Contents

Fri. Feb 10 2023

1 Vector Spaces

Suppose that V is a finite dimensional vector space over F, with $\dim(V) = n$.

V may have *many different* bases, we know that they all have the same size n.

Say $\mathcal{B} = \{\alpha_1, ..., \alpha_n\}$ is a basis fix the ordering of $\mathcal{B}$.

Fix the ordering of $\mathcal{B}$.

> **Theorem**.
>
> For any $\alpha \in V$, there is a unique n tuple $(x_1, ..., x_n) \in F^n$ such that
>
> $$\alpha = x_1\alpha_1 + \cdots + x_n\alpha_n$$

Existence is immediate, since $\mathcal{B}$ is a basis, thus $\mathcal{B}$ spans V.

Uniqueness

Say $\alpha = x_1\alpha_1 + \cdots + x_n\alpha_n$ and $\alpha = y_1\alpha_1 + \cdots + y_n\alpha_n$.

Then we have that

$x_1\alpha_1 + \cdots + x_n\alpha_n - y_1\alpha_1 + \cdots + y_n\alpha_n = 0$, so $(x_1 - y_1)\alpha_1 + \cdots + (x_n - y_n)\alpha_n = 0$

But since $\{\alpha_1, ..., \alpha_n\}$ is linearly independent, all coefficients must be 0.

What this means is that, for a vector space V, there is an associated mapping in F^n. Notice that we know nothing about the vectors α_i.

We define $[\alpha]_{\mathcal{B}}$ to be the *coordinates* of α with respect to $\mathcal{B}$.

Check: The mapping $\alpha \mapsto [\alpha]_{\mathcal{B}} \in F^n$ satisfies

1. One to one-ness
2. Onto-ness
3. "Additive", for any $\alpha, \beta \in V$, if $\alpha = x_1\alpha_1 + \cdots + x_n\alpha_n$ and $\beta = y_1\alpha_1 + \cdots + y_n\alpha_n$. Then

$$[\alpha + \beta]_{\mathcal{B}} = \begin{bmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = [\alpha]_{\mathcal{B}} + [\beta]_{\mathcal{B}}$$

4. $[c\alpha]_{\mathcal{B}} + c[\alpha]_{\mathcal{B}}$

There exists an *isomorphism* between V and F^n.

> **Example**.
>
> Let $\mathbb{P}$ be the space of al polynomials. Let $f(x) = x^3$, and $g(x) = x^5$. Then, let
>
> $$V = \text{Span}\{f, g\} = \{\text{all } ax^3 + bx^5 : a, b \in F\}$$

then, $\dim(V) = 2$, since f and g are linearly independent.

Typical $h(x) \in V$, say $h(x) = 10x^3 - 2x^5$.

$$[h]_{\mathcal{B}} = \begin{bmatrix} 10 \\ -2 \end{bmatrix}$$

Author Note.

$[h]_{\mathcal{B}}$ is the mapping of h to F^n. **TODO** is this right?

Now let $k(x) = 2x^3 + 4x^5$ and $l(x) = x^3 + 3x^5$. Since k, l are linearly independent, they form another basis of V.

$$\mathcal{B}' = \{k(x), l(x)\}$$

1.1 Change of Basis

Given $\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$, and $\mathcal{B}' = \{\alpha_1', ..., \alpha_n'\}$ bases for V.

We want to describe the map going from $[\alpha]_{\mathcal{B}} \mapsto [\alpha]_{\mathcal{B}'}$.

Author Note.

We want to find The $\mathcal{B}$ coordinate of $\alpha \mapsto$ the $\mathcal{B}'$ coordinate of α

Step 1.

Compute the $\mathcal{B}$ coordinate of $\alpha_1', ..., \alpha_n'$, *old* coordinates of the *new* basis elements.

Step 2.

For an $n \times m$ matrix

$$P = \Big[[\alpha_1']_{\mathcal{B}}, \ldots, [\alpha_n']_{\mathcal{B}}\Big]$$

Check: for any $\alpha \in V$

$$[\alpha]_{\mathcal{B}} = P[\alpha]_{\mathcal{B}'}$$

Ans: This is what we actually want

$$[\alpha]_{\mathcal{B}'} = P^{-1}[\alpha]_{\mathcal{B}}$$

<u>Mon. Feb 13 2023</u>

TODO Missing *some* info

Want: Describe the mapping $T : F^n \to F^n$

$$T([\alpha]_{\mathcal{B}_{\text{old}}}) = [\alpha]_{\mathcal{B}'_{\text{new}}}$$

Author Note.

If we switch the basis for some reason, we want to see what the new coordinates are.

To do this: For each α'_j, compute $[\alpha'_j]_{\mathcal{B}_{\text{old}}}$. Let

$$P = \begin{bmatrix} [\alpha'_1]_{\mathcal{B}_{\text{old}}} \cdots [\alpha'_n]_{\mathcal{B}_{\text{old}}} \end{bmatrix}$$

be an $n \times n$ matrix.

Claim: For any $\alpha \in V$

$$P \cdot [\alpha]_{\mathcal{B}'_{\text{new}}} = [\alpha]_{\mathcal{B}_{\text{old}}}$$

How?

$$P \cdot [\alpha'_1]_{\mathcal{B}'_{\text{new}}} = P \cdot \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = [\alpha'_1]_{\mathcal{B}_{\text{old}}}$$

This is the 1^{st} column of P, and similarly for all columns.

Thus: For any $\alpha \in V$,

$$[\alpha]_{\text{new}} = P^{-1} \cdot [\alpha]_{\text{old}}$$

Example.

In practice, we have the following.

$V = \text{Span}(\{x^3, x^5\})$ subspace of $\mathbb{P}$, the set of all polynomials. Let $f(x) = x^3, g(x) = x^5, \mathcal{B} = \{x^3, x^5\}$. Let $h(x) = 10x^3 - 2x^5 \in V$.

Question: What are the coordinates of h with respect to $\mathcal{B}$?

Answer:

$$[h]_{\mathcal{B}} = \begin{bmatrix} 10 \\ -2 \end{bmatrix}$$

Let's now see what happens when we create a new basis $\mathcal{B}'$.

Example.

Let $k(x) = 2x^3 + 5x^5$, $l(x) = x^3 + 3x^5$.

Let $\mathcal{B}' = \{k(x), l(x)\} = \{2x^3 + 5x^5, x^3 + 3x^5\}$ be another basis of V, still with $\mathcal{B} = \{f(x), g(x)\} = \{x^3, x^5\}$.

Question: What are the coordinates of $h(x) = 10x^3 - 2x^5$ with respect to $\mathcal{B}'$ now?

Answer:

Well we know that $[k(x)]_{\mathcal{B}} = \begin{bmatrix} 2 \\ 5 \end{bmatrix}$ and $[l(x)]_{\mathcal{B}} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$, these are just the coordinates of k, and l with respect to $\mathcal{B}$.

So now we can construct our P matrix

$$P = \Big[[k(x)]_{\mathcal{B}}, [l(x)]_{\mathcal{B}}\Big] = \begin{bmatrix} 2 & 1 \\ 5 & 3 \end{bmatrix}$$

notice that P's columns are constructed from $k(x)$ and $l(x)$, expressed in terms of our standard basis $\mathcal{B}$.

Check:

$$P^{-1} = \begin{bmatrix} 3 & -1 \\ -5 & 2 \end{bmatrix}$$

Then

$$\begin{bmatrix} 3 & -1 \\ -5 & 2 \end{bmatrix} \cdot \begin{bmatrix} 10 \\ -2 \end{bmatrix} = \begin{bmatrix} 32 \\ -54 \end{bmatrix}$$

This means:

$$h(x) = 32k(x) - 54l(x) = 10x^3 - 2x^5$$

Which is what we expect.

Example.

Let $V = \mathbb{R}^2$. Standard basis $\mathcal{B} = \{\varepsilon_1, \varepsilon_2\} = \{(1,0),(0,1)\}$

$$[(5,4)]_{\mathcal{B}} = \begin{bmatrix} 5 \\ 4 \end{bmatrix}$$

Fix angle θ, Let

$$\mathcal{B}' = \{(\cos(\theta), \sin(\theta)), (-\sin(\theta), \cos(\theta))\}$$

Question: What is $\begin{bmatrix} 5 \\ 4 \end{bmatrix}_{\mathcal{B}'_{\text{new}}}$?

Answer:

1. Form P

$$[(\cos(\theta), \sin(\theta))]_{\mathcal{B}} = \begin{bmatrix} \cos(\theta) \\ \sin(\theta) \end{bmatrix}$$

$$[(-\sin(\theta), \cos(\theta))]_{\mathcal{B}} = \begin{bmatrix} -\sin(\theta) \\ \cos(\theta) \end{bmatrix}$$

Then

$$P = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}$$

Fact:

$$P^{-1} = \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix}$$

so we have

$$\begin{aligned} [(5,4)]_{\mathcal{B}'_{\text{new}}} &= P^{-1} \begin{bmatrix} 5 \\ 4 \end{bmatrix} \\ &= \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} 5 \\ 4 \end{bmatrix} \\ &= \begin{bmatrix} 5\cos(\theta) & 4\sin(\theta) \\ -5\sin(\theta) & 4\cos(\theta) \end{bmatrix} \end{aligned}$$

2 Linear Transformations

Say V, W are both vector spaces over the same field F.

Definition.

A **Linear Transformation** $T : V \to W$ is a function satisifying two rules

1. For all $\alpha, \beta \in V$,

$$T(\alpha + \beta) = T(\alpha) + T(\beta)$$

   Note that the first $+$ is addition in V, but the second is addition in W.

2. For all $\alpha \in V$ and $c \in F$,

$$T(c\alpha) = cT(\alpha)$$

The book combines the two definitions above into one, like this,

$$T(c\alpha + \beta) = cT(\alpha) + T(\beta)$$

Let's quickly take some time to understand what V and W are here. Suppose we have a transformation $T : V \to W$, then V is the **domain** and W is the **codomain**.

Here, T is just a function, which means that it *must* use all of T, but it *does not* have to use all of W. For example, the following is a perfectly valid transformation.

Example.

Let $T : \mathbb{P}^3 \to \mathbb{P}^2$ be the transformation that takes all degree 3 polynomials to the space of degree 2 polynomials, with

$$T(f) = \mathbf{0}$$

for all $f \in \mathbb{P}^3$.

Its obvious that there are more degree 2 polynomials in the world than just the $\mathbf{0}$ polynomial. So here, we say that the $\text{Range}(T) = \{\mathbf{0}\}$, and that

$$\text{Range}(T) \subseteq W$$

but maybe we are getting ahead of ourselves.

2.1 Basic Facts

Suppose that $T : V \to W$ is a linear transformation

1. $T(0) = 0$

   Proof:

   $$T(0 + 0) = T(0) + T(0) = \mathbf{0} + \mathbf{0} = \mathbf{0}$$

   Note: 0 lives in the field, and $\mathbf{0}$ lives in W, the **codomain** of the transformation T.

   Author Note.

   Always be aware of where the 0 lives

2. For all $\{\alpha_1, ..., \alpha_n\} \subseteq V$, all $\{c_1, ..., c_n\} \in F$,

   $$c_1 T(\alpha_1) + \cdots + c_n T(\alpha_n)$$

   Proof Easy induction on n, just follows from part (2) of the definition.

2.2 Examples

Let's look at multiple examples of linear transformations to get an idea of how they behave.

Example.

We already know that each matrix A has an associated linear transformation T_A. Let's look at this in more detail now.

Let $A \in F^{m \times n}$ be an $m \times n$ matrix with entries from a field F.

Then, let $T_A : F^n \to F^m$ be defined by

$$T_A(\mathbf{x}) = A\mathbf{x}$$

where $\mathbf{x}$ is a vector in F^n.

Let's check that this is indeed a linear transformation.

Chose any $\mathbf{x}, \mathbf{y} \in F^n$, then

1. **Author Note**.

   Check that $T_A(\mathbf{x}+\mathbf{y}) = T_A(\mathbf{x}) + T_A(\mathbf{y})$

   Let $\mathbf{x}, \mathbf{y} \in V$, then

   $$T_A(\mathbf{x}+\mathbf{y}) = A(\mathbf{x}+\mathbf{y}) = A\mathbf{x} + A\mathbf{y} = T_A(\mathbf{x}) + T_A(\mathbf{y})$$

   so this works as we expect.

2. **Author Note**.

   Check that $T_A(c\mathbf{x}) = cT_A(\mathbf{x})$ for $c \in F$.

   let $c \in F$, then we have

   $$T_A(cX) = A \cdot (cX) = cAX = cT_A(X)$$

   which is also what we expect.

so we have proved that T_A is a linear transformation!

Example.

Consider $\mathbb{P}$ the set of all polynomials $a_0 + a_1 x + \cdots + a_n x^n$.

Let's define $D : \mathbb{P} \to \mathbb{P}$ which takes a function $f \in \mathbb{P}$ to $f' \in \mathbb{P}$, where f' is the *derivative* of f.

$$D(f) = f'$$

Claim:

D is a linear transformation.

Proof:

Take two functions $f, g \in \mathbb{P}$, then by definition of D, we have

$$D(f+g) = (f+g)' = f' + g' = D(f) + D(g)$$

and for $c \in F$,

$$D(cf) = (cf)' = c \cdot f' = cD(f)$$

so the derivative is a linear transformation!

Example.

Let $C(\mathbb{R})$ be the set of all continuous functions $f : \mathbb{R} \to \mathbb{R}$.

Let's define $I : C(\mathbb{R}) \to C(\mathbb{R})$ which takes a function $f \in C(\mathbb{R})$ to $F \in C(\mathbb{R})$, where F is the *antiderivative* of f.

$$I(f) = \int_0^x f(t)dt$$

Author Note.

Note that the integral exists because you can always integrate a continuous function.

The result is also continuous and differentiable by the Fundamental Theorem of Calculus.

$$D(I(f)) = f$$

Is the **Fundamental Theorem of Calculus**.

Therefore $I(f)$ really *is* continuous, $I(f) \in C(\mathbb{R})$.

Claim:

I is a linear transformation.

Proof:

Take two functions $f, g \in \mathbb{P}$, then by definition of I, we have

$$\begin{aligned} I(f+g) =& \int_0^x (f(t)+g(t))dt \\ =& \int_0^x f(t)dt + \int_0^x g(t)dt \\ =& I(f) + I(g) \end{aligned}$$

and

$$I(cf) = \int_0^x cf(t)dt = c\int_0^x f(t)dt = cI(f)$$

so the integral is a linear transformation!

Fri. Feb 15 2023

Recall: A linear transformation $T : V \to W$ is a function between two vector spaces over the same field F, satisifying

1. For all $\alpha, \beta \in V$,

$$T(\alpha + \beta) = T(\alpha) + T(\beta)$$

   Note that the first $+$ is addition in V, but the second is addition in W.

2. For all $\alpha \in V$ and $c \in F$,

$$T(c\alpha) = cT(\alpha)$$

For all $\alpha_1, ..., \alpha_k \in V$, and $c_1, ..., c_k \in F$, it breaks nicely into

$$T(c_1\alpha_1 + \cdots + c_k\alpha_k) = c_1T(\alpha_1) + \cdots + c_kT(\alpha_k)$$

Example.

$I^* : C(\mathbb{R}) \to \mathbb{R}$ (all continuous functions from $\mathbb{R}$ to $\mathbb{R}$)

$$I^*(f) = \int_0^1 f(x)dx$$

$$I^*(x^2) = \int_0^1 x^2dx = \frac{x^3}{x}\Big|_0^1 = \frac{1}{3}$$

Note that the output of $I*$ is just a number here. Additionally, I^* is linear: you can split integrals up for polynomials, and you can take constants outside.

For any V, W, we also have

$$X : V \to W$$

Is the zero transformation. It takes any $\alpha \in V$ to the 0 of W. We'll use this to prove theorems about linear transformations later.

Theorem.

Let's prove existence and uniqueness of linear transformations.

1. Linear Transformations $T : V \to W$ are **determined** by their behavior on a basis $\mathcal{B}$ of V. More precisely,

   Suppose that $\mathcal{B} = \{\alpha_1, ..., \alpha_n\}$ is a basis for V and suppose that $T, U : V \to W$ are both linear transformations (and they agree on a basis), such that

$$T(\alpha_1) = U(\alpha_1), T(\alpha_2) = U(\alpha_2), ..., T(\alpha_n) = U(\alpha_n)$$

   Then $T = U$

2. For **any map** $T_0 : \mathcal{B} \to W$, there s a unique linear transformation $T : V \to W$ with $T \supseteq T_0$. In other words,

   Let $\mathcal{B} = \{\alpha_1, ..., \alpha_n\}$ be **any basis** for V and let $\beta_1, ..., \beta_n$ be **any vectors** in W.

   Then, there is a **unique** linear transformation $T : V \to W$ such that

$$T(\alpha_1) = \beta_1, T(\alpha_2) = \beta_2, ..., T(\alpha_n) = \beta_n$$

1. **Uniqueness**: Chose any $\alpha \in V$, since $\mathcal{B}$ is a basis,

   Author Note.

   Will show that $T = U \Leftrightarrow$ For any $\alpha \in V$, $T(\alpha) = U(\alpha)$

$$\alpha = c_1\alpha_1 + \cdots + c_n\alpha_n$$

   for some **unique** $c_1, ..., c_n \in F$.

   Since T is a linear transformation,

$$T(\alpha) = c_1T(\alpha_1) + \cdots + c_nT(\alpha_n)$$

   Likewise with U,

$$U(\alpha) = c_1U(\alpha_1) + \cdots + c_nU(\alpha_n)$$

   But, since $T(\alpha_1) = U(\alpha_1), ..., T(\alpha_n) = U(\alpha_n)$, $T(\alpha) = U(\alpha)$.

   Author Note.

   Essentially, if T, U work the same for all α_i, then their sum will obviously be the same, and so they'll give the same result for the same α.

   Note that this theorem *still* works for infinite dimensional vector spaces.

2. **Existence**: Chose any $\alpha \in V$. **Author Note**.

   We must define $T(\alpha)$

   Since $\mathcal{B}$ is a basis, we can write

$$\alpha = c_1\alpha_1 + \cdots + c_n\alpha_n$$

   which is unique.

   Define

$$T(\alpha) := c_1\beta_1 + \cdots + c_n\beta_n \in W$$

   Check: Is T linear?

Say $\gamma = d_1\alpha_1 + \cdots + d_n\alpha_n$, $\delta = e_1\alpha_1 + \cdots + e_n\alpha_n$.

In V, we have that $\gamma + \delta = (d_1 + e_1)\alpha_1 + \cdots + (d_n + e_n)\alpha_n$.

By our definition of T, we have

$$\begin{aligned} T(\gamma + \delta) =&(d_1 + e_1)\beta_1 + \cdots + (d_n + e_n)\beta_n \\ =&(d_1\beta_1 + \cdots + d_n\beta_n) + (e_1\beta_1 + \cdots + e_n\beta_n) \\ =&T(\gamma) + T(\delta) \end{aligned}$$

Check: $T(c\gamma) = cT(\delta)$

So such a tranformation T exists. Additionally by part (1), it is unique.

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

Definition.

Range$(T) = \{T(\alpha) : \alpha \in V\} \subseteq W$ is the set of all vectors in W hit by T.

Fact: Range(T) is a **subspace** of W.

1. 0 is in it. This is because $T(0) = 0$, obviously.
2. **Combinations of α_i are in it**

   Say that $\beta_1, \beta_2 \in$ Range(T). **Author Note**.

   must show that $\beta_1 + \beta_2 \in$ Range(T)

   Since $\beta_1 \in$ Range(T), there is some $\alpha_1 \in V$ such that

   $$T(\alpha_1) = \beta_1$$

   similarly for β_2. Now $T(\alpha_1 + \alpha_2) = T(\alpha_1) + T(\alpha_2) = \beta_1 + \beta_2$, since T is linear. So $T(\alpha_1 + \alpha_2) = \beta_1 + \beta_2$ so $\beta_1 + \beta_2 \in$ Range(T) **Author Note**.

   since $\alpha_1, \alpha_2 \in V$ means that $\alpha_1 + \alpha_2 \in V$, because it's a vector space!
3. **Scaling Works:** Say $\beta \in$ Range(T), and $c \in F$. Chose $\alpha \in V$ such that $T(\alpha) = \beta$. Then $T(c\alpha) = cT(\beta)c\beta$, therefore $c\alpha \in$ Range(T).

In other books this space is also called the **image** of T.

Definition.

The **Null Space** of $T : V \to W$ is the set

$$\text{Null}(T) = \{\alpha \in V | T(\alpha) = \mathbf{0}\} \subseteq V$$

Author Note.

In other words, this is the set of all vectors α in V that, after a transformation T is applied, go to $\mathbf{0}$. Note that $\mathbf{0}$ here is the zero of the vector space $W \subseteq V$.

This is also sometimes called the **Kernel** of T.

Theorem.

Let $T : V \to W$ be a linear transformation. Null(T) is a subspace of V.

Let $\alpha, \beta \in \text{Null}(T)$ and $c \in F$. Then,

$$T(c\alpha + \beta) = cT(\alpha) + T(\beta) = c\mathbf{0} + \mathbf{0} = \mathbf{0} \Rightarrow c\alpha + \beta \in \text{Null}(T)$$

It's pretty easy to see from this (and it should make sense) that the Null Space for a transformation T is itself a vector space.

Definition.

The **Nullity** of T is the dimension of the Null space of T.

Definition.

The **Rank** of T is the dimension of Range(T). Is this is equal to the dimension of W, T is said to have **full rank**.

Note again that this comes back to our definition of W for our transformation T. Earlier, we saw that W was the *codomain* of T. If you think about how functions behave, this is like having a *surjective* function.

Example.

Let $\mathbb{P}_2$ be the set of all polynomials of degree 2 or less over a field F. Then, we have $\dim(\mathbb{P}_2) = 3$.

Consider the linear transformation $D : \mathbb{P}_2 \to \mathbb{P}_2$, the differentiation operator. Then

$$\text{Range}(D) = \text{Span}(\{D(1), D(x), D(x^2)\}) = \text{Span}(\{1, 2x\}) \Rightarrow \text{rank}(D) = 2$$

In other words, the Range of D is the Span of a basis of $\mathbb{P}_2$ (in this case $\{1, x, x^2\}$) after being evaluated through D, so $\{1, 2x\}$. So the rank of D here is 2.

For the Null Space of D, we have that

$$\text{Null}(D) = \{c \in F\} \Rightarrow \text{nullity}(D) = 1$$

The Null Space is the set of all constant functions since those are the function that, on D, go to $\mathbf{0}$.

2.3 The Rank-Nullity Theorem

Theorem.

Let V be a vector space with $\dim V = n$. Let $T : V \to W$.

$$\text{rank}(T) + \text{nullity}(T) = \dim V = n$$

Proof.

First, choose $\{\alpha_1, \alpha_2, \cdots, \alpha_k\}$ to be a basis for Null(T). This set is necessarily linearly independent in V. So, we can choose an additional $\{\alpha_{k+1}, \cdots, \alpha_n\}$ so that $\{\alpha_1, \alpha_2, \cdots, \alpha_n\}$ is a basis of V.

Certainly, $k \leq n$, since Null(T) is a subspace of V.

We claim $A = \{T(\alpha_{k+1}), \cdots, T(\alpha_n)\}$ is a basis for Range(T). From this we have our theorem.

Clearly, $A \subseteq$ Range(T). We also have, that since $\{\alpha_1, \alpha_2, \cdots, \alpha_n\}$ is a basis of V, $\{T(\alpha_i)\}$ spans Range(T).

However, $T(\alpha_1) = T(\alpha_2) = \cdots = T(\alpha_k) = \mathbf{0}$, since they are in the null space, and hence do not contribute to the span. Thus, A spans Range(V). Now we need only show A is linearly independent. We choose constants such that

$$c_{k+1}T(\alpha_{k+1}) + \cdots + c_nT(\alpha_n) = \mathbf{0}$$

Let

$$\alpha^* = c_{k+1}\alpha_{k+1} + \cdots + c_n\alpha_n \in V$$

We then have

$$T(\alpha^*) = c_{k+1}T(\alpha_{k+1}) + \cdots + c_nT(\alpha_n) = \mathbf{0} \Rightarrow \alpha^* \in \text{Null}(T)$$

So, we then have that, since α^* is in the null space,

$$\alpha^* = d_1\alpha_1 + d_2\alpha_2 + \cdots + d_k\alpha_k = c_{k+1}\alpha_{k+1} + \cdots + c_n\alpha_n$$

$$d_1\alpha_1 + d_2\alpha_2 + \cdots + d_k\alpha_k - c_{k+1}\alpha_{k+1} - \cdots - c_n\alpha_n = \mathbf{0} \in V$$

But since $\{\alpha_1, \alpha_2, \cdots, \alpha_n\}$ is a basis of V, all the constants are zero, and in particular all of the c_i are zero. So, $\{T(\alpha_{k+1}), \cdots, T(\alpha_n)\}$ is linearly independent and is thus a basis of Range(T).

■

Now that we have the rank-nullity theorem, we can analyze transformations and their matrices.

Definition.

Let A be a matrix in $F^{m\times n}$.

The **Column Space** is the vector space spanned by the n columns of A. This is precisely Range(T_A).

The **Row Space** is the vector space spanned by the m rows of A.

Theorem.

Let A be a matrix, that when row-reduced has n unknowns and r non-zero rows. nullity$(T_A) = n - r$

This follows from the fact that elementary row operations preserve the row space, and that solving a linear system in r equations with n unknowns will have $n - r$ degrees of freedom.

TODO I guess I can believe this but some more info would be nice.

Note.

Let A be a matrix. Then the following are equal

- The dimension of the row space of A
- The dimension of the column space of A
- The number of nonzero rows in the row-reduced form of A
- rank(T_A)

This follows immediately from the above and the Rank-Nullity Theorem.

Mon. Feb 20 2023

Suppose that A is an $m \times n$ matrix. Now suppose that we row reduce A, let's call this matrix A^{rr}. Then we have that

$$\text{RowSpace}(A) = \text{RowSpace}(A^{rr})$$

And we know that $\text{rank}(A)$ is the number of non-zero rows of A^{rr} which we call r.

Moreover, the solution set of the homogeneous system $A\mathbf{x} = \mathbf{0}$ has dimension $n - r$, where n is the number of columns subtract the number of redundant equations.

Now, we know that for a matrix A, there is an associated linear transformation $T_A : F^n \to F^m$.

Last time, we also saw that

1. $\text{Range}(T_A) = \text{ColSpace}(A)$,
2. $\text{Null}(T_A)$ is the solution set of $A\mathbf{x} = \mathbf{0}$.

Now we can put everything together. Recall the Rank-Nullity theorem, then we have that, for any linear transformation T_A,

1. $\text{rank}(T_A) + \text{nullity}(T_A) = \dim(F^n) = n$
2. $\text{rank}(T_A) := \dim(\text{Range}(T_A))$
3. $\text{nullity}(T_A) = \dim(\text{Null}(A)) = n - r$, which is exactly the dimension of the set of all solutions to the homogeneous.
4. Finally we have that

$$\begin{aligned}\text{rank}(A) &= \dim(\text{RowSpace}(A)) = \dim(\text{ColSpace}(A)) \\ &= \dim(\text{RowSpace}(A^{rr})) \\ &= \text{rank}(T_A) \\ &= r\end{aligned}$$

Recall also that $\text{nullity}(T_A) = \dim(\text{Null}(T_A)) = n - r$.

Consider a matrix A where

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 6 & 8 \\ 1 & 0 & 1 & 1 \end{bmatrix}$$

Then

$$A^{rr} = \begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 3/2 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Is the row reduced matrix.

A basis for the row space of A is

$$\{(1, 0, 1, 1), (0, 1, 1, 1/3)\}$$

but another is

$$\{(1, 2, 3, 4), (1, 0, 1, 1)\}$$

We have $T_A : \mathbb{R}^4 \to \mathbb{R}^3$, and $\text{rank}(T_A) = 2$.

Basis for Range(T_A) equals the basis for Col Space(A)

There are many more linear transformations than the ones given by a matrix, for instance the derivative or integrals.

Let $T : V \to W$ be a linear transformation. From here there are two questions we can now ask.

1. Is T onto?

   It is if and only if Range(T) $= W$. We saw this earlier. In terms of dimension, this means that $\text{rank}(T) = \dim(W)$.

   Note that here, V, W must be **finite dimensional**.

2. Is T one to one?

   This requires some more work.

> **Theorem**.
>
> $T : V \to W$ is one to one if and only if $\text{Null}(T) = \{\mathbf{0}\}$.
>
> **Author Note**.
>
> In other words, the Null space must only contain the zero vector.

Assume that T is one to one. We know that $T(\mathbf{0}_V) = \mathbf{0}_W$. Chose any $\alpha \in \text{Null}(T)$, then $T(\alpha) = \mathbf{0}_W$, by definition of being in the Null Space. Since T is one to one, α must equal $\mathbf{0}_V$.

Now assume that Null(T) is just $\mathbf{0}_W$. To see that T is one to one, chose any $\alpha, \alpha' \in V$, with $T(\alpha) = T(\alpha')$. Then $T(\alpha - \alpha') = T(\alpha) - T(\alpha')$ by linearity, but then since $\alpha = \alpha'$, $T(\alpha - \alpha') = \mathbf{0}$ so $T(\alpha - \alpha')$ must be in the Null space of T, and since $\text{Null}(T) = \{\mathbf{0}\}$, and $\alpha - \alpha' = 0$, so $\alpha = \alpha'$ and thus T is one to one.

> **Definition**.
>
> T is called **non-singular** if T is one to one.
>
> This is just another term for something we already know.

> **Theorem**.
>
> Now suppose that $T : V \to W$ is a linear transformation with $\dim(V) = \dim(W)$. Then T is one to one if and only if T is onto.

By the Rank-Nullity theorem from last time, we have that

$$\text{rank}(T) + \text{nullity}(T) = \dim(V)$$

Now, assume that T is one to one, then $\text{nullity}(T) = 0$, but then $\text{rank}(T) = \dim(V) = \dim(W)$.

Now conversely, assume that T is onto. Then

$$\text{rank}(T) = \dim(W) = \dim(V)$$

Therefore $\text{nullity}(T) = 0$, and so T is one to one.

We are now starting to get a pretty good understanding of linear transformations, but suppose that we now want to combine them.

2.4 Combining Linear Transformations

Say $T : V \to W$ and $U : W \to Y$ are linear transformations over F.

Author Note.

then $U \circ T : V \to Y$ is a function.

Check the following:

1. $U \circ T$ is a linear transformation.

   Author Note.

   You know how to do this, just check that they scale and add as we expect.

2. If both T and U are one to one, then the composition is also one to one.
3. If both T and U are onto, the composition is also onto.

Note.

$T \circ U$ would **not** be a linear transformation, assuming that Y and V are not the same vector space.

Author Note.

Linear transformations don't commute nicely like that.

Let's now look at T again.

Definition.

A linear transformation $T : V \to W$ is called **invertible** if there is a linear transformation $U : W \to V$ such that

1. $U \circ T : V \to V$ is the identity from V. In other words

$$U(T(\alpha)) = \alpha$$

   For any $\alpha \in V$.

2. $T \circ U : W \to W$

$$T(U(\alpha)) = \alpha$$

   For any $\alpha \in W$.

Note.

It might be interesting for you to prove that, if one of the above applies, the other automatically applies as well.

If T is invertible, we call such a U T^{-1}, the inverse transformation of T.

Note.

Inverse transformations are unique, if they exist.

Author Note.

> We didn't talk about this in class but it *has* to be true.

Proposition: If $T : V \to W$ is an *invertible* linear transformation if and only if T is both one to one and onto.

> **Note**.
>
> If T has an inverse, then it must be the case that $\dim(V) = \dim(W)$.
>
> If this is surprising, just consider that this follows from the fact that T must be both one to one, and onto in order to have an inverse.

Wed. Feb 22 2023

Today was exam review. As such, everything for today is written as examples.

What to expect

1. Short answer **True** / **False**. Then write a sentence explaining your choice, doesn't need to be a proof.
2. Matrix stuff. Row Space, Col Space, Col Space, Rank of Matrix, Solution set of homogeneous system, etc...
3. Linear Transformations. Polynomials, $\{e^{ix}, e^{-ix}\}$ stuff, etc...
4. Full Proofs of statements. Rather straightforward, just involving linear independence, or spanning, or dimensions, etc...

The exam *is* 50 minutes, problems *will* be reasonable. At this point, we have definitely covered everything that will be on exam 1.

3 Example Problems

Example.

Problem 3.1.4

Want: A linear transformation $T : \mathbb{R}^3 \to \mathbb{R}^2$ such that $T(1, -1, 1) = (1, 0)$, and $T(1, 1, 1) = (0, 1)$

Firstly, *is there one*?

The two vectors passed to T are linearly independent, so **they can be expanded to a basis**. Say for example,

$$\mathcal{B} = \{(1, -1, 1), (1, 1, 1), (0, 0, 1)\}$$

Where $\mathcal{B}$ is a basis of $\mathbb{R}^3$. Now, there will be a *unique* linear transformation that will take it to any two points in $\mathbb{R}^2$ (even if those two points are not linearly independent.)

Fundamental Fact:

If V, W are vector spaces, and $\mathcal{B} = \{\alpha_1, \dots, \alpha_n\}$ is a basis of V, then for any $\beta_1, ..., \beta_n \subseteq W$, there is *exactly* one linear transformation T for which

$$T(\alpha_1) = \beta_1, \dots, T(\alpha_n) = \beta_n$$

Note: β_i do *not* have to be linearly independent! For instance consider the 0 transformation, then $\beta_1 = \cdots = \beta_n = \mathbf{0} \in W$.

Question: How do we define such a T?

Take any $\alpha^* = c_1\alpha_1 + \cdots + c_n\alpha_n$. Then by linearity,

$$T(\alpha^*) = c_1\beta_1 + \cdots + c_n\beta_n$$

Note: This shows that such a T *does* exist, but it's not *constructive*. We don't *actually* know what it is; we only know what properties it has, and that it exists.

NOTE: Before we expanded $\{(1,-1,1),(1,1,1)\}$ to a basis, there were an *infinite* family of linear transformations. This is because we got to choose the last vector of $\mathcal{B}$, in our case $(0,0,1)$. If we have chosen, say, $(0,1,0)$ instead, then we would have gotten an entirely different transformation T.

Example.

Problem 2.4.6a

Question: How do we show that e^{ix} and e^{-ix} are linearly independent?

Let $f_1(x) = 1, f_2(x) = e^{ix}, f_3(x) = e^{-ix}$.

Claim: f_1, f_2, f_3 are *linearly independent.*

Recall what this means. For *any* x,

$$c_1 f_1(x) + c_2 f_2(x) + c_3 f_3(x) = \mathbf{0}$$

Implies that c_1, c_2, c_3 must all be 0. Where $\mathbf{0}$ is the zero function: $z(x) = 0$ for all x.

$$c_1 + c_2 e^{ix} + c_3 e^{-ix} = \mathbf{0}$$

Now let $x = -100i$, then $e^{i(-100i)} = e^{100}$, then $e^{-i(-100i)} = e^{-100}$, so we have

$$c_1 + c_2 e^{100} + c_3 e^{-100} = \mathbf{0}$$

Therefore, c_1, c_2, c_3 *must* all be 0, since clearly all the functions are positive for $x = -100i$. This shows that the functions are linearly independent.

Since $c_1 f_1(x) + c_2 f_2(x) + c_3 f_3(x) = \mathbf{0}$ *has* to work for *any* value of x, if you just find *one* counterexample (like $x = -100i$ above), that shows that they are linearly independent.

Example.

Question: Suppose that the field F is $\mathbb{C}$. Is

$$W = \{g_1(x) = 1, g_2(x) = \sin x, g_3(x) = \cos x\}$$

Linearly independent?

Answer: Yes!

Suppose that for some $x = \theta$ we have

$$\begin{aligned} c_1 g_1(x) + c_2 g_2(x) + c_3 g_3(x) &= \mathbf{0} \\ c_1 + c_2 \sin x + c_3 \cos x &= \mathbf{0} \end{aligned}$$

Then

$$c_2 \sin x + c_3 \cos x = -c_1$$

Then look at $\theta + \pi$,

$$c_2 \sin(\theta + \pi) + c_3 \cos(\theta + \pi) = -c_1$$

So then we have the set of equations

$$c_2 \sin\theta + c_3 \cos\theta = c_2 \sin(\theta + \pi) + c_3 \cos(\theta + \pi)$$

But since $\sin\theta = -\sin(\theta + \pi)$ and $\cos\theta = -\cos(\theta + \pi)$, we have

$$\begin{aligned} c_2 \sin\theta + c_3 \cos\theta &= -c_2 \sin\theta - c_3 \cos\theta \\ 2c_2 \sin\theta + 2c_3 \cos\theta &= 0 \end{aligned}$$

But in order for this to be true for all θ, c_1 and c_2 must both be 0.

NOTE: We know that $\sin^2 x + \cos^2 x = 1$, but we're not allowed to square g_2 and g_3 here. That's not a linear operation, so it's not relevant. However, it *is* true that $\{1, \cos^2 x, \sin^2 x\}$ *are* linearly dependent.

NOTE: $f_2 \in \text{Span}(\{g_1, g_2, g_3\})$.

$$\begin{aligned} e^{-ix} &= \cos(-x) + i \sin(-x) \\ e^{-ix} &= \cos(x) = i \sin(x) \end{aligned}$$

So $f_1 \in \text{Span}(\{1, \sin x, \cos x\})$, but $\dim(W) = 3$, and we already know that f_1, f_2, f_3 are linearly independent from the previous example, so we just found another basis for W

$$\{f_1, f_2, f_3\} = \text{Span}(W)$$

Example.

Problem 2.4.4d

Let $W = \text{Span}(\{(1, 0, i), (1 + i, 1, -1)\})$.

Question: Is this set linearly independent?

Yes, neither is 0, or a multiple of the other.

Question: What is $\dim(W)$?

2

So a basis could be

$$\mathcal{B} = \{(1, 0, i), (1 + i, 1, -1)\}$$

The vectors themselves.

Question: Let $\beta_1 = (1, 1, 0)$. Is $\beta_1 \in W$?

Let's try to make it.

$$(1, 1, 0) = c_1(1, 0, i) + c_2(1 + i, 1, -1)$$

Need: $1 = c_1 + c_2(i + i)$, and $1 = c_1(0) + c_2(1)$ so $c_2 = 1$, and finally, $0 = c_1(i) + c_2(-1)$.

So $\beta_1 = (-i)(1, 0, -i) + 1(1 + i, 1, -1)$

Example.

4d on Homework 3

If $\dim(V) = \dim(W)$, and $\mathcal{B} = \{\alpha_1, \ldots, \alpha_k\}$ is a basis for V, must

$$T(\alpha_1) + \cdots + T(\alpha_k)$$

Be a basis for W?

Answer: No!

Let T be the 0 transformation, then everything is mapped to 0.

NOTE: Here it's worth reiterating the difference between the *codomain* and the *range* of T.

Suppose that $T : V \to W$ is a linear transformation that maps vectors from V to W. Then, we say that V is the *domain* of T, and that W is the *codomain* of T.

We know that a linear transformation is just a function, and one thing that we know about functions is that they *must* use their entire domain (so everything from V has to map *somewhere* in W), but they *don't* have to use all of their codomain.

The nuance here is the word *"somewhere"*. Suppose that we have the function $f : \mathbb{R} \to \mathbb{R}$ with $f(x) = 5$. Then the domain of f is $\mathbb{R}$ and so is the codomain. Here, we can see that 5 is *somewhere* in the codomain of f, but it's certainly *not all* of it. Then, we say that the *range* of f is $\{5\}$. This is all just by definition.

Let's look at another example. Suppose that $g : \mathbb{R} \to \mathbb{Z}$ with $g(x) = \lfloor x \rfloor$ this time. In this case, g *does* use up all of its codomain, since every value in $\mathbb{R}$ will be mapped to a value in $\mathbb{Z}$.

Linear transformation behave in exactly the same way. When we talk about $\dim(W)$, we're talking about the size of a basis of W, but T makes no promises about mapping elements from V to all of it.

Mon. Feb 27 2023

From this point, the course is going to become much more abstract.

Fix V, W vector spaces over the same field F.

Let $L(V, W)$ consist of all linear transformations $T : V \to W$.

Theorem.

$L(V, W)$ is a vector space.

Say that T, U are each linear transformations. Let $T + U : V \to W$ be defined by

$$(T + U)(\alpha) = T(\alpha) + U(\alpha)$$

For $T \in L(V, W)$ and $c \in F$. Let $cT : V \to W$ be the linear transformation

$$(cT)(\alpha) = cT(\alpha)$$

It's easy to check that these linear transformations satisfy the properties of being a vector space.

> **Theorem**.
>
> Suppose that $\dim(V) = n$ and $\dim(W) = m$, then $\dim(L(V, W)) = nm$.

Recall that a linear transformation $T : V \to W$ is determined by what it does to a basis of V.

Chose a basis $\mathcal{B} = \{\alpha_1, \dots, \alpha_n\}$ of V, and $\mathcal{B}' = \{\beta_1, \dots, \beta_n\}$ of W.

For any $1 \leq p \leq m$ and $1 \leq q \leq n$,

Let $E^{pq} : V \to W$ be determined by

$$E^{pq}(\alpha_i) = \begin{cases} \beta_j & \text{if } i = q \\ 0 & \text{if } i \neq q \end{cases} = \delta_{iq} \cdot \beta_j$$

Where δ_{iq} is the "Kroneker δ function" defined by

$$\delta_{iq} = \begin{cases} 1 & \text{if } i = q \\ 0 & \text{if } i \neq q \end{cases}$$

Claim: $E^{pq} : 1 \leq p \leq m, 1 \leq q \leq n$ is a basis for $L(V, W)$.

Proof: **Author Note**.

Why does $\{E^{pq}\}$ span $L(V, W)$?

For $1 \leq p \leq m, 1 \leq q \leq n$, $E^{pq}(\alpha_q) = \beta_p$ but $E^{pq}(\alpha_{q'}) = 0$ for all $q' \neq q$.

Choose any $T \in L(V, W)$, i.e. $T : V \to W$ is a linear transformation. **Author Note**.

What does this T do to V?

For $1 \leq q \leq n$, say $T(\alpha_q) = A_{1q\beta_1} + A_{2q\beta_2} + \cdots + A_{mq\beta_m}$, for some $A_{1q}, \dots, A_{mq}$

Author Note.

Here, we're building an $m \times n$ matrix!

Subclaim:

$$T = \sum_{p=1}^{m} \sum_{q=1}^{n} A_{pq} E^{pq}$$

Author Note.

The subclaim shows that T is in the span of E^{pq}

Proof of Subclaim

Let $U = \sum_{p=1}^{m}\sum_{q=1}^{n} A_{pq}E^{pq}$. **Author Note**.

We're going to show that T, and U do the same thing to every basis element. Linear transformations are equal if and only if they agree on a basis.

Fix $1 \leq q \leq n$

$$\begin{aligned} U(\alpha_q) =& \sum_{p=1}^{m}\sum_{q=1}^{n} A_{pq}E^{pq}(\alpha_q) \\ =& \sum_{p=1}^{m} A_{pq}\beta_p \\ =& A_{1q}\beta_1 + A_{2q}\beta_2 + \cdots + A_{mq}\beta_m \\ =& T(\alpha_q) \end{aligned}$$

So we see that T and U agree on every $\alpha \in \mathcal{B}$, so $T = U$.

Author Note.

Now we need to show that they are linearly independent.

Subclaim 2: E^{pq} are linearly independent.

Chose $\{c_{pq}\} \in F$ such that

$$\sum_{p=1}^{m}\sum_{q=1}^{n} c_{pq}E^{pq} = 0$$

Author Note.

Now we must show that all c_{pq} must be 0.

Fix any $\leq q \leq n$. Then

$$\begin{aligned} U(\alpha_q) =& \sum_{p=1}^{m}\Big(\sum_{q=1}^{n} c_{pq}E^{pq}(\alpha_q)\Big) \\ =& \sum_{p=1}^{m} c_{pq}\beta_p & \text{By cheatsheet} =0 \end{aligned}$$

Since U is the zero transformation. Thus

$$c_{1q}\beta_1 + c_{2q}\beta_2 + \cdots + c_{mq}\beta_m = 0$$

Since $\{\beta_1, \ldots, \beta_m\}$ is a basis for W.

This means that $c_{1q} = c_{2q} = \cdots = c_{mq} = 0$.

This holds for every $1 \le q \le n$, therefore all c_{pq} must be 0, so $\{E^{pq}\}$ is linearly independent.

Wed. Mar 1 2023

> **Definition**.
>
> Suppose that V, W are vector spaces over the same field F. An **isomorphism** is a linear transformation T that has an inverse $U : W \to V$ satisfying
>
> $$U \circ T = I_V$$
>
> and
>
> $$T \circ U = I_W$$
>
> Write T^{-1} for this U if it exists.
>
> Note that T here is *necessarily* a bijection.
>
> Note that V and W must be over the **same field** for T to be an isomorphism.

> **Definition**.
>
> The vector spaces V, W are called **isomorphic** if there exists an *isomorphism* T from V to W.
>
> Note that there may be *many different* isomorphisms.

What does this all mean? Well if V and W are isomorphic, then, even if they are very different, they will behave in very similar ways.

The following hold

- $\dim V = \dim W$
- If $V' \subseteq V$ is a subspace, then $T(V') = W' \subseteq W$ is a subspace of W with $\dim(V') = \dim(W')$. Basically if V and W are isomorphic, then V' and W' are also isomorphic.

Recall: A linear transformation $T : V \to W$ is an isomorphism if and only if T is both one to one and onto (in other words, if T is a bijection)

Specical Case: If we're lucky and $\dim(V) = \dim(W)$, then there's an easier test. If $T : V \to W$ is a linear transformation, then the following are equivalent

- T is an isomorphism
- T is one to one
- T is onto

Let's look at examples!

> **Example**.
>
> Let V be a vector space over F of dimension $\dim(V) = n$. Let $\mathcal{B} = \{\alpha_1, \dots, \alpha_n\}$ be an ordered basis for V. The coordinate transformation
>
> $$C_{\mathcal{B}} : V \to F^n$$

is an isomorphism, sending $\alpha \mapsto [\alpha]_\mathcal{B}$. In other words it translates α to the language of $\mathcal{B}$.

If $V = \mathbb{P}^2$, $\mathcal{B} = \{1, x, x^2\}$, for any $f \in \mathbb{P}^2$, $f(x) = a_0 + a_1 x + a_2 x^2$,

$$C_\mathcal{B}(f) = [f]_\mathcal{B} = \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} \in F^3$$

is an isomorphism.

Given any $\begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} \in F^3$, Let $\alpha = a_0 + a_1 x + a_2 x^2 \in \mathbb{P}^2$, then $C_\mathcal{B}(\alpha) = [\alpha]_\mathcal{B} = \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix}$ is **onto**.

Conclusion:

If V has dimension n, then

$$V \cong F^n$$

and we say that V is *isomorphic* to F^n

This is great! Now we can compare V directly to some F^n.

Corrolary

If V, W are vector spaces over the same field F and $\dim(V) = \dim(W)$, then $V \cong W$.

This is huge! As long as linear transformations have the same dimension, they behave in the same way.

Proof

There exists a $C_\mathcal{B} : V \to F^n$ and $C_{\mathcal{B}'} : W \to F^n$, so there *must* exist a linear transformation $C_{\mathcal{B}'}^{-1} \circ C_\mathcal{B}$ is a linear transformation going from $V \to W$, this is an isomorphism!

Check

- The composition of any 2 isomorphisms *is* an isomorphism
- If T is an isomorphism going from V to W, then $T^{-1} : W \to V$ is an isomorphism

Let's recall some things

For any field F and any $m, n \geq 1$. If $F^{m \times n}$ consists of all $m \times n$ matrices over F, $F^{m \times n}$ is a vector space of dimension mn

$$\dim(F^{m \times n}) = mn$$

After all, there are mn free variables in the basis.

Now, let V, W be vector spaces over F. Let $\dim(V) = nb$ and $\dim(W) = m$. Now let $L(V, W)$ be the set of all linear transformations $T : V \to W$.

We saw that $L(V, W)$ is a vector space of dimension mn, and that

$$\{E^{pq} : 1 \leq q \leq n, 1 \leq p \leq m\}$$

is a basis for $L(V, W)$. But notice: $L(V, W)$ is a vector space of dimension mn, but so is $F^{m\times n}$. So they must be isomorphic!

$$L(V, W) \cong F^{m\times n}$$

Question: What is a linear transformation giving this isomorphism?

Sure, they're isomorphic, but how do we get from one to the other?

So the input here is just a linear transformation $T : V \to W$ (an element of $L(V, W)$), and the output is some m by n matrix (an element of $F^{m\times n}$).

Chose a basis $\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$ for V with $\dim(V) = n$, and a basis $\mathcal{B}' = \{\beta_1, \ldots, \beta_m\}$ for W with $\dim(W)$. Then there is an isomorphism $C_{\mathcal{B}}$ Taking V to F^n.

But there is also a coordinate isomorphism $C_{\mathcal{B}'}$ from W to F^m.

But recall that T goes from V to W, so the diagram commutes.

Let $M^{\mathcal{B}}_{\mathcal{B}'}(T) = [[T(\alpha_1)]_{\mathcal{B}'}, [T(\alpha_2)]_{\mathcal{B}'}, \ldots, [T(\alpha_n)]_{\mathcal{B}'}]$.

Then, we propose that, for any $\alpha \in V$,

$$M^{\mathcal{B}}_{\mathcal{B}'}(T) \cdot [\alpha]_{\mathcal{B}} = [T(\alpha)]_{\mathcal{B}'}$$

3.1 Diagram

$$\begin{tikzcd}
\alpha \in V \arrow[r, "C_{\mathcal{B}}"] \arrow[d, "T"'] & {[\alpha]_{\mathcal{B}} \in F^n} \arrow[d, dashed, "M^{\mathcal{B}}_{\mathcal{B}'}(T)"] \\
T(\alpha) \in W \arrow[r, "C_{\mathcal{B}'}"'] & {[T(\alpha)]_{\mathcal{B}'} \in F^n}
\end{tikzcd}$$

Let's give a concrete example for this.

Example.

Let $D : \mathbb{P}_2 \to \mathbb{P}_1$

$$D(a_0 + a_1x + a_2x^2) = a_1 + 2a_2x$$

$\mathcal{B} = \{1, x, x^2\}$ be a basis for $\mathbb{P}_2$, and $\mathcal{B}' = \{1, x\}$ be a basis for P_1.

Question

What is the matrix $M^{\mathcal{B}}_{\mathcal{B}'}(D)$? Well

$D(1) = 0,\ D(x) = 1, D(x^2) = 2x$, then $[D(1)]_{\mathcal{B}'} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, $[D(x)]_{\mathcal{B}'} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $[D(x^2)]_{\mathcal{B}'} = \begin{bmatrix} 0 \\ 2 \end{bmatrix}$, So

$$M^{\mathcal{B}}_{\mathcal{B}'}(D) = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

But what does this mean? Well, chose any $f \in P_2$, say

$$f(x) = 5 + 3x - x^2$$

What are the coordinates of f with respect to $\mathcal{B}$?

$$[f]_{\mathcal{B}} = \begin{bmatrix} 5 \\ 3 \\ -1 \end{bmatrix}$$

Then

$$M^{\mathcal{B}}_{\mathcal{B}'}(D) \cdot [f]_{\mathcal{B}} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} 5 \\ 3 \\ -1 \end{bmatrix} = \begin{bmatrix} 3 \\ -2 \end{bmatrix}$$

But what does this final vector mean? Well, it's just $[D(f)]_{\mathcal{B}'}$, in other words, it's the derivative of f with respect to the basis $\mathcal{B}'$!

Really, all we're doing here is moving around in the diagram. All we need to do is to apply T to every α_i, living in V.

Note: Look at Homework number 13 on 3.4.

To show that $\{E^{pq}\}$ span $L(V, W)$, chose $T : V \to W$. Think about $M^{\mathcal{B}}_{\mathcal{B}'}(E^{pq})$

Fri. Mar 3 2023

Last time, we saw that, given a transformation $T : V \to W$, and bases $\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$ of V, and $\mathcal{B}' = \{\beta_1, \ldots, \beta_n\}$ of W.

The matrix of T with respect to $\mathcal{B}, \mathcal{B}'$ is the $m \times n$ matrix

$$M^{\mathcal{B}}_{\mathcal{B}'}(T) = [[T(\alpha_1)]_{\mathcal{B}'}, \ldots, [T(\alpha_n)]_{\mathcal{B}'}]$$

For any $\alpha \in V$,

$$M^{\mathcal{B}}_{\mathcal{B}'}(T) \cdot [\alpha]_{\mathcal{B}} = [T(\alpha)]_{\mathcal{B}'}$$

Let's now look at a special case.

Definition.

Let V be a vector space over a field F, then a **Linear Operator** $T : V \to V$ is any linear transformation from V to itself.

These are extremely applicable, even in the real world. Let's look at some examples.

Example.

Let V be $\mathbb{R}^2$. Possible linear operators are $T_\theta : \mathbb{R}^2 \to \mathbb{R}^2$ where T_θ *rotates* points by θ radians.

Another $U : \mathbb{R}^2 \to \mathbb{R}^2$ would be to *stretch* x by a factor of 5, and y by a factor of 2. Notice that the unit square would be stretch by a factor of 5×2 **Author Note**.

Some will notice that this is the information that the determinant of T encodes!

A third example might be in differential equations. Let V be the set of all pairs of foxes and rabbits, and T encodes the number of foxes and rabbits one generation later.

Moreover, linear transformations are also used in physics! Let V be an "Electron cloud" in Quantum mechanics. Heisenberg's uncertainty principle tells us that, an observation on V is a linear operator!

Author Note.

This is beyond the scope of the class, but the point is that this *is* extremely useful in real life!

Previously we defined $L(V, W)$ as being the vector space of all vector spaces V and W. Now let's define $L(V, V)$ as the space of all linear operators.

If $\dim(V) = n$, then $\dim(L(V, V)) = n^2$. This is all matrices representing $T : V \to V$ will be square.

For $T, U \in L(V, V)$, then $UT \in L(V, V)$ is the linear operator which "does T first, then U." In other words

$$UT(\alpha) = U(T(\alpha)) = U \circ T(\alpha)$$

Conversely

$$TU(\alpha) = T(U(\alpha)) = T \circ U(\alpha)$$

So these operations are read from *right to left*.

Note.

Typically $UT \neq TU$. The order in which you put on shoes and socks matters.

Example.

Suppose that $T, U : \mathbb{R}^2 \to \mathbb{R}^2$, with

$$T(\begin{bmatrix} x \\ y \end{bmatrix}) = \begin{bmatrix} 5x \\ 2y \end{bmatrix}$$

and

$$U(\begin{bmatrix} x \\ y \end{bmatrix}) = \begin{bmatrix} y \\ x \end{bmatrix}$$

swaps x and y. So then

$$UT(\begin{bmatrix} x \\ y \end{bmatrix}) = U(\begin{bmatrix} 5x \\ 2y \end{bmatrix}) = \begin{bmatrix} 2y \\ 5x \end{bmatrix}$$

> and
>
> $$TU(\begin{bmatrix} x \\ y \end{bmatrix}) = T(\begin{bmatrix} y \\ x \end{bmatrix}) = \begin{bmatrix} 5y \\ 2x \end{bmatrix}$$
>
> which are not the same.

Let's look at what it means to raise a linear transformation to a power.

Let $T^2 : V \to V$, then

$$T^2(\alpha) = T(T(\alpha))$$

$$T^{10}(\alpha) = T(\cdots T(T(\alpha))\cdots)$$

10 times.

$$(T - U)(T + U) = T^2 + TU - UT + U^2$$

But note that TU and UT *cannot* be canceled out here, since they might not be the same!

We often call the identity operator $I : V \to V$. It just "does nothing".

$$I(\alpha) = \alpha$$

for any $\alpha \in V$.

> **Note**.
>
> I commutes with everything!
>
> $$T \circ I = I \circ T$$
>
> for any linear operator T.

If T is invertible, T^{-1} "undoes" T

$$T^{-1}T = TT^{-1} = I$$

But note the domain and codomain of T, they must be the same! **Author Note**.

If this isn't the case, the inverse might only work in *one direction* .

Let's simplify this more

We say that T is invertible

- if and only if T is onto
- if and only if T is one to one

Author Note.

If you have one of these facts, the others come for free!

Definition.

Fix V a vector space of dimension n, and fix *one* basis $\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$ of V. Then,

$$M_{\mathcal{B}}^{\mathcal{B}}(T) = \Big[[T(\alpha_1)]_{\mathcal{B}}, \ldots, [T(\alpha_n)]_{\mathcal{B}}\Big]$$

Is an $n \times n$ matrix. Let's look at what it does.

Let's take some $\alpha \in V$, then

$$M_{\mathcal{B}}^{\mathcal{B}}(T) \cdot [\alpha]_{\mathcal{B}} = [T(\alpha)]_{\mathcal{B}}$$

Where $[\alpha]_{\mathcal{B}}$ and the result are $n \times 1$ column vectors. What this matrix does then, is take a vector $\alpha \in V$, written in the basis of $\mathcal{B}$, and output the result of applying T to α, still in the basis of $\mathcal{B}$.

Note.

The book refers to $M_{\mathcal{B}}^{\mathcal{B}}(T)$ as $[T]_{\mathcal{B}}$. So we would have

$$[T]_{\mathcal{B}}[\alpha]_{\mathcal{B}} = [T(\alpha)]_{\mathcal{B}}$$

If you input the $\mathcal{B}$ coordinates of α, you get the $\mathcal{B}$ coordinates of $T(\alpha)$. What this means is that we have an isomorphism T going from $L(V, V)$ to $F^{n\times n}$

$$L(V,V) \xrightarrow{\quad T \quad} F^{n\times n}$$

In fact, there are many different isomorphisms $L(V, V) \to F^{n\times n}$. Fix any basis $\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$ of V, then we get an isomorphism using $\mathcal{B}$

$$T : V \to V \xleftrightarrow{\text{Isomorphism}} [T]_{\mathcal{B}}$$

This isomorphism takes T and writes it *in the language* of $\mathcal{B}$. We'll see later what this isomorphism actually is, because we can construct it!

In the mean time, let's invoke a basis $\mathcal{B}' = \{\alpha_1', \ldots, \alpha_n'\}$ of V, different of $\mathcal{B}$.

Question: How are $[T]_{\mathcal{B}}$ and $[T]_{\mathcal{B}'}$ related?

We actually already know how to do this from section 2.4. We had a "change of basis" matrix. Let's look at it again.

Given bases $\mathcal{B} = \{\alpha_1, \dots, \alpha_n\}$ and $\mathcal{B}' = \{\alpha'_1, \dots, \alpha'_n\}$ of V, the book gave us the matrix

$$P = \Big[[\alpha'_1]_{\mathcal{B}}, \dots, [\alpha'_n]_{\mathcal{B}} \Big]$$

In other words, P is just the $\mathcal{B}$ representation of the basis vectors of $\mathcal{B}'$, written as the column vectors of a matrix.

TODO Question: How do we get the $\mathcal{B}$ representation of α'_i? What basis is α'_i written in when it's in the basis of $\mathcal{B}' = \{\alpha'_1, \dots, \alpha'_n\}$? I presume it's the standard basis, but is this correct?

So for any $\alpha \in V$,

$$P[\alpha]_{\mathcal{B}'} = [\alpha]_{\mathcal{B}}$$

What this means is that if we have a vector $\alpha \in V$, written using the coordinates of $\mathcal{B}'$, we can translate it *to* the coordinates of $\mathcal{B}$ by multiplying it by P.

> **Note.**
>
> Recall what we mean when we talk about $[\alpha]_{\mathcal{B}}$. This just means "The coordinates of α, written using the basis $\mathcal{B}$".

In our notation P is just $M_{\mathcal{B}}^{\mathcal{B}'}(I)$. It's the matrix that translates *from* $\mathcal{B}'$ *to* $\mathcal{B}$. Why do we pass in the identity to M? Well, recall the definition of $M_{\mathcal{B}}^{\mathcal{B}'}(T)$

$$M_{\mathcal{B}}^{\mathcal{B}'}(T) = \Big[[T(\alpha'_1)]_{\mathcal{B}}, \dots, [T(\alpha'_n)]_{\mathcal{B}} \Big]$$

So if the transformation is the identity, it becomes

$$M_{\mathcal{B}}^{\mathcal{B}'}(I) = \Big[[\alpha'_1]_{\mathcal{B}}, \dots, [\alpha'_n]_{\mathcal{B}} \Big]$$

Then, in our notation, we have

$$M_{\mathcal{B}}^{\mathcal{B}'}(I)[\alpha]_{\mathcal{B}'} = [\alpha]_{\mathcal{B}}$$

Question: How are $M_{\mathcal{B}}^{\mathcal{B}'}(I)$ and $M_{\mathcal{B}'}^{\mathcal{B}}(I)$ related?

Well, we know that

$$M_{\mathcal{B}}^{\mathcal{B}'}(I)[\alpha]_{\mathcal{B}'} = [\alpha]_{\mathcal{B}}$$

So

$$\left(M_{\mathcal{B}}^{\mathcal{B}'}(I) \right)^{-1} [\alpha]_{\mathcal{B}} = [\alpha]_{\mathcal{B}'}$$

So they are inverses of each other!

Note.

We can do this because $M_{\mathcal{B}}^{\mathcal{B}'}(I) \cdot [\alpha]_{\mathcal{B}'}$ is nothing more than matrix multiplication. We have a vector α written in the basis of $\mathcal{B}'$, and we multiply it by the matrix $M_{\mathcal{B}}^{\mathcal{B}'}(I)$.

But how do we know that $M_{\mathcal{B}}^{\mathcal{B}'}(I)$ has an inverse? This is a fair question and the answer might not be immediately obvious. Recall once more then definition of $M_{\mathcal{B}}^{\mathcal{B}'}(I)$, we have

$$M_{\mathcal{B}}^{\mathcal{B}'}(I) = \Big[[\alpha_1']_{\mathcal{B}}, \ldots, [\alpha_n']_{\mathcal{B}}\Big]$$

First, notice that each α_i' is a vector of size n, since $\dim(V) = n$, so here, we're working with an $n \times n$ matrix. Secondly, we know that $\{\alpha_1', \ldots, \alpha_n'\}$ form a basis of V, so they *must* be linearly independent. But if they're linearly independent, that means that each column of $M_{\mathcal{B}}^{\mathcal{B}'}(I)$ is linearly independent, so this is a full rank matrix and so it must have an inverse.

Example.

Given $[T]_{\mathcal{B}}$, we want $[T]_{\mathcal{B}'}$. This is a 3 step process.

Author Note.

We want $[T]_{\mathcal{B}'}[\alpha]_{\mathcal{B}'} = [T(\alpha)]_{\mathcal{B}'}$

We first start with a vector α in the language of $\mathcal{B}'$, notated as $[\alpha]_{\mathcal{B}'}$

1. Change $[\alpha]_{\mathcal{B}'}$ to $[\alpha]_{\mathcal{B}}$

   We first multiply $[\alpha]_{\mathcal{B}'}$ by $M_{\mathcal{B}}^{\mathcal{B}'}(I)$, essentially translating it from the basis of $\mathcal{B}'$ to $\mathcal{B}$

$$M_{\mathcal{B}}^{\mathcal{B}'}(I) \cdot [\alpha]_{\mathcal{B}'} = [\alpha]_{\mathcal{B}}$$

2. Apply $[T]_{\mathcal{B}} \cdot [\alpha]_{\mathcal{B}} = [T(\alpha)]_{\mathcal{B}}$

   We then apply the transformation $[T]_{\mathcal{B}}$ to this.

$$[T]_{\mathcal{B}} \cdot \Big(M_{\mathcal{B}}^{\mathcal{B}'}(I) \cdot [\alpha]_{\mathcal{B}'}\Big) = [T(\alpha)]_{\mathcal{B}}$$

   $[T]_{\mathcal{B}}$ is the linear operator working with vectors *in the language* of $\mathcal{B}$, which is why we *can't* apply it to a vector $[\alpha]_{\mathcal{B}'}$. We first had to translate α to the language of $[T]_{\mathcal{B}}$.

3. Change $[T(\alpha)]_{\mathcal{B}}$ to $[T(\alpha)]_{\mathcal{B}'}$

   We now have a vector $[T(\alpha)]_{\mathcal{B}}$ which is the result of applying T to α, all in the language of $\mathcal{B}$. In order for us to get it in the language of $\mathcal{B}'$, we only have to multiply it by $M_{\mathcal{B}'}^{\mathcal{B}}(I)$, the matrix which takes us from $\mathcal{B}$ to $\mathcal{B}'$.

$$M_{\mathcal{B}'}^{\mathcal{B}}(I) \cdot \left([T]_{\mathcal{B}} \cdot \Big(M_{\mathcal{B}}^{\mathcal{B}'}(I) \cdot [\alpha]_{\mathcal{B}'}\Big)\right) = [T(\alpha)]_{\mathcal{B}'}$$

Mon. Mar 6 2023

Recall our discussion about linear operators last time.

$$T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} -y \\ x \end{bmatrix}$$

There is a nice visual intuition for these. In this case, this operation flips values over the x axis.

Let's look at other examples.

$$T_F\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} y \\ x \end{bmatrix}$$

This linear transformation (or more precisely, linear operator) flips values across the $y = x$ line.

Let's define a basis $S = \left\{\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}\right\}$, the standard basis in $\mathbb{R}^2$.

Question: What is the matrix associated with T_F? Well, T_F sends the first basis vector to the second, and the second to the first, so we get the linear transformation

$$[T_F]_S = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

Now, we can compute where a vector ends up on T by multiplying it by its associated matrix.

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 5 \\ 3 \end{bmatrix} = \begin{bmatrix} 3 \\ 5 \end{bmatrix}$$

Let $\mathcal{B} = \left\{\begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix}\right\}$ be a different basis (we know they form a basis because the two vectors are linearly independent)

Want: $[T]_\mathcal{B}$

We want $[T]_\mathcal{B}$, and we know that $[T]_\mathcal{B} = [[T(\alpha_1)]_\mathcal{B}, [T(\alpha_2)]_\mathcal{B}]$.

$T(\alpha_1) = T\left(\begin{bmatrix} 1 \\ 2 \end{bmatrix}\right) = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$. Notice that the coordinates of the output vector is with respect to the standard basis S. We want it expressed in terms of $\mathcal{B}$. But we know how to do that

$$\begin{bmatrix} -2 \\ 1 \end{bmatrix} = c_1\begin{bmatrix} 1 \\ 2 \end{bmatrix} + c_2\begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

By inspection, we see that $-1 = 3c_1$, so $c_1 = -\frac{1}{3}$, and so $c_2 = -\frac{5}{3}$.

So we have that

$$\begin{bmatrix} -2 \\ 1 \end{bmatrix}_S = \begin{bmatrix} -1/3 \\ -5/3 \end{bmatrix}_\mathcal{B}$$

Similarly, we can find that

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix}_S = \begin{bmatrix} 2/3 \\ 1/3 \end{bmatrix}_{\mathcal{B}}$$

Putting everything together, we have

$$[T]_{\mathcal{B}} = [[T(\alpha_1)]_{\mathcal{B}}, [T(\alpha_2)]_{\mathcal{B}}] = \begin{bmatrix} -1/3 & 2/3 \\ -5/3 & 1/3 \end{bmatrix}$$

So the two matrices $\begin{bmatrix} -1/3 & 2/3 \\ -5/3 & 1/3 \end{bmatrix}$ and $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ represent the *same* operator in $\mathbb{R}^2$. Hopefully at this point, we can see that certain matrices are easier to work with than others, even if they do the same thing.

Comic Relief

Let $c \in \mathbb{R}$

Claim: $T - cI$ is invertible.

With respect to the standard basis S,

$$T - cI = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} - c \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

So for any 2×2 matrix $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$, it's invertible if and only if $ad - bc \neq 0$.

> **Definition**.
>
> Two $n \times n$ matrices are **similar** if they represent the same linear transformation, but with respect to different bases.

Let $\mathcal{B}' = \left\{ \beta_1 = \begin{bmatrix} a \\ b \end{bmatrix}, \beta_2 = \begin{bmatrix} c \\ d \end{bmatrix} \right\}$ be *any* basis for $\mathbb{R}^2$.

Find $[T]_{\mathcal{B}'}$

Well we have

$$[T]_{\mathcal{B}'} = [[T(\beta_1)]_{\mathcal{B}'}, [T(\beta_2)]_{\mathcal{B}'}]$$

and we have that

$$[T(\beta_1)]_S = T(\begin{bmatrix} a \\ b \end{bmatrix}) = \begin{bmatrix} -b \\ a \end{bmatrix}_S$$

$$[T(\beta_2)]_S = T(\begin{bmatrix} c \\ d \end{bmatrix}) = \begin{bmatrix} -d \\ c \end{bmatrix}_S$$

Need: $[T(\beta_1)]_{\mathcal{B}'}, [T(\beta_2)]_{\mathcal{B}'}$

For this, we need $M^S_{\mathcal{B}'}(I)$, but this is kind of a pain, instead, let's ask:

What is $M^{\mathcal{B}'}_S(I)$? Well this is easy:

$$M_S^{\mathcal{B}'}(I) = [[\beta_1]_S, [\beta_2]_S] = \begin{bmatrix} a & c \\ b & d \end{bmatrix}$$

But then, finding the inverse is easy!

$$M_{\mathcal{B}'}^S(I) = \left(M_S^{\mathcal{B}'}(I)\right)^{-1} = \frac{1}{ad-bc}\begin{bmatrix} d & -c \\ -b & a \end{bmatrix}$$

Notice: $ad-bc$ will never be zero, if they were, and $ad = bc$, the rows would not be linearly independent, so $\mathcal{B}$ would not be a basis.

Finally, we have

$$\begin{aligned}
[T]_{\mathcal{B}'} =& M_{\mathcal{B}'}^S \cdot [T]_S \cdot M_S^{\mathcal{B}'}(I) \\
=& \frac{1}{ad-bc}\begin{bmatrix} d & -c \\ -b & a \end{bmatrix}\left(\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} a & c \\ b & d \end{bmatrix}\right) \\
=& \frac{1}{ad-bc}\begin{bmatrix} d & -c \\ -b & a \end{bmatrix}\left(\begin{bmatrix} -b & -d \\ a & c \end{bmatrix}\right) \\
=& \frac{1}{ad-bc}\begin{bmatrix} -db-ca & -d^2-c^2 \\ b^2+a^2 & bd+ac \end{bmatrix} = [T]_{\mathcal{B}'}
\end{aligned}$$

Now notice: $-d^2 - c^2 = 0$ only if $c = d = 0$. Secondly, $b^2 + a^2 = 0$ only if $a = b = 0$

For any 2×2 matrix, you can look at the trace, the sum of the diagonal.

Later, we will study similar matrices and we will show that $\text{trace}(A) = \text{trace}(A')$ if A, A' are similar. (The converse does not hold)

TODO Get Notes for March 8

Fri. Mar 10 2023

For $n \times n$ matrices

The trace(A) is defined as the sum of the main diagonal. In other words, the trace is a linear functional from $F^{n\times n} \to F$, and we then have the following operations.

- $\text{trace}(A + B) = \text{trace}(A) + \text{trace}(B)$
- $\text{trace}(cA) = c(\text{trace}(A))$

We also have the following fact, proved on the homework.

$$\text{trace}(AB) = \text{trace}(BA)$$

And the following corollary

If A and A' are similar, then

$$\text{trace}(A) = \text{trace}(A')$$

The proof is that, if $A' = P^{-1}AP$, then $\text{trace}(A') = \text{trace}((P^{-1}A)P) = \text{trace}((AP^{-1})P) = \text{trace}(A)$

We also have another corollary,

For any square matrices in fields of characteristic 0, we have that

$AB - BA = I$ is impossible, proved on the homework. The proof is that, if matrices are equal, then they have the same trace, but

$$\text{trace}(I) = n \neq \text{trace}(AB - BA) = \text{trace}(AB) - \text{trace}(BA) = 0$$

Recall the dual space.

If the $\dim(V) = n$ with basis $\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$, and $\dim(V^*) = n$ with basis $\{f_1, \ldots, f_n\}$ with the following property

$$f_i(\alpha_j) = \begin{cases} 1 \text{ if } i = j \\ 0 \text{ if } i \neq j \end{cases}$$

Definition.

The Annihilator of S denoted by $\text{Ann}(S)$ or S^0, defined by

$$\text{Ann}(S) = S^0 = \{f \in V^* : f(\beta) = 0 \text{ for every } \beta \in S\ \}$$

By this definition, we can see that S^0 is a subspace of V^* for any $S \subseteq V$.

Note that if $S \subseteq T$, then $\text{Ann}(T) \supseteq \text{Ann}(S)$. Why?

Choose some $f \in \text{Ann}(T)$, then $f(\beta) = 0$ for all $\beta \in T$. Since $S \subseteq T$, this implies that all $S = 0$ for all $\beta \in S$. Therefore **TODO**

Let's look at some examples

Example.

$\text{Ann}(\{0\}) = \{f \in V^* : f(0) = 0\} = V^*$ since V^* is a vector space.

$\text{Ann}(V) = \{f \in V^* : f(\beta) = 0 \text{ for all } \beta \in V\ \}$ is the zero transformation in V^*.

Theorem.

If $\dim(V)$ is finite, say n, then for any subspace $W \subseteq V$

$$\dim(W) + \dim(\text{Ann}(W)) = n$$

Let $\{\alpha_1, \ldots, \alpha_k\}$ be a basis for $W \subseteq V$. This basis can be extended to a basis of V. Let $\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$ be an extension to a basis for all of V.

Let $\mathcal{B}^* = \{f_1, \ldots, f_n\}$ be the dual basis of V^*.

Claim: $\{f_{k+1}, \ldots, f_n\}$ form a basis for $\text{Ann}(W)$.

Given the claim, the theorem is proved, since the numbers would add up.

There are three steps.

1. Step 1 is to show that $\{f_{k+1}, \dots, f_n\} \subseteq \text{Ann}(W)$.

   Chose any $\beta \in W$, say $\beta = c_1\alpha_1 + \cdots + c_k\alpha_k$. Now chose any $i \geq k+1$. Then $f_i(\beta) = f_i(c_1\alpha_1 + \cdots + c_k\alpha_k) = c_1 f(\alpha_1) + \cdots + c_k f(\alpha_k)$. But since $i \geq k$

   $$c_1 f(\alpha_1) + \cdots + c_k f(\alpha_k) = 0$$

   so what we've shown is that $f_i(\beta) = 0$, therefore $f_i \in \text{Ann}(W)$.

2. Step 2: $\{f_{k+1}, \dots, f_n\}$ are linearly independent.

   This is obvious since this is how they were defined, all $f_1, \dots, f_n$ is a basis for V^*.

3. $\{f_{k+1}, \dots, f_n\}$ spans $\text{Ann}(W)$.

   Chose any $g \in \text{Ann}(W) \subseteq V^*$. Since $\mathcal{B}^*$ spans V^*

   $$g = d_1 f_1 + \cdots + d_n f_n$$

   for some $d_1, \dots, d_n$

   then for any j,

   $$g(\alpha_j) = d_1 f_1(\alpha_j) + \cdots + d_n f_n(\alpha_j) = d_j(1) = d_j$$

   But $g \in \text{Ann}(W)$, therefore for $1 \leq j \leq k$

   $$g(\alpha_j) = 0$$

   since $\alpha_1, \dots, \alpha_k$ are in W.

   Therefore, $d_1 = 0, d_2 = 0, \dots, d_k = 0$, therefore $g \in \text{Span}(\{f_{k+1}, \dots, f_n\})$

<u>**Mon. Mar 13 2023**</u>

Last time, we saw that we have a vector space V and a subspace $W \subseteq V$ with $\dim(V) = n$ and $\dim(W) = k$.

Let $\{\alpha_1, \dots, \alpha_k\}$ be a basis for W. Extend to a basis of $\mathcal{B} = \{\alpha_1, \dots, \alpha_n\}$ of V.

Let $\mathcal{B}^* = \{f_1, \dots, f_n\}$ be a basis for $\text{Ann}(W)$ so $\dim(V) = n = \dim(W) + \dim(\text{Ann}(W))$, then we have that $\dim(\text{Ann}(W)) = n - k$

Corrolary

If $W \subseteq V$ is a subspace $\beta \in V \backslash W$, the nthere is some $f : V \to F \in V^*$ such that $f \in \text{Ann}(W)$, but $f(\beta) \neq 0$

Proof

Write $\beta = c_1\alpha_1 + \cdots + c_n\alpha_n$

Claim. At least one of $c_{k+1}, c_{k+2}, \dots, c_n \neq 0$. If not, $\beta \in c_1\alpha_1 + \cdots + c_k\alpha_k$ implies that $\beta \in W$.

Say that $c_i \neq 0$ with $i \geq k+1$. Then, $f_i \in \text{Ann}(W)$, but $f_i(\beta) = c_i \neq 0$.

Corrolary

If W_1, W_2 are both subspaces of V, then $W_1 = W_2$ if and only if $\text{Ann}(W_1) = \text{Ann}(W_2)$

Proof

The forwards direction is obvious, if they are the same, they have the same annihilator space.

For the backwards direction, there are two cases

1. There is some $\beta \in W_1 \backslash W_2$.

   Apply the first corollary to W_2, then there is some $f \in \text{Ann}(W_2)$ with $f(\beta) \neq 0$, therefore $f \in \text{Ann}(W_1)$ and so $\text{Ann}(W_2) \neq \text{Ann}(W_1)$.

2. There is some $\beta \in W_2 \backslash W_1$

   Apply the first corollary to W_1, the argument is similar.

Say that $f : V \to F$ is any linear functional, but $f \neq 0$. Then $\text{nullity}(f) = \dim(V) - 1$.

Proof. $\text{Range}(f) \subseteq F'$, but $\text{Range}(F) \neq \{0\}$ therefore the range of F must equal F. So $\text{rank}(f) = 1$.

$$\text{nullity}(f) = n - 1$$

> **Definition**.
>
> A **Hyperspace** is a subspace $W \subseteq V$ with $\dim(W) = \dim(V) - 1$.
>
> We say that W has "co-dimension" 1.

$\text{Null}(f)$ is a hyperspace.

Corollary

Let $W \subseteq V$ with $\dim(W) = k$, $\dim(V) = n$, and $k < n$.

Then W is the **intersection** of $n - k$ hyperspaces.

Author Note.

What this is saying is that you can shrink down from V to W in $n - k$ steps.

Proof

Claim.

$$W = \text{Null}(f_{k+1}) \cap \text{Null}(f_{k+1}) \cap \cdots \cap \text{Null}(f_n)$$

Author Note.

Note that f_{k+1} can't be 0 because they are part of a basis.

Proof. Chose $\beta \in W$. For $i \geq k+1$, $f_i \in \text{Ann}(W)$, therefore $f_i(\beta) = 0$, and so $\beta \in \text{Null}(f_i)$, and so β is in the intersection.

Next, chose $\beta \in \text{Null}(f_{k+1}) \cap \text{Null}(f_{k+2}) \cap \cdots \cap \text{Null}(f_n)$. Then

$$f_{k+1}(\beta) = 0, f_{k+2}(\beta) = 0, \ldots, f_n(\beta) = 0,$$

But since $\{f_{k+1}, \ldots, f_n\}$ is a basis of $\text{Ann}(W)$, we have that for any $g \in \text{Ann}(W)$, $g(\beta) = 0$. What we can conclude from this is that β must be in W.

Author Note.

There is no g that kills off everything in W but does not kill β.

Say $\dim(V) = n$, $\dim(W) = m$, and $T : V \to W$ is any linear transformation.

We now know that each of V and W have dual spaces V^* and W^* respectively. Note that V and W are not necessarily the same size, so their dual spaces might not either.

Definition.

The **transpose** of T, written T^t is a linear transformation going from $W^* \to V^*$, defined as

$$T^t(g) = f$$

Where $g \in W^*$, and $f \in V^*$ is a linear functional defined as $f : V \to F$. Take f, and apply it to any $\alpha \in V$

$$T^t(g)(\alpha \in V) := g(T(\alpha)) \in F$$

Check. $T^t : W^* \to V^*$ is a linear transformation.

TODO cleanup

Theorem.

Say V, W are finite dimensional vector spaces with $\dim(V) = n, \dim(W) = m$, and $T : V \to W$ is any linear transformation where $\text{rank}(T) = r$.

Then, the following are true

1. $\text{Null}(T^t) = \text{Ann}(\text{Range}(T))$
2. $\text{rank}(T^t) = \text{rank}(T) = r$
3. $\text{Range}(T^t) = \text{Ann}(\text{Null}(T))$

1. $\text{Null}(T^t) = \text{Ann}(\text{Range}(T))$

   The proof of this is largely definition unwinding.

$$\begin{aligned} & g \in \text{Null}(T^t) \\ \Leftrightarrow & T^t(g) = \mathbf{0} && \text{By definition of Null Space} \\ \Leftrightarrow & \forall \alpha \in V, (T^t(g))(\alpha) = 0 && \text{Expanding definition further} \end{aligned}$$

   Let's take a second to understand what's happening here. T^t is a map from W^* to V^*. So it takes in elements of W^*, which are linear functionals $g : W \to F$, and produces elements of V^*: linear functionals $f : V \to F$.

   What this means is that when we take a linear functional $g : W \to F$, and a linear transformation $T^t : W^* \to V^*$, $T^t(g)$ is itself now a map going from $V^* \to F$. In other words, T^t took our map g going from W to F and transformed it into a map f going from V to F.

Let's keep going.

$$
\begin{aligned}
&\forall \alpha \in V, (T^t(g))(\alpha) = 0 \\
\Leftrightarrow &\forall \alpha \in V, g(T(\alpha)) = 0 && \text{By definition of } T^t \\
\Leftrightarrow &g \in \text{Ann}(\text{Range}(T)) && \text{By definition of Annihilator}
\end{aligned}
$$

So what we have is that, if g is in the Null Space of T^t, it also must be in the Annihilator of the Range of T.

2. $\text{rank}(T^t) = \text{rank}(T) = r$

   TODO Diagram chasing

3. $\text{Range}(T^t) = \text{Ann}(\text{Null}(T))$

   Check. $\text{Range}(T^t) \subseteq \text{Ann}(\text{Null}(T))$

   Given this, $\dim(\text{Ann}(\text{Null}(T))) = r$. Now we know that $\text{Range}(T^t) = r$, and $\dim(\text{Ann}(\text{Null}(T))) = r$

   Therefore

$$
\text{Range}(T^t) = \text{Ann}(\text{Null}(T))
$$

Wed. Mar 15 2023

Let A be the $m \times n$ matrix for T with respect to $\mathcal{B}, \mathcal{B}'$, i.e.

$$
A \cdot [\alpha]_{\mathcal{B}} = [T(\alpha)]_{\mathcal{B}'}
$$

Let A^t be the $n \times m$ matrix for T^t with respect to $(\mathcal{B}')^*$ and $\mathcal{B}^*$.

But it's incredibly easy to go from A to A^t,

$$
\left(A^t\right)_{ij} = (A_{ji})
$$

For example

$$
A = \begin{bmatrix} 2 & 3 & 1 \\ 4 & 0 & 5 \end{bmatrix}
$$

then

$$
A^t = \begin{bmatrix} 2 & 4 \\ 3 & 0 \\ 1 & 5 \end{bmatrix}
$$

This is kind of profound! How on earth is this whole T^t thing related to flipping a matrix on its diagonal?

Last time, we saw that

- $\mathrm{rank}(T^t) = \mathrm{rank}(T)$
- $\mathrm{Range}(T^t) = \mathrm{Ann}(\mathrm{Null}(T))$

What does this mean when

$$V = \mathbb{R}^n \text{ and } W = \mathbb{R}^m$$

with standard bases

$$\mathcal{B} \text{ and } \mathcal{B}'$$

Well we saw that for a linear system

$$\begin{gathered} 2x_1 + 3x_2 + 1x_3 \\ 4x_1 + 0x_2 + 5x_3 \end{gathered}$$

Each of these lines is a linear functional!

$$f\left(\begin{bmatrix} a \\ b \\ c \end{bmatrix}\right) = 2(a) + 3(b) + 1(c)$$

When we had a linear transformation $T : \mathbb{R}^3 \to \mathbb{R}^2$, say for example

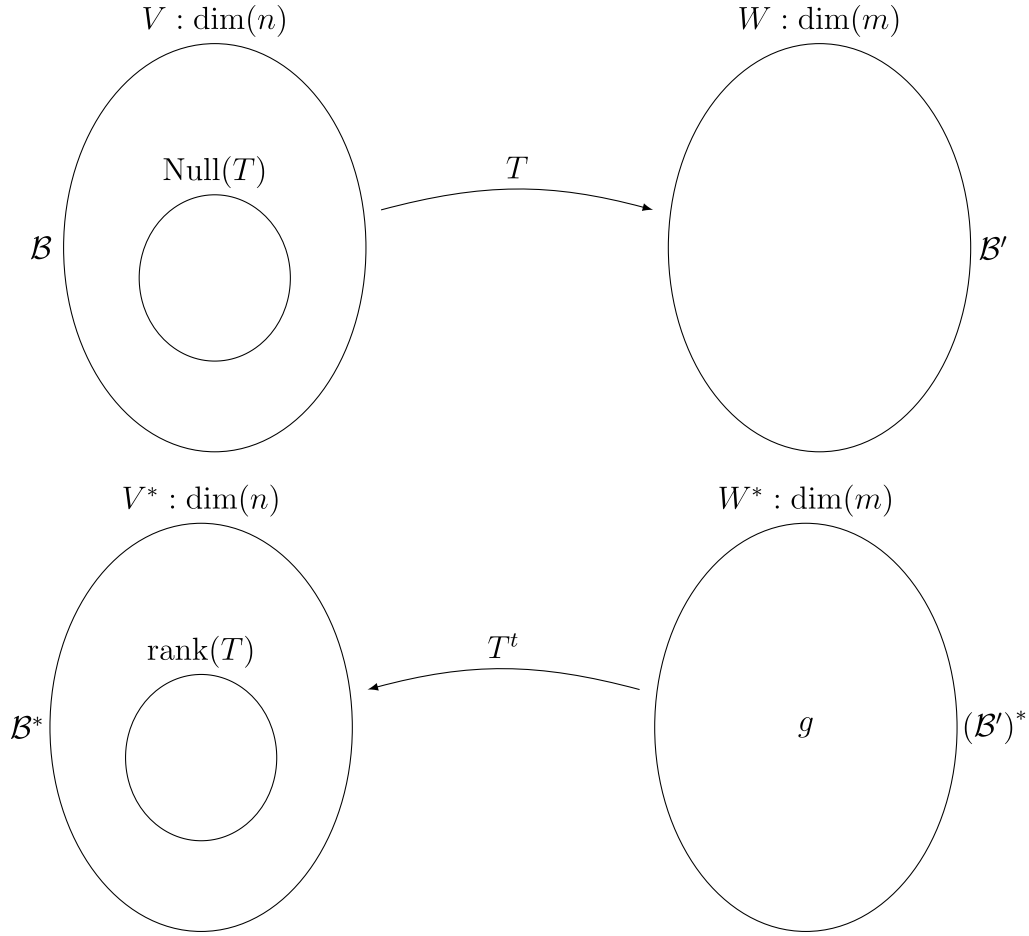
$$T\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = (2x_1 + 3x_2 + x_3, 4x_1 + 0x_2 + 5x_3) = \begin{bmatrix} 2x_1 + 3x_2 + x_3 \\ 4x_1 + 0x_2 + 5x_3 \end{bmatrix}$$

We have the matrix associated with T is

$$A = \begin{bmatrix} 2 & 3 & 1 \\ 4 & 0 & 5 \end{bmatrix}$$

We also saw that $\mathrm{rank}\left((T_A)^t\right) = \mathrm{rank}(T_A) = \dim(\mathrm{ColSpace}(A))$

But we know that $\mathrm{rank}\left((T_A)^t\right)$ is the same as $\dim(\mathrm{ColSpace}((T_A)^t))$ which is the same as $\dim(\mathrm{ColSpace}(A^t))$ which is the same as $\dim(\mathrm{RowSpace}(A))$.



This marks the end of Chapter 3. Chapter 4 will focus on polynomial, and chapter 5 on determinants.

We will study linear operators for some time, as they have a lot of real world significance.

We will find a magic polynomial: the characteristic polynomial, which will give us a lot of information about the linear transformation.

3.2 Chapter 4

Our discussion of polynomials begins.

Note: We will mostly skip 4.1 and 4.3 from the book

Let F be a field ($\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$). Which field we choose is going to matter here.

Definition.

$F[x]$ is all polynomials of the form

$$a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$$

Where $a_0, a_1, \dots, a_n \in F$

We saw this before, we just called it $\mathbb{P}$ before. We also had that $\mathbb{P}$ had the basis $\{1, x, x^2, \dots\}$

Note: We can multiply two polynomials together and get a polynomial!

You can always add, always subtract, always multiply, but not always divide! For instance you cant divide by the zero polynomial. There are restrictions about when you can divide.

$f(x) = x^2$ is a perfectly fine polynomial, but $\frac{1}{x^2}$ is not a polynomial, since there is no multiplicative inverse. **TODO** verify

We say that $F[x]$ is a **ring**.

Note: Every field is a ring, but not every ring is a field.

As a side note, we can see that $\mathbb{Q}$ is a field, but $\mathbb{Z}$ is not: *it* is a ring.

What we have in mind here, is to study linear operators $T : V \to V$. Polynomials are a tool for understanding these Ts.

Given, say $f(x) = x^2 + 2x - 3$, we can apply f to T, getting $f(T)$

$$f(T) = T^2 + 2T - 3(I)$$

Author Note.

Remind yourself what T^2 means: it's $T \circ T$.

Which is itself a different linear transformation going from V to V.

Given such a T, we are interested in the set J of all polynomials $f \in F[x]$ such that $f(T)$ is the zero transformation.

$$J = \Big\{ \text{polynomials } f \in F[x] \text{ such that } f(T) = \mathbf{0} \Big\}$$

We'll see that, for any T, this is an **ideal** of $F[x]$.

Within such a J, there is a particular polynomial called the **characteristic polynomial**.

We'll see later that, how the characteristic polynomial factors will tell us a lot of information about T.

But what do we mean by factoring?

> **Definition**.
>
> A polynomials $d(x)$ divides $f(x)$ if there is some other $q(x) \in F[x]$ such that
>
> $$f = d \cdot q$$

We'll later try to factor this all the way down, similarly to the unique prime factorization of the natural numbers. We'll see that for a given polynomial p, we'll get a unique factorization.

Danger: For a particular $f(x)$, whether or not f factors may depend on the field F. For example

let $f(x) = x^2 + 1$. In $\mathbb{R}[x]$, f is irreducible. But in $\mathbb{C}[x]$, f does factor!

$$x^2 + 1 = (x - i)(x + i)$$

Notation: The zero polynomial is always a special case. If we're trying to prove something for all polynomials, we tend to handle zero separately.

For $f \in F[x], f \neq 0$, the **degree** of $a_0 + a_1 x + \cdots + a_m x^m$ is the largest n such that $a_i \neq 0$.

We say that f is **monic** if $a_n = 1$, where n is the degree of 2. For example

$$f(x) = x^2 + 2x - 3$$

f is a monic of degree 2, since the leading coefficient is 1.

In the integers, we have the idea of the GCD, the largest number that divides all members of a set. For example

$$\mathrm{GCD}(24, 36) = 12$$

For polynomials, we can do something similar. Given $\{f_1, f_2, f_3\}$, $\mathrm{GCD}(f_1, f_2, f_3)$ is the polynomial h of largest degree such that h divides f_1, f_2, and f_3.

Fri Mar. 17 2023

Looking at 3 on the homework

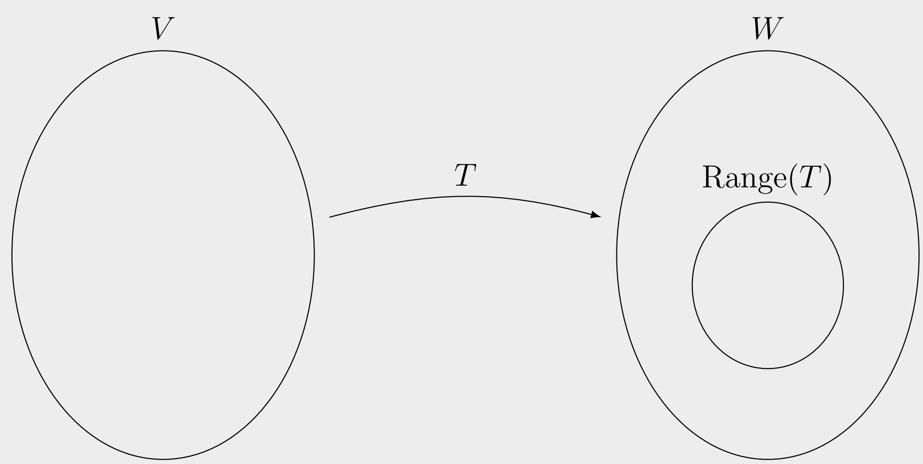
Problem 3

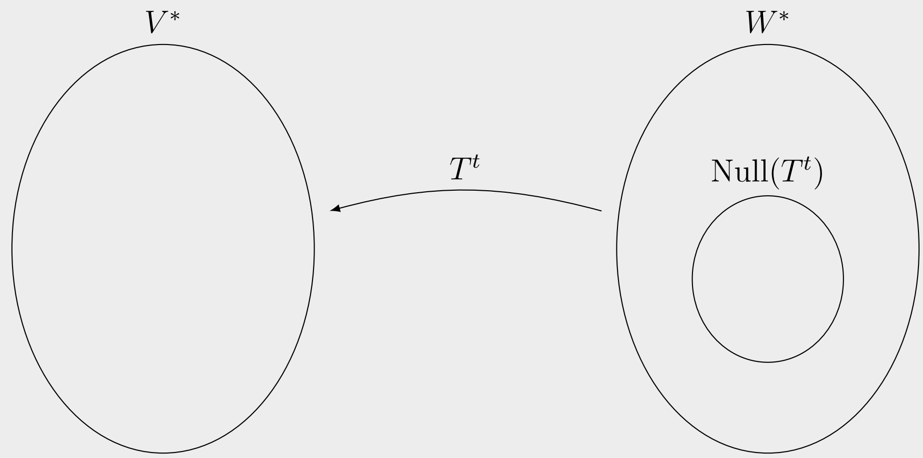
Suppose that V, W are finite dimensional vector spaces and $T : V \to W$ is a linear transformation.

1. Prove that if there is some $f \in W^*$ such that $\mathrm{Null}(T^t) = \mathrm{Span}(\{f\})$, then $\mathrm{Range}(T) = \mathrm{Null}(\{f\})$
2. Prove that if there is some $f \in V^*$ such that $\mathrm{Range}(T^t) = \mathrm{Span}(\{f\})$, then $\mathrm{Null}(T) = \mathrm{Null}(\{f\})$

Solution

Recall our picture





From class, we saw that $\text{Ann}(\text{Range}(T)) = \text{Null}(T^t)$.

For part a, since we have that $\text{Null}(T^t) = \text{Span}(\{f\})$ for some $f \in W^*$. Thus we have that

$$\text{Ann}(\text{Range}(T)) = \text{Span}(\{f\}) \tag{1}$$

Claim: $\text{Range}(T) = \text{Null}(\{f\})$

Proof: Chose any $\beta \in \text{Range}(T)$, by (1), we must have that $f(\beta) = 0$, because f is in the Annihilator of the Range of T, that means that f kills anything in the Range of T, but β is in the range of T, so $f(\beta) = 0$. But this means that

$$\beta \in \text{Null}(\{f\})$$

so we have that $\text{Range}(T) \subseteq \text{Null}(\{f\})$

For the other direction, its easier to just count the dimension.

Say that $f \neq 0$. Since $f : W \to F$ is a linear transformation,

$$\dim(\text{Null}(f)) = m - 1$$

By the Rank Nullity Theorem (since the output is of dimension 1) (but how do we know its not less than $m - 1$, oh because it outputs a single number, the dimensions just work out)

Since $\text{Null}(T^t) = \text{Span}(f)$ by assumption, $\dim(\text{Null}(T^t)) = 1$

By Rank Nullity of T^t,

$$\dim(\text{Range}(T^t)) + \dim(\text{Null}(T^t)) = m$$

So $\dim(\text{Range}(T^t)) = m - 1$ but from the diagram chase from class we saw that $\dim(\text{Range}(T^t)) = \dim(\text{Range}(T))$

3b is similar to this

■

Let's look at other problem

Problem 4.2.3

Let A be an $n \times n$ diagonal matrix over the field F, i.e., a matrix satisfying $A_{ij} = 0$ for $i \neq j$. Let f be the polynomial over F defined by

$$f = (x - A_{11}) \cdots (x - A_{nn})$$

What is the matrix $f(A)$?

> Solution
>
> Let's start by asking: what would $(A - a_{11}(I))$ look like? It would simply be A without the top left entry. What about $(A - a_{22}(I))$? That would be the matrix with an entry in each element of the diagonal, except for the second.
>
> When we multiply all these guys together, we get zeroes along the diagonal, since each matrix removes an element from the diagonal.
>
> All in all, this is then the zero matrix!
>
> ■

Let's look at ideals. We want to factor polynomials, and prove theorems about factorization, so we want to do this carefully.

3.2.1 Euclidean Algorithm

Lemma: For any polynomials $f, d \in F[x]$ with $d \neq 0$ and $\mathsf{deg}(d) \leq \mathsf{deg}(f)$, there is a polynomial $g \in F[x]$ such that

$$\mathsf{deg}(f - gd) < \mathsf{deg}(f) \text{ or } f - gd = 0$$

Proof:

Write $f = a_m x^m + (\text{stuff of degree} < m)$, similarly, write $d = b_n x^n + (\text{stuff of degree} < n)$. By assumption, $n \leq m$.

Let $g(x) = (\frac{a_m}{b_n} x^{m-n})$. Then $f - gd$ has degree $< m$ or is equal to zero.

Note: we use the fact that we're in a field to do this division.

> **Theorem**.
>
> Suppose that f and d are two different polynomials in $F[x]$ and $d \neq 0$. There are $q, r \in F[x]$ such that
>
> $$f = d \cdot q + r$$
>
> Where q is the quotient, and r is the remainder. Additionally, $\mathsf{deg}(r) < \mathsf{deg}(d)$ or $r = 0$.

By induction.

Fix any $d \in F[x]$ with $d \neq 0$. Let $(*)_m$: For every $f \in F[x]$ of degree less than m (or $f = 0$), there are $q, r \in F[x]$ such that $f = d \cdot q + r$ with $\mathsf{deg}(r) < \mathsf{deg}(d)$ or $r = 0$.

So the theorem says that $(*)_m$ holds for all m, we assume that it holds for m and we will prove that $(*)_{m+1}$ holds. (This is just induction)

To do this, chose $f \in F[x]$ with $\mathsf{deg}(f) = m$. Now there are two cases.

1. If $m < \mathsf{deg}(d)$, then we're trivially done! Simply take $f = r$, and $q = 0$, and $\mathsf{deg}(r) < \mathsf{deg}(d)$.
2. If $m \geq \mathsf{deg}(d)$. But then, we can use the Lemma. We apply it to get some $g(x) \in F[x]$ such that $f - dg = 0$ or $\mathsf{deg}(f - dg) < \mathsf{deg}(f) = m$. But now we can apply $(*)_m$ to $f - dg$, since its degree is strictly smaller than m. So we get $q, r \in F[x]$ such that

$$(f - dg) = dq + r$$

   Where $r = 0$ or $\mathsf{deg}(r) < \mathsf{deg}(d)$. Now let $q^* = g + q$, then $f = d \cdot q^* + r$, where $\mathsf{deg}(r) < \mathsf{deg}(d)$ or $r = 0$. So we're done.

Let's look at a fun Corollary from this.

Corollary: For any $f \in F[x]$ and $c \in F$. $(x - c)$ is a factor of f if and only $f(c) = 0$ (if c is a root of f)

Proof: Say $f(x) = (x - c)h(x)$ for some $h(x) \in F[x]$, then just apply c, so we have

$$f(c) = (c - c)h(c) = 0$$

More interestingly is the other direction. Assume that $f(c) = 0$. Apply the theorem to f and $d(x) = (x - c)$, and we get $q, r \in F[x]$ such that

$$f(x) = (x - c)q(x) + r(x)$$

with $\mathsf{deg}(r) < \mathsf{deg}(x - c)$ or $r = 0$. But since $\mathsf{deg}(x - c) = 1$, that tells us that r must be a constant (possibly zero).

Plugging in c, we get

$$f(c) = (c - c)q(c) + r(c) = r(c)$$

But $f(c) = 0$, so $r(c) = 0$, but since r was a constant function, r must have been the zero polynomial this whole time. Rolling back, this tells us that there was no remainder in $f(x)$, so $(x - c)$ is a factor!

$$f(x) = (x - c)q(x)$$

Definition.

An **Ideal** $I \subseteq F[x]$ satisfies

1. $0 \in I$
2. if $f, f' \in I$, then $f + f' \in I$
3. If $f \in I$ and $g \in F[x]$, then $fg \in I$

Ex: $\{0\}$ is an ideal, called the zero ideal.

Another fact, maybe more surprising, is the following. Suppose that I is an ideal, and $1 \in I$, then $I = F[x]$ by (3)

Mon Mar. 27 2023

TODO Talk about the Euclidean Algorithm for $F[x]$.

$F[x]$ is a *commutative ring* if, for any $f, g \in F[x]$

$$f \cdot g = g \cdot f$$

Say $S = \{f_1, \ldots, f_n\} \subseteq F[x]$, the ideal generated by S is

$$I(S) = \{f_1 g_1 + \cdots + f_n g_n : g_1, \ldots, g_n \in F[x]\}$$

For example, $I = \{f \in F[x] : f(2) = 0 \text{ and } f(4) = 0\}$

In a commutative ring, we want to know what the ideals look like. For $F[x]$ as with the integers, they turn out to be incredibly nice. We'll even show that every ideal is generated by a single polynomial

> **Theorem**.
>
> In $F[x]$, every ideal $I \subseteq F[x]$ is generated by a single polynomial $d \in F[x]$.
>
> $I = I(\{d\}) = d \cdot F[x]$ is the ideal generated by d.

We say that $F[x]$ is a "principal ideal domain" **Author Note**.

"Principal" because there is one guy that controls everything. This is also true of the integers.

Chose any ideal I. If $I = \{0\}$, take $d = 0$. If I has a non-zero element, chose $d \in I - \{0\}$ of *least possible degree.*

Claim. $I = d \cdot F[x]$

Proof.

$\supseteq$ is clear. Chose any $g(x) \in F[x]$ Since I is an ideal, and $d \in I$, $d \cdot g \in I$ by definition of ideal.

$\subseteq$. Chose any $f \in I$. Since $d \neq 0$, apply the Euclidean Algorithm to get $f = dq + r$ where $r = 0$ or $\mathsf{deg}(r) < \mathsf{deg}(d)$. But notice that $d \in I$, so $d \cdot q \in I$, $f \in I$, and so $f - d \cdot q \in I$ by definition of I, but that means $r \in I$. But r must be 0, since we chose d to be of smallest degree.

This all means that f can be written as $d \cdot q$.

Aside: Why do we care about all this? I thought I signed up for 405, not 403

Ideals show up in Linear Algebra.

Let A be any $n \times n$ matrix. Let $I = \{f \in F[x] : f(A) = 0\}$. This is an ideal which depends on a matrix. Now we know that this ideal I is generated by one single polynomial.

Author Note.

End of Aside

To find d for a particular ideal I, we need to find the **gcd**: the Greatest Common Divisor. How do we find a gcd?

In $\mathbb{Z}$, what we did was use the Euclidean Algorithm!

> **Definition.**
>
> A non-constant $f \in F[x]$ is **irreducible** if $f \neq g \cdot h$ where $\texttt{deg}(g), \texttt{deg}(h) \geq 1$.
>
> These are the "primes" of $F[x]$.

As in the integers, every non-constant $f \in F[x]$ can be written as $f = c \cdot p_1^{k_1} \cdot p_2^{k_2} \cdots p_l^{k_l}$, where c is a constant, and $p_1, \ldots, p_l$ are *monic, irreducible* polynomials.

> **Example.**
>
> Let $f(x) = 3(x^2+1)^2(x-3)^4(x+2)$ and $g(x) = -4(x^2+1)(x-3)^2(x+2)^5$ in $\mathbb{R}[x]$.
>
> The gcd of f and g is now trivial!
>
> $$\gcd(f, g) = (x^2+1)(x-3)^2(x+2)$$
>
> In fact, this gcd is our polynomial d. This is always the case.

> **Definition.**
>
> f, g are **relatively prime** if
>
> $$\gcd(f, g) = 1$$
>
> In other words, they are relatively prime if they share not factors.

Which polynomials are irreducible depends on F. For example

x^2+1 is irreducible in $\mathbb{Q}[x]$, $\mathbb{R}[x]$, but not $\mathbb{C}[x]$. If you allow complex numbers, we can factor it into

$$(x^2+1) = (x-i)(x+i)$$

Similarly, (x^2-2) is irreducible in $\mathbb{Q}[x]$ but not $\mathbb{R}[x]$.

$$(x^2-2) = (x-\sqrt{2})(x+\sqrt{2})$$

In $\mathbb{C}[x]$, the only monic, irreducible polynomials are of the form $x-a$.

Author Note.

This is it. This is all that we needed from chapter 4. We'll go back to linear algebra land now.

Wed Mar. 28 2023

4 The Determinant Function

TODO Write notes for that day

(My computer broke so I couldn't write them on the day, sorry about that)

Fri Mar. 31 2023

> **Note.**
>
> We're interested in the determinant function! On Wednesday we showed the existence of the function, and today we're going to show the uniqueness.

Fix a commutative ring K

For each $n \geq 1$, a **determinant function**a $D : (K^n)^n \to K$ satisfies three things

1. D is n **linear**
2. D is **alternating**
3. $D(I_{n\times n}) = 1$

What we're aiming for is: for every n, there is **exactly one** determinant function.

For the moment, just trust that this is important! We'll see later exactly what we can do with the determinant and it's properties.

4.1 Existence

Last time, we showed this for $n = 1$.

$$D(a) = a \text{ for all } a \in K$$

is a determinant function.

For $n = 2$, we had

$$D\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = ad - bc$$

But we want this for an arbitrary n, how do we do this?

We do it by induction.

4.2 General Method

Let's look at how we can get from an $(n-1) \times (n-1)$ determinant function D, to a $n \times n$ determinant function E

This is sometimes called "expansion by minors".

For any $1 \leq i, j \leq n$ for any $n \times n$ matrix A.

$A[i|j]$ is the $(n-1) \times (n-1)$ matrix formed by *deleting* the i^{th} row and j^{th} column.

TODO draw examples with barred matrices

We also define $D_{ij}(A) = D(A[i|j])$, where A is an $n \times n$ matrix.

Definition.

For any $1 \leq j \leq n$, let $E_j : (K^n)^n \to K$ defined as

$$E_j(A_{n\times n}) = \sum_{i=1}^{n} \underbrace{(-1)^{i+j}}_{\in K} \underbrace{A_{ij}}_{\in K} \underbrace{D_{ij}(A)}_{\in K}$$

Of course, we already know how to take an $n \times n$ determinant of a matrix, so this shouldn't feel completely arbitrary to the reader.

Theorem.

For every $1 \leq j \leq n$, E_j is a determinant function on $n \times n$ matrices.

TODO He didn't write the proof for this in class but the book probably has it in 5.2

Example.

Let

$$A = \begin{bmatrix} (x-2) & x^2 & x^3 \\ 0 & (x-2) & x^2+5 \\ 0 & 0 & x+6 \end{bmatrix}$$

Compute $E_i(A) = \sum_{i=1}^{n} (-1)^{i+1} A_{i1} D_{i1}(A)$

$i = 1$

$A[1|1] = \begin{bmatrix} x-4 & x^2+5 \\ 0 & x+6 \end{bmatrix}$

So $D_{1,1}(A) = (x-4)(x+6) - (x^2+5) \cdot 0 = (x-4)(x+6)$

So $(-1)^{1+1} A_{1,1} D_{1,1}(A) = 1(x-2)(x-4)(x+6)$

$i = 2$

$A[2|1] = \begin{bmatrix} x^2 & x^3 \\ 0 & x+6 \end{bmatrix}$

$D_{2,1}(A) = x^2(x+6) - x^3 \cdot 0$

$(-1)^{2+1} \cdot A_{2,1} D_{2,1}(A) = (-1)^3 \cdot 0(x^2)(x+6) = 0$

$i = 3$

$A[3|1] = \begin{bmatrix} x^2 & x^3 \\ x-4 & x^2+5 \end{bmatrix}$

TODO $(-1)^{3+1} \cdot A_{3,1} D_{3,1}(A) = (-1)^4 \cdot 0 \cdot * = 0$

So $E_1(A) = (x-2)(x-4)(x+6) + 0 + 0 = (x-2)(x-4)(x+6)$

Note.

When we write 0, it's important to keep in mind where it comes from. The answer is that it comes from K, moreover, it's a polynomial in this case.

TODO Ok to me I dont see why this process creates a determinant function, it's what we did in 240, but why does THIS process create a function with the 3 properties above? unless I'm missing something, THAT wasn't explained.

The answer is Theorem 1 in 5.2, it will not be proved in class because it's pretty involved.

> **Theorem.**
>
> We've now proved that for every $n \geq 1$, a determinant function $D : (K^n)^n \to K$ exists.

The proof is by induction on n.

4.3 Uniqueness

> **Theorem.**
>
> For any $n \geq 1$, there is **exactly one** determinant function $D : (K^n)^n \to K$.

Fix $n \geq 1$. Suppose $A = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}$ is the $n \times n$ matrix with rows $\alpha_1, \ldots, \alpha_n$.

Let $D(A) = D(\alpha_1, \ldots, \alpha_n)$. Use the fact that D is linear with $i = 1$. (after all, it's linear in all terms since it's n linear)

We know that $\alpha_1 = (A_{1,1}, A_{1,2}, \ldots, A_{1,n})$. We can also write it as $\alpha_1 = (A_{1,1}\varepsilon_1, A_{1,2}\varepsilon_2, \ldots, A_{1,n}\varepsilon_n)$. Recall that $\varepsilon_1 = (1, 0, \ldots, 0), \varepsilon_2 = (0, 1, 0, \ldots, 0), \ldots$ **Author Note.**

ε_i is the ith row of the identity matrix.

Then we have

$$
\begin{aligned}
&D(\alpha_1, \alpha_2, \ldots, \alpha_n) && (2)\\
=&D((A_{1,1}\varepsilon_1 + \cdots + A_{1,n}\varepsilon_n), \alpha_2, \ldots, \alpha_n) && (3)\\
=&A_{1,1}D(\varepsilon_1, \alpha_2, \ldots, \alpha_n) + A_{1,2}D(\varepsilon_2, \alpha_2, \ldots, \alpha_n) + \cdots + A_{1,n}D(\varepsilon_n, \alpha_2, \ldots, \alpha_n) && (4)\\
=&\sum_{j=1}^{n} A_{1,j}D(\varepsilon_j, \alpha_2, \ldots, \alpha_n) && (5)
\end{aligned}
$$

For (4), recall that $A_{i,j}$ is just a number in K, so we can factor it from D by n linearity.

Let's look at $D(\varepsilon_j, \alpha_2, \ldots, \alpha_n)$ in more detail.

$$
\begin{aligned}
&D(\varepsilon_j, \alpha_2, \ldots, \alpha_n) && (6)\\
=&D(\varepsilon_j, (A_{2,1}\varepsilon_1 + \cdots + A_{2,n}\varepsilon_n), \alpha_3, \ldots, \alpha_n) && (7)\\
=&A_{2,1}D(\varepsilon_j, \varepsilon_1, \alpha_2, \ldots, \alpha_n) + A_{2,2}D(\varepsilon_j, \varepsilon_2, \alpha_3, \ldots, \alpha_n) + \cdots + A_{2,n}D(\varepsilon_j, \varepsilon_n, \alpha_3, \ldots, \alpha_n) && (8)\\
=&\sum_{k=1}^{n} A_{2,k}D(\varepsilon_j, \varepsilon_k, \alpha_3, \ldots, \alpha_n) && (9)
\end{aligned}
$$

We can keep exploding these terms on and on. Since D is n linear, we end up with this

$$
\begin{aligned}
D(A) =& D(\alpha_1, \dots, \alpha_n) \\
=& \sum_{(k_1,\dots,k_n)} A_{1,k_1} A_{2,k_2} \cdots A_{n,k_n} D(\varepsilon_{k_1}, \varepsilon_{k_2}, \dots, \varepsilon_{k_2}) \\
=& \sum_{k_1=1}^{n} \sum_{k_2=1}^{n} \cdots \sum_{k_n=1}^{n} A_{1,k_1} A_{2,k_2} \cdots A_{n,k_n} D(\varepsilon_{k_1}, \varepsilon_{k_2}, \dots, \varepsilon_{k_2})
\end{aligned}
$$

This is a lot of products.

For $n = 3$: Sum of $3^3 = 27$ products.

For $n = 10$: Sum of $10^{10} = 10$ billion products.

I should get paid for texing this.

Mon Apr. 3 2023

We're in the midst of determining that there is a unique determinant for any $n \geq 1$. Today, we will finish proving uniqueness.

Let's recall what this means.

K is any commutative ring, $n \geq 1$. For any $n \times n$ matrix A with **rows** $\alpha_1, \dots, \alpha_n$, we can either write

- $D : K^{n \times n} \to K$ for which we have $D(A)$
- $D : (K^n)^n \to K$ for which we have $D(\alpha_1, \dots, \alpha_n)$

These mean the same thing, and they are just different ways of thinking about what we're doing.

We also had the following properties for the determinant

1. D is n **linear**
2. D is **alternating**

   With this property, we ensure that if two rows are the same, the determinant is zero.
3. **"normalization"**, $D(I) = D(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n) = 1$

We want to show that **there is only one** function which satisfies these three properties.

Last time, we saw that if D is n linear, we had this property

$$
\begin{aligned}
D(A) = D(\alpha_1, \dots, \alpha_n) &= \sum_{(k_1,\dots,k_n)} A_{1,k_1} A_{2,k_2} \cdots A_{n,k_n} D(\varepsilon_{k_1}, \varepsilon_{k_2}, \dots, \varepsilon_{k_2}) \\
&= \sum_{k_1=1}^{n} \sum_{k_2=1}^{n} \cdots \sum_{k_n=1}^{n} A_{1,k_1} A_{2,k_2} \cdots A_{n,k_n} D(\varepsilon_{k_1}, \varepsilon_{k_2}, \dots, \varepsilon_{k_2})
\end{aligned}
$$

In other words, this is the sum of n^n products.

The question now is, what happens to $D(\varepsilon_{k_1}, \varepsilon_{k_2}, \dots, \varepsilon_{k_n})$?

Let's assume that D is n linear and alternating. Then, by definition, we have that $D(\varepsilon_{k_1}, \varepsilon_{k_2}, \dots, \varepsilon_{k_n}) = 0$ **unless** $k_1, \dots, k_n$ are all distinct!

Definition.

A **permutation** of $\{1, \dots, n\}$ is a one to one (therefore onto) function $\sigma : \{1, \dots, n\} \to \{1, \dots, n\}$.

Author Note.

This function just shuffles things around

We know that it's onto simply by pigeonhole principle! (Also this isn't exactly true for infinite sets)

TODO clean up and explain examples.

Example.

$\sigma(1) = 2$

$\sigma(2) = 4$

$\sigma(3) = 5$

$\sigma(4) = 1$

$\sigma(5) = 3$

Let S_n be the set of all permutations $\sigma : \{1, \dots, n\} \to \{1, \dots, n\}$.

Question: How big in S_n? Well,

- How many places can you send 1 to? n places of course
- What about 2? Well you fixed a place for 1, so $n - 1$ places are left for 2.
- For 3, we have $n - 2$ places left, and so on...

So S_n is $n!$ big.

If D is both n linear and alternating, then

$$D(\alpha_1, \dots, \alpha_n) = \sum_{\sigma \in S_n} A_{1,\sigma(1)} A_{2,\sigma(2)} \cdots A_{n,\sigma(n)} D(\varepsilon_{\sigma(1)}, \varepsilon_{\sigma(2)}, \dots, \sigma_{\sigma(n)})$$

TODO explain notation here

Example.

$\tau(1) = 1$

$\tau(2) = 2$

$\tau(3) = 5$

$\tau(4) = 4$

$\tau(5) = 3$

Definition.

A **transposition** $\tau_{i,j}$ is the permutation swapping only i and j.

Given any 2 permutations σ and τ, we can "compose them".

$$\sigma \circ \tau \in S_n$$

Example.

Given

$\sigma(1) = 2$

$\sigma(2) = 4$

$\sigma(3) = 5$

$\sigma(4) = 1$

$\sigma(5) = 3$

and

$\tau(1) = 1$

$\tau(2) = 2$

$\tau(3) = 5$

$\tau(4) = 4$

$\tau(5) = 3$

$\sigma \circ \tau(1) = 2$

$\sigma\tau(2) = 4$

$\sigma\tau(3) = 3$

$\sigma\tau(4) = 1$

$\sigma\tau(5) = 5$

Every $\sigma \in S_n$ has an "inverse".

Example.

Following from the examples from earlier, we have

$\sigma^{-1}(1) = 4$

$\sigma^{-1}(2) = 1$

$\sigma^{-1}(3) = 5$

$\sigma^{-1}(4) = 2$

$\sigma^{-1}(5) = 3$

Note.

If τ is a transposition, $\tau \circ \tau$ is the same as doing nothing!

$$\tau \circ \tau = I$$

Transpositions are called "idempotent" from this property.

Note.

As a sidenote, S_n is a **group** with operation of composition and identity element being the "do nothing" permutation.

Fact: In S_n, every $\sigma \in S_n$ is a composition of *transpositions*.

We might see $\sigma = \tau_k \circ \tau_{k-1}, \circ \cdots \circ \tau_1$ where each τ_j is a transposition.

Proof: Fix any permutation $\sigma \in S_n$. Choose any $a \in \{1, \ldots, n\}$. Now keep applying σ to a.

$$a, \sigma(a), \sigma(\sigma(a)), \ldots, \sigma^k(a) = a$$

Eventually, we'll get back to a, since there's only finitely many ways to permute finitely many things.

The cycle $C = \{a, \sigma(a), \ldots, \sigma^{k-1}(a)\}$ is a subset of $\{1, \ldots, n\}$.

Choose any $b \in \{1, \ldots, n\} \setminus C$. Then

$$C_b = \{b, \sigma(b), \ldots, \sigma^{k'-1}(b)\}$$

So $\{1, \ldots, n\}$ is partitioned into **cycles**.

$$\sigma = \tau_{3,5} \circ \tau_{4,1} \circ \tau_{1,2}(1)$$

What we get is that each cycle is a product of transpositions.

Even though the permutations might not be the shortest, their parity will always be right.

Definition.

A permutation σ is called **even** if σ is the product of of an **even number** of transpositions.

σ is odd otherwise.

Note.

The identity is even. Any transposition is τ is odd.

In our example, σ is odd.

$$\sigma = \tau_{3,5} \circ \tau_{4,1} \circ \tau_{1,2}(1)$$

The number of even transpositions is $n!/2$ and so is the number of odd transpositions.

Now if σ is even, we have

$$1 \cdot D(\varepsilon_{\sigma(1)}, \varepsilon_{\sigma(2)}, \ldots, \sigma_{\sigma(n)})$$

Author Note.

We can get from σ to the identity in an even number of flips, but every time we flip, we switch the sign.

But if σ is odd, we have

$$-1 \cdot D(\varepsilon_{\sigma(1)}, \varepsilon_{\sigma(2)}, \ldots, \sigma_{\sigma(n)})$$

Author Note.

We can get from σ to the identity in an odd number of flips, but every time we flip, we switch the sign.

> **Definition**.
>
> We define $\operatorname{sgn}(\sigma)$ as a function $\operatorname{sgn} : S_n \to \pm 1$
>
> $$\operatorname{sgn}(\sigma) = \begin{cases} 1 & \text{if } \sigma \text{ is even} \\ -1 & \text{if } \sigma \text{ is odd} \end{cases}$$
>
> Which represents the parity of permutation σ.

Finally,

If D is n linear and alternating, then

$$D(\alpha_1, \ldots, \alpha_n) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) A_{1,\sigma(1)} \cdots A_{n,\sigma(n)} \cdot D(\varepsilon_1, \ldots, \varepsilon_n)$$

Corollary:

There is *exactly one* determinant function

$$\det(A) = \det(\alpha_1, \ldots, \alpha_n) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) A_{1,\sigma(1)} A_{2,\sigma(2)} \cdots A_{n,\sigma(n)}$$

Since $D(\varepsilon_1, \ldots, \varepsilon_n) = 1$ by "normalization".

Wed. Apr 5 2023

Last Time: For any n linear, alternating D, we had

$$D(\alpha_1, \ldots, \alpha_n) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) A_{1,\sigma(1)} \cdots A_{n,\sigma(n)} D(I)$$

And if we insist that $D(I) = 1$, it shrinks down.

We define

$$\det(\alpha_1, \ldots, \alpha_n) = D(\alpha_1, \ldots, \alpha_n) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) A_{1,\sigma(1)} \cdots A_{n,\sigma(n)} D(I)$$

so $\det(I) = 1$

Author Note.

This is not even by definition, we just see it as a byproduct of the definition.

4.4 Properties of the Determinant

1. $\det(AB) = \det(A) \cdot \det(B)$

   Proof

   Fix an $n \times n$ matrix B. Say that $A = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}$ is any $n \times n$ matrix.

   Check

$$AB = \begin{pmatrix} \alpha_1 \cdot B \\ \vdots \\ \alpha_n \cdot B \end{pmatrix}$$

   Author Note.

   These are the rows of AB

   Now, we define $D(A) = \det(AB) = \begin{pmatrix} \alpha_1 \cdot B \\ \vdots \\ \alpha_n \cdot B \end{pmatrix}$.

   Check: D is n linear and alternating.

   Therefore, by what we had last time, we have that

$$D(A) = \underbrace{\Big(\sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) A_{1,\sigma(1)} \cdots A_{n,\sigma(n)} \Big)}_{\det(A)} D(I)$$

   But now notice

$$D(I) = \det(I \cdot B) = \det(B)$$

   So we have

$$D(A) = \Big(\sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) A_{1,\sigma(1)} \cdots A_{n,\sigma(n)} \Big) D(I) = \det(A) \cdot \det(B)$$

2. $\det(A^t) = \det(A)$

   Recall A^t is defined as $(A^t)_{i,j} := A_{j,i}$. Again, by definition of the determinant, we have

$$\det(A^t) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) (A^t)_{1,\sigma(1)} \cdots (A^t)_{n,\sigma(n)}$$

   Now fix any $\sigma \in S_n$, we have

$$(A^t)_{i,\sigma(i)} := A_{\sigma(i),i}$$

Then we have

$$\det(A^t) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) A_{\sigma(1),1} \cdots A_{\sigma(n),n}$$

But remember that every permutation has an inverse! If we have

$$i = \sigma^{-1}(j)$$

Then,

$$A_{\sigma(i),j} = A_{i,\sigma^{-1}(j)}$$

Additionally, the sign doesn't change, since

$$\sigma = \tau_1 \tau_2 \cdots \tau_k$$

$$\sigma^{-1} = \tau_k \tau_{k-1} \cdots \tau_1$$

Author Note.

We know these are inverses, because look what happens when we compose them.

$$\sigma\sigma^{-1} = \tau_1\tau_2\tau_3 \cdots (\tau_k\tau_k) \cdots \tau_3\tau_2\tau_1 = \mathrm{id}$$

So we have

$$\operatorname{sgn}(\sigma^{-1}) = \operatorname{sgn}(\sigma)$$

So finally, we have

$$\begin{aligned}\det(A^t) &= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma^{-1}) A_{1,\sigma^{-1}(1)} \cdots A_{n,\sigma^{-1}(n)} \\ &= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) A_{\sigma(1),1} \cdots A_{\sigma(n),n} = \det(A)\end{aligned}$$

3. $\det(cA) = c^n \det(A)$
4. If A is upper or lower triangular, then

$$\det(A) = \prod_{i=1}^{n} A_{i,i}$$

We can expand by minors for this

TODO Write proof

To show this for the lower triangular case, simply notice that the transpose of an upper triangular matrix is lower triangular, and then use (2)

5. If A has a zero row, or zero column, then $\det(A) = 0$

   The proof of this is that every single one of the products in

$$D(A) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) A_{1,\sigma(1)} \cdots A_{n,\sigma(n)} \Big) D(I)$$

   Is going to be zero. So if a row is all zeroes, the determinant is zero.

   To show this for the columns, we use the transpose fact again.

6. Let $r + s = n$. Suppose that we can decompose A as follows

   Let $A = \begin{bmatrix} B & C \\ 0 & D \end{bmatrix}$

   Where

   - $B : r \times r$
   - $D : s \times s$
   - 0 is the $s \times r$ zero matrix
   - $C : r \times s$

   Then $\det(A) = \det(B)\det(D)$. You can see this if you once again expand by minors.

7. $A : n \times n$. Let's look at elementary row operations.

$$A \to A^{\mathrm{rr}}$$

   Question: What do these row operations do to the determinant?

   - **Swap two rows**: $\det(A)$ switches signs
   - **Multiply α_i by $c \neq 0$**: $\det(A)$ gets multiplied by c.
   - **Replace α_j by $c\alpha_i + \alpha_j$**: You can check that $det(A)$ is unchanged

   So

$$\det(A) = 0 \Leftrightarrow \det(A^{\mathrm{rr}}) = 0$$

Fri. Apr 7 2023

Author Note.

Exam 2 is next Friday. We'll be covering new material today though.

Today, we're going to use our results about polynomials and determinants to better understand linear operators (remember, these are linear transformations going from V to V) over a field F.

$$T : V \to V$$

It's been a few weeks since we've been in the world of linear operators, and it's time we jumped back in.

In 6.2, 6.3, we'll define two polynomials that help describe the behavior of T. One is the *characteristic polynomial*, the other is the *minimal polynomial*.

Recall: For any linear operator $T : V \to V$, and for any basis $\mathcal{B}$ of V, $[T]_{\mathcal{B}}$ is the matrix of T with respect to $\mathcal{B}$. Additionally, for any $\alpha \in V$,

$$[T]_{\mathcal{B}}[\alpha]_{\mathcal{B}} = [T(\alpha)]_{\mathcal{B}}$$

As we vary $\mathcal{B}$, we get *different* matrices $[T]_{\mathcal{B}'}$, where $\mathcal{B}'$ is another basis for V.

Additionally, recall the definition of *similar matrices*: two matrices $A, B : n \times n$ are similar if they represent the same linear transformation (possibly under different bases.)

$$B = P^{-1}AP$$

for some invertible matrix P.

Fact: If A, B are similar $n \times n$ matrices, then

$$\det(A) = \det(B)$$

Proof: Say $B = P^{-1}AP$, then

$$\det(B) = \det(P^{-1}AP) = \det(P^{-1})\det(A)\det(P) = \det(P^{-1}P)\det(A) = 1 \cdot \det(A)$$

> **Definition**.
>
> For a linear operator $T : V \to V$, *define*
>
> $$\det(T) = \det(A)$$
>
> for any/some $n \times n$ matrix A representing T for some basis $\mathcal{B}$.
>
> $$A = [T]_{\mathcal{B}}$$

Let A be any $n \times n$ matrix over a field F.

Then $\det(A) \neq 0$ if and only if A is invertible.

Proof: Row reduce $A \to A^{\text{rr}}$. From last time, we saw that $\det(A)$ might not equal $\det(A^{\text{rr}})$, but

$$\det(A) \neq 0 \Leftrightarrow \det(A^{\text{rr}}) \neq 0$$

If A is invertible, then A^{rr} has no completely zero rows (or columns). Moreover, all the rows will contain ones, so $\det(A^{\text{rr}}) = 1$. So again, $\det(A) \neq 0$.

Conversely, if A is not invertible, then there *is* a fully zero row, so $\det(A^{\text{rr}}) = 0$, so $\det(A) = 0$.

Thus: $T : V \to V$ is invertible (the book calls this "non-singular") if and only if $\det(T) \neq 0$.

This is great, but there is more.

> **Definition**.
>
> A linear operator $T : V \to V$ is **diagonalizable** if there is some basis $\mathcal{B}$ of V such that $[T]_{\mathcal{B}}$ is **diagonal**.
>
> i.e. $A_{i,j} = 0$ whenever $i \neq j$.

Similarly

> **Definition**.
>
> An $n \times n$ matrix A is **diagonalizable** if it is **similar** to a diagonal matrix.
>
> i.e. if and only if T_A is diagonalizable.

Suppose that T is diagonalizable and $\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$ is a basis such that

TODO

T is invertible if and only if **none** of $c_1, \ldots, c_n$ are 0. More generally,

- $\text{rank}(T)$ is the number of non-zero c_is.
- $\text{nullity}(T)$ is the number of zero c_is.

Let's look at some examples.

> **Example**.
>
> Is
>
> $$\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$$
>
> diagonalizable?
>
> **Answer**: Yes
>
> We'll also see that
>
> $$\begin{bmatrix} 1 & 4 \\ 0 & 1 \end{bmatrix}$$
>
> is **not** diagonalizable.

Take A to be the matrix

$$A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

Then there is **no** matrix of the form

$$\begin{bmatrix} c_1 & 0 \\ 0 & c_2 \end{bmatrix}$$

With $c_1, c_2 \in \mathbb{R}$ similar to A, but in $\mathbb{C}$, there is!

What this tells us is that, whether a matrix is diagonalizable might depend on the field that we're using.

5 Eigenvalues

Definition.

An element $c \in F$ is an **eigenvalue** of T if there is some non-zero $\alpha \in V$ such that

$$T(\alpha) = c\alpha$$

In such a case, α is called the "eigenvector".

Note.

For any $T : V \to V$, and $c \in F$, let $W_c = \{\alpha \in V : T(\alpha) = c\alpha\}$.

Claim: W_c is always non-empty (after all, the zero vector is in it)

$$T(0) = c \cdot 0$$

More interestingly though, W_c is a **subspace** of V.

$$= \{\alpha \in V : (cI - T)(\alpha) = 0\} = \text{Null}(cI - T)$$

QUESTION woah woah woah what's happening here?

Note.

For most $c \in F$, $W_c = \{0\}$.

So $c \in F$ is an eigenvalue of T if and only if $\text{Null}(cI - T) \neq \{0\}$ if and only if $(cI - T)$ is **not invertible** if and only if $\det(cI - T) = 0$.

5.1 Finding Eigenvalues

To find the eigenvalues of T, let A be a matrix of T with respect to $\mathcal{B}$. The **characteristic polynomial** of A is

$$\det(xI - A)$$

Lemma.

In fact, if A and B are similar, they have the same characteristic polynomial.

Proof.

Say $B = P^{-1}AP$. Then

$$\begin{aligned} & xI - B \\ =& xI - (P^{-1}AP) \\ =& x(P^{-1}P) - P^{-1}AP \\ =& P^{-1}(xI)P - P^{-1}AP \\ =& P^{-1}(xI - A)P \end{aligned}$$

■

Definition.

The characteristic polynomial of T is

$$\det(xI - A)$$

for some/every matrix A representing T.

TODO write down examples of characteristic polynomial for the two matrices from earlier.

Mon. Apr 10 2023

Author Note.

Exam is Friday. Everything in class is fair game to be covered.

5.4, 5.5, not the adjoint of A, or Cramer's Rule.

Material for today is also fair game.

We study linear operators $T : V \to V$ where V is a vector space where $\dim(V) = n$.

Last time, we learned about the **characteristic polynomial**, defined as

$$f(x) = \det(xI - A)$$

Where A is any $n \times n$ matrix representing T. We also saw that f does not depend on our choice of A. If B is similar to A, $\det(xI - A) = \det(xI - B)$.

$$xI - A = \begin{bmatrix} x & 0 & \cdots & 0 & 0 \\ 0 & x & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & x & 0 \\ 0 & 0 & \cdots & 0 & x \end{bmatrix} - \begin{bmatrix} A_{1,1} & 0 & \cdots & 0 & 0 \\ 0 & A_{2,2} & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & A_{n-1,n-1} & 0 \\ 0 & 0 & \cdots & 0 & A_{n,n} \end{bmatrix}$$
$$= \begin{bmatrix} x - A_{1,1} & 0 & \cdots & 0 & 0 \\ 0 & x - A_{2,2} & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & x - A_{n-1,n-1} & 0 \\ 0 & 0 & \cdots & 0 & x - A_{n,n} \end{bmatrix}$$

> **Note**.
>
> $\det(xI - A)$ is a **monic** of degree n

QUESTION Why is it monic?

$$\det(xI - A) = \sum_{\sigma \in S_n}$$

Sometimes, T, might not have eigenvalues. Consider for example

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

which doesn't have eigen values over $\mathbb{R}$. (but it does over $\mathbb{C}$.)

Proposition

Any $T : V \to V$ where V is a vector space over $\mathbb{C}$ has **at least one** eigenvalue.

> Proof
>
> Let $f(x)$ be the characteristic polynomial. By the Fundamental Theorem of Algebra, f has a root c. Therefore c is an eigenvalue of T.
>
> ■

For each eigenvalue $c \in F$ of $T : V \to V$, the corresponding **Eigenspace** $W_c = \{\alpha \in V : T(\alpha) = c\alpha\} = \text{Null}(cI - T)$ is a subspace of V.

Say $c_1 \neq c_2$, then $W_{c_1} \cap W_{c_2} = \{0\}$ Since $T(0) = 0 = c_1 \cdot 0 = c_2 \cdot 0$.

Then $T(\alpha) = c_1\alpha$, and $T(\alpha) = c_2\alpha$, we have that $(c_1 - c_2)\alpha = 0$. Since $c_1 \neq c_2$, α must be 0.

W_{c_i} only share $\{0\}$.

Moreover, we have the following theorem.

Theorem.

If $c_1, \ldots, c_k$ are distinct eigenvalues of some transformation $T : V \to V$. Then for any $\beta_i \in W_{c_i} \setminus \{0\}$ for $1 \leq i \leq k$. We have that

$$\{\beta_1, \ldots, \beta_k\}$$

are linearly independent in V.

Proof.

By induction on k.

Base Case. $k = 1$. So $\beta_1 \in W_{c_1}$. We obviously see that $\{\beta_1\}$ is linearly independent.
Inductive Hypothesis.
Assume that the result holds for k.
Inductive Step.
Chose $\beta \in W_{c_1} \setminus \{0\}, \ldots, \beta_{k+1} \in W_{c_{k+1}} \setminus \{0\}$.

We must show that $\{\beta_1, \ldots, \beta_{k+1}\}$ is linearly independent.

Chose $a_1, \ldots, a_{k+1} \in F$, such that

$$a_1\beta_1 + \cdots a_{k+1}\beta_{k+1} = 0 \qquad (*)$$

1. Multiply (*) by c_{k+1}

$$c_{k+1}a_1\beta_1 + \cdots c_{k+1}a_{k+1}\beta_{k+1} = 0$$

2. Apply T to (*)

$$T(a_1\beta_1 + \cdots + a_{k+1}\beta_{k+1}) = T(0) = 0 = \quad T(a_1\beta_1) + \cdots + T(a_{k+1}\beta_{k+1}) = c_1\beta_1 + \cdots + c_{k+1}\beta_{k+1}$$

Subtract (2) from (1).

$$(c_{k+1} - c_1)\alpha_1\beta_1 + \cdots (c_{k+1} - c_k)\alpha_k\beta_k = 0$$

By the Inductive Hypothesis, $\{\beta_1, \ldots, \beta_k\}$ are linearly independent.

■

Given $T : V \to V$, is T diagonalizable? In other words, is there a basis $\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$ consisting of eigenvectors?

$$[T]_{\mathcal{B}} = \begin{bmatrix} c_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & c_n \end{bmatrix}$$

where $\alpha_i \in W_{c_i}$, in other words $T(\alpha_i) = c_i\alpha_i$ for all i.

Corollary

If T has n distinct eigenvalues, then T is diagonalizable. Since $\{\alpha_1, \ldots, \alpha_n\}$ are linearly independent by the theorem, they form a basis.

Proof

No

■

The characteristic polynomial doesn't tell you the dimension of the subspace

Let $T : V \to V$ be any linear operator. Let I be the collection of all $f \in F[x]$ such that $f(T) = 0$. Then I is an ideal.

$L(V, V)$ is a vector space of dim n^2, so $\{1, T, T^2, \ldots, T^{n^2}\}$ are linearly dependent.

Say $c_0 + c_1 T + c_2 T^2 + \cdots + c_{n^2} T^{n^2} = 0$ with not all 0.

Let $g(x) = c_0 + c_1 x + c_2 x^2 + \cdots + c_{n^2} x^{n^2}$, then $g(T) = 0$.

Definition.

The **minimum polynomial** of T $m(x)$ is the non-zero element of I of *least degree*.

By Chapt. 4

$$I = m(x)F[x]$$

■

Fact: Cayley-Hamilton Theorem. The min polynomial divides the characteristic polynomial.

For instance if f is the characteristic polynomial, then

$$f(T) = 0$$

For the matrix $\begin{bmatrix} 1 & 4 \\ 0 & 1 \end{bmatrix}$ is $m(x) = (x-1)^2$. For the identity, its $(x-1)$.

Mon. Apr 17 2023

Suppose that $T : V \to V$ is a linear operator, where V is a vector space over F of dimension n.

Goal: Decompose T and V into smaller understandable bits. To do this, the tools we have are the characteristic polynomial, and the minimal polynomial.

We're going to need some notions.

1. T-invariant subspaces. This is section 6.4 f the book, and what we're going to talk about today.
2. Direct sum decompositions of V. (Chapter 6.6)
3. Direct sum decompositions into T invariant subspaces. (Chapter 6.7)
4. The primary decomposition theorem. (Chapter 6.8)

Let's do a simple example.

Example.

Say $\{c_1, \ldots, c_k\}$ are the eigenvalues of T. Then for each eigenvalue c_i, we have a subspace W_{c_i} such that

$$\begin{aligned} W_{c_i} =& \{\alpha \in V \mid T(\alpha) = c_i\alpha\} \\ =& \{\alpha \in V \mid (T - c_i I)(\alpha) = 0\} \end{aligned}$$

If we let $W = W_{c_1} + W_{c_2} + \cdots + W_{c_k}$, we'll see that W is a subspace of V, and more.

The minimal polynomial of T $m(x)$ will really help us here.

Note.

Recall what the minimal polynomial $m(x)$ is: it's the least polynomial that annihilates the matrix associated with T.

Check: If A is any matrix representing T, then $m(T) = 0$ if and only if $m(A) = 0$.

Let's start easy.

Definition.

A subspace $W \subseteq V$ is T**-invariant** if

$$T(\alpha) \in W \text{ for all } \alpha \in W$$

■

If W is T-invariant, we get a **restriction operator** $T_W : W \to W$ such that

$$T_W(\alpha) = T(\alpha)$$

Two boring examples.

Example.

- V is always T-invariant
- $\{0\}$ is always T-invariant

Let's look at more interesting examples.

Example.

Say c is an eigenvalue of T. Then

$$W_c = \{\alpha \in V \mid T(\alpha) = c\alpha\}$$

is T-invariant.

Proof.

Take any $\alpha \in W_c$. Then we know that $T(\alpha) = c\alpha$ but now, is $c\alpha \in W_c$? Well

$$T(c\alpha) = cT(\alpha) = c(c\alpha)$$

Which is in W_c.

Therefore W_c is T-invariant.

Let's look at another example

Example.

Let $F[x]$ be the vector space of all polynomials with coefficients in F.

For each n, let P_n be the set of all polynomials with degree at most n. So P_n is a subspace of $F[x]$.

Let $D : F[x] \to F[x]$ be the differentiation operator.

Question: Is P_n D-invariant?

Answer: Yes. When you differentiate a polynomial, the degree only gets lower, so it's in P_n.

Example.

Let $W = W_{c_1} + \cdots + W_{c_k}$. Is W T-invariant?

Take any $\beta \in W$. Then, β is of the form

$$\beta = a_1\alpha_1 + a_2\alpha_2 + \cdots + a_k\alpha_k$$

Where each $\alpha_i \in W_{c_i}$ and each $a_i \in F$. So

$$\begin{aligned} T(\beta) =& T(a_1\alpha_1 + a_2\alpha_2 + \cdots + a_k\alpha_k) \\ =& a_1T(\alpha_1) + \cdots + a_kT(\alpha_k) \\ =& a_1(c_1\alpha_1) + \cdots + a_k(c_k\alpha_k) \end{aligned}$$

So $T(\beta) \in W$.

Let's look at a big source of examples.

Example.

Given $T : V \to V$. Let $U : V \to V$ be such that

$$TU = UT$$

Then, Range(U) and Null(U) are T-invariant.

Proof.

Choose any $\alpha \in$ Range(U), then we must show that $T(\alpha) \in$ Range(U).

Choose any $\beta \in V$ such that $U(\beta) = \alpha$. Then

$$T(\alpha) = T(U(\beta)) \underset{(TU=UT)}{=} U(T(\beta))$$

Therefore, $T(\alpha) \in \text{Range}(U)$ by definition of the range!

Now choose any $\alpha \in \text{Null}(U)$, we must show that $T(\alpha) \in \text{Null}(U)$.

Compute $U(T(\alpha)) = T(U(\alpha)) = T(0) = 0$. Therefore $T(\alpha) \in \text{Null}(U)$.

Thus: If $g(x) \in F[x]$, then $\text{Range}(g(T))$ and $\text{Null}(g(T))$ and both T-invariant.

Proof. By our previous example, we only need to show that $g(T)$ commutes with T

$$g(T) \cdot T = T \circ g(T)$$

Obvious: Since the polynomials $g(x) \cdot x = x \circ g(x)$.

Say $T : V \to V$ is a linear operator and $W \subseteq V$ is T-invariant. Choose $\mathcal{B}_W = \{\alpha_1, \dots, \alpha_r\}$ a basis for W. Let $\mathcal{B} = \{\alpha_1, \dots, \alpha_r, \alpha_{r+1}, \dots, \alpha_n\}$ be a basis for V, extending $\mathcal{B}_W$.

Let $\mathcal{B} = [T]_{\mathcal{B}_W}$ be the $r \times r$ matrix representing T_W.

Then

$$[T]_{\mathcal{B}} = \begin{bmatrix} B & C \\ 0 & D \end{bmatrix} = [T(\alpha_1), T(\alpha_2), \dots, T(\alpha_n)]$$

Where D is an $(n-r) \times (n-r)$ matrix. Since for $1 \leq i \leq r$,

$$T(\alpha_i) = T_W(\alpha_i) \in W = \text{Span}(\alpha_1, \dots, \alpha_r)$$

But recall the determinant of a block matrix of this form!

$$\det\left([T]_{\mathcal{B}}\right) = \det\left(\begin{bmatrix} B & C \\ 0 & D \end{bmatrix}\right) = \det(B) \cdot \det(D)$$

Similarly: $\det(xI - T) = \det(xI - T_W) \cdot \det(xI - D)$. Therefore the characteristic polynomial of T_W divides the characteristic polynomial of T.

Similarly: The minimum polynomial of T_w divides the minimal polynomial of T. Notice

$$\left([T]_{\mathcal{B}}\right)^k = \begin{bmatrix} B^k & C^k \\ 0 & D^k \end{bmatrix}$$

“If you don’t believe the above, just trust me that it works” -Laskowski (kinda)

The minimal polynomial $m(x)$ gives that $m(T) = 0$. In particular: $m(T)(\alpha) = 0$.

Fix $\alpha \in V$, and $W \subseteq V$ T-invariant.

Definition.
The **stuffer ideal**

$$I = \{g(x) \in F[x] \mid g(T)(\alpha) \in W\}$$

I (which really depends on α, W) is an ideal of $F[x]$.

■

Definition.
The T**-conductor** of α into W is the monic $g(x) \in I$ of **least degree**.

■

Clearly, the T-conductor of $\alpha \in W$ **divides** the minimal polynomial $m(x)$.

Proof.

Since $m(T)(\alpha) = 0 \in W$, we get that $m(x) \in I$, but $g(x)$ divides every polynomial in I.