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

$$\alpha = x_1\alpha_1 + \cdots + x_n\alpha_n$$

Proof

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 $\mathcal{P}$ 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 $h(x) = 10x^3 - 2x^5$.

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

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

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 $[\alpha]_{\mathcal{B}} \mapsto [\alpha]_{\mathcal{B}'}$.

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

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 $T : F^n \to F^n$

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

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

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

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

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

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

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

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 **0** lives in W, the **codomain** of the transformation T.
   ⟨ 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. ⟨ 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. ⟨ 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 $\mathcal{P}$ the set of all polynomials $a_0 + a_1 x + \cdots + a_n x^n$.

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

$$D(f) = f'$$

Claim:

D is a linear transformation.

Proof:

Take two functions $f, g \in \mathcal{P}$, 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.

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

$\langle$ Note that the integral exists because you can always integrate a continuous function. $\rangle$

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 $f, g \in \mathcal{P}$, 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 $\mathbb{R}$)

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

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

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

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

Proof

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

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

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

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

   $\langle$ 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. $\rangle$

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

2. **Existence**: Chose any $\alpha \in V$. $\langle$ We must define $T(\alpha)$ $\rangle$

   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

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

Fact: $\text{Range}(T)$ 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 $\beta_1, \beta_2 \in \text{Range}(T)$. $\langle$ must show that $\beta_1 + \beta_2 \in \text{Range}(T)$ $\rangle$
   Since $\beta_1 \in \text{Range}(T)$, 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 $\beta_1 + \beta_2 \in \text{Range}(T)$ $\langle$ since $\alpha_1, \alpha_2 \in V$ means that $\alpha_1 + \alpha_2 \in V$, because it's a vector space! $\rangle$
3. **Scaling Works:** Say $\beta \in \text{Range}(T)$, and $c \in F$. Chose $\alpha \in V$ such that $T(\alpha) = \beta$. Then $T(c\alpha) = cT(\beta)c\beta$, therefore $c\alpha \in \text{Range}(T)$.

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

Definition

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

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

$\langle$ 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 $W \subseteq V$. $\rangle$

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

Theorem

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

Proof

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

Consider the linear transformation $D : \mathcal{P}_2 \to \mathcal{P}_2$, 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 $\mathcal{P}_2$ (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

Rank-Nullity 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, $A \subseteq \text{Range}(T)$. 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 $\text{Range}(T)$.

■

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 $\text{Range}(T_A)$.

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. $\text{Nullity}(T_A) = n - r$

Proof

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
- $\text{Rank}(T_A)$

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 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. Range(T_A) = ColSpace(A),
2. 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 $\text{Range}(T_A)$ 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 $\text{Range}(T) = W$. 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 $\text{Null}(T) = \{\mathbf{0}\}$.

$\langle$ In other words, the Null space must only contain the zero vector. $\rangle$

Proof

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 $\text{Null}(T)$ 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.

Proof

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.

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

Check the following:

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

⟨ 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.
>
> $\langle$ We didn't talk about this in class but it *has* to be true. $\rangle$

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

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

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

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

So $f_1 \in \mathrm{Span}(\{1, \sin x, \cos x\})$, 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

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

Example

Problem 2.4.4d

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

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.

Proof

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 $\dim(L(V,W)) = nm$.

Proof

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

Chose a basis $\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$ of V, and $\mathcal{B}' = \{\beta_1, \ldots, \beta_n\}$ 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: $\langle$ Why does $\{E^{pq}\}$ span $L(V,W)$? $\rangle$

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. $\langle$ What does this T do to V? $\rangle$

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 $A_{1q}, \ldots, A_{mq}$

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

Subclaim:

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

$\langle$ The subclaim shows that T is in the span of E^{pq} $\rangle$

Proof of Subclaim

Let $U = \sum_{p=1}^{m} \sum_{q=1}^{n} A_{pq} E^{pq}$. $\langle$ 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. $\rangle$

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 $T = U$.

⟨ Now we need to show that they are linearly independent. ⟩

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

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

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

⟨ 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 $1 \leq q \leq n$, 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 $V = \mathcal{P}^2$, $\mathcal{B} = \{1, x, x^2\}$, for any $f \in \mathcal{P}^2$, $f(x) = a_0 + a_1x + a_2x^2$,

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

is an isomorphism.

Given any $\begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} \in F^3$, Let $\alpha = a_0 + a_1x + a_2x^2 \in \mathcal{P}^2$, 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 $D : \mathcal{P}_2 \to \mathcal{P}_1$

$$D(a_0 + a_1 x + a_2 x^2) = a_1 + 2a_2 x$$

$\mathcal{B} = \{1, x, x^2\}$ be a basis for $\mathcal{P}_2$, 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 $\mathcal{B} = \{\alpha_1, \dots, \alpha_n\}$ of V, and $\mathcal{B}' = \{\beta_1, \dots, \beta_n\}$ of W.

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

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

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 $\langle$ Some will notice that this is the information that the determinant of T encodes! $\rangle$

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!

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

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! $\langle$ If this isn't the case, the inverse might only work in *one direction* $\rangle$.

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

$\langle$ If you have one of these facts, the others come for free! $\rangle$

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

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

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 $\mathcal{B}' = \{\alpha'_1, \ldots, \alpha'_n\}$? 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 $M_{\mathcal{B}}^{\mathcal{B}'}(T)$

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

So if the transformation is the identity, it becomes

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

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.

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

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.

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

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.

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

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

Let's define a basis $S = \{\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}\}$, 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 $\mathcal{B} = \{\begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix}\}$ 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 $[T]_{\mathcal{B}} = [[T(\alpha_1)]_{\mathcal{B}}, [T(\alpha_2)]_{\mathcal{B}}]$.

$T(\alpha_1) = T(\begin{bmatrix} 1 \\ 2 \end{bmatrix}) = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$. 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 $\mathcal{B}' = \{\beta_1 = \begin{bmatrix} a \\ b \end{bmatrix}, \beta_2 = \begin{bmatrix} c \\ d \end{bmatrix}\}$ 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:

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

But then, finding the inverse is easy!

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

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 7

Fri. Mar 10 2023

For $n \times n$ matrices

The $\text{trace}(A)$ 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 $\text{trace}(A') = \text{trace}(PA^{-1}P) = \text{trace}(AP^{-1}P) = \text{trace}(A)$

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

TODO explain why

Recall the dual space.

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

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

By this definition, we can see that S 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

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

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

Theorem

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

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

Proof

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 $\{f_{k+1}, \ldots, f_n\} \subseteq \text{Ann}(W)$.
   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: $\{f_{k+1}, \ldots, f_n\}$ are linearly independent.
   This is obvious since this is how they were defined, all $f_1, \ldots, f_n$ is a basis for V^*.
3. $\{f_{k+1}, \ldots, f_n\}$ 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 $d_1, \ldots, d_n$

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 $\alpha_1, \ldots, \alpha_k$ are in W.

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

■

Mon Mar. 13 2023

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

Let $\mathcal{B}^* = \{f_1, \ldots, f_n\}$ 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 $c_{k+1}, c_{k+2}, \ldots, c_n \neq 0$. 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.

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

Proof

Claim.

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

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

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 $\beta \in \text{Null}(f_{k+1}) \cap \text{Null}(f_{k+1}) \cap \cdots \cap \text{Null}(f_n)$. 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 $\text{Ann}(W)$. Therefore, for any $g \in \text{Ann}(W)$, $g(\beta) = 0$. What we can conclude from this is that β must be in W.

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

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 $\alpha \in V$

$$T^t(g)(\underbrace{\alpha}_{\in V}) := \underbrace{g(T(\alpha))}_{\in F}$$

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

Proof

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

   $g \in \text{Null}(T^t)$, if and only if $T^t(g) = \mathbf{0}$, if and only if $\forall \alpha \in V, (T^t(g))(\alpha) = 0$, if and only if $\forall \alpha \in V, T(g(\alpha)) = 0$, if and only if g annihilates $\text{Range}(T)$, if and only if $g \in \text{Ann}(\text{Range}(T))$.

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

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

■