Math 350H Notes

Nilay Tripathi

Spring 2023

Contents

Lecture 1

Basic Abstract Algebra & Vector Spaces

1.1 Basic Abstract Algebraic Structures

Before beginning, we will present some definitions related to abstract algebra

Definition 1.1: Binary Operation
Let S be a set. A binary operation $\star$ on S is a function $\star : S \times S \to S$

In this class, we will use the notation $\star(s_1, s_2)$ for $s_1 \star s_2$. This notation emphasizes that the binary operation is a function on the operands.

Example 1.1
Consider the set of integers $\mathbb{Z}$. The binary operation $+ : \mathbb{Z} \times \mathbb{Z} \to \mathbb{Z}$ is a binary operation on the set of integers where $+(m, n) = m + n$.

1.1.1 Groups

Now we will present the concept of a group. Here is the definition

Definition 1.2: Group
Let $\star$ be a binary operation on a set S. We say that $(S, \star)$ is a group provided that it satisfies the following three properties:

1. The operation $\star$ is associative. We have $a \star (b \star c) = (a \star b) \star c$ for all $a, b, c \in S$.
2. There exists an element $e \in S$ such that for all $s \in S$, $e \star s = s \star e = s$.
3. For all $s \in S$, there exists $s' \in S$ such that $s \star s' = s' \star s = e$.

There are thus four requirements that a group and its operation must satisfy.

- The group is closed under its respective operation
- The group operation is associative
- There is an element in the group that serves as the identity under that operation. We will prove later that this identity element is unique.

- All elements of the group have an inverse that is also in the group. As with the uniqueness of the identity, we will prove that the inverse of any particular element is unique.

Another important property of groups is closure. That is, for any two elements, the result of the group operation must be contained in the group. This is implied in the definition of a binary operation on a set.

Example 1.2

We have that $(\mathbb{Z}, +)$ presented in the example above is a group. This is because we have that addition on integers is associative. The element 0 is an identity of the group and for all integers m, the additive inverse is $-m$.

Proposition 1.1

The identity element of a group is unique.

Proof. Let e_1 and e_2 both be identity elements of a group. Then we have that $e_1 \star e_2 = e_2$ since e_1 is an identity of the group. However, we also have that $e_1 \star e_2 = e_1$ since e_2 is also an identity of the group. Thus, we have $e_1 = e_2$, as desired. □

Proposition 1.2

Suppose that $(S, \star)$ is an arbitrary group. Then for all members of the group, the inverse of that member is unique.

Proof. Let $s \in S$ and let s_1 and s_2 be inverses of s. Also let e be the identity of the group. We have that $s_1 = s_1 \star e$. Since s_2 is an inverse of s, we have $e = s \star s_2$. So we have $s_1 = s_1 \star (s \star s_2)$. By associativity, we have $s_1 = (s_1 \star s) \star s_2$. Since s_1 is an inverse of s, we have $s_1 = s_2$, as desired. □

Example 1.3

Here are some non-examples of groups.

- $(\mathbb{R}, \times)$ is not a group since 0 does not have an inverse.
- $(\mathbb{N}, +)$ is not a group since it has no identity.
- $(\mathbb{N}, \times)$ is not a group since not all elements have an inverse that is also a natural number (only 1 does)
- Let S be a non-empty set. Then $(\mathcal{P}(S), \setminus)$ is not a group (where "\" denotes set difference). This is because set difference is not associative.

Example 1.4

Here is a more interesting example of a group. For a non-empty set S, $(\mathcal{P}(S), \triangle)$ is a group (here $\triangle$ is the symmetric difference). Recall that the symmetric difference of two sets A and B is defined as

$$A \triangle B = (A \setminus B) \cup (B \setminus A) \tag{1.1}$$

We have that $\varnothing$ is the identity element and for all sets $A \in \mathcal{P}(S)$, we have that A is the inverse of itself (since $A \triangle A = \varnothing$. The symmetric difference is also associative since set union is associative.

Note that the axioms of a group do not specify that the group operation needs to be commutative; in fact, it may not be. In most of the examples above, the operation was commutative, but this need not be the case. This fact motivates the following definition.

Definition 1.3: Abelian Groups

For a group $(S, \star)$ we say the group is abelian if $\star$ is commutative (i.e. if $a \star b = b \star a$). If the operation is not commutative, then the group is non-abelian.

Example 1.5

The integers under addition is an abelian group. Here are some examples of non-abelian groups.

- The set $\text{Aut}(\mathbb{R})$ denotes the set of all bijetions of $\mathbb{R}$. Then, $(\text{Aut}(\mathbb{R}), \circ)$ is a group. The identity is the identity function and bijections are invertible. We also have that function composition is associative but it is not commuative.
- $GL_n(\mathbb{R})$ is the general linear group. It is the set of all invertible $n \times n$ matrices. This set under matrix multiplication does not form a group. The identity is the identity matrix and the matrices are invertible. However, matrix multiplication is not commutative.
- Define $\mathbb{N}_3 = \{1, 2, 3\}$. We have that $\text{Aut}(\mathbb{N}_3)$ is a non-abelian group. We define $S_3 \coloneqq \text{Aut}(\mathbb{N}_3)$.

1.1.2 Fields

In linear algebra we will primarily study vector spaces which exist over a field. With the notion of a group established, we will now define a field.

Definition 1.4: Field

Let S be a nonepty set with binary operations $+$ and $\times$. Then, we have $(S, +, \times)$ is a field provided that

1. $(S, +)$ and $(S', \times)$ are abelian groups. Here $S' = S \setminus \{0\}$
2. For all $s_1, s_2, s_3 \in S$, we have $s_1 \times (s_2 + s_3) = s_1 \times s_2 + s_1 \times s_3$. So the two operations obey a distributive law (multiplication distributes over addition).

Example 1.6

The two relevant examples of a field for this class are $(\mathbb{R}, +, \times)$ and $(\mathbb{C}, +, \times)$. Here, $\mathbb{C}$ is the set of complex numbers.

1.2 Vector Spaces

We now present the definition of a vector space, which will be central to linear algebra.

Definition 1.5: Vector Space

A vector space V over a field $(F, +, \cdot)$ is an abelian group $(V, +)$ such that there is an operation $\star : F \times V \to V$ with the properties

1. $a \star (u + v) = a \star u + a \star v$ for all $a \in F$ and for all $u, v \in V$.
2. We have $1 \star v = v$ for all $v \in V$.
3. $(a \cdot b) \star v = a \star (b \star v)$
4. $(a + b) \star v = (a \star v) + (b \star v)$

Example 1.7

We have that $\mathbb{R}$ is a field. $\mathbb{R}^3$ is the set of all 3-tuples with real numbered entries. We define the operation $\star : \mathbb{R} \times \mathbb{R}^3 \to \mathbb{R}^3$ as follows

$$\left(r, \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} \right) = \begin{bmatrix} ra_1 \\ ra_2 \\ ra_3 \end{bmatrix}$$

It can be shown that this function does satisfy the axioms of a vector space.

1.3 End of Class Problems

For these problems, an element with an arrow over it denoted an element of a vector space while those without an arrow is from the field.

Proposition 1.3

$0 \star \vec{v} = \vec{0}$.

Proof. We have

$$\begin{aligned} 0 \star \vec{v} &= (0+0) \star \vec{v} \\ &= 0 \star \vec{v} + 0 \star \vec{v} \end{aligned}$$

Adding the additive inverse of $0 \star \vec{v}$ to both sides, we get $0 \star \vec{v} = \vec{0}$, as desired. □

Proposition 1.4

For all $a \in F$, we have $a \star \vec{0} = \vec{0}$.

Proof. We have

$$\begin{aligned} a \star 0 &= a \star (0+0) \\ &= a \star 0 + a \star 0 \end{aligned}$$

Adding the additive inverse of $a \star 0$ gives $a \star 0 = 0$, as desired. □

Proposition 1.5

$(-1) \star \vec{v} = -\vec{v}$. Here, $-\vec{v}$ is the inverse of $\vec{v}$ under the operation $+$.

Proof. We want to show $(-1) \star \vec{v}$ is the additive inverse of $\vec{v}$. We will show $(-1) \star \vec{v} + \vec{v} = \vec{0}$. We have

$$\begin{aligned} (-1) \star \vec{v} + \vec{v} &= (-1) \star \vec{v} + 1 \star \vec{v} \\ &= (-1+1) \star \vec{v} \\ &= 0 \star \vec{v} \\ &= \vec{0} \end{aligned}$$

And so the desired result is obtained. □

Lecture 2

Subspace, Linear Combinations & Span

Here is a review of the properties of a vector space. If $(V, +_V)$ is an abelian group, $(F, +_F, \cdot_F)$ is a field and $\cdot : F \times V \to V$ is a map such that

1. $1 \cdot v = v$
2. $(a_1 +_F a_2) \cdot v = a_1 \cdot v +_V a_2 \cdot v$
3. $(a_1 \cdot_F a_2) \cdot v = a_1 \cdot (a_2 \cdot v)$

Then, $((V, +_V), (F, +_F, \cdot_F))$ is a vector space over F and the operation $\cdot$ is called *scalar multiplication*. This is the more rigorous/formal definition of a vector space. The most common vector spaces involve the field $\mathbb{R}$ and $\mathbb{C}$ and the corresponding vector space is $\mathbb{R}^n$ and $\mathbb{C}^n$.

Here is another example of a vector space that is presented in the book.

Example 2.1

Take the field to be $\mathbb{R}$ and let $P(\mathbb{R})$ be the set of all polynomials with real coefficients. We will show that there is an operation that makes this set a vector space.

First, we define the group structure. The operation $+_P$ denotes the sum of two polynomials. So for $f, g \in P(\mathbb{R})$, we have the polynomial sum is $f +_P g$. Note that this definition of polynomial addition is commutative. We define the scalar multiplication of the polynomial as $(r \cdot f)(x) = rf(x)$ for all $r \in \mathbb{R}$ and $f \in P(\mathbb{R})$.

It should also be noted that this vector space is infinite-dimensional.

Example 2.2

For the field $\mathbb{R}$ let $n \in \mathbb{N}$. We define $P_n(\mathbb{R})$ be the set of all polynomials with coefficients of degree $\leq n$, Then $P_n(\mathbb{R})$ is a vector space of $\mathbb{R}$. We will prove later that this vector space is finite-dimensional.

2.1 Subspaces

Below, we present the definition of subspaces.

Definition 2.1: Subspace

Let V be a vector space over F and let W be a subset of V. Then the set W is called a subspace of V if W is a vector space over F with the restrictions of the operations for the vector space V.

From the restrictions on the operations we see that for all $w \in W$ and for all $a \in F$, we see that $w_1 + w_2 \in W$ (so it is closed under vector addition). Also $a \cdot w \in W$ (so it is closed under scalar multiplication). So the inverse of any w will also be in W. Hence, we see that W is abelian since it is a vector space.

Example 2.3

For any vector space V there will always be two trivial subspaces. One of them is V itself and the other is the subspace $\{0\}$.

Example 2.4

Let $V = \mathbb{R}^2$. Then two subspaces of V include $W_x = \mathbb{R} \times \{0\}$ and $W_y = \{0\} \times \mathbb{R}$. These subspaces represent the x-axis and y-axis respectively. This can be generalized for any $\mathbb{R}^n$.

Example 2.5

From earlier, we can show that $P_n(\mathbb{R})$ is a subspace of $P(\mathbb{R})$.

Theorem 2.1

Let V be a vector space over F and $W \subseteq V$ is a subspace of V if and only if the following three conditions hold

1. $0 \in W$
2. For all $x, y \in W$, we have $x + y \in W$
3. For all $x \in W$ and for all $c \in F$, we have $cx \in W$

Proof. **Proof of ($\implies$):** assume that W is a subspace of V. We will show the three properties hold. Since W is a vector space, we know that it is an abelian group and so it is closed under addition, so condition (2) holds. Likewise, from the definition of a vector space, condition (3) also holds. Since W is a vector space, we know there is a member $0' \in W$ such that $w + 0' = w$ for all $w \in W$. But since $W \subseteq V$, we have that $w, 0' \in V$. Since the identity of a vector space must be unique (proved in the last class), we must have that $0 = 0'$. Hence, we have $0 \in W$.

Proof of ($\impliedby$): assume that the three conditions hold. We must show that W is a subspace. So we must show that W is a vector space. We have that $0 \in W$, so the identity is in W. By condition (2), we have that W is closed under addition. By condition (3), we have that $(-1)w \in W$. From the theorem proved yesterday, we have that $(-1)w = -w$ and so the inverse is also in W. Hence, we have that W is an abelian group. □

We will present another example of a subspace that relates to matrices. Before we introduce it, here is a definition.

Definition 2.2: Symmetric Matrices

Let A be an $m \times n$ matrix and let A^T denote the transpose of the matrix A. Then we say the matrix A is symmetric if $A = A^T$.

Definition 2.3: Trace

Let A be an $n \times n$ matrix. Then the trace of A, denoted $\text{tr}(A)$, is the sum of its diagonal elements.

Before we present the subspaces, we will introduce the overarching vector space. We let $M_n(\mathbb{R})$ be the set of all $n \times n$ matrices with real-valued entries. Then we can show that $M_n(\mathbb{R})$ is a vector space over $\mathbb{R}$ with scalar multiplication defined on matrices element-wise.

Example 2.6

Let $W_1 = \text{Sym}_n(\mathbb{R})$ be the set of all symmetric $n \times n$ matrices. Then W_1 is a subspace of $M_n(\mathbb{R})$ since we have the property $(A+B)^T = A^T + B^T$.

Let W_2 be all $n \times n$ matrices with trace 0. We can verify that W_2 is a subspace by using the property $\text{tr}(A+B) = \text{tr}(A) + \text{tr}(B)$.

Here is a theorem that allows us to construct new subspaces from existing ones.

Theorem 2.2

Let $\mathcal{C}$ be a collection of subspaces of the vector space V. Then the intersection of the subspaces in $\mathcal{C}$ is also a subspace of V.

Proof. Let $\{W_i\}_{i \in I}$ be an arbitrary collection of subspaces from $\mathcal{C}$. We must show $\bigcap_{i \in I} W_i$ is a subspace. We will show that the three properties hold.

1. For all $i \in I$, we have that W_i is subspace of V. Hence, we have for all $i \in I$, $0 \in W_i$. Thus, it follows that, $0 \in \bigcap_{i \in I} W_i$.
2. Let $x, y \in \bigcap_{i \in I} W_i$. Then by the definition of intersection, we have $x, y \in W_i$ for all $i \in I$. Since all W_i are subspaces, we have $x + y \in W_i$ for all $i \in I$. Hence, we have $x + y \in \bigcap_{i \in I} W_i$.
3. Let $x \in \bigcap_{i \in I} W_i$ and let $c \in F$. From the definition of intersection, we have $x \in W_i$ for all $i \in I$. Since all W_i are subspaces, we have $cx \in W_i$. Thus, we have $cx \in \bigcap_{i \in I} W_i$.

□

2.2 Linear Combinations

We will now give the definition of a linear combination, which is central to linear algebra.

Definition 2.4: Linear Combination

Let V be a vector space over F. Then for $v_1, v_2 \in V$ and $c_1, c_2 \in F$, a linear combination of v_1 and v_2 is a vector of the form $c_1 v_1 + c_2 v_2$. The scalars c_1 and c_2 are the coefficients of the vectors v_1 and v_2 respectively.

While the definition only gives the linear combination of two vectors, we can use induction to rigorously define a linear combination of a finite set of vectors. Then we have that for vectors $v_1, v_2, ..., v_n \in V$ and scalars $c_1, c_2, ..., c_n \in F$, the linear combination of a finite set of vectors is $c_1 v_1 + c_2 v_2 + \cdots + c_n v_n$. Here, $c_1, c_2, ..., c_n$ are the coefficients of the vectors.

2.2.1 Span

The next step to move forward will be to define sets of linear combinations. We present the definition below

Definition 2.5: Span

Let S be a non-empty subset of a vector space V. Then the span of S, denoted $\text{span}(S)$ is defined to be collection of all finite linear combinations of the elements from S.

Example 2.7: L

t V be the collection of all sequences of real numbers. For sequences (a_n) and (b_n) define addition component-wise so $(a_n + b_n)_i = (a_n)_i + (b_n)_i$ and define scalar multiplication to be $r(a_n) = (ra_n)$. Then V is a vector space over $\mathbb{R}$.

Let $e^i \in V$ be the sequence where $e^i_n = 1$ if $i = n$ and 0 if $i \neq n$. So we have $e^1 = (1, 0, 0, ...)$, $e^2 = (0, 1, 0, ...)$, $e^3 = (0, 0, 1, ...)$. Then we have that $(1, 1, 1, ...) \notin \text{span}(S)$.

The span of a set has some nice properties. The properties are given in the following theorems.

Theorem 2.3
Let V be a vector space over F and let $S \subseteq V$. Then we have

1. $\text{span}(S)$ is a subspace of V.
2. If W is a subspace of V containing S, then $\text{span}(S) \subseteq W$. In other words, $\text{span}(S)$ is the smallest subspace of V that contains S.

Proof. □

2.3 End of Class Problems

The following example is an exercise. It is exercise 10 in section 1.4

Example 2.8
Define the matrices as

Lecture 3

Linear Independence, Basis, and Dimension

3.1 Linear Independence & Dependence

We define the concept of linearly independent and linearly dependent below.

Definition 3.1: Linear Independence

Let V be a vector space over F. Let S be a collection of vectors from V. Then the set S is said to be linearly independent if $a_1x_1 + a_2x_2 + \cdots + a_nx_n = 0$ for all finite subcollections $\{x_1, \ldots, x_n\}$ in S and scalars $a_1, a_2, ..., a_n \in F$, then $a_1 = a_2 = \cdots = a_n = 0$.

Definition 3.2: Linear Dependence

If a set $S \subseteq V$ is not linearly independent, then we say that is linearly dependent.

Example 3.1

Consider the vector space $\mathbb{R}^n$ over the field $\mathbb{R}$. Then we have that

$$S_m = \{e_1, e_2, \ldots, e_n\} \qquad m \leq n$$

Then the set S_m is linearly independent. Note that e_i is the vector of all 0s except for the i-th component, which has a value of 1.

Now consider the set $S = \{e_1, e_2, e_1 + e_2\}$. This set is not linearly independent since we have that $(-1)e_1 + (-1)e_2 + 1(e_1 + e_2) = 0$. Thus, there is a non-trivial linear combination of the vectors in S that equals 0.

Example 3.2

Consider the vector space $P(\mathbb{R})$ which is all polynomials with real number coefficients. Then consider the set $S = \{x^i : i \in \mathbb{N} \cup \{0\}\}$. This set is linearly independent in $P(\mathbb{R})$. This is an infinite set that is linearly independent.

Here is another example of an infinite set that is linearly independent. Consider the vector space of all sequences with real valued entries. Then $T = \{e_i : i \in \mathbb{N}\}$ is linearly independent.

3.2 Bases

With the notion of linear independence established, we now move on to the definition of a basis.

Definition 3.3: Basis
Let V be a vector space over F. Let B be a set of vectors from V. Then the set B is a basis of V if B is linearly independent and span $B = V$.

A basis is essentially just a linearly independent spanning set for a subspace. Here are some examples.

Example 3.3
The set $S_n = \{e_1, e_2, \ldots, e_n\}$ is a basis for the subspace $\mathbb{R}^n$. Note that the basis for this set is finite.

The set S defined in a previous example is a basis for $P(\mathbb{R})$. Consider $f \in P(\mathbb{R})$. Then $f = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n$ for $a_1, a_2, \ldots, a_n \in \mathbb{R}$. So we have that $f \in \operatorname{span} S$.

Example 3.4
Consider the set $S = \{e_i : i \in \mathbb{N}\}$ defined in a previous example. Even though the set is linearly independent, this set is not a basis. For a counterexample, consider the sequence $\mathbf{1} : \mathbb{N} \to \mathbb{R}$ which is given by $\mathbf{1}(n) = 1$ for all $n \in \mathbb{N}$. We have that $\mathbf{1} \notin \operatorname{span} S$.

Now we will prove some important theorems regarding basis. The first one will lead to the concept of dimension later on. We formally introduce the terminology of finite-dimensional vector spaces.

Definition 3.4: Finite Dimensional Vector Space
Let V be a vector space over F. Then we say V is finite dimensional if any basis for V is finite.

The definition will also lead to the important result that will be proved in this theorem.

Theorem 3.1
Let V be a finite dimensional vector space over F. Let $B = \{b_1, b_2, \ldots, b_n\}$ be a basis for V. Then if $S \subseteq V$ has m vectors where $m > n$, then S is linearly dependent.

Proof. Let $S = \{x_1, x_2, \ldots, x_m\}$. Then to show S is linearly dependent, we must find scalars $c_1, c_2, \ldots, c_m \in F$, not all 0, such that $c_1 x_1 + \cdots c_m x_m = 0$. Then since B is a basis for V, we have

$$\begin{aligned} x_1 &= \sum_{j=1}^{n} \alpha_{1j} b_j \qquad \text{for } \alpha_{1j} \in F,\ j = 1, ..., n \\ x_2 &= \sum_{j=1}^{n} \alpha_{2j} b_j \qquad \text{for } \alpha_{2j} \in F,\ j = 1, ..., n \\ &\vdots \\ x_m &= \sum_{j=1}^{n} \alpha_{mj} b_j \qquad \text{for } \alpha_{mj} \in F,\ j = 1, ..., n \end{aligned} \tag{3.1}$$

Then we have the original equation is

$$c_1 \sum_{j=1}^{n} \alpha_{1j} b_j + c_2 \sum_{j=2}^{n} \alpha_{2j} b_j + \cdots c_m \sum_{j=1}^{n} \alpha_{mj} b_j = 0 \tag{3.2}$$

Then we can collect the coefficients in front of each b_j. This gives us that

$$(c_1\alpha_{11} + c_2\alpha_{21} + \cdots + c_m\alpha_{m1})b_1 + \cdots + ()b_j = 0 \tag{3.3}$$

Then since B was a basis, we have that it is linearly independent. Thus, the individual coefficients must be 0. This gives us that

$$\begin{aligned} \sum_{i=1}^{m} \alpha_{i1}c_i &= 0 \\ \sum_{i=1}^{m} \alpha_{i2}c_i &= 0 \\ &\vdots \\ \sum_{i=1}^{m} \alpha_{in}c_i &= 0 \end{aligned} \tag{3.4}$$

Finish proof in PSS session: from elementary linear algebra, we know that since $m > n$, the null space of this system is non-trivial. Pick any nonzero vector from the null space, which will be a set of scalars that will suffice. □

From the proof of the statement above, we have the following corollary.

Corollary 3.1
If V is a finite dimensional vector space over F. Then, any two bases of V must have the same number of vectors.

Here are some other propositions that were introduced in an elementary linear algebra course. Note that we prove these theorems in the next class and use some results from the next class to aid in the proofs.

Proposition 3.1
Let V be a finite dimensional vector space and $S \subseteq V$ is linearly independent. Then, V has a basis B and $B \supseteq S$.

Proof. Assume the dimension of V is equal to n. We have two cases to consider: either $|S| = n$ or $|S| < n$. If $|S| = n$, then S is a basis for V and the result is proved. So assume that $|S| < n$. Then we must have that S does not span V from the theorem above. Then using Theorem 1.7 (in the book), there exists a $v \in V$ that is not in the span of S. Thus, we have that $S \cup \{v\}$ is a linearly independent set with size $|S| + 1$. This process can repeat until we arrive at a generating set for V, which (by the previous problem), will occur when there are n vectors. □

Proposition 3.2
Let V be a finite dimensional vector space over F. Let S be a subset of V such that $\operatorname{span} S = V$. Then S contains a basis of V.

Proof. Since V is finite dimensional, then it must have a finite basis that we'll call $B = \{b_1, b_2, \ldots, b_n\}$. Now let $\Sigma = \{A \subseteq S : A \text{ is finite and linearly independent}\}$. Note that V cannot have an infinite linearly independent subset. This is because if this were the case, then all subsets of V also be linearly independent. But then, we can find a linearly independent subset of size $n + 1$, which contradicts PSS2.

Now for all $A \in \Sigma$, we must have $|A| \leq n$. We claim that there exists $A \in \Sigma$ such that $|A| = n$. Let $\mathbb{N}_\Sigma = \{|A| : A \in \Sigma\}$. Note that for all $m \in \mathbb{N}_|Sigma$, we must have $m \leq n$. So n is an upper bound of $\mathbb{N}_\Sigma$. So $\mathbb{N}_\Sigma$ has a maximum element since it is a finite subset of $\mathbb{N}$. Note that if $|A| < n$, then $\operatorname{span} A \neq V$. So there exists a $v \in S$ such that $A \cup \{v\}$ is linearly independent (otherwise it would contradict the fact that $\operatorname{span} S = V$). This contradicts the assumption that $|A|$ is a maximum and so there is a set of size n that follows.

Now we claim that the set A from the previous claim is a basis for V. We know that A must be linearly independent, so we must have that $\operatorname{span} A \neq V$. So there exists $v \in V \setminus \operatorname{span} A$ so A is linearly independent and $v \notin \operatorname{span} A$. By theorem 1.7, we have that $A \cup \{v\}$ is linearly independent. But since $A \cup \{v\}$ has $n+1$ members, this contradicts PSS2. So A is a basis of V.

So since $A \subseteq S$ we have proved that S contains a basis of V. □

The two propositions above lead to an important corollary.

Corollary 3.2

Every non-trivial vector space has a basis.

Proof. Let $x \in V$ such that $x \neq 0$. Then the set $\{x\}$ is linearly independent. By proposition 1.1, we have that $\{x\}$ is contained in a basis of V. □

3.3 End of Class Problems

The proofs of the propositions are the in-class problems. The other problem is to construct a basis for $V_S(\mathbb{R})$.

Lecture 4

Dimension of a Vector Space

In today's class, we mostly revisited work from last lecture. Hence, a lot of notes from today are in the section for last lecture

4.1 Dimension

We will present an introductory example and introduce the definition of dimension.

Definition 4.1: Dimension
Let V be a vector space over F. Then the dimension of V, denoted $\dim V$ is the number of vectors in any basis of V.

Example 4.1
We have that $\mathbb{C}$ is a vector space over $\mathbb{R}$. Then we have that V is a two dimensional vector space over $\mathbb{R}$. Then we denote this as $\dim_{\mathbb{R}}(\mathbb{C}) = 2$. We have that $\{1, i\}$ is a basis for this vector space.

We can also consider that $\mathbb{C}$ is a vector space over $\mathbb{C}$ (in general, any field is a vector space over itself). We have that $\dim_{\mathbb{C}}(\mathbb{C}) = 1$ since any basis for $\mathbb{C}$ over this field will have only one member. An example of such a basis is $\{1\}$ (although any nonzero member will work as well).

The following theorem will be very useful in proving later theorems.

Theorem 4.1
Let V be a vector space over F. Let $S \subseteq V$ be linearly independent. Then for $v \in V$, the set $S \cup \{v\}$ is linearly independent if and only if $v \notin \operatorname{span} S$.

Proof. **Proof of ($\implies$):** assume $v \in V$ such that $S \cup \{v\}$ is linearly independent. Assume, for contradiction, that $v \in \operatorname{span} S$. Then we have that $v = c_1u_1 + c_2u_2 + \cdots + c_nu_n$ where $c_1, ..., c_2 \in F$, for $u_1, u_2, ..., u_n \in S$. So we have that $c_1u_1 + c_2u_2 + \cdots + c_nu_n - v = 0$. But this is a contradiction to the fact that $S \cup \{v\}$ is linearly independent. Thus, we must have that $v \notin \operatorname{span} S$.

Proof of ($\impliedby$): let $v \notin \operatorname{span} S$. For contradiction, assume that $S \cup \{v\}$ is not linearly independent. Therefore, there must be a subset of $S \cup \{v\}$ that is linearly dependent. So we have $c_1u_1 + \cdots + c_nu_n = 0$ where not all $c_i = 0$. Then, there must exist a $j \in \{1, ..., n\}$ such that $u_j = 0$, otherwise we would have a linearly dependent subset in S (a contradiction to the fact that S is linearly independent). Without loss of generality, take $u_1 = v$. Then $c_1v + c_2u_2 + \cdots + c_nu_n = 0$. Note that we must have that $c_1 \neq 0$; otherwise, it would contradict the fact that S is linearly independent. Since c_1 is nonzero, it must have a multiplicative inverse. So we have that $v = (-c_2c_1^{-1})u_2 + \cdots + (-c_nc_1^{-1})u_n$. □

The theorem below is a stronger version of a theorem that was stated in the previous class.

Theorem 4.2
Let V be a vector space and $S \subseteq V$ is linearly independent. Then S is contained in a basis of V.

The proof will be done later on, as it uses the concepts of maximally linearly independent sets. Here's another theorem that is fairly easy to prove.

Theorem 4.3
Let V be a vector space over F with a basis $\{b_1, b_2, ..., b_n\}$. Then each $v \in V$ can be written as a unique linear combination of $b_1, ..., b_n$

Proof. Let $v = a_1b_1 + a_2b_2 + \cdots + a_nb_n$ and $v = c_1b_1 + c_2b_2 + \cdots + c_nb_n$. We must show $a_n = c_n$. Then, we subtract these two to get $(a_1 - c_1)b_1 + (a_2 - c_2)b_2 + \cdots + (a_n - c_n)b_n = 0$. Then since the basis is linearly dependent, we must have that $a_i - c_i = 0$ for all $i \in \{1, ..., n\}$. It follows that $a_i = c_i$ for all $i \in \{1, ..., n\}$. $\square$

Lecture 5

Infinte Dimensional Vector Space Business

5.1 Extra Results About Finite Dimensional Vector Spaces

Here, we will state a theorem that we looked at in the previous class.

Theorem 5.1
Let V be a finite-dimensional vector space over F. Then B is a basis of V if and only if each $v \in V$ can be uniquely represented as a linear combination of vectors from B.

Proof. **Proof of** ($\implies$)**:** let $B = \{v_1, v_2, ..., v_n\}$ be a basis of V and let $v \in V$ be an arbitrary member of V. Now, since $\operatorname{span} B = V$ so we have $v \in \operatorname{span} B$. So v is a linear combination of vectors from B. Let $v = a_1v_1 + a_2v_2 + \cdots + a_nv_n = b_1v_1 + b_2v_2 + \cdots + b_nv_n$ where $a_i, b_i \in F$ for all $i \in \{1, ..., n\}$. Now, if we subtract these two representations, we get

$$(a_1 - b_1)v_1 + (a_2 - b_2)v_2 + \cdots + (a_n - b_n)v_n = 0 \tag{5.1}$$

But since we have that $v_1, v_2, ..., v_n$ are linearly independent, we must have $a_i - b_i = 0$. It follows that $a_i = b_i$ and the representations are unique.

Proof of ($\impliedby$)**:** since each $v \in V$ is a unique linear combination of vectors from B, it follows that $\operatorname{span} B = V$. Now we have to show that B is linearly independent. Assume that $a_1v_1 + \cdots + a_nv_n = 0$ where $a_i, ..., a_n \in F$. Note that $0 = 0v_1 + \cdots + 0v_n$. By uniqueness of the representations, we must have $a_i = 0$ for all $i \in \{1, ..., n\}$. So B is linearly independent. Since B is linearly independent and spans V, we have that B is a basis of V. □

Here is another theorem that one should look at on their own time.

Theorem 5.2
If a vector space V is generated by a finite set S, then S contains a basis of V. Consequently, V has a finite basis.

Proof. To be done as take home reading □

From the theorems above, we state and define the following corollary, that we have used in Math 250 before.

Corollary 5.1

Let V be a finite dimensional vector space of dimension n over F. Then

1. Any linearly independent set of vectors has less than or equal to n elements. Furthermore, if there are exactly n elements, the set is a basis of V.
2. Any generating set of V has at least n vectors. If there are exactly n vectors, then the generating set is a basis of V.

Proof. More take home reading □

Now, we will state and prove this (very important) theorem.

Theorem 5.3

Let V be a finite dimensional vector space over F and let W be a subspace of V. Then $\dim W \leq \dim V$. Furthermore, if $\dim W = \dim V$, then $W = V$.

Proof. Let $\dim V = n$ and let B_W be a basis for W. Since B_W is a set of linearly independent vectors from V, we have that B_W is finite and $|B_W| \leq n$. By definition of dimension, we have $\dim W \leq n = \dim V$.

Now assume $\dim W = \dim V$. Let B_W be a basis of W. Then B_W is a linearly independent set of vectors from V with cardinality equal to $\dim V$. Then by the corollary, we have that B_W is a basis of V. So $\operatorname{span} B_W = V$. Since $\operatorname{span} B_W = W$, we must have that $W = V$. □

Here is a corollary to this theorem.

Corollary 5.2

Let V be a finite dimensional vector space over F and let W be a subspace of V. Then any basis of W can be extended to a basis of V.

5.2 Maximal Linearly Independent Sets

Here are some definitions that we will use.

Definition 5.1: Maximal Set

Let $\mathcal{F}$ be a collection of sets. Then a set $M \in \mathcal{F}$ is said to be maximal (with respect to the set inclusion order) if M is contained in no member of $\mathcal{F}$ other than M itself.

Example 5.1

Take A to be a nonempty set. Take $\mathcal{F} = \mathcal{P}(A)$ (the power set of A, which is the collection of all subsets of A). Then A is a maximal element of $\mathcal{F}$ (since A is contained in no subset of A other than itself).

Here is a non-example: let A be an infinite set and take $\mathcal{F} = \{S \subseteq A : S \text{ is finite}\}$. Note that this set does not have a maximal elements. Assume for contradiction that $M \in \mathcal{F}$ was a maximal element. Then M must be finite. So there exists $a \in A \setminus M$. Then we have that $M \cup \{a\} \in \mathcal{F}$, since it is a finite subset of A. But this contradicts that M is maximal since this set contains M. So there is no maximal element in this set.

Now we will introduce the concept of a chain, which was discussed in Math 300H.

Definition 5.2: Chain

A family of sets $\mathcal{F}$ is called a chain (or nest or tower, with respect to set inclusion) if given any two $S_1, S_2 \in \mathcal{F}$ either $S_1 \subseteq S_2$ or $S_2 \subseteq S_1$

Now, we will discuss a very important theorem in this area of math. We will assume this theorem without proof.

Theorem 5.4: Zorn's Lemma
Let $\mathcal{F}$ be a family of sets. If for every chain $\mathcal{C} \subseteq \mathcal{F}$ there exists a $U \in \mathcal{F}$ such that $S \subseteq U$ for all $S \in \mathcal{C}$, then $\mathcal{F}$ has a maximal element.

Here is a theorem that will utilize these concepts.

Theorem 5.5
Let V be a vector space and $V = \operatorname{span} S$. If B is a maximal linearly independent subset of S, then B is a basis of V.

Proof. The hypothesis of the theorem states that B is linearly independent. So we will only need to prove that $S \subseteq \operatorname{span} B$. If not, then there exists $s \in S \setminus \operatorname{span} B$. So $s \notin \operatorname{span} B$ which implies that $B \cup \{s\}$ is a linearly independent set which contradicts the assumption that B is a maximal linearly independent set. Thus, we have $S \subseteq \operatorname{span} B$ and consequently, we have $V = \operatorname{span} B$. □

Here is the theorem that will utilize Zorn's lemma.

Theorem 5.6
Let S be a linearly independent subset of vector space V. Then there exists a maximal linearly independent subset of V that contains S.

Proof. Let $\mathcal{F}$ be the family that contains all linearly independent subsets of V containing S. Let $\mathcal{C}$ be a chain such that $\mathcal{C} \subseteq \mathbf{F}$. Now consider $U = \bigcup_{C \in \mathcal{C}} C$. Note that since $S \subseteq \mathcal{C}$for all $C \in \mathcal{C}$, we have $S \subseteq U$. So U is an upper bound for an arbitrary chain. Now, we must show $U \in \mathcal{F}$, which involves showing that U is linearly independent. Let $a_1u_1 + a_2u_2 + \cdots + a_nu_n = 0$ for $a_i \in F$ and $u_i \in U$ for all $i \in \{1, ..., n\}$. Since $u_i \in U$ we have $u_i \in C_j$ for some $C_j \in \mathcal{C}$. Since $\mathcal{C}$ is a chain, there exists a $k_0 \in \{1, ..., n\}$ such that $C_i \subseteq C_{k_0}$ for all $i \in \{1, ..., n\}$. (prove this as an exercise using induction). Since C_{k_0} is linearly independent we get $a_1 = a_2 = \cdots = a_n = 0$. Since $\{u_1, u_2, ..., u_n\}$ is arbitrary for U, we have U is linearly independent. So $U \in \mathcal{F}$ and U is an upper bound of $\mathcal{C}$.

By Zorn's lemma, we have that $\mathcal{F}$ has a maximal element and so there exists a maximal element linearly independent subset of V that contains S. By the previous theorem, we have that this maximal member is a basis of V. □

Here is a corollary to this theorem, that will establish a well known fact that we knew for finite dimensional vector spaces.

Corollary 5.3
Every vector space has a basis.

Lecture 6

Linear Transformations

6.1 Linear Transformations Definitions

We will now define the notion of a linear transformation.

Definition 6.1: Linear Transformation
Let V and W be vector spaces over F. Then a function $L : V \to W$ is called a linear transformation if $L(cv_1 + v_2) = cL(v_1) + L(v_2)$ for all $c \in F$ and $v_1, v_2 \in V$.

Here is another related definition.

Definition 6.2: Linear Operator
A linear transformation $L : V \to V$ is a linear operator

Example 6.1
Let $L : V \; toW$ be a linear transformation. Then we note that

$$\begin{aligned} L(0) &= L(0+0) \\ &= L(0) + L(0) \end{aligned}$$

And from this fact, it follows that $L(0) = 0$

Example 6.2
Here is a slightly harder example. Recall that $P(\mathbb{R})$ is the set of polynomials with real-valued coefficients. Then $D : P(\mathbb{R}) \to P(\mathbb{R})$ is the derivative. Then D is linear since

1. $D(f) = f'$
2. $D(cf + g) = cf' + g' = cD(f) + D(g)$

Example 6.3
The identity function and 0 functions are both examples of a linear transformation. They are defined as $\mathbf{1}_V : V \to V$ where $\mathbf{1}_V(v) = v$ and $\mathbf{0} : V \to V$ where $\mathbf{0}(v) = 0$

Example 6.4

Here is an example from Math 250. Let $V = \mathbb{F}^n$ and $W = \mathbb{F}^m$. Then $L : V \to W$ defined as

$$L\left(\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}\right) = \begin{bmatrix} a_1 \\ \vdots \\ a_m \end{bmatrix} \tag{6.1}$$

is called a projection and is a linear transformation.

For the same fields, we can define additional linear transformations by using matrix multiplication. Let L_A :

Now we will define linear transformations by introducing subspaces for linear transformations.

Definition 6.3: Null Space (Kernel)

Let $L : V \to W$ be a linear transformation. Then we define the null space (or kernel) of L to be the set

$$\operatorname{Null} L = \{v \in V : L(v) = 0\} \tag{6.2}$$

Definition 6.4: Range

Let $L : V \to W$ be a linear transformation. Then the range of L is defined as

$$\operatorname{Range} L = \{L(v) : v \in V\} \tag{6.3}$$

Proposition 6.1

For $L : V \to W$ a linear transformation, $\operatorname{Null} L$ is a subspace of V and $\operatorname{Range} L$ is a subspace of W.

Proof. First we show that $\operatorname{Null} L$ is a subspace of V. We will show the three requirements for a subspace

1. Since $0v = 0$ for any $v \in V$, we have $0 \in \operatorname{Null} L$
2. Let $x, y \in \operatorname{Null} T$. Then $Lx = 0$ and $Ly = 0$. Then, we have $L(x + y) = L(x) + L(y) = 0 + 0 = 0$. So $x + y \in \operatorname{Null} T$.
3. Let $x \in \operatorname{Null} T$ and $c \in F$. Then $Lx = 0$ and so $L(cx) = cL(x) = c(0) = 0$. So $cx \in \operatorname{Null} T$.

So we have $\operatorname{Null} L$ is a subspace of V. Now we will show that $\operatorname{Range} L$ is a subspace of W.

1. Since $0 \in V$, we have $L(0) = 0$. So $0 \in \operatorname{Range} L$
2. Let $x, y \in \operatorname{Range} L$. Then there exists $v_1 \in V$ such that $L(v_1) = x$ and there exists $v_2 \in V$ such that $L(y) = v_2$. Then we have $L(x + y) = L(x) + L(y) = v_1 + v_2$. So

□

This theorem will only work for finite dimensional vector spaces.

Theorem 6.1

Let V, W be finite dimensional vector spaces over F. Let $B = \{v_1, ..., v_n\}$ be a basis of V. Then given any $w_1, ..., w_n \in W$, there exists a unique linear transformation $L : V \to W$ such that $L(v_i) = w_i$ for all $i \in \{1, ..., n\}$.

Proof. Let $v \in V$. Then we know there exists unique scalars $a_1, ..., a_n$ such that $v = a_1v_1 + \cdots + a_nv_n$. Now define $L(v) = a_1w_1 + \cdots + a_nw_n$. We need to confirm that L is linear. So let $v, v' \in V$ and let $c \in F$. Then there exist unique scalars $a_1, ..., a_n$ such that $v = \sum_{i=1}^{n} a_iv_i$ and $a'_1, ..., a'_n$ such that $v' = \sum_{i=1}^{n} a'_iv_i$. Then $cv+v' = (ca_1+a'_1)v_1+\cdots+(ca_n+a'_n)v_n$. Note that this representation is unique in terms of $v_1, ..., v_n$. So by definition , we have $L(cv+v') = (ca_1+a'_1)w_1+\cdots+(ca_n+a'_n)w_n = c(a_1w_1+\cdots+c_nw_n)+(a'_1w_1+\cdots+a'_nw_n) = cL(v) + L(v')$. Thus, we have L is linear.

Now we show the uniqueness part. So let $L_1(v_i) = L_2(v_i)$ for all $i \in \{1, ..., n\}$. Then we have $L_1(v_i) - L_2(v_i) = 0$. This gives us that $(a_1w_1 + \cdots + a_nw_n) - (a_1w_1 + \cdots + a_nw_n) = 0$. Since $a_1v_1 + \cdots + a_nv_n$ is the unique linear combination, we have $L_1 = L_2$ since v is arbitrary. □

In order to further our study of linear transformations, we will introduce the concept of an ordered basis.

Definition 6.5: Ordered Basis
Let V be a finite dimensional vector space over F. Then an ordered basis of V is an n-tuple $(v_1, ..., v_n)$ where $n = \dim V$ and $\{v_1, ..., v_n\}$ is a basis of V.

Let $L : V \to W$ be a linear transformation and $B' = (w_1, ..., w_n)$ be an ordered basis of W. Then, given any $v \in V$, $L(v) = a_1w_1 + \cdots + a_nw_n$ for unique scalars $a_1, ..., a_n \in F$ (since $L(v)$ is in W, any vector can be represented in this manner). Using this, we define the notation the notation

$$[L(v)]_{B'} \coloneqq \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} \tag{6.4}$$

This vector in F^n is the B'-coordinate vector of $L(v)$.

Lecture 7

Properties Of The Null Space And Range

Recall that a linear transformation between vector spaces V and W over a field F, is a function $T : V \to W$ such that

1. $T(x + y) = T(x) + T(y)$ for all $x, y \in V$
2. $T(kx) = kT(x)$ for all $x \in V$ and $k \in F$

Mainly, a linear transformation preserves linear combinations between vectors. So we have $T(ax + by) = aT(x) + bT(y)$. Here are some examples of linear transformations.

- From linear algebra: matrix-vector products are linear transformations. So $A(u + v) = Au + Av$ and $A(cu) = c(Au)$.
- From calculus: derivatives and anti-derivatives are linear transformations.
- From geometry: rotations and reflections are linear transformations.

7.1 Null Spaces & Ranges

Two important spaces associated with a linear transformation are the null space (kernel) and range. Two important facts about these subspaces are given below

- The null space of T is a subspace of V while the range is a subspace of W.
- The range of a linear transformation is $\text{Range}\, T = \text{span}\{T(v_1), T(v_2), ..., T(v_n)\}$ where $\{v_1, v_2, ..., v_n\}$ is a basis of V. In other words, we only need to know the images of the basis vectors to generate the range of the transformation.

While the null space and range seem unrelated, they are actually somewhat connected. This will be the goal of today's lecture. We will introduce two new terms in this part.

Definition 7.1: Rank & Nullity
For a linear transformation $T : V \to W$, we have

1. The rank of a transformation is the dimension of its range.
2. The nullity of a transformation is the dimension of its null space.

With this in mind, we now present the following theorem, which details a very simple, but incredibly useful, relationship between the two terms.

Theorem 7.1
Let $T : V \to W$ be an arbitrary linear transformation where V and W are finite. Then we have that

$$\operatorname{rank} T + \operatorname{nullity} T = \dim V \tag{7.1}$$

Before presenting a formal proof, we will present a concrete example.

Example 7.1

Consider a linear transformation $T : \mathbb{R}^3 \to \mathbb{R}^2$ given by the rule $T(x_1, x_2, x_3) = (x_1 - x_2, 2x_3)$. The null space is given by $x_3 = 0$ and $x_1 = x_2$. We can see that the dimension is 1 since a basis for the null space is $\{(1, 1, 0)\}$.

We claim that the range of the transformation is all of $\mathbb{R}^2$. For arbitrary (y_1, y_2), we can argue that there is always a preimage of this vector in $\mathbb{R}^3$. This can be done by solving a linear system.

So here is a proof of the theorem

Proof. Assume $\dim V = n$ and $\dim(\operatorname{Null} T) = k$. Note that by a previous theorem, we know that $k \leq n$. To find the rank of T, we will generate a basis for the range. Let $\{v_1, v_2, ..., v_k\}$ be a basis for $\operatorname{Null} T$. Now, we can extend this basis to $\{v_1, ..., v_k, v_{k+1}, v_{k+2}, ..., v_n\}$, which is a basis of V. Note that there are $n - k$ vectors that were added to the basis. Now consider the vectors $\{T(v_{k+1}), T(v_{k+2}), ..., T(v_n)\}$. Now we claim that this set is a basis for the range of T.

So first we show that $\operatorname{Range} T = \operatorname{span}\{T(v_1), T(v_2), ..., T(v_k), T(v_{k+1}, ..., T(v_n)\}$. But by choice, we have $\{v_1, ..., v_k\}$ is a basis for the null space. So we get $\operatorname{Range} T = \operatorname{span}\{T(v_{k+1}, ..., T(v_n)\}$, as desired.

Now we need to show linear independence. So assume $\sum_{k+1}^{n} b_i T(v_i) = 0$. We will show that $b_i = 0$. Since T is linear, we have

$$\sum_{k+1}^{n} b_i T(v_i) = T\left(\sum_{k+1}^{n} b_i v_i\right) = 0 \tag{7.2}$$

So we have that $\sum b_i v_i$ is in the null space of T. So we observe that there should be a way to represent this vector as a linear combination of the basis vectors of the null space. This implies

$$\sum_{i=k+1}^{n} b_i v_i - \sum_{i=1}^{k} c_i v_i = 0 \tag{7.3}$$

And since basis vectors are linearly independent, we have $b_i = 0$ and $c_i = 0$ □

7.2 Injections & Surjections

Recall that we have the definitions of one-to-one functions and onto functions from Math 300.

Definition 7.2: Injective Functions & Surjective Transformations
et $T : V \to W$ be an arbitrary linear transformation. Then, we say T is

1. *One-to-one* (or injective) if $T(v_1) = T(v_2)$ implies $v_1 = v_2$ for all $v_1, v_2 \in V$
2. *Onto* (or surjective) if for all $w \in W$, there exists $v \in V$ such that $T(v) = w$

We can use the null space and range to determine whether or not a transformation is one-to-one or onto. The results are stated in the following theorems.

Theorem 7.2
A linear transformation $T : V \to W$ is one-to-one if and only if $\operatorname{Null} T = \{0\}$.

Proof. Proof of ($\implies$): assume T is one-to-one. Then assume $T(x) = 0$. But from the definition of a linear transformation, we have that $T(0) = 0$. Since T is one-to-one, we have $x = 0$ and so $\operatorname{Null} T = \{0\}$.

Proof of ($\impliedby$): assume $\operatorname{Null} T = \{0\}$. Assume that $T(x) = T(y)$. Then, observe that we have

$$\begin{aligned} T(x) &= T(y) \\ T(x) - T(y) &= 0 \\ T(x - y) &= 0 \end{aligned}$$

Then, we have that $x - y \in \operatorname{Null} T$. But the only vector in the null space is 0. So we must have $x - y = 0$. It follows that $x = y$. □

From this fact, we have a simple corollary.

Corollary 7.1
Let $T : V \to W$ be a linear transformation with V and W finite and with the same dimension. Then the following are equivalent.

1. T is one-to-one
2. T is onto
3. $\operatorname{Range} T = W$

Proof. Proof of (1) $\implies$ (3): if T is one-to-one so $\operatorname{Null} T = \{0\}$. Therefore, $\operatorname{nullity} T = 0$ and so $\operatorname{rank} T = n$. By the rank-nullity theorem, we have that $\dim V = n = \dim W$. Now observe that by definition, $\operatorname{rank} T = n = \dim(\operatorname{Range} T)$. Now observe that $\operatorname{Range} T$ is a subspace of W. By a previous result, we have $\operatorname{Range} T = W$, as desired.

Proof that (2) $\implies$ (3): assume that T is onto. □

Here is one last theorem, that will be used to connect this topic to the next section.

Theorem 7.3
Let $\{v_1, v_2, ..., v_n\}$ be a basis of V. Let $\{w_1, ..., w_n\}$ be n vectors in W (they don't have to be distinct). Then, there exists a unique linear transformation $T : V \to W$ such that $T(v_i) = w_i$.

Lecture 8

Matrix Representations of Linear Transformations

8.1 Matrix Representations

We will revisit the idea of coordinate vectors in a particular basis. Let V be a finite dimensional vector space over F. Let $B = (b_1, b_2, ..., b_n)$ be an ordered basis of V. Then given $v \in V$, we let the notation

$$[v]_B = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \in F^n \tag{8.1}$$

where $v = \sum_{i=1}^{n} a_i b_i$. From the properties of a basis, this representation is unique. We call this vector the B-coordinate vector of v. We will now define a notation that can help us express any linear transformation.

Definition 8.1: Matrix Representation

Let V, W be finite dimensional vector spaces over F. Let $B = (b_1, b_2, ..., b_n)$ be an ordered basis of V and $\Gamma = (g_1, g_2, ..., g_m)$ be an ordered basis of W. Let $L : V \to W$ be a linear transformation. Then, we define the matrix denoted $[L]_B^\Gamma$ where the jth column is given as $[L(b_j)]_\Gamma$ for $j = 1, ..., n$. This matrix is called the matrix representation of L.

Note that both V and W are isomorphic to F^n and F^m. From Math 250, we know that a linear transformation from these two vector spaces can be represented as a $m \times n$ matrix. Some more examples are given below

Example 8.1

Let $V = P_3(\mathbb{R})$ and $W = P_2(\mathbb{R})$. The differentiation operation (taking derivatives) is a linear transformation. So $D : V \to W$ given by $D(f) = df/dx$. Then take $B = (1, x, x^2, x^3)$ and $\Gamma = (1, x, x^2)$ be ordered

bases for V and W, respectively. Then we have that

$$D(1) = 0$$
$$\left[D(1)\right]_\Gamma = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$
$$D(x) = 1$$
$$\left[D(x)\right]_\Gamma = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$
$$D(x^2) = 2x$$
$$\left[D(x^2)\right]_\Gamma = \begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix}$$
$$D(x^3) = 3x^2$$
$$\left[D(x^3)\right]_\Gamma = \begin{bmatrix} 0 \\ 0 \\ 3 \end{bmatrix}$$

So we can conclude that the matrix representation of this transformation in these two bases is

$$\left[D\right]_B^\Gamma = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}$$

Here is a proposition that can show us to how to use the generated matrix

Proposition 8.1

Let V, W be finite dimensional vector spaces over F. Let $B = (b_1, b_2, ..., b_n)$ be an ordered basis of V and $\Gamma = (g_1, g_2, ..., g_m)$ be an ordered basis for W. Let $L : V \to W$ be a linear transformation. Then given $v \in V$, we have

$$\left[L(v)\right]_\Gamma = \left[L\right]_B^\Gamma \left[v\right]_B \tag{8.2}$$

Proof. □

We can represent this using a commutative diagram.

$$V$$

Here are two important definitions that will be used to define a vector space of linear transformations

Definition 8.2: Addition & Scalar Multiplication of Linear Transformations

Let V, W be vector spaces over F. Let L_1, L_2 be linear transformations from $V \to W$. Define $(L_1 + L_2) : V \to W$ as $(L_1 + L_2)(x) = L_1(x) + L_2(x)$ for all $x \in V$. Define $\alpha L_1 : V \to W$ as $(\alpha L_1)(x) = \alpha L(x)$ for all $\alpha \in F$ and $x \in V$

Proposition 8.2

The functions $L_1 + L_2$ and αL_1 defined above are linear transformations.

Proof. We have

$$\begin{aligned}(L_1 + L_2)(x + y) &= L_1(x + y) + L_2(x + y) \\ &= L_1(x) + L_1(y) + L_2(x) + L_1(y) \\ &= L_1(x) + L_2(x) + L_1(x) + L_1(y) \\ &- (L_1 + L_2)(x) + (L_1 + L_2)(y)\end{aligned}$$

Also, we have

$$(\alpha(L_1 + L_2))(x + y) =$$

□

Now we can define the vector space that contains all linear transformations between two vector spaces.

Definition 8.3: Vector Space of Linear Maps

Let $\mathcal{L}(V, W)$ be the set of all linear transformations from V to W, where V, W are vector spaces over F.

Proposition 8.3

The set $\mathcal{L}(V, W)$ as defined above is a vector space over F.

Proof. First we show that the set is an abelian group under vector addition.

1. The 0 map where $L(v) = 0$ for all $v \in V$ is the identity map
2. By definition, we have $(L_1 + L_2)(x) = L_1(x) + L_2(x) = L_2(x) + L_1(x) = (L_2 + L_1)(x)$ for all $x \in V$. So the group is abelian
3.
4.

□

From above, if V and W are finite, we know that every linear transformation has a matrix representation. So we can realize the set $\mathcal{L}(V, W)$ as the set of $M_{m\times n}(F)$, the set of all $m \times n$ matrices

Lecture 9

Invertibility Of Linear Transformations & Isomorphisms

9.1 Invertibility Of Linear Transformations

We define the inverse of a linear transformation below

Definition 9.1: Inverse Of A Linear Transformation
Let $L : V \to W$ be a linear transformation. If there exists $U : W \to V$ such that $U \circ L = \mathbb{1}_V$ and $L \circ U = \mathbb{1}_W$ then we say L is invertible.

For simplicity of notation, we shorten $U \circ L$ as UL. Also note that we have that L is invertible if and only if L is bijective (both injective and surjective).

Example 9.1

Let $T : P_1(\mathbb{R}) \to \mathbb{R}^2$ given by the rule $T(a+bx) = (a, a+b)$. Note that T is indeed a linear transformation (check it later). Then the inverse of this transformation is $U(c, d) = c + (d - c)x$. Check (later) that this is indeed the inverse by showing TU and UT both equal the identity.

Before moving forward, here are some notations. For a linear transformation $L : V \to W$ then the inverse $U : W \to V$ is unique. We denote this function as L^{-1}.

Here are some important propositions regarding the inverse of a linear transformation.

Proposition 9.1
If V, W are vector spaces over a field F and let $L : V \to W$ be an invertible linear transformation. Then we have that $L^{-1} : W \to V$ is a linear transformation

Proof. We need to show $L^{-1}(w_1 + cw_2) = L^{-1}(w_1) + cL^{-1}(w_2)$. Then, it is sufficient to show that

$$\begin{aligned} L(L^{-1}(w_1 + cw_2)) &= L(L^{-1}(w_1)) + cL(L^{-1}(w_2)) \\ &= w_1 + cw_2 \end{aligned}$$

Since L is bijective, the equalities above are valid. □

Here is a corollary to this proposition.

Corollary 9.1
Let $L : V \to W$ be a linear transformation that is invertible. Then V is finite dimensional if and only if W is finite dimensional. Furthermore, if both are finite dimensional, then $\dim V = \dim W$.

9.1.1 Matrix Representations

Now we will present some results concerning the matrix representation of a linear transformation.

Theorem 9.1
Let V, W be finite dimensional vector spaces over F. Let $T : V \to W$ be a linear transformation. Also, let β and γ be ordered bases for V and W respectively. Then T is invertible if and only if the matrix $[T]_\beta^\gamma$ is invertible. Moreover, the matrix $[T^{-1}]_\gamma^\beta = \left([T]_\beta^\gamma\right)^{-1}$.

Proof. Since T is invertible, we know the inverse exists. Now, we know that there is a matrix $B = [T^{-1}]_\gamma^\beta$. We claim that this matrix is the inverse of $[T]_\beta^\gamma$. Note that matrix multiplication corresponds to composition of the linear transformations. Then, we have $BA = [T^{-1}T]_\beta^\beta = I$ and $AB = [TT^{-1}]_\gamma^\gamma = I$ since $T^{-1}T$ and TT^{-1} are the identity functions. □

Example 9.2
Let F^n be a vector space over field F (has dimension n). Let A be an $n \times n$ matrix with entries in F. Define $L_A : F^n \to F^n$ using the rule

$$L_A\left(\begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}\right) = A \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}$$

Note that L_A is invertible if and only if A is invertible (corollary below)

Corollary 9.2
The linear transformation corresponding to multiplying by a matrix A is invertible if and only if A is invertible.

Proof. Let $\beta = (e_1, e_2, ..., e_n)$ be the standard basis for F^n. Then observe that $[L_A]_\beta^\beta = A$. Then, the result follows from the previous theorem. □

9.2 Isomorphisms

We now define an isomorphism between two vector spaces.

Definition 9.2: Isomorphic Vector Spaces
Let V and W be vector spaces over F. Then we say V and W are isomorphic if and only if there exists an inveritble linear transformation $L : V \to W$.

We use the notation $V \cong W$ to say that V and W are isomorphic. Here is a theorem that establishes isomorphic vector spaces, if they are finite dimensional.

Theorem 9.2
Let V and W be finite dimensional vector spaces over F. Then V and W are isomorphic if and only if $\dim V = \dim W$.

Lecture 10

Isomorphism Between Linear Transformations & Matrices

Recall that we have defined the vector space $\mathcal{L}(V,W) = \{L : V \to W \mid L \text{ is linear}\}$. Here is the first proposition about this vector space

Theorem 10.1
Let V, W be finite dimensional vector spaces over a field F and let β and γ be ordered basis of V and W respectively. Then, define the function $\Phi_\beta^\gamma : \mathcal{L}(V,W) \to M_{m\times n}(F)$ as follows

$$\Phi_\beta^\gamma(T) = [T]_\beta^\gamma$$

Then, we have that Φ_β^γ is an isomorphism and thus, $\dim \mathcal{L}(V,W) = mn$.

Proof. We prove that Φ_β^γ is linear as follows

$$\begin{aligned}\Phi_\beta^\gamma(cT_1 + T_2) &= [cT_1 + T_2]_\beta^\gamma \\ &= [cT_1]_\beta^\gamma + [T_2]_\beta^\gamma \\ &= c[T_1]_\beta^\gamma [T_2]_\beta^\gamma \\ &= c\Phi_\beta^\gamma(T_1) + \Phi_\beta^\gamma(T_2)\end{aligned}$$

And so we have Φ_β^γ is linear.

Now we will define $\phi : M_{m\times n} \to \mathcal{L}(V,W)$. For $A \in M_{m\times n}(F)$, define a linear transformation $\phi(A) : V \to W$ as follows: for $v \in V$, we have $\phi(A)(v) = A[v]_\gamma$ where $[v]_\gamma$ is the $\square$

Lecture 11

Change Of Basis & Dual Spaces

Last time we saw that $\mathcal{L}(V,W) \cong M_{m\times n}(F)$. In other words, any linear transformation can be represented by a matrix. The proof utilizes the commutative diagram for linear transformations. Namely, we have ϕ_β is an isomorphism from V to F^n and ϕ_γ is an isomorphism from W to F^n. The functions are given by

$$\phi_\beta(v) = [v]_\beta$$

11.1 Change Of Basis

Here we define the change of basis

Definition 11.1: Change Of Basis
Let V be a finite dimensional vector space over F. Let β and β' be two (distinct) bases of V. Recall that we have the identity transformation $\mathbf{1}_V : V \to V$ and define $Q \coloneqq [\mathbf{1}_V]_\beta^{\beta'}$

Theorem 11.1
Let V, β, and β' be defined as above. Then we have

1. Q is invertible
2. $[v]_\beta = Q[v]_{\beta'}$

Proof. 1. Define $P = [\mathbf{1}_V]_\beta^{\beta'}$. We will show $PQ = QP = I$. We have

$$PQ = [\mathbf{1}_V]_\beta^{\beta'}[\mathbf{1}_V]_{\beta'}^{\beta} = [\mathbf{1}_V \circ \mathbf{1}_V]_{\beta'}^{\beta'} = [\mathbf{1}_V]_{\beta'}^{\beta'}$$

The other direction for $QP = I$ is similar.

2. PSS 1.

□

The theorem above relates the matrices of the change of basis. Now, we will consider the transformations themselves.

Theorem 11.2
Let V be a finite dimensional vector space over F. Let $T : V \to V$ be a linear transformation and β and β' be ordered bases of V. Define $Q = [\mathbf{1}_V]_{\beta'}^{\beta}$. Then, we have $[T]_{\beta'}^{\beta'} = Q^{-1}[T]_\beta^\beta Q$.

Proof. It is enough to show $Q[T]_{\beta'}^{\beta'} = [T]_{\beta}^{\beta}Q$. Note that $Q = [\mathbf{1}_V]_{\beta'}^{\beta}$. So we have

$$\begin{aligned} Q[T]_{\beta'}^{\beta'} &= [\mathbf{1}_V]_{\beta'}^{\beta}[T]_{\beta'}^{\beta'} \\ &= [\mathbf{1}_V \circ T]_{\beta'}^{\beta} \\ &= [T]_{\beta'}^{\beta} \end{aligned}$$

And similarly, we have

$$\begin{aligned} [T]_{\beta}^{\beta}Q &= [T]_{\beta}^{\beta}[\mathbf{1}_V]_{\beta'}^{\beta} \\ &= [T \circ \mathbf{1}_V]_{\beta'}^{\beta} \\ &= [T]_{\beta'}^{\beta} \end{aligned}$$

So the desired result is obtained. □

The form of the theorem above is very common in linear algebra and it has a special name.

Definition 11.2: Similarity/Conjugates

Let $P, Q \in M_{n\times n}(F)$. If there exists $R \in GL_n(F)$ such that $P = RQR^{-1}$, they P and Q are said to be similar or they are conjugates of each other.

Proposition 11.1

Similarity of matrices is an equivalence relation in $M_{n\times n}(F)$

Proof. PSS 2 □

11.2 Dual Spaces

Now we will introduce the notion of a dual space.

Definition 11.3: Dual Space

Let V be a vector space over a field F. Let $\mathcal{L}(V, F)$ be the set of all linear transformations from V to its field of scalars F. Then the set $V^\star = \mathcal{L}(V, F)$ is called the dual space of V.

If V is a finite dimensional vector space, then we can see $\dim(V^\star) = \dim V$. This is because we proved that $\dim \mathcal{L}(V, W)$ is $\dim V \cdot \dim W$. But the dimension of F over itself is just 1 (since a basis would just be $\{1\}$). So the result follows.

This next theorem proposes a basis for the dual space.

Theorem 11.3

Let V be a vector space over F and let B be a basis of V. For all $b \in B$, define $b^\star : V \to F$ as

$$b^\star(v) = \begin{cases} 1 & \text{if } v = b \\ 0 & \text{if } v \neq b \end{cases}$$

Then the set $B^\star = \{b^\star : b \in B\}$ is a basis of $V^\star$ if V is finite dimensional.

Proof. Note that any $T \in V^\star$ is determined by its values on the basis vectors in B. Let $B = \{b_1, ..., b_n\}$. Then we have

$$T = T(b_1)b_1^\star + T(b_2)b_2^\star + \cdots + T(b_n)b_n^\star$$

So we see that $\operatorname{span} B^\star = V^\star$. Now we will prove that the set $B^\star$ is linearly independent (PSS 3). □

Lecture 12

More Dual Spaces & Rank Of Matrices

12.1 Dual Spaces (Continued)

In the last lecture, we ended our discussion of dual spaces by proposing a basis for the dual space. Now we will prove some more theorems regarding dual spaces.

Theorem 12.1
Let V and W be finite dimensional vector spaces over F. Let $\beta = (b_1, b_2, ..., b_n)$ and $\gamma = (g_1, g_2, ..., g_m)$ be ordered bases of V and W respectively. Suppose $T \in \mathcal{L}(V, W)$ and define $T^\star : W^\star \to V^\star$ as $T^\star(f) = f \circ T$. Then we have

1. $T^\star \in \mathcal{L}(W^\star, V^\star)$. So $T^\star$ is a linear transformation between the dual spaces $W^\star$ and $V^\star$.
2. $[T^\star]_{\gamma^\star}^{\beta^\star} = \left([T]_\beta^\gamma\right)^T$

Proof. 1. $T^\star$ is linear since it is the composition of two linear transformations.

2. Denote $[T]_\beta^\gamma$ as A. We need to show that the (i,j)-th element of $[T^\star]_{\gamma^\star}^{\beta^\star}$ is A_{ji}. Note that the j-th column of $[T^\star]_{\gamma^\star}^{\beta^\star}$ corresponds to the coefficients for $T^\star(g_j^\star) = g_j^\star \circ T$. Now, we get that $g_i^\star \circ T = \sum_{k=1}^n c_k b_k^\star$ for some $c_1, ..., c_n \in F$ (since this element is in the dual space and so it has a unique representation in the dual basis). Then, we have that the i-th column of $[T^\star]_{\gamma^\star}^{\beta^\star}$ is

$$\begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$$

Now for $f \in V^\star$, we have $f = \sum_{k=1}^n c_k b_k^\star$. Now observe that we have

$$\begin{aligned} f(b_i) &= \left(\sum_{k=1}^n c_k b_k^\star\right)(b_j) \\ &= \sum_{k=1}^n c_k b_k^\star(b_j) \\ &= \sum_{k=1}^n c_k \delta_{ki} = c_i \end{aligned}$$

So we have $c_i = f(b_i)$. So the (i,j)-th element of $[T^\star]_{\gamma^\star}^{\beta^\star}$ is $(g_i^\star \circ T)(b_i) = g_i^\star(T(b_i))$. Note that $T(b_i) = \sum_{k=1}^m d_k g_k$ (since it is in W). So we have $g_i^\star(T(b_i)) = d_j$, which is the j-th element of the i-th column of A, which is precisely A_{ji}. Thus, the desired result is obtained.

□

12.2 Double Dual

Now we will describe the double dual of a vector space.

Definition 12.1: Double Dual Space

For a finite dimensional vector space V over a field F, the double dual space of V, denoted $V^{\star\star}$, is defined to be $\mathcal{L}(V^\star, F)$.

To realize the elements of the double dual, consider $x \in V$. We want to define $\hat{x} : V^\star \to F$, a linear map. We will define $\hat{x} = f(x)$ where $f \in V^\star$ (where $f : V^\star \to F$, is in the dual space of V). The following theorem characterizes some nice properties about this function.

Theorem 12.2

Let V be a finite dimensional vector space over F. Then $\phi : V \to V^{\star\star}$ defined as $\phi(x) = \hat{x}$ (defined in the preceding discussion). Then the function ϕ is an isomorphism.

Proof. First, show ϕ is linear. We need to show $\phi(c_1 x_2 + x_2) = c_1\widehat{x_1 + x_2} = c_1\hat{x_1} + \hat{x_2}$.

Now note that $\dim(V) = \dim(V^\star) = \dim(V^{\star\star})$. So we only need to show that ϕ is one-to-one (since rank-nullity implies it is onto for equal dimensional vector spaces). So assume $\phi(x) = 0$. So $\hat{x} : V^\star \to V$ is a zero map. For contradiction, assume $x \neq 0$. Since $\{x\}$ is linearly independent, extended $\{x\}$ to an ordered basis $(b_1, ..., b_n)$ where $b_1 = x$ and $n = \dim V$. So we have that $\hat{x}(b_1^\star) = b_1^\star(x) = b_1^\star(b_1) = 1 \neq 0$, which contradicts the fact that $\hat{x}$ is a zero map. So we must have that $x = 0$. □

The following corollary is to be proved as an exercise.

Corollary 12.1

Let V be a finite dimensional vector space over F. Then every ordered basis of $V^\star$ is dual to an ordered basis of V.

12.3 Rank Of Matrices

We have seen the rank of a linear transformation; now we will define the rank of a matrix.

Definition 12.2: Rank Of A Matrix

Let $A \in M_{n\times n}(F)$. Then the rank of A is the dimensional of the span of its columns. So $\operatorname{rank} A = \dim(\operatorname{span}\{a_1, ..., a_n\})$.

Theorem 12.3

Let $A \in M_{n\times n}(F)$, $P \in GL_n(F)$, and $Q \in GL_n(F)$. Then, we have

1. $\operatorname{rank}(AQ) = \operatorname{rank} A$
2. $\operatorname{rank}(PA) = \operatorname{rank} A$
3. $\operatorname{rank}(PAQ) = \operatorname{rank} A$

Proof. Here is a proof of (2): define $L_A : F^n \to F^n$. Then, we have that

$$\begin{aligned} \operatorname{rank} A &= \dim(\operatorname{Range} L_A) \\ &= \dim(L_A(F^n)) \end{aligned}$$

Similarly, we have

$$\begin{aligned} \operatorname{rank}(PA) &= \dim(\operatorname{Range} L_{PA}) \\ &= \dim(L_{PA}(F^n)) \\ &= \dim(L_P(L_A(F^n))) \end{aligned}$$

So we have $L_P : K_A(F^n) \to F^n$. Since P is invertible, we have nullity $L_P = 0$. So since P is invertible, we have $\dim(L_A(F^n)) = \dim(L_P(L_A(F^n)))$, which is the desired result. □

Note that from elementary row operations, we can view it as left multiplying an elementary row matrix. But by the theorem above, we see that this does not change the rank. In other words, applying an elementary row operation to a matrix preserves its rank. Since a reduced row echelon form of a matrix is obtained through repeated elementary row operations, the reduced row echelon form of a matrix has the same rank as the original.