate of v under the basis $\mathcal{B}$.

> **Problem 6.1.** Let $V = \mathscr{P}_2$. Find the coordinate of $x^2 + x + 1$ under basis $\{x^2, x, 1\}$, and under the basis $\{x^2, x, 1 - x^2\}$.

Solution. For the first basis, we have the coordinate $(1, 1, 1)$.

For the second basis, we have the coordinate $(2, 1, 1)$. ∎

6.3.2   Definitions of Matrix

Suppose for $T : V \to W$, we have the basis B and C for V and W respectively.

Then, the matrix of T under BC has the following definitions:

1. $[T]_C^B \cdot [v]_B = [Tv]_C$.

2. $[T]_C^B = [[Tv_1]_C, [Tv_2]_C, \cdots, [Tv_n]_C]$

> **Problem 6.2.** Find the matrix of $\frac{\mathrm{d}}{\mathrm{d}x}$, with the basis $\{x^2, x, 1\}$.

Solution. We see that the matrix is:

$$\begin{bmatrix} 0 & 0 & 0 \\ 2 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

Other than the usual way (of using definition 2), we instead note that we can do the following to find the matrix:

$$[v]_B = (a, b, c)$$

Then, we note that $Tv = \frac{\mathrm{d}}{\mathrm{d}x}(ax^2 + bx + c) = 2ax + b$. Then, we see that $[Tv]_C = (0, 2a, b)$.

So, we want to find some matrix M such that

$$M \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0 \\ 2a \\ b \end{bmatrix}$$

The matrix that yields this is indeed the same matrix

$$\begin{bmatrix} 0 & 0 & 0 \\ 2 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

∎

> **Problem 6.3.** Let V, W be finite-dimensional vector spaces, with basis $\{v_1, \ldots, v_m\}$ and $\{w_1, \ldots, w_n\}$ respectively.
>
> Use $\mathcal{M} : \mathcal{L}(V, W) \to M_{n \times m}$ for the matrix of a linear map under the basis chosen. Show that $\mathcal{M}$ is a linear map.

Solution. For $\mathcal{M}$ to be a linear map, we require that $\mathcal{M}(\alpha T_1 + \beta T_2) = \alpha \mathcal{M}(v_1) + \beta \mathcal{M}(v_2)$.

We first observe that each column i in $M_{n \times m}$ will be represented as the coefficients of Tv_i in the basis of W.

Then, we observe the following:

$$\mathcal{M}(T_1 + T_2) = \begin{bmatrix} \alpha a_{1,1} + \beta b_{1,1} & \cdots & \alpha a_{1,m} + \beta b_{1,m} \\ \vdots & \ddots & \vdots \\ \alpha a_{n,1} + \beta b_{n,1} & \cdots & \alpha a_{n,m} + \beta b_{n,m} \end{bmatrix}$$

∎

> **Problem 6.4.** Given matrices A, B such that AB makes sense, show that there exists a matrix C with $A = ABC$ if and only if the rank of A and AB are equal.

Solution. We observe that for $A = ABC$ to be true, then it follows then that we must have $BC = I$, where I is the identity matrix to get $ABC = AI = A$.

Then, we observe that $CBA = A$ as well. However, we note here that

Now, from here, it follows then that B must be invertible, with C being its inverse. ∎

WEEK 7

DUALITY

7.1 Lecture - 10/3/2023

Recall that our setting is some finite-dimensional vector space V, over a field $\mathbb{F} = \mathbb{R} \vee \mathbb{C}$.

And last lecture, we were defining $V' = \mathcal{L}(V, \mathbb{F})$; each elements in this V' are called linear functionals. In our example, we considered $V = \mathscr{P}_2(\mathbb{R})$, and V' to be the dual of V.

We recall that we took the functions $\delta_0, \delta_0', \delta_0'$ such that

$$\begin{aligned} \delta_0 &: f \mapsto f(0) \\ \delta_0' &: f \mapsto f'(0) \\ \delta_0'' &: f \mapsto f''(0) \end{aligned}$$

Last time, we got the Gramian matrix:

$$\begin{bmatrix} \delta_0(1) & \delta_0'(1) & \delta_0''(1) \\ \delta_0(x) & \delta_0'(x) & \delta_0''(x) \\ \delta_0(x^2) & \delta_0'(x^2) & \delta_0''(x^2) \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

We note the that the Gramian matrix is invertible, and thus we can conclude that $\delta_0, \delta_0', \delta_0''$ are linearly independent. In fact, we observe the following:

Theorem 7.1. Suppose that $\dim V = n$, and V is over $\mathbb{R}$ or $\mathbb{C}$. Now, suppose we have n vectors $v_1, \ldots, v_n \in V$ and linear functions $\varphi_1, \ldots, \varphi_n \in V'$.

We further suppose that the Gramian matrix, consisting of

$$\left[\varphi_j(v_i) \right], \quad j = 1, \ldots, n; i = 1, \ldots, n$$

is invertible.

Then, the following is true:

1. $\varphi_1, \ldots, \varphi_n$ are a basis for V', the dual of V.

2. $v_1, \ldots, v_n$ are a basis for V.

42

Proof. We need to check only for the linearly independence $v_1, \ldots, v_n$ and $\varphi_1, \ldots, \varphi_n$.

Now, if we suppose that there exists some linear combination of $a_1\varphi_1 + \ldots + a_n\varphi_n = 0$, then we observe that

$$\sum_{j=1}^{n} a_j \varphi_j(v_i) = 0 \quad \forall i = 1, \ldots, n.$$

Then, we observe that $\varphi_j v_i$ form our coefficients, and the a_j's form our unknowns. Thus, we have in fact that this is our Gramian matrix that we know is invertible.

This shows that $a_j = 0$ for all $j = 1, \ldots, n$ is the only solution to this linear system.

Similarly, if $\sum_{i=1}^{n} b_i v_i = 0$, then we see that

$$\sum_{i=1}^{n} b_i \varphi_j(v_i) = 0, \quad \forall j = 1, \ldots, n$$

Then, similarly, we see that the only solution is if $b_i = 0$ for all $i = 1, \ldots, n$. ∎

Theorem 7.2. The special case when the Gramian is in fact the identity matrix

$$\begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \cdots & 1 \end{bmatrix}$$

We note that this has a special name/term attached to it: we say then that the two bases $v_1, \ldots, v_n$ and $\varphi_1, \ldots, \varphi_n$ are dual to each other.

Remark 7.3. We note that to modify our original example, we can make the bases duals by changing x^2 to $\frac{1}{2}x^2$ in V.

Similarly, we can also change $\delta_0'' : f \mapsto f''(0)$ to $\delta_0'' : f \mapsto \frac{1}{2}f''(0)$ in V'.

Example 7.4. Let $V = \mathscr{P}(\mathbb{R})$. Find a basis for V' such that the Gramian is not the identity matrix.

To do this, we can look at $\delta_0, \delta_1, \delta_2$.

The Gramian is then

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 1 & 4 \end{bmatrix}$$

We note that this matrix is actually called the Vandermonde matrix. Now, to check for invertibility, we can just row reduce it to see that it is, indeed, invertible.

Remark 7.5 (Quadrature Rule). Say we have some basis for the dual space V': $\varphi_1, \ldots, \varphi_n$, where V is some function space with $\dim V < \infty$.

Then any other functional such as integration over a finite interval can be written as a linear combination of the linear functionals in our basis.

Concretely, let us suppose that $V = \mathscr{P}_2(\mathbb{R})$, and we take $\delta_0, \delta_1, \delta_2$.

Now, take $\varphi : f \mapsto \int_0^2 f(t)\mathrm{d}t$. We note that this φ is indeed an element in V'.

Now, to find φ as a combination of $a\delta_0 + b\delta_1 + c\delta_2$, we just need to evaluate this prospective combination at a basis of $V = \mathscr{P}_2(\mathbb{R})$.

$$\varphi(1) = (a\delta_0 + b\delta_1 + c\delta_2)(1)$$
$$2 = a + b + c$$
$$\varphi(x) = (a\delta_0 + b\delta_1 + c\delta_2)(x)$$
$$1 = b + 2c$$
$$\varphi(x^2) = (a\delta_0 + b\delta_1 + c\delta_2)(x^2)$$
$$\frac{8}{3} = b + 4c$$

Therefore, we see that $c = \frac{1}{3}, b = \frac{4}{3}, a = \frac{1}{3}$.

Remark 7.6. Suppose that $v_1, \ldots, v_n$ is a basis for V. Then, is there necessarily a unique basis $\varphi_1, \ldots, \varphi_n$ of V' which is dual to $v_1, \ldots, v_n$.

In this case, we observe that a linear map is defined by the action of the map on the basis V.

First, to get a dual basis for $v_1, \ldots, v_n$, we can define $\varphi_j(v_i) = \delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$

7.2   Lecture - 10/5/2023

7.2.1   Dual Maps

Definition 7.7 (Dual Maps). Given some $T \in \mathcal{L}(V, W)$, we can define the dual map to be:

$$T' : W' \to V' : T' : \varphi \mapsto \varphi \circ T$$

We note that $\varphi \in W'$, and $\varphi \circ T \in V'$.

Example 7.8. Suppose we have $V = \mathscr{P}_2(\mathbb{R})$, and $W = \mathscr{P}_2(\mathbb{R})$.

Now, let $T : f(x) \mapsto xf(x)$. Furthermore, we define φ such that $\varphi : p \mapsto p'(2)$.

Then, we observe that $T'(\varphi) = \varphi \circ T$. From here, we see that $\varphi \circ T = (xp(x))'(2) = p(2) + 2p'(2)$.

Example 7.9. Suppose we have $T \in \mathcal{L}(V, W)$ and $T' \in \mathcal{L}(W', V')$.

Now, what is the null space of T' as it relates to T. Similarly, what is $\operatorname{range} T'$ as it relates to $\operatorname{range} T$.

To do this, we require the use of annihilators.

Definition 7.10 (Annihilator)**.** Given any subset S (not necessarily a subspace) of some vector space V, we can define its annihilator S^0 as follows:

$$S^0 := \{\varphi \in V' : \varphi(v) \mapsto 0, \forall v \in S\}.$$

Example 7.11. Suppose $V = \mathbb{R}^3$. Suppose we have the subset $S = \{(x, 1, z) : x, z \in \mathbb{R}\}$.

Note that for any functional on $\mathbb{R}^3$ looks like $(x, y, z) \mapsto ax + by + cz$, for some $a, b, c \in \mathbb{R}$.

Here, we consider the canonical basis for $\mathbb{R}^3$, and the following:

$$(x, y, z) \mapsto x$$
$$(x, y, z) \mapsto y$$
$$(x, y, z) \mapsto z$$

is a dual basis in $\mathbb{R}^3$.

Now, we observe that in our context, we have $(x, y, z) \mapsto ax + b + cz$, for fixed a, b, c.

We want to find what a, b, c such that $ax + b + cz = 0$. We see that in this case, $a = b = c = 0$. Thus, only the zero functional is contained in the annihilator.

Example 7.12. Suppose that instead, we changed S to be $S := \{(x, 1, 2) : x \in \mathbb{R}\}$.

So, S^0 will consist of all functions of the type $(x, y, z) \mapsto ax + by + cz$.

Then, in this case, we observe that for $ax + b + 2z = 0$, we first must have that $a = 0$, and $b = -2c$.

Now, we note here that S^0 is actually a subspace of V'.

Generally, if S is a subset of V, then $0 \in V'$ is in S^0 as well. Furthermore, if $\varphi, \psi \in S^0$, then so is any combination $a\varphi + b\psi$, since $(a\varphi + b\psi)(v) = a\varphi(v) + b\psi(v) = 0$, whenever $v \in S$.

What is interesting is that although S doesn't necessarily have to be a subset of V, S^0 is.

Example 7.13. Now, suppose that U is a subspace of V. Our goal here is to compute the dimension $\dim U^0$, in terms of $\dim U$.

Here, we claim that it appears as though $\dim V = \dim U + \dim U^0$.

Theorem 7.14. The relation between the dimensions of U, U^0, V is as follows:

$$\dim V = \dim U + \dim U^0.$$

Proof. Suppose we have the inclusion map $I : U \to V$ such that $u \to u$.

Similarly, suppose we have $I' : V' \mapsto U'$. I' maps φ to $\varphi \circ I$.

Now, we see then that $\operatorname{null} I' = U^0$; this comes from the fact that I' consists of all functions $\varphi \circ I$ such that $(\varphi \circ I)(u) = 0$, for all $u \in U$. Thus, it is all the functionals φ that sends u to zero. This is precisely the annihilator.

Now, we observe that $\operatorname{range} I' \cong U'$, as every φ on U can be realised as $\varphi \circ I$, for some larger $\varphi \in V'$.

Then, we observe that:

$$
\begin{aligned}
\dim \operatorname{null} I' + \dim \operatorname{range} I' &= \dim U^0 + \dim U' \\
&= \dim U^0 + \dim U \\
&= \dim V' \\
&= \dim V
\end{aligned}
$$

So, from here, we observe that for

$\blacksquare$

> **Remark 7.15.** Incidentally, T' being in $\mathcal{L}(W', V')$ means that the requisite check is as follows:
>
> We take $\varphi, \psi \in W'$ and check that:
>
> $$
> \begin{aligned}
> a\varphi + b\psi \mapsto (a\varphi + b\psi) \circ T \\
> = a\varphi \circ T + b\psi \circ T \\
> = aT'(\varphi) + bT'(\psi)
> \end{aligned}
> $$

MIDTERM PAIN

8.1 Lecture - 10/12/2023

This lecture, we explore the connection between the two notions – the annihilator and the dual map.4

First, recall that we have

$$\dim U^0 = \dim V - \dim U.$$

An interesting question to ponder is what would happen if U was not a subspace. Then, we observe that we can in fact use the above equation and modify it a bit to get

$$\dim U^0 = \dim V - \dim \operatorname{span} U$$

With this in mind, we observe the following:

> **Corollary 8.1.** If the dimension $\dim U^0 = \{0\}$, then it follows that $\dim U = \dim V$.
>
> Similarly, if $U^0 = V'$, then $U = \{0\}$.

In other words, if we happen to come across a subspace such that its annihilator is the entirety of the dual space V', then we can conclude that U must consist of 0.

Then, let us consider the null space of T', $\operatorname{null} T'$, in terms of $\operatorname{range} T$. Now, consider some $w \in T(v)$. Then, $(\varphi \circ T)(v) = 0$ whenever $\varphi \in \operatorname{null} T'$.

Then, the null space T' is related related to the annihilator of $\operatorname{range} T$. In fact, we have

$$\operatorname{null} T' = (\operatorname{range} T)^0.$$

Then, we can proceed further to observe the following:

$$\begin{aligned}
\dim \operatorname{null} T' &= \dim(\operatorname{range} T)^0 \\
&= \dim W - \dim \operatorname{range} T \\
&= \dim W - \dim V + \dim \operatorname{null} T
\end{aligned}$$

Next, we observe that $\operatorname{range} T' = \dim(\operatorname{null} T)^0$. To do this, we first observe that if we have some $v \in \operatorname{null} T$, then $Tv = 0$. This means then that $(\varphi \circ T)(0) = 0$; we observe then that $\operatorname{range} T' \subseteq (\operatorname{null} T)^0$.

Then, to prove equality, we observe the following:

$$\dim \operatorname{range} T' = \dim W' - \dim \operatorname{null} T'$$
$$= \dim W - (\dim W - \dim \operatorname{range} T)$$
$$= \dim \operatorname{range} T$$
$$= \dim V - \dim \operatorname{null} T$$
$$= \dim(\operatorname{null} T)^0$$

8.1.1 Matrix Representations v2

Now, concretely, what does this mean exactly? To do this, we can look at matrix representations.

First, consider $T \in \mathcal{L}(V, W)$. We note consider some bases for V and W.

Then, for $T' \in \mathcal{L}(W', V')$, we consider some dual bases for W' and V'.

Now, we consider (T); in other words, a matrix with entries A_{ij}.

Now, consider the basis for V, $v_1, \ldots, v_n$, and the basis for W consisting of $w_1, \ldots, w_m$.

Then, we note that $A_{ij} = \psi_i(Tv_j)$.

Then, what about T'? What is the $(j, i)^{th}$ entry in (T')? To do this, we can simply get $(T'\psi_i)(v_j) = \psi_i \circ T(v_j) = \psi_i T(v_j) = A_{ij}$. So, in fact, we see that they're transposes of each other.

In other words, we have:
$$(\mathcal{T})^\mathsf{T} = (\mathcal{T}')$$

POLYNOMIALS

9.1   Lecture - 10/17/2023

To begin with, what is the simplest "building blocks" for polynomials viewed multiplicatively?

The goal for today's lecture is to factor polynomials as much as possible.

> **Remark 9.1.** Using Galois Theory, we can determine which polynomials can be solved using radicals. In this case, we can't create a formula for solving polynomials of degree five.

First, let's consider the case of a degree two polynomial. We have three different scenarios:

1. Two real roots

2. One double root

3. No real roots (we have a complex solution!)

In the first scenario, we observe that for roots at x_1, x_2, we can write our polynomial as $(x - x_1)(x - x_2)$.

In the case of a double root at x_1, we have $(x - x_1)^2$.

In the case where we don't have any real roots, we can still factor our polynomial. However, it will occur over $\mathbb{C}$ rather than $\mathbb{R}$. We can factor our polynomial as follows: $(x - z_1)(x - \overline{z_1})$.

> **Theorem 9.2.** Suppose we have $f(x) \in \mathscr{P}(\mathbb{R})$, and $z \in \mathbb{C}$ is a root of f. In other words, $f(z) = 0$. Then, we note that $f(\overline{z}) = 0$.

Proof. Suppose we have some $f(x) \in \mathscr{P}(\mathbb{R})$ such that $f(z) = 0$, for some $z \in \mathbb{C}$. Now, we observe that this

means, for $a_0, \ldots, a_n \in \mathbb{R}$, we have:

$$f(x) = a_0 x^n + a_1 x^{n-1} + \ldots + a_{n-1} x + a_n$$
$$f(z) = 0$$
$$a_0 z^n + a_1 z^{n-1} + \ldots + a_{n-1} z + a_n = 0$$
$$\overline{a_0 z^n + a_1 z^{n-1} + \ldots + a_{n-1} z + a_n} = \overline{0}$$
$$\overline{a_0 z^n} + \overline{a_1 z^{n-1}} + \ldots + \overline{a_{n-1} z} + \overline{a_n} = 0$$
$$\overline{a_0} \cdot \overline{z}^n + \overline{a_1} \cdot \overline{z}^{n-1} + \ldots + \overline{a_{n-1}} \cdot \overline{z} + \overline{a_n} = 0$$
$$a_0 \overline{z}^n + a_1 \overline{z}^{n-1} + \ldots + a_{n-1} \overline{z} + a_n = 0$$
$$f(\overline{z}) = 0$$

∎

Remark 9.3. We note that for some $z \in \mathbb{R}$ or $\mathbb{C}$, it is a root of a polynomial $f \in \mathscr{P}(\mathbb{R})$ or $\mathscr{P}(\mathbb{C})$ if and only if $f(x) = (x - z)g(x)$ for some $g \in \mathscr{P}$.

Proof. We begin with the backwards direction. We note that if $f(x) = (x - z)g(x)$, then of course $f(z) = (z - z)g(z) = 0g(z) = 0$.

For the forward direction, we observe the following:

Theorem 9.4. For any two polynomials with real or complex coefficients, f, g, we can always find two more polynomials q, r such that:
$$f(x) = q(x)g(x) + r(x).$$

We note that:

$$\deg q = \begin{cases} \deg f - \deg g & \deg f \geq \deg g \\ -\infty & \text{otherwise} \end{cases}$$

And we note that $\deg r < \deg g$.

Example 9.5. To illustrate that $\deg r < \deg g$, we observe the following:

Suppose we have $f(x) = x^3 + 3x^2 + 2x + 1$, and $g(x) = x^2 + 2x + 1$. Then, we observe the following:

$$f(x) = q(x)g(x) + r(x)$$
$$q(x) = x + 1$$
$$r(x) = -x$$

Now, is it possible to have some other $f(x) = \tilde{q}(x)g(x) + \tilde{r}(x)$? We note that if this was the case, we have:

$$\tilde{q}(x)g(x) + \tilde{r}(x) = q(x)g(x) + r(x)$$
$$q(x)g(x) - \tilde{q}(x)g(x) = \tilde{r}(x) - r(x)$$

Then, we note that $\deg \tilde{r}(x) - r(x) \leq \deg g$. Furthermore, we note that $\deg(q(x) - \tilde{q}(x))g(x) \geq \deg g(x)$.

Then...

Now, if $f(z) = 0$, we can always get $f(x) = (x - z)q(x) + r(x)$. Then, we note that $\deg r < \deg(x - z) = 1$. In other words, we have $\deg r < 1$. Moreover, because we have that $f(z) = 0$, then $f(z) = r(x) = 0$. Thus, we must have that $r(x) = 0$. ∎

Theorem 9.6 (Fundamental Theorem of Algebra)**.** Suppose we have $f(x) = \mathscr{P}(\mathbb{C})$ has degree $n \geq 1$. Then, there exists a number $z \in \mathbb{C}$ such that $f(z) = 0$.

Lemma 9.7. Any polynomial in $\mathscr{P}(\mathbb{C})$ of degree $n \geq 1$ factors as:

$$f(x) = a(x - z_1) \cdots (x - z_n),$$

where $z_1, \ldots, z_n \in \mathbb{C}$, and $a \in \mathbb{C} \setminus \{0\}$.

Remark 9.8. Over $\mathbb{R}$, we have

$$f(x) = a(x - z_1)(x - \overline{z_1}) \cdots (x - z_k)(x - \overline{z_k}) \cdot (x - x_1) \cdots (x - x_{n-2k}),$$

where $z_1, \ldots, z_k \in \mathbb{C} \setminus \mathbb{R}$, and $x_1, \ldots, x_{n-2k} \in \mathbb{R}$.

9.2   Lecture - 10/19/2023

Remark 9.9 (Proof for Fundamental Theorem of Algebra)**.** We note that the Fundamental Theorem of Algebra, despite it name, doesn't have a truly algebraic proof.

One of the proofs was by d'Alambert. In LADR, Axler lists two proofs, the second of which requires a result in complex analysis about analytic functions.

The proof uses the fact that if a function f of a complex variable is analytic in $\mathbb{C}$, and is bounded (i.e. $|f(z)| \leq c$, for all $z \in \mathbb{C}$), then $f(z) \equiv c$, where c is some constant.

Then, suppose we have any polynomial p such that $p \in \mathscr{P}(\mathbb{C})$, with $\deg p \geq 1$. Then, suppose that $p(z) \neq 0$ for all $z \in \mathbb{C}$.

Then, we consider $g(z) = \frac{1}{p(z)}$. Then, $g(z)$ is analytic in $\mathbb{C}$, and is also bounded globally. Then, $g(z)$ must be constant as well. From here, it follows that $p(z)$ must be constant too.

However, if this is the case, then $\deg p \not\geq 1$, and thus we have a contradiction.

Thus, we note that $p(z)$ must be equal to zero for some value $z \in \mathbb{C}$.

9.2.1   Eigenvalues and Eigenvectors

Definition 9.10 (Linear Operator)**.** Suppose we have $T \in \mathcal{L}(V)$, where $\dim V < \infty$. We say then that T is a linear operator.

Definition 9.11 (Invariance)**.** We call a subspace $U \subseteq V$ invariant with respect to T if $T(U) \subseteq U$.

Example 9.12. Suppose we have $V = \mathbb{R}^2$, and $T : (x, y) \mapsto (y, x)$.

Then, we observe that all the invariant subspaces are:

1. 0,

2. $\mathbb{R}^2$,

3. $\operatorname{span} \{(1, 1)\}$,

4. $\operatorname{span} \{(1, -1)\}$.

Now, the subquestion of interest to us is: what are the one-dimensional invariant subspaces of a linear operator T?

Theorem 9.13 (Invariance of T)**.** We can express invariance with the following formula:

$$Tv = \lambda v \hspace{4cm} \text{(for some } v \in V \setminus \{0\} , \lambda \in \mathbb{F}\text{)}$$

Example 9.14. Suppose we are in $\mathbb{R}^2$, and we have $T : (x, y) \mapsto (-y, x)$.

Then, this transformation is rotating everything by $90°$ counterclockwise. Thus, we see that we don't have invariance for any one-dimensional subspace.

However, we can think of it in another way: since $Tv = \lambda v$, then it follows that $T^2 v = \lambda^2 v$.

However, we note that $T^2 : (x, y) \mapsto (-x, -y)$. Then, we observe that $\lambda^2 = -1$; however, we note that $\lambda \in \mathbb{R}$, which is thus a contradiction.

Remark 9.15. Note here that if we were working in $\mathbb{C}$, then T has eigenvalues and eigenvectors.

Namely, we note that $T(1, i) = (-i, 1)$. Then, we see that $T(1, i) = -i(1, i)$. Similarly, $T(1, -i) = (i, 1) = i(1, i)$.

Then, we observe that these two vectors in fact form a basis for $\mathbb{C}^2$, which consist of eigenvectors of T such that $\mathcal{T}$ is diagonal in that basis.

We observe here that the eigenvectors are indeed linearly independent.

Proof. To show this, suppose that $\lambda_1, \ldots, \lambda_k$ are distinct eigenvalues of $T \in \mathcal{L}(V)$. Furthermore, suppose that $v_1, \ldots, v_k$ are corresponding eigenvectors.

We proceed by induction. Suppose that our base case works.

Now, suppose that $av_1 + \ldots + kv_k = 0$. Then,

$$T(av_1 + \ldots + kv_k) = T(0)$$
$$aT(v_1) + \ldots + kT(v_k) = 0$$
$$a\lambda_1 v_{k-1} + \ldots + k\lambda_k v_k = 0$$

Then, we note that $(\lambda_k)(av_1 + \ldots + kv_k) = 0$. Then, if we subtract this from the previous line, we have:

$$a(\lambda_1 - \lambda_k)v_1 + \ldots + j(\lambda_1 - \lambda_k - 1)v_{k-1} = 0$$

Then, we observe that each $\lambda_i - \lambda_k \neq 0$, so it follows then that all $a, \ldots, j$ must be equal to zero as well. Hence, we see that $a_k = 0$ as well. ∎

9.3   Discussion - 10/20/2023

> **Problem 9.1.** Let $f \in \mathscr{P}(\mathbb{R})$, with $\deg f \geq 3$. Then, prove that there exists $g, h \in \mathscr{P}(\mathbb{R})$, both non-constant, such that $f = g \cdot h$.

Solution. We first note that since .   ∎

> **Problem 9.2.** Let p be a polynomial over $\mathbb{C}$. Show that p has distinct roots if and only if p and p' has no common roots.

Solution. First, we will proceed with the forward direction. Suppose that p has distinct roots.

Since p has distinct roots, it follows then that we can rewrite p as:

$$p(z) = c(z - \lambda_1) \cdots (z - \lambda_n)$$

And since the roots are all distinct, it follows then that $\lambda_1, \ldots, \lambda_n$ are all not equal to each other.

From here, we note then that $p'(z)$ can be written as:

$$p'(z) = c \left(\sum_{i=1}^{n} \prod_{j=1, j \neq i}^{n} (z - \lambda_j) \right)$$

Then, from here, we observe that at each λ_i's such that $p(\lambda_i) = 0$, we have that $p'(z) \neq 0$; we can see this from the fact that λ_i will eliminate all terms except one: the term which doesn't contain $(z - \lambda_i)$ in its product.

So, it follows that p and p' have no common roots.

Now, let us proceed with the backwards direction. To do this, let us proceed by contraposition.

Suppose instead that p has a root of multiplicity greater than 1. Then, it follows that we have:

$$p(z) = c(z - \lambda_1)^{n_1} \cdot (z - \lambda_n)^{n_n},$$

where $n_i \geq 1$, and each λ_i. Furthermore, since p has some non-distinct root, at least one of n must be greater than or equal to 1.

From here, when we differentiate it, we have:

$$p'(z) = c \left(\sum_{i=1}^{n} n_i (z - \lambda_i)^{n_i - 1} \prod_{j=1, j \neq i}^{n} (z - \lambda_j)^{j} \right).$$

Then, we observe that for the $n_k > 1$, we have $(z - \lambda_k)^{n_k - 1}$, and $n_k - 1 \geq 1$ somewhere in the polynomial p', and the other terms also contains $z - \lambda_k$ in it. Then, it follows that for $z = \lambda_k$, we have that $p'(z) = 0$; thus, p and p' have common roots.

Therefore, we have shown the backwards direction, and thus can conclude that the claim is indeed true.   ∎

> **Problem 9.3.** Suppose that $f(x) \geq 0$ for all real numbers if and only if there exists real polynomials g, h such that $f(x) = g^2(x) + h^2(x)$.

Solution. Let us proceed with the backwards direction. Suppose that $f(x) = g^2(x) + h^2(x)$.

We note then that $g^2(x) \geq 0$ and $h^2(x) \geq 0$. Then, it follows that their sum $f(x) = g^2(x) + h^2(x) \geq 0$ as well.

For the forward direction, let us suppose that $f(x) \geq 0$ for all real numbers. Now, we note that for each real root of $f(x)$, it must have even multiplicity; if the multiplicity was odd then $f(x)$ would cross the x-axis, and thus violate $f(x) \geq 0$. For the remaining factors of $f(x)$, it must be of some quadratic of the form $x^2 + bx + c$, with $b^2 < 4c$; it must have complex roots.

Additionally, for any number $a \in \mathbb{R}$, we can rewrite it as the square of some other number $c \in \mathbb{R}$; $c^2 = a$.

Then, with this in mind, we can write $f(x)$ as follows:

$$f(x) = c^2(x - \lambda_0)^2 \cdots (x - \lambda_m)^2 [(x - \mu_0)(x - \overline{\mu_0})] \cdots [(x - \mu_n)(x - \overline{\mu_n})],$$

where we have $c, \lambda_0, \ldots, \lambda_m \in \mathbb{R}$ and $\mu_0, \ldots, \mu_n \in \mathbb{C}$.

From here, let us define the functions g, h as follow:

$$s(x) = c^2(x - \lambda_0) \cdots (x - \lambda_m)$$
$$t(x) = (x - \mu_0) \cdots (x - \mu_n)$$

Furthermore, we can split $t(x)$ into a "real" and "imaginary" component as follows:

$$t(x) = t_{\mathrm{Re}}(x) + it_{\mathrm{Im}}(x)$$

Then, from here, we note that we can rewrite $f(x)$ as follows:

$$
\begin{aligned}
f(x) &= s^2(x)t(x)\overline{t(x)} \\
&= s^2(x)(t_{\mathrm{Re}}(x) + it_{\mathrm{Im}}(x))(t_{\mathrm{Re}}(x) - it_{\mathrm{Im}}(x)) \\
&= s^2(x)(t_{\mathrm{Re}}^2(x) + t_{\mathrm{Im}}^2(x)) \\
&= s^2(x)t_{\mathrm{Re}}^2(x) + s^2 t_{\mathrm{Im}}^2(x) \\
&= (s(x)t_{\mathrm{Re}}(x))^2 + (s(x)t_{\mathrm{Im}}(x))^2
\end{aligned}
$$

Therefore, if we let

$$g(x) = s(x)t_{\mathrm{Re}}(x)$$
$$h(x)0s(x)t_{\mathrm{Im}}(x),$$

then we can rewrite $f(x)$ as

$$f(x) = g^2(x) + h^2(x)$$

$\blacksquare$

WEEK 10
EIGENWHATEVERTHEFUCK

10.1 Lecture - 10/24/2023

10.1.1 Recap

Remember from last time that we established that for some $T \in \mathcal{L}(V)$, an eigenvalue λ is one such that

$$Tv = \lambda v,$$

where $v \neq 0, \lambda \in \mathbb{F}$. Furthermore, we observe that we have then the following:

$$(T - \lambda I)v = 0.$$

This gives rise to the following statements which are all equivalent:

Theorem 10.1. If λ is an eigenvalue of T, then the following are equivalent:

1. $T - \lambda I$ is not invertible.

2. $T - \lambda I$ is not injective.

3. $T - \lambda I$ is not surjective.

Note that this only applies to finite-dimensional vector spaces V.

Furthermore, recall from last lecture that we have the following:

Theorem 10.2. Each eigenvector corresponding to distinct eigenvalues are necessarily linearly independent.

This leads to the following corollary:

Corollary 10.3. Any $T \in \mathcal{L}(V)$ has at most $\dim V$ distinct eigenvalues.

Now, we will begin thinking of the following question: can we always find eigenvalues/vectors if V is over $\mathbb{C}$? What about over $\mathbb{R}$?

And a related question is: what do eigenvectors have to do with polynomials?

To do this, let us first proceed with the following topic:

10.1.2   Applying Polynomials to Operators

Suppose that we have $T \in \mathcal{L}(V)$. Then, we can consider

$$T^n = T \circ T \circ \cdots \circ T \qquad \text{(n times)}$$

Then, we can understand any polynomial in T, in other words $p(T)$ where $p(x) = a_0 + a_1 x_1 + \ldots + a_n x^n$. We can just think of this as replacing x with T.

> **Example 10.4.** Suppose we have $T : (x, y, z) \mapsto (z, x, y)$.
>
> What is $T^2 + T + I = p(t)$, where $p(x) = x^2 + x + 1$. Then, we have:
>
> $$p(T) : (x, y, z) \mapsto (y, z, x) + (z, x, y) + (x, y, z) = (x + y + z, x + y + z, x + y + z)$$

Now, suppose that we have (λ, v) which is an eigenpair of T. What is $p(T)v$? In this case, we can simply find $p(\lambda)v$

Indeed, we observe that $T^n v = \lambda^n v$. Then, we have:

$$
\begin{aligned}
p(T) &= a_0 I + a_1 T + \ldots + a_n T^n \\
&= a_0 \lambda^0 + a_1 \lambda + \ldots + a_n \lambda^n \\
&= p(\lambda)v.
\end{aligned}
$$

Now, let's see if we can always find eigenvalues and eigenvectors for some finite-dimensional vector space V over $\mathbb{C}$.

The answer is yes.

> **Theorem 10.5.** We can always find eigenvalues/eigenvectors if V is a finite-dimensional vector space over $\mathbb{C}$.

Proof. Suppose that we have a finite-dimensional vector V with dimension $\dim V = n$.

Let us fix some $v \in V \setminus \{0\}$.

Now, consider the list $v, Tv, \ldots, T^n v$. This list then contains $n + 1$ vectors, and thus they must be linearly dependent.

There, we know that there exists a non-trivial combination which is equal to zero. Moreover, there is a smallest power k such that $v, Tv, \ldots, T^k$ is dependent.

In other words, we have:

$$a_0 v + a_1 Tv + \ldots + a_k T^k v = 0$$

Then, we know that $a_k \neq 0$; otherwise, k would not be the smallest power which makes the list dependent. Furthermore, we know that $k > 0$ (as we start with a non-zero vector).

So, without loss of generality, we can let $a_k = 1$. Then, note in fact that this can be rewritten as $p(T)v = 0$, where $p(x) = a_0 + a_1 x + \ldots + a_k x^k$.

Then, the eigenvalues are ones such that

$$T - \lambda I = 0.$$

Then, since we are over $\mathbb{C}$, we can in fact split our polynomial up into degree 1 factors. Then, we can rewrite $p(x)$ as:

$$p(x) = (x - \lambda)q(x),$$

for some $\lambda \in \mathbb{C}$ and $q \in \mathscr{P}(\mathbb{C})$.

Then, we note that λ is an eigenvalue corresponding to $q(T)v$:

$$p(T)v = (T - \lambda I)q(T)v = 0$$

We should also check that $q(T)v \neq 0$, else it wouldn't be valid. However, we recall that since $\deg q(x) = k - 1$, and we note that k is the *smallest* degree such that we have a linear dependence, it follows then that $q(T)v \neq 0$, else it would violate this condition.

So, we have found an eigenvalue and eigenvector as desired. $\blacksquare$

Example 10.6. Let $V = \operatorname{span}\{1, \cos x, \sin x, \cos 2x, \sin 2x\}$, and define T to be $T := D^3 - D$.

Now, let $v(x) = \cos 2x$. Then, we have:

$$(Tv)(x) = 10 \sin 2x$$

Then, $(T^2 v)x = -100 \cos(2x)$. So, we observe that $100v + T^2 v = 0$.

In other words, we have $p(T)v = 0$, where $p(t) = 100 + t^2$. So, we have:

$$p(t) = (t + 10i)(t - 10i).$$

Now, from this construction, we observe that the eigenvector corresponding to $\lambda = 10i$ is $[(T+10iI)v](x) = 10 \sin 2x + 10i \cos 2x$.

This is the same as $10e^{2ix}$.

Then, to check our answer, we note that ...

Remark 10.7. A better basis to work with, at least for differential operators, is the exponential basis:

$$1, e^{ix}, e^{-ix}, e^{2ix}, e^{-2ix}.$$

10.2   Lecture - 10/26/2023

10.2.1   Minimal Polynomials

Theorem 10.8. Suppose that we have $T \in \mathcal{L}(V)$ with $\dim V < \infty$. Now, V is over some field $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$. Then, there exists a monic polynomial $p \in \mathscr{P}(\mathbb{F})$ such that $p(T) = 0$ (on V).

We recall that previously, we were working over $\mathbb{C}$. However, this isn't necessarily the case for here. Furthermore, last lecture, we had $p(T)(v) = 0$ for some fixed v.

Today, we will find a monic polynomial such that $p(T) = 0$; in other words, we don't have to fix v.

Proof. We shall proceed by induction on $\dim V$.

Base Case: Suppose that $\dim V = 1$. Then, we note that $Tv = \lambda v$. Then, we want the minimal polynomial to be $p(t) = t - \lambda$, since $T - \lambda I = 0$.

Inductive Step: Suppose that $\dim V > 1$. Then, we observe that for V, we can take some $v \in V$ (where $v \neq 0$). We know from last lecture that there exists some $g \in \mathscr{P}(\mathbb{F})$ such that $g(T)v = 0$. Furthermore, $1 \leq \deg g \leq \dim V$

Let us denote $\deg g = k$. Then, due to the minimality of k with this property, we know then that $v, Tv, \ldots, T^{k-1}v$ are linearly independent.

Then, let us consider $\operatorname{null} g(T)$. Then, we observe that $v \in \operatorname{null} g(T)$. Furthermore, we note that $Tv, \ldots, T^{k-1}v$ are also included in $\operatorname{null} g(T)$.

We note then that this follows from the fact that:

$$g(T)(T^j v) = T^j(g(T)v) = 0$$

So, we know that $\dim \operatorname{null} g(T) \geq k \geq 1$. Furthermore, we know then that $\dim \operatorname{range} g(T) = \dim V - \dim \operatorname{null} g(T)$.

Then, we have:

$$\dim \operatorname{range} g(T) = \dim V - \dim \operatorname{null} g(T)$$
$$\leq \dim V - k$$
$$\leq \dim V - 1$$

However, we note that $\operatorname{range} g(T)$ is T-invariant. Then, we can use the induction hypothesis on $\operatorname{range} g(T)$ to conclude that we can create a monic polynomial $q \in \mathscr{P}(\mathbb{F})$ whose degree $\deg q \leq \dim \operatorname{range} g(T)$ such that $q(T \mid_{\operatorname{range} g(T)}) = 0$ on the range of $g(T)$.

Then, consider the product $g(t)q(t)$. Then, we note that

$$q(T)g(T) = 0$$

This follows from the fact $q(T)g(T)v = q(T)w$, where $w \in \operatorname{range} g(T)$. Then, we note that any vector $w \in \operatorname{range} g(T)$ gets sent to 0 by $q(T)$. So, we see that this polynomial $q(t)g(t) \leq \dim \operatorname{range} g(T) + \deg g(t) \leq \dim V$.

Thus, we have an existence of some polynomial which satisfies our condition.

However, this does not necessarily guarantee that $g(t)q(t)$ is the minimal polynomial. But, we note that since there are only finitely many degrees below, there exists a polynomial p of minimal degree.

Thus, we have $p(T) = 0$, and p is in fact monic. ∎

Then, what are the connections between the minimal polynomial and eigenvalues? We claim the following:

> **Theorem 10.9.** The zeroes of the minimal polynomial are precisely the eigenvalues of T.

Proof. First, suppose that λ is an eigenvalue of T. Then, we note that $p(T)v = p(\lambda)v = 0$, for any eigenvector associated with λ.

In other words, we have $p(\lambda) = 0$; λ is a zero of our minimal polynomial.

On the other hand, if we have some minimal polynomial p and we have $p(\lambda) = 0$, for some $\lambda \in \mathbb{F}$. Then, we note that we can factor our a degree one term from $p(t)$ to get $p(t) = (t - \lambda)q(t)$ for some $q \in \mathscr{P}(\mathbb{F})$.

Then, we observe that $p(T) = (T - \lambda I)q(T)$. Then, by the minimality of p, we have that $q(T) \neq 0$. So, we note that there exists some v such that $q(T)v \neq 0$. Then, say $q(T)v = w$. Then, we have $p(T)w = (T - \lambda I)w$. So, $q(T)v$ is an eigenvector associated with an eigenvalue λ. ∎

Then, this leads to the following crucial idea:

> **Theorem 10.10.** If V is a finite-dimensional vector space over $\mathbb{R}$, we may not necessarily have eigenvalues. More specifically, only when $\dim V$ is odd is it that we have an eigenvalue.

10.2.2 Triangularisable Case

This is when computing the minimal polynomial is easy.

> **Definition 10.11** (Triangularisability)**.** When T is triangularisable, it means that we have a matrix representation $\mathcal{M}(T)$ such that
> $$\mathcal{M}(T) = \begin{bmatrix} * & * & \cdots & * \\ 0 & * & \cdots & * \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & * \end{bmatrix}$$
> or a lower triangular matrix idk. ill tex it out later.

Consider a basis for V: $v_1, \ldots, v_n$.

We note that in the lower-triangular matrix, if we apply T to the first vector, we get a linear combination of all of the vectors. On the second vector, we get a linear combination of the second vector and beyond.

Meanwhile, the upper-triangular matrix is simply applying T onto our basis for V, but in reverse.

Then, say we have

$$\begin{bmatrix} \lambda_1 & * & \cdots & * \\ 0 & \lambda_2 & \cdots & * \\ \vdots & \ddots & \ddots & * \\ 0 & \cdots & \cdots & \lambda_n \end{bmatrix}$$

Then, we note that the polynomial $(x - \lambda_1)(x - \lambda_2) \cdots (x - \lambda_n)$ is an annihilating polynomial for T. We see that if we apply $p(T)$ onto v_n, then it can be written as a linear combination of the previous $n - 1$ vectors. We repeat this process until we get to the first vector. Then, we see that $p(T)v_1 = 0$, as desired.

10.3 Discussion - 10/27/2023

> **Problem 10.1.** Define $T \in (\mathbb{R}^2)$ as $T(x, y) = (y, x)$. Find the eigenvalues and eigenvectors of T. Furthermore, find the minimal polynomial of T.

Solution. We observe that we want the following to hold true:

$$Tv = \lambda v$$

Then, we observe that for $v = (x, y)$, we have:

$$Tv = \lambda v$$
$$(y, x) = (\lambda x, \lambda y)$$

So, we see that $\lambda = 1$ for $(x, y) = (1, 1)$ satisfies this. ∎

Problem 10.2. If $T \in \mathcal{L}(V)$ satisfies $T^2 = I$, then there is a basis of V formed by eigenvectors of T.

Solution. We observe that if $T^2 = I$, then it follows that T is in fact its own inverse. This further means that T is in fact invertible. ∎

DIALGANALISATION

11.1 Lecture - 10/31/2023

11.1.1 Triangular Form

We recall from last lecture that if we have the following matrix representation for T

$$\mathcal{M}(T) = \begin{bmatrix} \lambda_1 & * & \cdots & * \\ 0 & \lambda_2 & \cdots & * \\ \vdots & \ddots & \ddots & * \\ 0 & \cdots & \cdots & \lambda_n \end{bmatrix},$$

then the polynomial $(x - \lambda_1) \cdots (x - \lambda_n)$ will annihilate T. However, we note that this polynomial is **not** necessarily the minimal polynomial.

Example 11.1. To begin with, let $V = \mathbb{R}^3$. Then, we note that any linear operator T such that

$$\mathcal{M}(T) = \begin{bmatrix} \lambda_1 & * & * \\ 0 & \lambda_2 & * \\ 0 & 0 & \lambda_3 \end{bmatrix},$$

where $\lambda_1, \lambda_2, \lambda_3$ are all distinct.

More concretely, let us define $T : (x, y, z) \mapsto (x, 2y + 3x, 3z - x - y)$. Then,

$$\mathcal{M}(T) = \begin{bmatrix} 1 & 3 & -1 \\ 0 & 2 & -1 \\ 0 & 0 & 3 \end{bmatrix}$$

Then, we claim that the diagonal entries of T are the eigenvalues. The easiest way to verify this is as follows:

Let us take $\mathcal{M}(T) - 2I$. Then, we note that the first and third diagonal entries are non-zero, but the second is zero. Then, we claim that $\mathcal{M}(T) - 2I$ is non-invertible; this follows from the fact that since the second diagonal entry is zero, the second column lives in the span of the first.

Thus, we observe that the minimal polynomial is $p(t) = (t - 1)(t - 2)(t - 3)$.

Example 11.2. On the other hand, let us consider an example for the other direction. In this case, let $T := I$. Then, we observe that

$$\mathcal{M}(T) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1. \end{bmatrix}$$

Then, of course, $p(t) = (t-1)^3$ annihilates T. However. the minimal polynomial is simply $(x-1)$.

Notice that if we had

$$\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

then the minimal polynomial will be $(x-1)^2$.

Now, we will introduce an important theorem for triangular matrix representations:

Theorem 11.3. An operator $T \in \mathcal{L}(V)$ has an upper-triangular matrix representation if and only if its minimal polynomial has the form $(x - \lambda_1) \cdots (x - \lambda_n)$, where $\lambda_i \in \mathbb{F}$ (more specifically, $\mathbb{R}$ or $\mathbb{C}$).

Proof. Let us proceed with the forward direction. Suppose that T is upper-triangular. Then, it follows that the polynomial $(x - \lambda_1)^k \cdots (x - \lambda_n)^k$, where each λ_i are the diagonal entries of T. We observe then that the minimal polynomial will consist of the same factors, but with powers $l \leq k$ for each term.

Let us proceed with the backwards direction. Suppose that the minimal polynomial has the form $(x - \lambda_1) \cdots (x - \lambda_n)$, where $n \leq \dim V$. We shall use induction on k:

First, if $k = 1$, then the minimal polynomial is simply $(x - \lambda_1)$. Then, we see that $T - \lambda_1 I = 0$. Then, we have

$$\mathcal{T} = \lambda_1 \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & 1 \end{bmatrix}.$$

Thus, T is upper triangular.

Now, suppose that $p_{\min}(x) = (x - \lambda_1)q(x)$, where $q(x) = (x - \lambda_2) \cdots (x - \lambda_n)$.

Then, we already have λ_1 as an eigenvalue of T. Then, we can split our space V into a direct sum $U_1 \oplus U_2$. Then, **INSERT PROOF HERE** ∎

Note here that we do not require the eigenvalues to be distinct!

Now, with this theorem in mind, we observe that we in fact have the following corollary:

Corollary 11.4. Any $T \in \mathcal{L}(V)$ will always have an upper-triangular matrix representation if $\mathbb{F} = \mathbb{C}$.

Proof. We recall this is because in $\mathbb{C}$, any polynomial we have can be factored completely. ∎

11.1.2 Diagonal Form

The following are all equivalent statements:

> **Theorem 11.5.** T is diagonalisable if $\mathcal{M}(T)$ is diagonal in some basis. Then, we observe the following about $\mathcal{M}(T)$:
>
> $$\mathcal{M}(T) = \begin{bmatrix} \lambda_1 I & & \\ & \ddots & \\ & & \lambda_n I \end{bmatrix}$$
>
> Then, we observe that, in fact, the following are all equivalent:
>
> 1. $V = E(\lambda_1, T) \oplus \ldots \oplus E(\lambda_n, T)$. Conversely, if V splits like this, then we can take an eigenbasis for each E_i, combine them together, and thus will get a diagonalisable matrix $\mathcal{M}(T)$.
>
> 2. $\dim E(\lambda_1, T) + \ldots + \dim E(\lambda_n, T) = \dim V$.

11.2 Lecture - 11/2/2023

We recall from last time that we discussed triangular matrices. Now, suppose we have the following matrix for T;

$$\mathcal{T} = \begin{bmatrix} \lambda_1 I & 0 & 0 \\ 0 & \lambda_2 I & 0 \\ 0 & 0 & \lambda_k I \end{bmatrix}$$

Now, note that $\mathcal{M}(p(T)) = p(\mathcal{M}(T))$. Then, we have:

$$\begin{bmatrix} p(\lambda_1) I & 0 & 0 \\ 0 & p(\lambda 2) I & 0 \\ 0 & 0 & p(\lambda_k) I \end{bmatrix}$$

Then, we observe that in fact, we have the minimal polynomial to be $(x - \lambda_1)(x - \lambda_2) \cdots (x - \lambda_k)$.

Now instead, let us suppose that $(x - \lambda_1)(x - \lambda_2) \cdots (x - \lambda_k)$, where λ_i are all distinct. Then, we make the following observation: for any λ_j, $\text{null}(T - \lambda_j I) \cap \text{range}(T - \lambda_j I) = \{0\}$.

We see that for any $v \in \text{range}(T - \lambda_j I) \cap \text{null}(T - \lambda_j I)$, we have $v = (T - \lambda_j I)w$, for some $w \in V$.

Then, we see that $(T - \lambda_1 I) \cdots (T - \lambda_k I)(T - \lambda_j I)w$. Then, we see that this is equal to zero.

On the other hand, suppose $v \in \text{null}(T - \lambda_j I)$, then $Tv = \lambda_j v$. So, we have the following:

$$\begin{aligned} (T - \lambda_1 I) \cdots (T - \lambda_k I)(T - \lambda_j I)w &= (T - \lambda_1 I) \cdots (T - \lambda_k I)v \\ &= (\lambda_j - \lambda_1) \cdots (\lambda_j - \lambda_k)v \\ &= 0 \end{aligned}$$

Now, we note that all of the $(\lambda_j - \lambda_i)$ are non-zero, and thus we can conclude that v then must be equal to zero.

Then, we note that, by the Rank-Nullity theorem, we have:

$$\text{null}(T - \lambda_j I) \oplus \text{range}(T - \lambda_j I) = V.$$

So, with respect to any basis of $\text{null}(T - \lambda_1 I)$ and a basis of $\text{range}(T - \lambda_1 I)$, we can put them together and are guaranteed to have a block diagonal form of T. Note that the non-diagonal blocks in this matrix will be zero; this comes from the fact that the range and nullspace are invariant subspaces under T.

Incidentally, for any choice of basis in $\text{null}(T - \lambda_1 I)$, the corresponding submatrix is simply $\lambda_1 I$. And on the range of $\text{range}(T - \lambda_1 I)$, we simply use the induction hypothesis to get a diagonal form. We can do this because $T \mid_{\text{range}(T - \lambda_1 I)}$ is annihilated by the other factors $(x - \lambda_2) \cdots (x - \lambda_k)$.

So, the minimal polynomial for $T \mid_{\text{range}(T - \lambda_1 I)}$ still splits into distinct, degree 1 factors.

Remark 11.6. The base case for where we only have one factor, then T is simply λI.

Example 11.7. Suppose we have $V = \mathscr{P}_2^2(\mathbb{C})$. This is the span of the polynomials $1, x, y, x^2, y^2, xy$. Let $T \in \mathcal{L}(V)$ such that $T : p \mapsto \frac{\partial p}{\partial x} + \frac{\partial p}{\partial y}$.

Then, is T triangularisable? And is it diagonalisable?

First, since V is over $\mathbb{C}$, we know that T must be triangularisable. To check for diagonalisability, we first find $\mathcal{M}(T)$:

$$\mathcal{T} = \begin{bmatrix} 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 & 1 \\ 0 & 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Now, we note that the only eigenvalue of T is 0, but $T \neq 0$, then T does not diagonalise.

To find the minimal polynomial, we note that if we apply T^2 onto x^2, we have non-zero result. However, T^3 is zero for everything.

11.2.1 Commutativity

Theorem 11.8. Suppose that two operators $T, S \in \mathcal{L}(V)$ commute. In other words, $ST = TS$. Then, any eigenspace E for T must be invariant under S, and vice versa.

Proof. If $Tv = \lambda v$, then $TSv = STv = S(\lambda v) = \lambda(Sv)$. So, we have that Sv is an eigenvector of T; that is, it belongs to the eigenspace of T. ∎

A consequence of this is the following:

Corollary 11.9. Over $\mathbb{C}$, any two commuting operators must share an eigenvector. Furthermore, any two commuting linear operator can be simultaneously triangularised.

Definition 11.10 (Simultaneously Triangularisation)**.** There exists some basis such that both S and T are triangular (not mixed though; both will be of the same form).

Proof. ... ∎

11.3 Discussion - 11/3/2023

11.3.1 How to Find Eigenvalues/Eigenvectors

In order to find the eigenvalues of some A, we want to first find $\det(A - \lambda I)$, yielding us the characteristic polynomial. Then, the roots of this polynomial are our eigenvalues.

Then, any $v \in \text{null}(A - \lambda I)$ such that $v \neq 0$ are our eigenvectors corresponding to λ.

11.3.2   Is it Diagonalisable?

To begin with, we will introduce the following definitions:

> **Definition 11.11** (Multiplicities of Eigenvalues)**.** Let λ be an eigenvalue.
>
> Then, we denote its algebraic multiplicity as the multiplicity of λ as a root of the characteristic polynomial.
>
> Meanwhile, its geometric multiplicity is simply $\dim \operatorname{null}(A - \lambda I)$.
>
> Note that the geometric multiplicity $\leq$ algebraic multiplicity.

If $T \in \mathcal{L}(V)$ is diagonalisable, then

1. The algebraic multiplicity and geometric multiplicity of every λ must be equal.

2. Let $\lambda_1, \ldots, \lambda_n$ be all distinct eigenvalues of T. Then, $(\lambda - \lambda_1) \cdots (\lambda - \lambda_n)$ is the minimal polynomial of T.

> **Problem 11.1.** If $T \in \mathcal{L}(\mathbb{C}^4)$ is diagonalisable, with three distinct eigenvalues $\lambda_1, \lambda_2, \lambda_3$, then the minimal polynomial of T is (λ)

Solution. Suppose that T is diagonalisable, with three distinct eigenvalues $\lambda_1, \lambda_2, \lambda_3$. Then, since T is diagonalisable, we note that there exists some basis $v_1, \ldots, v_4$ consisting of the eigenvectors of T.

From here, we note that for each v_i, there exists a λ_j such that

$$T(v_i) = \lambda_j v_i$$
$$(T - \lambda_j I) v_i = 0$$

Then, we note that we have:

$$(T - \lambda_1)(T - \lambda_2)(T - \lambda_3) v_i = 0.$$

Thus, we see that, indeed, the minimal polynomial for T is $(\lambda - \lambda_1)(\lambda - \lambda_2)(\lambda - \lambda_3)$. ∎

> **Problem 11.2.** If λ is an eigenvalue of T, then $p(\lambda)$ is an eigenvalue of $p(T)$.

Solution. Suppose that λ is an eigenvalue of T. Then, by definition, we have that $Tv = \lambda v$, for some $v \neq 0$.

Now, we observe that

$$\begin{aligned}
p(T)(v) &= (a_0 I + a_1 T + \cdots + a_n T^n)(v) \\
&= a_0 v + a_1 T v + \ldots + a_n T^n v \\
&= a_0 \lambda^0 v + a_1 \lambda^1 v + \cdots + a_n \lambda^n v
\end{aligned}$$

Now, we note that we have for the $v \neq 0$:

$$\begin{aligned}
p(\lambda) &= a_0 \lambda^0 + a_1 \lambda + \cdots + a_n \lambda^n \\
p(\lambda)(v) &= (a_0 \lambda^0 + a_1 \lambda + \cdots + a_n \lambda^n)(v) \\
&= a_0 \lambda^0 v + a_1 \lambda^1 v + \cdots + a_n \lambda^n v \\
&= p(T)(v)
\end{aligned}$$

Thus, we see that since for $v \neq 0$, we have $p(T)(v) = p(\lambda)v$, it follows then that, indeed, if λ is an eigenvalue of T, then $p(\lambda)$ is an eigenvalue of $p(T)$. ∎

Problem 11.3. If $T \in \mathcal{L}(V)$ is upper-triangular with respect to the basis $v_1, \ldots, v_n$, show then that any polynomial of T is upper-triangular with respect to the basis of $v_1, \ldots, v_n$.

Solution. Suppose that T is upper-triangular with respect to some basis $v_1, \ldots, v_n$ of V. Then, we have the following:

$$\mathcal{M}(T) = \begin{bmatrix} \lambda_1 & * & \cdots & * \\ 0 & \lambda_2 & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

■

INNER PRODUCTS

12.1   Lecture - 11/7/2023

12.1.1   Inner Products

This week, we will be discussing the concept of "inner products."

> **Definition 12.1** (Inner Product)**.** An inner product space is a vector with an additional operation which is denoted by $\langle \cdot, \cdot \rangle$. The $\cdot$ denotes where the arguments go.
>
> We note that the inner product is a function from $V \times V \to \mathbb{F}$.

Before we continue, recall that we are working in a field $\mathbb{R}$ or $\mathbb{C}$.

Now, an inner product has the following properties, with linearity in the first slot:

1. **Positivity** (Non-Negativity): $\langle v, v \rangle \geq 0$, for all $v \in V$.

2. **Definiteness**: $\langle v, v \rangle = 0 \iff v = 0$.

3. **Additivity**: $\langle v_1 + v_2, w \rangle = \langle v_1, w \rangle + \langle v_2, w \rangle, \forall v_1, v_2, w \in V$.

4. **Homogeneity**: $\langle \lambda v, w \rangle = \lambda \langle v, w \rangle, \forall v, w \in V$ and $\forall \lambda \in \mathbb{F}$.

5. **Conjugation "Symmetry"**: $\langle v, w \rangle = \overline{\langle w, v \rangle}, \forall v, w \in V$.

> **Remark 12.2.** Note here that additivity holds in the second slot, but homogeneity is replaced with some "homogeneity with conjugation." More specifically, we have:
> $$\begin{aligned} \langle v, \lambda w \rangle &= \overline{\langle \lambda w, v \rangle} \\ &= \overline{\lambda \langle w, v \rangle} \\ &= \overline{\lambda} \langle v, w \rangle. \end{aligned}$$

> **Remark 12.3.** Such a product is called sesquilinear, where "sesqui" means "one and a half."

Subordinate to any inner product, we have the norm, which is defined to be:

> **Definition 12.4** (Norm).
> $$\|v\| = \sqrt{\langle v, v \rangle}.$$

Now, we will look at the following example

> **Example 12.5** (Dot Product). Suppose we have $V = \mathbb{R}^n$, with the feld $\mathbb{F} = \mathbb{R}$.
>
> Then, a standard inner product is the following:
>
> $$\langle (x_1, \ldots, x_n), (y_1, \ldots, y_n) \rangle = x_1 y_1 + \ldots + x_n y_n.$$
>
> This is called a "dot product."
>
> To check that this is indeed an inner product, we proceed as follows:
>
> 1. Positive Definiteness: We note that $\langle x, x \rangle = x_1^2 + \ldots + x_n^2$, which is non-negative, and is equal to zero only when $x = 0$.
>
> 2. Additivity: We can just do some computations.
>
> 3. Homogeneity: As above.
>
> 4. Conjugation "Symmetry": We observe that since $a = \bar{a}$ for all $a \in \mathbb{R}$, we have then that $x_1 y_1 = y_1 x_1 = \overline{y_1 x_1}$.

> **Example 12.6** (Modifying the Dot Product for $\mathbb{C}$). However, we note that if we are working in $V = \mathbb{C}^n$ (over $\mathbb{C}$), we no longer have an inner product using the dot product. Consider having a vector v where all of the entries are equal to i; then, $\langle v, v \rangle < 0$.
>
> Furthermore, if we have, say, $\langle (1 + 3i, i), (1 + 3i, i) \rangle = 6i - 9$, which is non-real.
>
> So, to fix this, we need to redefine the dot product as such:
>
> $$x_1 \overline{y_1} + \ldots + x_n \overline{y_n} = \langle x, y \rangle.$$
>
> Then, we can check all of our requirements and see that this is indeed an inner product.

> **Example 12.7.** Suppose we have $V = C[0, 1]$, such that $\{ f : [0, 1] \to \mathbb{C}, \text{ where } f \text{ is continuous on } [0, 1] \}$.
>
> Then, an inner product would be:
>
> $$\langle f, g \rangle = \int_0^1 f(t) \overline{g(t)} \mathrm{d}t.$$
>
> This can be generalised to
>
> $$\int_0^1 f(t) \overline{g(t)} w(t) \mathrm{d}t,$$
>
> where $w : [0, 1] \to \mathbb{R}$.

12.1.2 Basic Consequences

Some basic consequences of our inner product is that

1. $\langle v, 0 \rangle = 0 = \langle 0, v \rangle$.

2. $\|v\| \geq 0$, for any $v \in V$.

3. $\|v\| = 0 \iff v = 0$.

4. $\|\alpha v\| = |\alpha| \, \|v\|$.

12.1.3 Orthogonality

Definition 12.8. We say that vectors v, w are orthogonal to each other if $\langle v, w \rangle = 0$.

With this definition, we note that the zero vector is orthogonal to the entire space of V. That is, $0 \perp V$. Conversely, if $v \perp V$, then $v \perp v$; in other words, $\langle v, v \rangle = 0 \implies v = 0$.

Thus, we note that the zero vector is the only one orthogonal to the entire space.

Theorem 12.9. Suppose $v \perp w$. Then,

$$\|v + w\|^2 = \|v\|^2 + \|w\|^2.$$

Proof. We note that

$$
\begin{aligned}
\|v + w\|^2 &= \langle v + w, v + w \rangle \\
&= \langle v, v \rangle + \langle w, v \rangle + \langle v, w \rangle + \langle w, w \rangle \\
&= \langle v, v \rangle + \langle w, w \rangle \\
&= \|v\|^2 + \|w\|^2.
\end{aligned}
$$

∎

We note that in fact, we also have:

$$\|v - w\|^2 = \|v\|^2 + \|w\|^2$$

Remark 12.10. The Pythagorean Theorem generalises to the case of multiple orthogonal vectors, $v_1, \ldots, v_k$ such that $v_i \perp v_j$, for all $i \neq j$.

So, we have

$$
\begin{aligned}
\|v_1 + \ldots + v_k\|^2 &= \sum_{i,j=1}^{k} \langle v_i, v_j \rangle \\
&= \sum_{i=1}^{k} \langle v_i, v_i \rangle \\
&= \sum_{i=1}^{k} \|v_i\|^2.
\end{aligned}
$$

12.1.4 Projections (Onto 1D Subspaces)

[insert diagram here]

Our goal is to write v as some kind of multiple of w added with another vector orthogonal to w.

So, we observe that we have:

$$v - \alpha w \perp w.$$

So,

$$\langle v - \alpha w, w \rangle = 0$$
$$\langle v, w \rangle - \alpha \langle w, w \rangle = 0$$
$$\alpha = \frac{\langle v, w \rangle}{\langle w, w \rangle}.$$

12.2   Lecture - 11/9/2023

12.2.1   Some Corollaries

Recall from where we left off last time. Then, some of the consequences of the observation we made last time include:

Theorem 12.11 (Cauchy-Schwarz Inequality)**.** Suppose we have two vectors u, v. Then, we have:

$$|\langle u, v \rangle| \leq \|u\| \cdot \|v\|.$$

Proof. First, in the case where $v = 0$, then it is trivial to see that the inequality holds (and in fact, we have equality of both sides).

In the case where $v \neq 0$, we see the following:

$$v = \frac{\langle u, v \rangle}{\langle u, u \rangle} u + w \qquad\qquad (u \perp w)$$

$$\|v\|^2 = \left\| \frac{\langle u, v \rangle}{\langle u, u \rangle} u \right\|^2 + \|w\|^2 \qquad\qquad \text{(By Pythagorean Theorem)}$$

$$\|v\|^2 \geq \left\| \frac{\langle u, v \rangle}{\langle u, u \rangle} u \right\|^2$$

$$\|v\|^2 \geq |\langle u, v \rangle|^2 \cdot \left(\frac{1}{\|u\|^4} \right) \|u\|^2$$

$$\|v\|^2 \geq |\langle u, v \rangle|^2 \cdot \frac{1}{\|u\|^2}$$

$$\|v\|^2 \cdot \|u\|^2 \geq |\langle u, v \rangle|^2$$

$$\|u\| \cdot \|v\| \geq |\langle u, v \rangle|$$

■

Corollary 12.12 (Triangle Inequality)**.**

$$\|v + w\| \leq \|v\| + \|w\|$$

Proof. We will begin by doing the following:

$$\begin{aligned}
\|v + w\|^2 &= \langle v + w, v + w \rangle \\
&= \langle v, v \rangle + \langle v, w \rangle + \langle w, v \rangle + \langle w, w \rangle \\
&= \|v\|^2 + 2\text{Re}\,\langle v, w \rangle + \|w\|^2 \\
&\leq \|v\|^2 + 2|\langle v, w \rangle| + \|w\|^2 \\
&\leq \|v\|^2 + 2\,\|v\|\,\|w\| + \|w\|^2 \\
&= (\|v\| + \|w\|)^2 \\
\|v + w\| &\leq \|v\| + \|w\|
\end{aligned}$$

■

Corollary 12.13 (Reverse Triangle Inequality)**.**

$$|\|v\| - \|w\|| \leq \|v - w\|$$

Exploration. We note that $(v - w) + w = v$. Then, using the Triangle Inequality, we have:

$$\begin{aligned}
\|v - w + w\| &\leq \|v - w\| + \|w\| \\
\|v\| - \|w\| &\leq \|v - w\| \\
\|w\| - \|v\| &\leq \|w - v\| \\
&= \|v - w\| \\
|\|v\| - \|w\|| &\leq \|v - w\|
\end{aligned}$$

■

Corollary 12.14 (Parallelogram)**.**

$$\|v + w\|^2 + \|v - w\|^2 = 2\left(\|v\|^2 + \|w\|^2\right)$$

12.2.2 Talks on Orthogonality

Theorem 12.15. Suppose we have $v_1, \ldots, v_k$ which are linearly independent. Then, we have $u_1 = v_1, u_2, \ldots, u_k$ which can be built by repeated projections so that each vector is orthogonal to each other.

Proof. Specifically, if we have $u_1 = v_1, u_2, \ldots, u_k$ which are already mutually orthogonal, and $\text{span}\,\{v_1, \ldots, v_j\} = \text{span}\,\{u_1, \ldots, u_j\}$, then $u_{j+1} = v_{j+1} - a_1 u_1 - \ldots - a_j u_j$, where we have $a_j = \frac{\langle v_{j+1}, u_j \rangle}{\langle u_j, u_j \rangle}$.

Thus, with this construction, we can guarantee that u_{j+1} is orthogonal is all the preceding vectors.

■

Theorem 12.16. Now, conversely, if $u_1, \ldots, u_k$ are non-zero vectors and are mutually orthogonal, it follows then that they must be linearly independent.

Proof. We observe that we have the following:

$$a_1 u_1 + \ldots + a_k u_k = 0$$
$$|a_1|^2 \|u_1\|^2 + \ldots + |a_k|^2 \|u_k\| = 0$$

However, since each $u_1, \ldots, u_k$ are non-zero, it follows then that their norms are ≥ 0. Then, for the equality to hold true, we must have that $a_1, \ldots, a_k = 0$. ∎

> **Definition 12.17.** We say that $u_1, \ldots, u_k$ are orthonormal if $\langle u_i, u_j \rangle = \delta_{ij}$ (this is the Kronecker delta).

> **Remark 12.18.** We remark that this is reminiscent with the dual maps; this suggests a connection with functionals which will be explored in future lectures.

Now, we note that we can in fact extend our list $u_1, \ldots, u_k$ of orthonormal vectors into a basis for V.

12.2.3   Orthornomal Bases Properties

The following are some good properties of orthonormal bases.

Say we have an orthonormal basis $e_1, \ldots, e_n$ for V. Then, for arbitrary vectors $v, w \in V$, we have:

1. $v = \langle v, e_1 \rangle e_1 + \ldots + \langle v, e_n \rangle e_n$.

2. $\langle v, w \rangle = \sum_{j=1}^{n} \langle v, e_j \rangle \langle e_j, w \rangle$. And in fact, if $w = v$, then we have:

3. $\|v\|^2 = \sum_{j=1}^{n} |\langle v, e_j \rangle|^2$ (Parsival's Equality).

Now, imagine if we didn't have a list of orthonormal vectors $e_1, \ldots, e_m$ that is not necessarily a basis. Then, we have

$$\|v\|^2 \geq \sum_{j=1}^{k} |\langle v, e_j \rangle|^2$$

12.2.4   Connection with Functionals

Suppose we have $\varphi \in V'$ and $e_1, \ldots, e_n$ is an orthonormal basis. Then, for any $v, w \in V$, we have: ...

MORE INNER PRODUCTS

13.1 Lecture - 11/16/2023

13.1.1 Recap

Recall that, given a subspace $U \subseteq V$, we have an orthogonal projector P_U onto U. The remaining properties are:

1. $I - P_U = P_{U^\perp}$; in other words, it's the orthogonal projector onto $U^\perp$. From here, it follows that the following holds:

2. $\operatorname{range}(I - P_U) = U^\perp$.

3. $(I - P_U)^2 = I - P_U$

4. $\|P_U v\| \leq \|v\|$. This is a consequence of the Pythagorean Theorem.

5. If we have an orthonormal basis $e_1, \ldots, e_k$ for U, then $P_U v = \sum_{j=1}^{k} \langle v, e_j \rangle e_j$.

> **Remark 13.1** (Uniqueness of Orthogonal Projector)**.** Our orthogonal projector splits the space into a direct sum, where the complement to our subspace is the orthogonal complement. Thus, we see that by the fact of V being a direct sum, each vector v has a unique representation using u and $u^\perp$.

13.1.2 Adjoints

> **Definition 13.2** (Adjoint)**.** Given two inner product vector spaces V, W and a map $T \in \mathcal{L}(V, W)$, we can define the adjoint T^* to be
> $$\langle Tv, w \rangle := \langle v, T^* w \rangle .$$
> We observe here that $T^* \in \mathcal{L}(W, V)$. Furthermore, the first inner product lives in W, whereas the second lives in V.

Now, is such a T^* well-defined?

First, if T and w are fixed, then the map $v \mapsto \langle Tv, w \rangle$ is in fact a linear functional. Then, by Riesz's Theorem, we say that there exists a representer $u \in V$ such that $\langle Tv, w \rangle = \langle v, u \rangle$ for all $v \in V$. Then, we say that $T^*w = u$.

Since the representer is unique, then we can conclude that indeed T^* is well-defined.

Next, we will check for linearity of T^*. First, take $w_1, w_2 \in W$ and $\alpha, \beta \in \mathbb{F}$. Then, we observe the following:

$$\begin{aligned}
\langle v, T^*(\alpha w_1 + \beta w_2) \rangle &= \langle Tv, \alpha w_1 + \beta w_2 \rangle \\
&= \langle Tv, \alpha w_1 \rangle + \langle Tv, \beta w_2 \rangle \\
&= \overline{\alpha} \langle Tv, w_1 \rangle + \overline{\beta} \langle Tv, w_2 \rangle \\
&= \overline{\alpha} \langle v, T^*w_1 \rangle + \overline{\beta} \langle v, T^*w_2 \rangle \\
&= \langle v, \alpha T^*w_1 \rangle + \langle v, \beta T^*w_2 \rangle \\
&= \langle v, \alpha T^*w_1 + \beta T^*w_2 \rangle .
\end{aligned}$$

Then, the difference between $T^*(\alpha w_1 + \beta w_2)$ and $\alpha T^*(w_1) + \beta T^*(w_2)$ is orthogonal to the entire space V. This means then that the difference is equal to zero. Thus, we see that, indeed, $T^*(\alpha w_1 + \beta w_2) = \alpha T^*w_1 + \beta T^*w_2$. Thus, $T^* \in \mathcal{L}(W, V)$; i.e. it's a linear map.

Example 13.3. Let $V = \mathbb{R}^3$ and $W = \mathbb{R}^2$, with the standard inner products. Let's take $T : V \to W$ to be $T(x, y, z) = (x + y, z)$. Then, we see the following:

$$\mathcal{M}(T) = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Then, $\mathcal{M}(T^*)$ would be:

$$\begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Another way to solve this is to first suppose that we have $\langle Tv, w \rangle$, with $v = (a, b, c)$ amd $w = (\alpha, \beta)$.

Then, we see that $\langle Tv, w \rangle = \langle (a + b, c), (\alpha, \beta) \rangle = (a + b)\alpha + c\beta$.

Then, we want this to be equal to $\langle v, T^*w \rangle = \langle (a, b, c), T^*w \rangle$. Then, we want $T^*(\alpha, \beta) = (\alpha, \alpha, \beta)$.

Theorem 13.4. The following is a property of the adjoint:

$$(T + S)^* = T^* + S^*.$$

Proof. We observe the following:

$$\begin{aligned}
\langle v, (T + S)^*w \rangle &= \langle (T + S)v, w \rangle \\
&= \langle Tv + Sv, w \rangle \\
&= \langle Tv, w \rangle + \langle Sv, w \rangle \\
&= \langle v, T^*w \rangle + \langle v, S^*w \rangle \\
&= \langle v, T^*w + S^*w \rangle .
\end{aligned}$$

∎

Theorem 13.5. Similarly, another property is:

$$(\alpha T)^* = \overline{\alpha} T^*$$

Proof. We observe that

$$\langle v, (\alpha T)^* w \rangle = \langle (\alpha T) v, w \rangle$$
$$= \alpha \langle Tv, w \rangle$$
$$= \alpha \langle v, T^* w \rangle$$
$$= \langle v, \overline{\alpha} T^* w \rangle .$$

∎

Theorem 13.6. We also have
$$(T^*)^* = T$$

Theorem 13.7. The adjoint of the identity is itself. In other words,

$$I^* = I.$$

Theorem 13.8. We have the following:
$$(ST)^* = T^* S^*$$

Theorem 13.9. For the inverse, we have

$$(T^{-1})^* = (T^*)^{-1}.$$

Proof. Suppose that T is invertible. First, we see that

$$T^*(T^{-1})^* = (T^{-1}T)^* = I^* = I$$
$$(T^{-1})^* T^* = (TT^{-1})^* = I^* = I$$

∎

ADJOINTS AND STUFF

14.1   Lecture - 11/21/2023

The following are basic results about the null spaces and ranges of adjoints:

1. $\operatorname{null} T^* = (\operatorname{range} T)^\perp$.

2. $\operatorname{range} T^* = (\operatorname{null} T)^\perp$.

3. $\operatorname{null} T = (\operatorname{range} T^*)^\perp$.

4. $\operatorname{range} T = (\operatorname{null} T^*)^\perp$.

Proof.   We will first prove 1.

To begin with, we note that by definition of T^*, this means that $w \in \operatorname{null} T^*$ is one such that $\langle Tv, w \rangle = \langle v, 0 \rangle$.

But, note that $\langle v, 0 \rangle = 0$.

So, we have that $\langle Tv, w \rangle = 0$. Then, we see that any vector $w \in T^*$ must be orthogonal to all Tv. In other words, it is in the orthogonal complement of Tv, which is the same as saying $w \in (\operatorname{range} T)^\perp$.

From here, to get 3, we can simply replace T with T^*. Then, using the fact that $(T^*)^* = T$, we have then that $\operatorname{null}(T^*)^* = \operatorname{null} T = (\operatorname{range} T^*)^\perp$.

To get 4, we observe that $(\operatorname{null} T^*)^\perp = \left((\operatorname{range} T)^\perp\right)^\perp = \operatorname{range} T$.

Then, we can do the same with 4 as we did to 1 to get 2. ∎

The last thing to tackle is the following: First, suppose that $e_1, \ldots, e_n$ is an orthonormal basis of V. Furthermore, suppose $f_1, \ldots, f_n$ is an orthonormal basis for W.

Then, what are the connection between $\mathcal{T}$ and $\mathcal{T}^*$?

> **Theorem 14.1.** Given an orthonormal basis $e_1, \ldots, e_k$ and $f_1, \ldots, f_m$ for V and W respectively, then
> $$\mathcal{M}(T) = \overline{\mathcal{M}(T)^\intercal}$$

Proof. To begin with, we see that the (j,k)th entry of $\mathcal{M}(T)$ is

$$\langle Te_k, f_j \rangle = \langle e_k, T^* f_j \rangle$$
$$= \langle T^* f_j, e_k \rangle.$$

Then, we see that the last line is the (k,j)th entry of $\mathcal{M}(T^*)$.

Thus, we have that $\mathcal{T}^* = \overline{\mathcal{M}(T)^\intercal}$. ∎

> **Remark 14.2.** However, we note that this only holds for when we have an orthonormal basis!
>
> If we only had orthogonality, we will have to adjust by the norm of the vectors.

14.1.1 Self-Adjoint Operators

We say an operator $T \in \mathcal{L}(V)$ is self-adjoint if $T = T^*$.

First, suppose $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. Given an eigenvalue $\lambda \in \mathbb{F}$ of $T \in \mathcal{L}(V)$ such that $T = T^*$, there exists an eigenvector $v \in V$ corresponding to λ. Then, we observe that

$$\lambda \langle v, v \rangle = \langle \lambda v, v \rangle = \langle Tv, v \rangle = \langle v, T^* v \rangle = \langle v, Tv \rangle = \langle v, \lambda v \rangle = \overline{\lambda} \langle v, v \rangle.$$

Thus, we see that since $\langle v, v \rangle \neq 0$, then it follows that $\lambda = \overline{\lambda}$. In other words, $\lambda \in \mathbb{R}$.

In fact, we can get a stronger version of this observation about the structure of any self-adjoint operator.

> **Theorem 14.3** (Complex Spectral Theorem)**.** The following are equivalent over $\mathbb{C}$ for a self-adjoint operator:
>
> 1. $T \in \mathrm{L}(V)$ is self-adjoint.
>
> 2. $\mathcal{M}(T)$ is diagonal with real entries with respect to some orthonormal basis vectors.
>
> 3. V has an orthonormal basis consisting of eigenvectors of T.

Proof. We note that b and c are equivalent.

Now, to show b implies a, we proceed as follows: suppose that $\mathcal{M}(T)$ is diagonal with real diagonal entries. Then, we note that $\overline{\mathcal{M}(T)^\perp} = \mathcal{M}(T)$. So, $T = T^*$ as $\mathcal{M}(T)$ is written with respect to an orthonormal basis.

To show the other way, suppose that T is self-adjoint. Then, $T = T^*$. Then, we note that $\mathcal{M}(T)$ has an upper-triangular representation in $\mathbb{C}$ with respect to some orthonormal basis.

Then, we note that since $T = T^*$, we have then that $\mathcal{M}(T^*) = \overline{\mathcal{M}(T)^\intercal} = \mathcal{M}(T)$. However, we note that since this is the case, it follows then that all of the entries in the strictly-upper triangular section of $\mathcal{M}(T)$ is equal to the strictly-lower triangular entries. In other words, these entries must be zero.

Furthermore, since the conjugate is equal, it follows that the diagonal entries must be rael.

Thus, we see that a implies b as desired. ∎

To prove this result for the reals, we must first investigate whether the minimal polynomial of $T = T^*$ can contain quadratic factors $x^2 + ax + b$ where $a^2 - 4b < 0$. This is because for $\mathbb{C}$, we used the fact that in $\mathbb{C}$, we always have a triangular representation with respect to some orthonormal basis.

Now, given $v \in V$, consider $\langle (T^2 + aT + bI)v, v \rangle$.

Blast from the Past

Quick side-track: the discriminant arises from completing the square. Note that

$$x^2 + ax + b = (x + \frac{a}{2})^2 + b - \frac{a^2}{4}$$
$$= (x + \frac{a}{2})^2 - \frac{1}{4}(a^2 - 4b)$$

Anyways, we see then that:

$$(T^2 + aT + bI) = (T + \frac{a}{2}I)^2 - \frac{1}{4}(a^2 - 4b)I$$
$$\langle (T^2 + aT + bI)v, v \rangle = \left\langle (T + \frac{a}{2}I)^2 v, v \right\rangle - \frac{1}{4}(a^2 - 4b)\langle v, v \rangle$$
$$= \left\langle (T + \frac{a}{2}I)v, (T + \frac{a}{2}I)v \right\rangle - \frac{1}{4}(a^2 - 4b)\langle v, v \rangle$$

However, we note that both parts are non-zero.

However, this means then that $(T^2 + aT + bI)v \neq 0$, meaning that the operator has a trivial null space. In other words, it is injective. However, this means then that it must be invertible.

From here, it follows that because the polynomial is invertible, it cannot occur in the minimal polynomial. Thus, we see that the minimal polynomial contains only real roots!

MORE ON OPERATORS

15.1 Lecture - 11/28/2023

15.1.1 Recap

Remember that over $\mathbb{C}$, if T is normal, them $\mathcal{M}(T)$ is diagonal under some orthonormal basis. This isn't necessarily the case for $\mathbb{R}$ though.

15.1.2 Nonnegative and Positive Operators

> **Definition 15.1.** We say that S is nonnegative if and only if S is self-adjoint and $\langle Sv, v \rangle \geq 0$ for all $v \in V$.

> **Definition 15.2.** We say that S is positive if and only if S is self-adjoint and $\langle Sv, v \rangle > 0$ for all v in $V \setminus \{0\}$.

> **Remark 15.3.** To check self-adjointness, remember to check if $\langle Tv, w \rangle = \langle v, Tw \rangle$, for all $v, w \in V$.

The following are characterisations of non-negativity:

1. T is non-negative.

2. T is self-adjoint and all of its eigenvalues are nonnegative.

3. $\mathcal{M}(T)$ is diagonal with non-negative entries over some orthonormal basis.

4. T has a non-negative square root.

5. T has a self-adjoint square root.

6. $T = R^*R$.

Before starting, we let S be the square root of T if $S^2 = T$. Then, we want to show that $1 \implies 2$:

Proof. Suppose that (λ, e) is an eigenpair of T. Then, we note that $\langle Tv, v \rangle = \langle \lambda v, v \rangle = \lambda \langle v, v \rangle \geq 0$. We note that $\langle v, v \rangle > 0$, and thus we have that $\lambda \geq 0$. ∎

From here, we see that $2 \implies 3$ by the Spectral Theorem. Then, to show $3 \implies 4$, we observe:

Proof. By 3, we have $e_1, \ldots, e_n$ by an orthonomal basis such that $Te_i = \lambda_i e_i$, for all i, $\lambda_i \geq 0$. Then, we define S by $e_i \mapsto \sqrt{\lambda_i} e_i$, which is non-negative. ∎

Then, to show $4 \implies 5$, we note that any non-negative operator is also self-adjoint.

From $5 \implies 6$, we see that any self-adjoint square root S can be defined as $R = S$. Then, we note that $R^* = R$. So, we see that $R^*R = RR = R^2 = S^2 = T$ by definition.

Then, from here, to show $6 \implies 1$, we simply observe that $(R^*R)^* = R^*(R^*)^* = R^*R$. Then, we see that $\langle R^*Rv, v \rangle = \langle Rv, Rv \rangle \geq 0$.

For a positive operator, we instead have the following characterisations:

1. T is positive.

2. T is self-adjoint and all of its eigenvalues are positive.

3. $\mathcal{M}(T)$ is diagonal with positive entries over some orthonormal basis.

4. T has a positive square root.

5. T has a self-adjoint square root, with T invertible.

6. $T = R^*R$ where R is invertible.

15.1.3 SVDs

Let $T \in \mathcal{L}(V, W)$. Then,

1. T^*T is a nonnegative operator on V.

2. $\operatorname{null} T^*T = \operatorname{null} T$.

3. $\operatorname{range} T^*T = \operatorname{range} T^*$.

4. $\dim \operatorname{range} T^* = \dim \operatorname{range} T = \dim \operatorname{range} T^*T$.

Proof. To begin with, we note that the first property has been shown previously.

To show the second property, we observe that for any $v \in \operatorname{null} T$, we have that $T^*Tv = 0$ as well. Meanwhile, we note that if $v \in \operatorname{null} T^*T$, then we have that $T^*Tv = 0$. However, we note that this means that $\langle T^*Tv, v \rangle = \langle Tv, Tv \rangle = \|Tv\|^2 = 0$. Thus, we see that, indeed, $Tv = 0$ and thus $v \in \operatorname{null} T$.

For the third property, we observe that $\operatorname{range} T^*T = (\operatorname{null} T^*T)^\perp = (\operatorname{null} T)^\perp = \operatorname{range} T^*$.

... ∎

We say that the singular values of $T \in \mathcal{L}(V, W)$ are the square roots of the eigenvalues of T^*T. We list these values in non-increasing order.

> **Theorem 15.4** (SVD Simple Version)**.** Let $T \in \mathcal{L}(V, W)$ and $s_1, \ldots, s_m$ are its positive singular values. Then, there exists two orthonormal lists $e_1, \ldots, e_m \in V$ and $f_1, \ldots, f_m \in W$ such that
>
> $$Tv = s_1 \langle v, e_1 \rangle f_1 + \ldots + s_m \langle v, e_m \rangle f_m$$

Proof. Let $e_1, \ldots, e_n$ be an orthonormal basis of V such that T^*T has a diagonal representation in that basis, with diagonal entries s_j^2.

Then, we have $v = \langle v, e_1 \rangle e_1 + \ldots + \langle v, e_n \rangle e_n$. Then, for the non-zero s_j, we set $f_j = \frac{1}{s_j} T e_j$.

Then, from here, we see that by this definition, we can indeed get to our claim above.

However, we will also need to prove the orthonormality of the f_j's. To do this, we note that we have:

$$
\begin{aligned}
\langle f_j, f_k \rangle &= \left\langle \frac{1}{s_j} T e_j, \frac{1}{s_k} T e_k \right\rangle \\
&= \frac{1}{s_j s_k} \langle T e_j, T e_k \rangle \\
&= \frac{1}{s_j s_k} \langle T^*T e_j, e_k \rangle \\
&= \frac{s_j}{s_k} \langle s_j^2 e_j, e_k \rangle \\
&= \langle s_j, s_k \rangle \langle e_j, e_k \rangle .
\end{aligned}
$$

From here, we note that since e_j, e_k are orthogonal to each other for $j \neq k$, it follows then that the inner product must be equal to zero. In other words, we have that each f_j, f_k are orthogonal to each other as well for $j \neq k$. If they're equal to each other, we have that they're equal to 1. Thus, we see that it is indeed orthonormal as desired. ∎

Example 15.5. Let $V = \mathscr{P}_2(\mathbb{R})$ and $W = \mathscr{P}_3(\mathbb{R})$, with $p(x) \mapsto xp(x) + p'(x)$. Define the inner product to be

$$
\langle f, g \rangle = \int_{-1}^{1} fg \, dx
$$

First, we will find its singular values. To do this, we first must consider what T^* does.

$$
\langle Tv, w \rangle = \langle xp(x) + p'(x), g(x) \rangle = \int_{-1}^{1} xp(x) + p'(x) dx = \int_{-1}^{1} xp(x) dx + \int_{-1}^{1} p'(x) dx
$$

Then, from here, we use integration by parts...

15.2   Lecture - 11/30/2023

Recall that for SVD, we have:
$$
Tv = s_1 \langle v, e_1 \rangle f_1 + \ldots + s_m \langle v, e_m \rangle f_m,
$$
where $e_1, \ldots, e_m$ and $f_1, \ldots, f_m$ are orthonormal in V and W respectively. Furthermore, we let $s_1 \geq s_2 \geq \cdots \geq s_m > 0$ be the singular values.

Then, we have the following:

1. T is injective $\iff$ 0 is not a singular value.

2. $\dim \operatorname{range} T$ is the number of positive singular values of T.

3. T is surjective $\iff$ $\dim W$ is equal to the number of positive singular values of T.

For b, we first note that the range of T must be contained in the span of $f_1, \ldots, f_m$. So, ...

For c, we note that if T is surjective, then $\dim \operatorname{range} T = \dim W$, and from the previous part, we know that $\dim W$ is the number of positive singular values of T.

To find the adjoint of T, we look at the definition of the adjoint. That is,

$$\langle Tv, w \rangle = \langle v, T^*w \rangle.$$

Then, we observe the following:

$$
\begin{aligned}
\langle s_1 \langle v, e_1 \rangle f_1 + \ldots + s_m \langle v, e_m \rangle f_m, w \rangle &= s_1 \langle v, e_1 \rangle \langle f_1, w \rangle + \ldots + s_m \langle v, e_m \rangle \langle f_m, w \rangle \\
&= \left\langle v, s_1 \overline{\langle f_1, w \rangle} e_1 \right\rangle + \ldots + \left\langle v, s_m \overline{\langle f_m, w \rangle} e_m \right\rangle \\
&= \langle v, s_1 \langle w, f_1 \rangle e_1 + \ldots + s_m \langle w, f_m \rangle e_m \rangle
\end{aligned}
$$

15.2.1 Isometry

$T \in \mathcal{L}(V, W)$ is called an isometry if $\|Tv\| = \|v\|$, for all $v \in V$.

In case $T \in \mathcal{L}(V)$ is an invertible isometry, we call it "unitary."

> **Remark 15.6.** If $T \in \mathcal{L}(V)$ is unitary, then any eigenpair (λ, v) satisfies:
>
> $$\|Tv\| = \|\lambda v\| = |\lambda| \, \|v\| = \|v\|.$$
>
> Thus, we have that $|\lambda| = 1$.

Note that if T is an isometry, then
$$\|Te_j\| = \|e_j\| = 1.$$

Furthermore, we have $\|Te_j\| = \|s_j f_j\| = s_j$. THus, the singular value s_j must be one.

We can also note that since $\operatorname{null} T^*T = \operatorname{null} T$, then if $v \in \operatorname{null} T$, then we have $Tv = 0$. Thus, $\|Tv\| = 0$. Then, $\|v\| = 0$. So, $v = 0$.

Thus, if we have an isometry, we must have that T is injective. However, it doesn't have to necessarily be surjective.

15.2.2 Jordan Normal Form

To begin with, we have the following setting: V is finite-dimensional inner product space over $\mathbb{C}$.

Our goal is that for $T \in \mathcal{L}(V)$, we get a matrix representation which is as "simple as possible."

In this case, we want to get a matrix which has as many zeroes as possible (i.e. as sparse as possible).

Our first stop is to decompose V into a direct sum $V = U \oplus W$, where U, W are T-invariant and T is nilpotent on U (meaning that $T^n\big|_U = 0$, for some $n \in \mathbb{N}$) and invertible on W.

To get this, we notice that $\{0\} \subseteq \operatorname{null} T \subseteq T^2 \cdots$. Now, since $\dim V$ is not infinite, this chain stabilises. In other words, there exists some $n \in \mathbb{N}$ where $\operatorname{null} T^n = \operatorname{null} T^{n_1}$.