Math 146 Notes

Eason Li

2024 W

Lecture 1 - Mon - Jan 8 - 2024

Vector Spaces

For F is a field, suppose V is an F- vector space, then

V is an abelian group under addition $(+)$ with

1. Commutativity
2. Associativity
3. Additive Identity
4. Additive Inverse

In addition, V **has scalar multiplication**:

$$\cdot : F \times V \to V$$

in particular, $\lambda \in F,\ v \in V \to \lambda v \in V$, and

1. $1 \cdot v = v$
2. $(\lambda_1 + \lambda_2) \cdot v = \lambda_1 v + \lambda_2 v$
3. $\lambda(v_1 + v_2) = \lambda v_1 + \lambda v_2$

Example 0.1

Let $X \neq \varnothing$ be a set, let F be a field, and let $V = \{f : X \to F\}$ is a vector space,

1. $+$ on V is pointwise addition

$$(f + g)(x) = f(x) + g(x)$$

2. 0_v is the function that $0_v(x) = 0, \forall\, x \in X$.
3. $\lambda \in F, f \in V$, then $(\lambda f)(x) = \lambda f(x)$.

Example 0.2

For $n \in \mathbb{N}$ and F field, $F^n = \{a_1, a_2, \ldots, a_n : a_1, \ldots, a_n \in F\}$ is a vector space,

1. $(a_1, \ldots, a_n) + (b_1, \ldots, b_n) = (a_1 + b_1, \ldots, a_n + b_n)$.
2. $\vec{0} = (0, 0, \ldots, 0)$.
3. $\lambda(a_1, \ldots, a_n) = (\lambda a_1, \ldots, \lambda a_n)$.

Example 0.3

If F is a field, $F \subseteq K$ where K is a field extension of F, then K is F- vector space.

Example 0.4

F field, then $V = F[x]$ is F- vector space.

Subspace

Definition 0.1

W is a subspace of V if $W \subseteq V$ is a subgroup and $\forall\, \lambda \in V$, $\forall\, w \in W$ we have $\lambda w \in W$.

More concretely, W is a subspace of V if:

1. $0 \in W$
2. $\forall\, w_1, w_2 \in W, \quad w_1 + w_2 \in W$
3. $\forall\, \lambda \in F$, $\forall\, w \in W, \quad \lambda w \in W$

Remark: We do not need to mention the additive inverse because we can get it for free from 3.

Example 0.5

Let $X = [0, 1]$, $F = \mathbb{R}$, and $V = \{f : X \to F\}$, then the following two are examples of subspaces of V:

1. $V \supseteq W = \{f : X \to F,\ f \text{ continuous}\}$
2. $V \supseteq U = \{f : X \to F,\ f \text{ continuously differentiable}\}$

Linear Mapping

For F is a field, V, W are F- vector spaces, and $T : V \to W$, then for T to preserve the vector spaces' structures, we want:

1. $T(v_1 + v_2) = T(v_1) + T(v_2)$
2. $T(\lambda v) = \lambda T(v),\ \forall\ \lambda \in F,\ v \in V$
3. $T(0_V) = 0_W$

Remark: We can deduce 3 from 1 and 2.

Definition 0.2: Linear map

A map T with properties 1 and 2 is called a **linear map**.

Example: Let V be the real vector space of continuous functions from $[0, 1]$ to $\mathbb{R}$. Let $T : V \to R$ be given by $T(f(x)) = \int_0^1 f(x)dx$. Show that T is a linear map.

Exercise: Let K be a field of characteristic p. Notice that K has a subfield $\{0, 1, 2, \ldots, p-1\}$ that we denote by $\mathbb{F}_p$ and so K is an $\mathbb{F}_p$-vector space. Show that $T : K \to K$ given by $T(x) = x^p$ is a linear map when we regard K as an $\mathbb{F}_p$-vector space.

Therefore, given a linear map

$$T : V \to W \qquad V, W \text{ are } F\text{- vector space}$$

For the kernel of T, $\ker(T) = \{v \in V,\ T(v) = 0_W\}$ and the image of T, $\text{im}(T) = \{T(v) :\ v \in V\}$, we have the following lemma:

Lemma 0.1

$\ker(T) \subseteq V$ is a subspace of V and $\text{im}(T) \subseteq W$ is a subspace of W.

Proof: Definition check lol. □

Lecture 2 - Wed - Jan 10 - 2024

Recall that from last lecture, for F is a field, V, W are F-v.s. and $T : V \to W$, then T is linear (or we call it F-linear) if

$$\begin{aligned} &1.\ T(v_1 + v_2) = T(v_1) + T \quad &&\text{for } v_1, v_2 \in V \\ &2.\qquad T(\lambda v) = \lambda T(v) &&\text{for } \lambda \in F, v \in V \end{aligned}$$

Example 0.6

Suppose $V = \mathbb{R}^3$ and $W = \mathbb{R}^2$, and $T : \mathbb{R}^3 \to \mathbb{R}^2$, $T(x, y, z) = (2x + y, 3x - z)$. Then T is linear.

Example 0.7

uppose $V = \mathbb{C}[x]$ (aka $\mathbb{C}$-v.s.), and $T : V \to V$, $T(p(x)) = p'(x)$, then T is linear.

Example 0.8

uppose K is a field of characteristic $p > 0$, p prime and $K \supseteq \{0, 1, \ldots, p-1\} =: \mathbb{F}_p$. Then K is a $\mathbb{F}_p$-v.s. Let $F : K \to K$, then $F = x^p$ is $\mathbb{F}_p$-linear.

Proposition 0.1

The following holds:

1. If $T : V \to W$ and $S : W \to U$ are linear, then $S \circ T : V \to U$.
2. If $T : V \to W$ is linear and bijective then $T^{-1} : W \to V$ is linear.
3. $I : V \to V$ (identity), $I(v) = v$ is linear.

Proof:

1. Notice

$$\begin{aligned} S \circ T(v_1 + v_2) &= S\left(T(v_1) + T(v_2)\right) \\ &= S\left(T(v_1)\right) + S\left(T(v_2)\right) \\ &= S \circ T(v_1) + S \circ T(v_2) \\ S \circ T(\lambda v) &= S\left(T(\lambda v)\right) \\ &= S\left(\lambda T(v)\right) \\ &= \lambda S(T(v)) \\ &= \lambda S \circ T(v) \end{aligned}$$

2. Notice

$$\begin{aligned} T\left(T^{-1}\left(w_1 + w_2\right)\right) &= w_1 + w_2 \\ T\left(T^{-1}(w_1)\right) + T\left(T^{-1}(w_2)\right) &= w_1 + w_2 \\ \Rightarrow\; T\left(T^{-1}\left(w_1 + w_2\right)\right) &= T\left(T^{-1}(w_1)\right) + T\left(T^{-1}(w_2)\right) \end{aligned}$$

Since T is one-to-one, $T^{-1}(w_1 + w_2) = T^{-1}(w_1) + T^{-1}(w_2)$. Moreover,

$$\begin{aligned} T(T^{-1}(\lambda w)) &= \lambda w \\ T(\lambda T^{-1}(w)) &= \lambda w \\ \Rightarrow \quad T^{-1}(\lambda w) &= \lambda T^{-1}(w) \end{aligned}$$

□

Example 0.9

et $F = \mathbb{C}$ and $V = \mathbb{C}[x]$, we say that $T : \mathbb{C}[x] \to \mathbb{C}[x]$, $T(p(x)) = p'(x)$ is linear, and we can show that $S(p(x)) = xp(x)$ is linear. Thus

$$S \circ T(p(x)) = xp'(x)$$
$$T \circ S(p(x)) = p(x) + xp'(x)$$

Exercise:

$$(T \circ S - S \circ T)(p(x)) = p(x) = I(p(x))$$

Moreover

$$\ker(T) = \{\text{ constant polynomial} \subseteq \mathbb{C}[x]\}$$
$$\operatorname{im}(T) = \mathbb{C}[x]$$
$$\ker(S) = \{0\}$$
$$\operatorname{im}(S) = x\mathbb{C}[x] = \{p(x) : p(0) = 0\}$$

Isomorphism

If there is an isomorphism from V to W, then we write $V \cong W$ and say that V is isomorphic to W or that V and W are isomorphic.

Example:

If V, W are F-v.s. and $T : V \to W$ is linear and bijective, then T is an isomorphism from V to W.

Remark: Isomorphism is reflective, symmetric, and transitive.

Example 0.10

Let $V = F[x]_{\leq n}$ and $W = F^{n+1} = \{(a_1, \ldots, a_n), a_i \in F\}$. Then $V \cong W$.

Linear Independence

Set-up:

Definition 0.3: Linear combination

F field, V is F-vector space, and $S \subseteq V$ is an F-linear combination of S is a sum of the form $\sum_{s \in S} \lambda_i S$ for $\lambda_i \in F$, and $\forall\ s \in S$.

Example 0.11

Let $V = F[x]$, what is a linear combination of $S = \{1, x, x^2, \ldots\}$

Answer:

$$\lambda_0 \cdot 1 + \lambda_1 \cdot x + \lambda_2 \cdot x^2 + \cdots \quad \lambda_i \in F$$

(note this is a polynomial).

Lecture 3 - Fri - Jan 12 - 2024

Recall the definition for linear combination. Alternatively, a linear combination of S can be obtained by taking a finite subset $\{v_1, \ldots, v_n\}$ of S and scalars $c_1, \ldots, c_n \in F$ and form the sum:

$$c_1 v_1 + \cdots + c_n v_n$$

Definition 0.4: Span

We define the **Span** of $S \subseteq V$, which we denote as $\operatorname{span}(S)$ to be the collection of all linear combinations of elements of S.

Example 0.12

et $V = \mathbb{R}^3$, let $S = \{(1,0,0),(0,2,0),(1,3,0)\}$, so

$$\operatorname{span}(S) = \Big\{(a,b,0) : a, b \in \mathbb{R}\Big\}$$

Remark: $\operatorname{span}(S) \neq \mathbb{R}^2$, but $\operatorname{span}(S) \cong \mathbb{R}^2$.

Example 0.13

et $V = F[x]$,
let $S = \{1, x, x^2, \ldots, x^n\}$, so $\operatorname{span}(S) = F[x]_{\leq n}$
let $T = \{1, x^2, x^4, x^6, \ldots\}$, so $\operatorname{span}(T) = F[x^2] \subseteq F[x]$.

Proposition 0.2

Let $S \subseteq V$, then $\operatorname{span}(S)$ is a subspace of V.

Proof: To show that $\operatorname{span}(S)$ is a subspace of V, we must show that

1. $w_1, w_2 \in \operatorname{span}(S) \;\Rightarrow\; w_1 + w_2 \in \operatorname{span}(S)$
2. $\lambda \in F,\; w \in \operatorname{span}(S) \;\Rightarrow\; \lambda w \in \operatorname{span}(S)$
3. $0 \in \operatorname{span}(S)$

:3 $\square$

Definition 0.5

If $\operatorname{span}(S) = V$, then we say S **spans** V.

Definition 0.6: Linear dependent

Stop staring at this and being confused. I really have nothing here.

Example 0.14

Which of the following subsets of $\mathbb{R}^3$ are linear independent?

1. $\{(1,2,3),(1,1,1)\}$ ✓
2. $\{(1,2,3),(-2,-4,-6)\}$ ✗
3. $\varnothing$ ✓
4. $\mathbb{R}^2$ ✗
5. $\{(0,0,0)\}$ ✗

Definition 0.7: Basis

A set $S \subseteq V$ is a **basis** if S **spans** V and S is linear independent.

Tutorial 1 - Mon - Jan 15 - 2024

Category

Definition 0.8

A category C consists of

1. A class (collection) of objects: Ob C
2. A class of morphisms for each $(A,B) \in \text{Ob C} \times \text{Ob C} : \text{Hom(A,B)}$ ("maps" from A to B).
3. Composition $\circ$ between compatible morphisms: $f \in \text{Hom(A,B)}$, $g \in \text{Hom(B,C)} \rightarrow$ $g \circ f \in \text{Hom(A,C)}$.
4. $(h \circ g) \circ f = h \circ (g \circ f)$
5. $\exists\ 1_A \in \text{Hom(A,A)}$ s.t. $g \circ 1_A = g$, $1_A \circ f = f$.

Remark: C is locally small if $\forall\ A,B \in \text{Ob C}$, $\text{Hom}(A,B)$ is a set.
Remark: In a category, objects have some **structures**, morphisms are often **structure-preserving**.

Functor

Definition 0.9: Functor

A functor between categories C and D denoted as $F : C \mapsto D$, consists of

1.
$$\begin{aligned} \text{map } F : \text{Ob C} &\mapsto \text{Ob D} \\ A &\mapsto FA \end{aligned}$$

2.
$$\begin{aligned} \text{map } F : \mathrm{Hom_C(A,B)} &\mapsto \mathrm{Hom_D(FA,FB)} \\ f &\mapsto F(f) \end{aligned}$$

subject to

1. $F(g \circ f) = F(g) \circ F(f)$
2. $F(1_A) = 1_{FA}$

Subcategory

Definition 0.10: Subcategory

Subcategory of C is a category D where

1. $\text{Ob D} \subseteq \text{Ob C}$
2. $\forall\ A, B$ we have $\mathrm{Hom_D(A,B)} \subseteq \mathrm{Hom_C(A,B)}$

If $\forall\ A, B$ we have $\mathrm{Hom_D(A,B)} = \mathrm{Hom_C(A,B)}$, then we say D is a full subcategory.

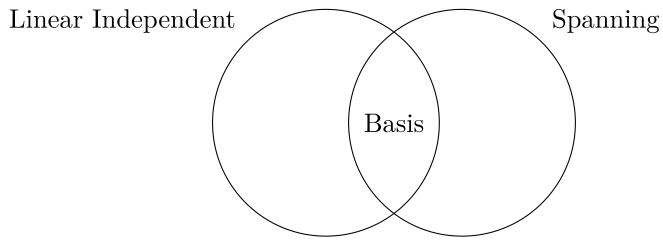
Example 0.15

1. Abelian group is a subcategory of groups (FULL)
2. Fields is a subcategory of rings (FULL)
3. Rng is a subcategory of Ring (NOT FULL)

Lecture 4 - Mon - Jan 15 - 2024

Remark:

1. If $S \subseteq V$, S spans V and $v \in V \backslash S$, then $S \cup \{v\}$ is linear dependent.
2. If $T \subseteq V$ is linear independent and $v \in T$, then $T \backslash \{v\}$ does not span.



Main Facts

1. Every vector space has a basis
2. If V is a vector space, either every basis for V is infinite, or there exists $n \in \mathbb{N} \cup \{0\}$ such that all basis have size n.

Remark: $\varnothing$ is a basis for (0).

Proposition 0.3

If V is a vector space, $S \subseteq T \subseteq V$ are subsets with S linear independent, T spanning V, $|T| < \infty$. Then there exists a basis $\mathcal{B}$ for V with $S \subseteq \mathcal{B} \subseteq T$.

Proof: Let $\mathcal{U}$ be the set of all linear independent subsets U with $S \subseteq U \subseteq T$. Notice $\mathcal{U} \neq \varnothing$ since $S \subseteq \mathcal{U}$. Now let $\mathcal{B}$ be an element of $\mathcal{U}$ of maximal size. **Claim:** $\mathcal{B}$ is a basis for V. STP: $\mathcal{B}$ spans V. SFAC $\text{span}(\mathcal{B}) \neq V$, then $\exists\, v \in T$ such that $v \notin \text{span}(\mathcal{B})$. Consider $\mathcal{B} \cup \{v\}$, it must be linear dependent because of the maximality of $\mathcal{B}$. From which we can conclude that $v \in \text{span}(\mathcal{B})$. Thus we can conclude that $\text{span}(\mathcal{B}) = V$. $\square$

Zorn's Lemma

We first define partially ordered set (poset)

Definition 0.11: Poset

Let P be an non-empty set with $\leq$ binary relation such that the following hold

1. $\forall\, a \in P$, $a \leq a$. (Reflective)
2. $\forall\, a, b, c \in P$, $a \leq b, b \leq c \ \Rightarrow\ a \leq c$. (Transitive)
3. $\forall\, a, b \in P$, $a \leq b, b \leq a \ \Rightarrow\ a = b$. (Anti-symmetric)

Inside poset, we can have **chains**.

Lecture 5 - Wed - Jan 17 - 2024

Recall the definition for poset from last lecture. We provide an example of a poset:

Example 0.16

We take a look at the set $\{1,2,3\}$ and all its subsets, and we define the binary operation to be `\subseteq` : $\subseteq$. It is easy to find that this forms a poset:

$$\begin{tikzcd}
 & \{1,2,3\} \arrow[dl] \arrow[d] \arrow[dr] & \\
\{1,2\} \arrow[d] \arrow[dr] & \{1,3\} \arrow[dl] \arrow[dr] & \{2,3\} \arrow[dl] \arrow[d] \\
\{1\} \arrow[dr] & \{2\} \arrow[d] & \{3\} \arrow[dl] \\
 & \varnothing &
\end{tikzcd}$$

Remark: Taking the proper subsets of $\{1,2,3\}$ would also form a poset with same binary operation.

Exercise: Put a binary relation on the set X of all living things that have ever lived, by declaring that $x \leq y$ if and only if x is an ancestor of y. Is this a partial order on X?

Proof: This really depends on how you think "I am an ancestor of myself" :3 $\square$

Exercise: Let $X = \mathbb{N}$ and declare that $x \leq y$ if $x \mid y$. Is this a partial order? Does X have a least element? Does it have a greatest element.

Proof:

1. It is reflective, since we know $x \mid x$
2. It is transitive, since $x \mid y$ and $y \mid z$ implies $x \mid z$
3. It is antisymmetric sicne $x \mid y$ and $y \mid x$ implies $x = y$

as desired. $\square$

Terminologies

To state the Zorn's Lemma, we need to clarify some terminologies:

1. **Maximal**: An element $m \in (P, \leq)$ is called the **maximal** if whenever $x \in P$ such that $x \geq m$, we have $x = m$.
2. **Chain**: If P is a poset, we say that the subset $C \subseteq P$ is a **chain** if for all $x, y \in C$, either $x \leq y$ or $y \leq x$.
3. **Upper bound**: Given a subset $C \subseteq (P, \leq)$, we say that $x \in P$ is an upper bound for C if $x \geq c$ for all $c \in C$.

Zorn's Lemma

Lemma 0.2: more like an axiom

Let $P \neq \varnothing$ be a poset and suppose that every chain $C \subseteq P$ has an upper bound, then P has at least one maximal element.

Theorem 0.1

Let V be vector space and let $S \subseteq T$ be subset such that S is linear independent and T spans V, then there exists n basis $\mathcal{B}$ for V with $S \subseteq \mathcal{B} \subseteq T$.

Proof:

Warning! The proof is LONG. 💀

Let $\mathcal{U}$ be the set of linear independent subsets U of V with $S \subseteq U \subseteq T$. Notice that $\mathcal{U} \neq \varnothing$ because $S \in \mathcal{U}$. Therefore we can view $\mathcal{U}$ as a poset by declaring

$$U \leq U' \quad \Leftrightarrow \quad U \subseteq U'$$

We will now show that every chain in $\mathcal{U}$ has an upper bound. Notice that a chain is just a collection of sets $\{U_\alpha\}_{\alpha \in Y}$ such that

1. $S \subseteq U_\alpha \subseteq T$, $\forall\, \alpha \in Y$
2. U_α is linear independent, $\forall\, \alpha \in Y$
3. $\forall\, \alpha, \beta \in Y$, $U_\alpha \subseteq U_\beta$ or $U_\beta \subseteq U_\alpha$

Thus we let $U = \bigcup\limits_{\alpha \in Y} U_\alpha$, we claim that $U \in \mathcal{U}$ and it is the upper bound for the chain $\{U_\alpha\}_{\alpha \in Y}$.

It is easy to show that it is the upper bound. However, it is a bit of a work to do to show that $U \in \mathcal{U}$.

First of all, we have $U = \bigcup\limits_{\alpha \in Y} U_\alpha \supseteq U_\alpha \supseteq S$ and each $U_\alpha \in T \ \Rightarrow\ U = \bigcup\limits_{\alpha \in Y} U_\alpha \subseteq T$, which implies that $S \subseteq U \subseteq T$.

Remark: If $X \subseteq V$ is linear dependent, then there exists finite subset $X_0 \subseteq X$ that is linear dependent.

To see that U is linear independent, SFAC that it is linear dependent. By remark, there then exists finite subset $\{u_1, \ldots, u_n\} \subseteq U$ that is linear dependent. Since $U = \bigcup\limits_{\alpha \in Y} U_\alpha$, thus $\forall\, i \in \{1, \ldots, n\}$, $\exists\, \alpha_i \in Y$ such that $u_i \in U_{\alpha_i}$. Hence, $u_1 \in U_{\alpha_1}, \ldots, u_n \in U_{\alpha_n}$. Since U_{α_i}'s form a chain, we know that $\exists i \in \{1, \ldots, n\}$ such that $U_{\alpha_i} \supseteq U_{\alpha_j}$ $\forall\, 1 \leq j \leq n$.

$$\Rightarrow \quad \{u_1, \ldots, u_n\} \in U_{\alpha_i}$$

However, we know that U_{α_i} is linear independent, which is a contradiction (linear dependent set being a subset of a linear independent set). Therefore we obtain that U is linear independent, and thus it is an upper bound for our chain.

By Zorn, we know there exists a maximal $\in \mathcal{U}$, we call it $\mathcal{B}$.

To finish the proof, we show that $\mathcal{B}$ is the basis for V. Since we know $\mathcal{B} \in \mathcal{U}$, $S \subseteq \mathcal{B} \subseteq T$, and $\mathcal{B}$ is linear independent. Additionally, $\mathcal{B}$ spans V, because if it does not, we would have $\text{span}(\mathcal{B}) \not\supseteq T$. (If $\text{span}(\mathcal{B}) \supseteq T \Rightarrow \text{span}(\mathcal{B}) \supseteq \text{span}(T) = V$). Therefore $\exists$ some $t \in T$ such that $t \notin \text{span}(\mathcal{B})$. However, this gives us that $\mathcal{B} \cup \{t\}$ is still linear independent, which contradicts the maximality of $\mathcal{B}$. It follows that $\mathcal{B}$ is linear independent and it spans V, so it is a basis of V. $\square$

Remark:

1. If we take $S \neq \varnothing$, then this says if T spans, then there exists a basis $\mathcal{B} \subseteq T$.
2. If we take $T = V$, then it says that if S is linear independent, then there exists basis $\mathcal{B}$ such that $\mathcal{B} \supseteq S$.

Trailor for next lecture:

Example: We will see that all basis for vector spaces have the same size, and we call the size **dimension**.

Lecture 6 - Fri - Jan 19 - 2024

Goal: show that if V is a vector space, then either

1. All basis for V are infinite, or
2. there exists $n \in \mathbb{N} \cup \{0\}$ such that all basis for V have size n.

We say that V is infinite-dimensional and write $\dim V = \infty$ if (1) folds. Otherwise we say it is finite-dimensional and write $\dim V = n$.

Once we establish this, we have facts:

Result 0.1

(a) If $W \subseteq V$, then $\dim W \leq \dim V$.

(b) If V has a infinite linear independent set, then $\dim V = \infty$.

(c) If V has a finite spanning set, then $\dim V << \infty$.

Proof:

(a) Let $\mathcal{B}$ be a basis for W, then $\mathcal{B} \subseteq V$ and it is linear independent. Thus $\mathcal{B}$ can be expanded to a basis $\mathcal{B}'$ of V and since $|\mathcal{B}| \leq |\mathcal{B}'|$, we have $\dim W \leq \dim V$.

(b) Let $S \subseteq V$ be infinite independent subset of V, then we can expand S to a basis $\mathcal{B}$ and since $\mathcal{B} \supseteq S$, we have $|B| = \infty \Rightarrow \dim V = \infty$.

(c) Similar to part (b).

$\square$

Corollary 0.1

If $V =$ set of all continuous functions from $\mathbb{R}$ to $\mathbb{R}$, then $\dim_{\mathbb{R}} V = \infty$ (sub $\mathbb{R}$ denote $\mathbb{R}$ vector space?).

Remark: We can prove this using HW1Q2 and Result 0.1(b)

Corollary 0.2

$\mathbb{R}$ is infinite-dimensional as $\mathbb{Q}$ - vector space.

Remark: We can prove this using HW1Q3 and Result 0.1(b)

Infinite Dimensional Case

Proposition 0.4

Let V be F - vector space. If V has a infinite basis $\mathcal{B}$, then every basis for V is infinite.

Proof: of prop'n

SFAC there is a finite basis $S = \{v_1, \ldots, v_n\}$. Since $\mathcal{B}$ spans, we can write each v_1 as a linear combination of $\mathcal{B}$, say

$$v_i = \sum_{b \in \mathcal{B}} \lambda_{ib} \cdot b \qquad \lambda_{ib} \in F$$

If we let $\mathcal{B}_i = \Big\{ b \in \mathcal{B} : \lambda_{ib} \neq 0 \Big\}$, then $\mathcal{B}_i$ is finite and $v_i \in \operatorname{span}(\mathcal{B}_i)$. Let $\mathcal{B}' = \mathcal{B}_1 \cup \cdots \cup \mathcal{B}_n$, we easily know that $\mathcal{B}'$ is also finite, and

$$\begin{aligned} \operatorname{span}(\mathcal{B}') \supseteq \operatorname{span}(\mathcal{B}_i) \ni v_i \quad &\Rightarrow \quad S \subseteq \operatorname{span}(\mathcal{B}') \\ &\Rightarrow \quad \operatorname{span}(S) \subseteq \operatorname{span}(\mathcal{B}') \\ &\Rightarrow \quad V \subseteq \operatorname{span}(\mathcal{B}') \\ &\Rightarrow \quad V = \operatorname{span}(\mathcal{B}') \qquad \because \operatorname{span}(\mathcal{B}') \text{ cannot be bigger} \end{aligned}$$

which contradicts the fact that $\mathcal{B}$ is the basis, bacause taking away elements from a basis should not span the whole vector space anymore. □

Finite Dimensional Case

Now we consider the case when V has some (hence all) finite basis.

Suppose $\mathcal{B}_1 = \{u_1, \ldots, u_m\}$, and $\mathcal{B}_2 = \{v_1, \ldots, v_n\}$, thus $|\mathcal{B}_1| = m$ and $|\mathcal{B}_2| = n$. WLOG $m \leq n$, we will show that $m = n$.

Lemma 0.3

Let V be vector space, and let $\{s_1, \ldots, s_p\}$ be linearly dependent set, then there exist $i \in \{0, 1, \ldots, p-1\}$ such that $\{s_1, \ldots, s_i\}$ is linearly independent and $s_{i+1} \in \operatorname{span}\Big(\{s_1, \ldots, s_i\}\Big)$.

Proof: Let $i \in \{0, 1, \ldots, p-1\}$ be the largest index such that $\{s_1, \ldots, s_i\}$ is a linearly independent set. We know that $i \leq p-1$. Because $\{s_1, \ldots, s_p\}$ is linearly dependent, then by the definition of i, the set

$\{s_1, \ldots, s_{i+1}\}$, is linearly dependent, so there exists an non-trivial linear combination of $\{s_1, \ldots, s_{i+1}\}$ equal to 0. Then we have λ_{i+1} is 0 because $\{s_1, \ldots, s_i\}$ is a linearly independent set, thus

$$\begin{aligned} \lambda_{i+1}s_{i+1} &= -\lambda_1 s_1 - \cdots - \lambda_i s_i \\ \Rightarrow \qquad s_{i+1} &= -\frac{\lambda_1}{\lambda_{i+1}}s_1 - \cdots - \frac{\lambda_i}{\lambda_{i+1}}s_i \\ \Rightarrow \qquad s_{i+1} &\in \operatorname{span}\Big(\{s_1, \ldots, s_i\}\Big) \end{aligned}$$

as desired. □

Tutorial 2 - Mon - Jan 22 - 2024

Theorem 0.2: Axiom of Choice

Let C be a collection of non-empty sets (could be infinite, even uncountable), then there exists $f : C \to \bigcup_{A \in C} A$ such that for all $A \in C$, $f(A) \in A$.

Example: If $C = \{A_\alpha\}_{\alpha \in I}$, then there exists I-tuple $(n_\alpha)_{\alpha \in I}$ such that $n_\alpha \in A_\alpha$.

Proposition 0.5

Let $g : A \to B$ be surjective, then there exists $f : B \to A$ such that $g \circ f = \mathrm{id}_B$.

Theorem 0.3: Krull

Let R be a ring, where $0 \neq 1$, then R has at least one maximal ideal.

Theorem 0.4: Tyohonoff

If $\{K_\alpha\}$ is a collection of compact topological spaces, then the product space $\prod_\alpha K_\alpha$ is also compact.

Example 0.17: Controversies with AC

With Krull, we proved the existence of a maximal ideal in a ring $R \neq \{0\}$.

However, we do not know what M (the maximal) looks like.

Example: Consider $\mathbb{R}$ over $\mathbb{Q}$, AC / Zorn tells us that there exists basis $\mathcal{B}$. What does $\mathcal{B}$ look like?

1. **Exercise:** $\mathcal{B}$ is uncountable
2. Take $S = \{\log p : p \text{ prime}\}$, we know S is linearly independent, and we can extend it into a basis $\mathcal{B}_s$.

Lecture 7 - Mon - Jan 22 - 2024

Theorem 0.5

Let V be a vector space, and suppose that V has a basis $\mathcal{B}$ of size $n < \infty$. Then every basis for V has size n.

Lemma 0.4: Exchange Lemma

Let $S = \{u_1, \ldots, u_m\}$ and $T = \{v_1, \ldots, v_n\}$ be subsets of a vector space V, with S spans V and T is linearly independent. Then for $i \in \{0, 1, \ldots, \min(n, m)\}$, there exists a subset S_i of S of size i such that

$$\{v_1, \ldots, v_i\} \cup \{S \backslash S_i\}$$

still spans.

How does the lemma give us the theorem???

Proof: How does the lemma imply the theorem?

Suppose that V has two basis of different sizes,

$$\mathcal{B}_1 = \{u_1, \ldots, u_m\} \quad \& \quad \mathcal{B}_2 = \{v_1, \ldots, v_n\}$$

WLOG, $m \leq n < \infty$.

Suppose that $m < n$, hence we have $\min(n, m) = m$. Taking $i = m$ in the exchange lemma, we see that $\{v_1, \ldots, v_m\}$ spans V. (In this case where we are applying the exchange lemma, we are taking $S = B_1$ and $T = B_2$, so $(S - B_1) = \varnothing$, so $\{v_1, \ldots, v_m\} \cup (S \backslash B_1) = \{v_1, \ldots, v_m\}$.) However, since $n > m$, this means that $\mathcal{B}_2 = \{v_1, \ldots, v_n\}$, is linearly dependent, contradicting the fact that it is a basis. $\square$

Proof: Proof of the Exchange Lemma

Suppose we have

$$S = \underbrace{\{u_1, \ldots, u_m\}}_{\text{spans}} \quad \& \quad T = \underbrace{\{v_1, \ldots, v_n\}}_{\text{lin. ind.}}$$

For $i \in \{0, 1, \ldots, \min(n, m)\}$, we want to find a subset $S_i \subseteq S$ of size i such that

$$\{v_1, \ldots, v_i\} \cup (S \backslash S_i)$$

still spans V.

Base case 1: $i = 0$, so $S_i = \varnothing$

In this case, we simply have

$$\{v_1, \ldots, v_i\} \cup (S \backslash S_i) = S \text{ spans } V v$$

Base case 2: $i = 1$

1. Case 1: $\exists\, j$ such that $u_j = v_1$

   We can take $S_1 = \{u_j\}$, then we would have

$$(S \backslash S_i) \cup \{v_1\} = S$$

2. Case 2: $\nexists\, j$ such that $u_j = v_1$

   Consider $\{v_1, u_1, \ldots, u_m\}$, we know that it is linearly dependent since $v_1 \notin S$ and S spans.

   By our criterion, there exists $i \in \{1, \ldots, m\}$ such that $\{v_1, u_1, \ldots, u_{i-1}\}$ is linearly independent and $\{v_1, u_1, \ldots, u_i\}$ is linearly dependent, so $u_i \in \operatorname{span}(\{v_1, u_1, \ldots, u_{i-1}\})$. Let $S_1 = \{u_i\}$. We claim that

$$\begin{aligned} \{v_1\} \cup (S \backslash \{u_i\}) \quad &\text{still spans} \\ X := &\{v_1, u_1, \ldots, u_{i-1}, u_{i+1}, \ldots, u_m\} \end{aligned}$$

   Notice that $u_1, \ldots, u_{i-1}, u_{i+1}, \ldots, u_m \in \operatorname{span}(X)$ and $u_i \in \operatorname{span}(\{v_1, u_1, \ldots, u_{i-1}\}) \subseteq \operatorname{span}(X)$, so

$$\begin{aligned} &\{u_1, u_2, \ldots, u_m\} \subseteq \operatorname{span}(X) \\ \Rightarrow\ &\operatorname{span}(\{u_1, u_2, \ldots, u_m\}) \subseteq \operatorname{span}(X) \\ \Rightarrow\ &\operatorname{span}(S) = V \subseteq \operatorname{span}(X) \subseteq V \\ \Rightarrow\ &V = \operatorname{span}(X) \end{aligned}$$

Now suppose the claim holds for $i \leq k$, $2 \leq k < \min(n, m)$

Inductino step: $i = k$ By induction hypothesis, there exists a subset $S_{k-1} \subseteq S$ of size $k-1$ such that

$$\{v_1, \ldots, v_{k-1}\} \cup (S \backslash S_{k-1}) \quad \text{spans } V$$

After relabelling, we may assume that

$$S_{k-1} = \{u_1, \ldots, u_{k-1}\}$$

so that $\{v_1, \ldots, v_{k-1}, u_k, \ldots, u_m\}$ spans V.

Now we apply the base case with v_k. By our criterioon, there exists $i \geq 0$ such that $\{v_k, v_1, \ldots, v_{k-1}, u_k, \ldots, u_{k+i}\}$ is linearly independent and

$$\{v_k, v_1, \ldots, v_{k-1}, u_k, \ldots, u_{k+i+1}\}$$

is linearly dependent and

$$u_{k+i+1} \in \operatorname{span}\{v_k, v_1, \ldots, v_{k-1}, u_k, \ldots, u_{k+i}\} \tag{$\dagger$}$$

since T is linearly independent.

So now, $S_k = (S \backslash S_k) \cup \{v_1, \ldots, v_k\}$ still spans.

To show this, since $\{v_1, \ldots, v_{k-1}, u_k, u_{k+1}, \ldots, u_m\}$ spans V , it suffices to show that each of these vectors is in

$$\operatorname{span}\{v_1, \ldots, v_k, u_1, \ldots, u_{k+i}, u_{k+i+2}, \ldots, u_m\}$$

It is straightforward to see that $v_1, \ldots, v_{k-1}, u_k, \ldots, u_{k+i}, u_{k+i+2}, \ldots, u_m$ are in the span. So it remains to show u_{k+i+1} is, which follows from ($\dagger$). The result follows by induction. $\square$

Result 0.2

If $W \subseteq V$ are vector spaces and W is a subspace, if $\dim V < \infty$, then $\dim W \leq \dim V$ with equality **if and only** if $W = V$.

Proof: Let $\mathcal{B}$ be a basis for W, then $\mathcal{B} \subseteq V$ and is linearly independent, so we can extend $\mathcal{B}$ to be a basis $\mathcal{B}'$ for V and since $\mathcal{B} \subseteq \mathcal{B}'$, we have $\dim W = |\mathcal{B}| \leq |\mathcal{B}'| = \dim V$. Suppose that $\dim V = |\mathcal{B}'| < \infty$, then we have following

$$\begin{aligned} \dim W = \dim V \quad &\Leftrightarrow \quad \mathcal{B} = \mathcal{B}' \\ &\Leftrightarrow \quad \operatorname{span}(\mathcal{B}) = \operatorname{span}(\mathcal{B}') \\ &\Leftrightarrow \quad W = V \end{aligned}$$

as desired. □

Example 0.18

What is $(n \in \mathbb{N})$

1. $\dim(F^n) = n$
2. $\dim F[x]_{\leq n} = n + 1$

Lecture 8 (Consolidate) - Wed - Jan 24 - 2024

Suppose V is a F-vector space, and suppose $\mathcal{B}$ is its basis, then all the basis of V have the same size as $\mathcal{B}$.

Example 0.19

Suppose $V = F^n$ for F is a field, for instance, $V = \mathbb{R}^3$. Notice if $S \subseteq \mathbb{R}^3$ and S either spans or linearly independent and $|S| = 3$, then we can conclude that S is a basis.

Proof:

1. If S is linearly independent, then we can expand it to a basis $\mathcal{B}$, since we know that $|\mathcal{B}| = 3$, so $\mathcal{B} = S$.
2. If S spans, then we can contract it to a basis $\mathcal{B}$, since we know that $|\mathcal{B}| = 3$, so $\mathcal{B} = S$.

□

Example 0.20

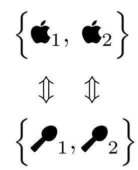
1. What is $\dim F^n$ as an F-vector space? `n`
2. What is $\dim \mathbb{C}$ as a $\mathbb{C}$-vector space? `1`
3. What is $\dim \mathbb{C}$ as a $\mathbb{R}$-vector space? `2`
4. What is $\dim \mathbb{C}$ as a $\mathbb{Q}$-vector space? ∞ (HW1Q3)

Remark: $\mathbb{C}$ and $\mathbb{R}^2$ are isomorphic with map $T : \mathbb{C} \to \mathbb{R}^2$, $T : a + bi \mapsto (a, b)$, and the map is $\mathbb{R}$-linear by definition check.

Off Script

— A note on "size"

Suppose we have a bijection between apples and spoons as shown to the right.



In general, we say that two sets S and T have the same size (cardinality) if there exists $f : S \to T$ that is one-to-one and onto.

A guy named Cantor noticed that there can be different sizes of infinite size. He showed that you can never find a one-to-one and onto map from $S = \mathbb{N}$ to $T =$ set of all (right) infinite binary strings, i.e. $\left\{\epsilon_1\epsilon_2\epsilon_3 \ldots : \epsilon_i \in \{0,1\}\right\}$

But we wonder why?

Suppose we can find such map f such that

$$
\begin{aligned}
&f : \mathbb{N} \to T \\
&f(1) = \boxed{0}0000\cdots \\
&f(2) = 1\boxed{0}000\cdots \\
&f(3) = 11\boxed{1}11\cdots \\
&f(4) = 010\boxed{1}0\cdots \\
&\qquad\vdots
\end{aligned}
$$

Now for i^{th} digit of the $f(i)$, we can construct a binary string that is not on the list by taking the digit opposite to the one we have selected, and we know that the binary string we just created is indeed not mapped by f.

Remark: However, notice that there is a bijection between

$$g : T \to \mathcal{P}(\mathbb{N}) \quad \text{(set of all subsets of } \mathbb{N}\text{)}$$

So Cantor shows that there does not exist an onto map from T to $\mathcal{P}(\mathbb{N})$ for $T \neq \varnothing$.

Result 0.3

Two vactor spaces are isomorphic if and only if their basis have the same size (i.e. there is a bijection between).

Theorem 0.6

Let F be a field, and let V and W be F-vector space with basis $\mathcal{B}$ and $\mathcal{C}$ respectively. If there exists a one-to-one and onto map $f : \mathcal{B} \to \mathcal{C}$, then $V \cong W$.

Discovery 0.1

Let's try to build a linear map $T : V \to W$ from f:

$$v \in V \rightsquigarrow v = \sum_{b \in \mathcal{B}} \lambda_b \cdot b; \quad \lambda_b \in F \text{ \& only finitely many } \lambda_b\text{'s are non-zero}$$

$$\begin{aligned} \Rightarrow \quad T(v) &= T\left(\sum_{b \in \mathcal{B}} \lambda_b \cdot b\right) \\ &= \sum_{b \in \mathcal{B}} \lambda_b \, T(b) \qquad T(b) := f(b) \end{aligned}$$

Proof: We define $T : V \to W$ as follows: If $v \in V$, we write $v = \sum_{b \in \mathcal{B}} \lambda_b \cdot b$; $\lambda_b \in F$ & only finitely many λ_b's are non-zero, and we define $T(v) = \sum_{b \in \mathcal{B}} \lambda_b \cdot f(b)$. Since we know that $\mathcal{B}$ is a basis, so the linear combination of each v is unique. This tells us that T is well-defined.

1. Proof showing it is **onto**

   Notice T is onto since if $w \in W$, we can write $w = \sum_{c \in \mathcal{C}} \gamma_c \cdot C$; $\gamma_c \in F$ and only finitely many γ_c's are non-zero. And since f is a bijection, thus this is

   $$\sum_{b \in \mathcal{B}} \gamma_{f(b)} \cdot f(b) = T\left(\sum_{b \in \mathcal{B}} \gamma_{f(b)} b\right)$$

   which implies that T is onto.

2. Proof showing it is **one-to-one**

   By HW2WU5, we know that a linear map T is one-to-one if and only if $\ker(T) = \{0\}$. Therefore, to prove that T is one-to-one, it suffices to prove that $\ker(T) = \{0\}$.

   If $T(v) = 0$, then we write $v = \sum_{b \in \mathcal{B}} \lambda_b \cdot b$; $\lambda_b \in F$ and only finitely many λ_b's are non-zero. Then we have

   $$\begin{aligned} T(v) = 0 \Rightarrow \quad & T\left(\sum_{b \in \mathcal{B}} \lambda_b \cdot b\right) = 0 \\ \Rightarrow \quad & \sum_{b \in \mathcal{B}} \lambda_b \cdot f(b) = 0 \\ \Rightarrow \quad & \lambda_b = 0 \ \ \forall\, b \in \mathcal{B} \\ \Rightarrow \quad & v = 0 \\ \Rightarrow \quad & \ker(T) = 0 \end{aligned}$$

as desired. $\square$

Terminologies

Now we need to introduce a bit more terminologies to progress further.

Recall that $T : V \to W$ linear, we have

$$\text{im}(T) \subseteq W \qquad \ker(T) \subseteq V$$

We call the dimension of the kernel of T as the nullity, and the dimension of the image of T as the rank. More specifically

Definition 0.12: rank & nullity

$$\dim\Big(\ker(T)\Big) = \text{nullity}(T)$$
$$\dim\Big(\text{im}(T)\Big) = \text{rank}(T)$$

Theorem 0.7: The rank-nullity theorem

Let V and W be finite dimensional F-vector spaces, $T : V \to W$ linear, then we have

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

Proof: Let d be the nullity of T and e be the rank of T, so there exist

$$\text{basis } \{u_1, \ldots, u_d\} \text{ for } \ker(T)$$
$$\text{basis } \{w_1, \ldots, w_e\} \text{ for } \text{im}(T)$$

Then there exist $v_1, \ldots, v_e \in V$ such that $T(v_i) = w_i$. Claim: $\{u_1, \ldots, u_d, w_1, \ldots, w_d\}$ is basis for V □

Lecture 9 - Fri - Jan 26 - 2024

Exercise: Prove that $M_n(F)$ is not commutative for $n \geq 2$
Proof: Given that both $A, B \in M_n(F)$, and suppose $C^1 = A \cdot B$, thus we can obtain that

$$C^1(i,j) = \sum_{k=1}^{n} a_{i,k} b_{k,j}$$

However, on the other hand, suppose $C^2 = B \cdot A$, we can obtain that

$$C^2(i,j) = \sum_{k=1}^{n} b_{i,k} a_{k,j}$$

Note that $C^1(i,j)$ and $C^2(i,j)$ are not necessarily equal, and thus we find that multiplication for matrices $\in M_n(F)$ is not commutative. □

Exercise: Show that if R is a ring and $n \geq 1$, then we can make a ring $M_n(R)$ of $n \times n$ matrices with entries in R.

Proof:

$\square$

Exercise: Let $D_n(F)$ and $U_n(F)$ denote respectively the set of $n \times n$ diagonal and upper-triangular matrices with entries in F. Show that $D_n(F)$ and $U_n(F)$ are subspaces and subrings of $M_n(F)$. What are their dimensions?
Proof:

$\square$

Exercise: Check directly that left multiplication by A is a linear map from F^n to F^m.
Proof:

$\square$

Exercise: Show that if $n \geq 2$ then there are nonzero nilpotent elements in $M_n(F)$.
Proof: Consider the matrices $S \in M_n(F)$ such that $S(i,j) = 0$ for all $i \geq j$. $\square$

Example 0.21

We first finish off the proof for the **rank-nullity** theorem.

Proof: Let $T : V \to W$ be linear, V, W are F-vector spaces. $\dim V, \dim W < \infty$. Also let rank of $T = \dim(\operatorname{im}(T) \subseteq W) < \infty$, and nullity of $T = \dim(\ker(T) \subseteq V) < \infty$. Suppose $d = \dim(\ker(T)) =$ nullity of T. Then there exists a basis $\{v_1, \dots, v_d\}$ for $\ker(T)$. Also Let $e = \operatorname{rank}(T)$, thus there exists a basis $\{w_1, \dots, w_e\}$ for $\operatorname{im}(T)$. Goal: show that $d + e = \dim V$.

Since we know that w_i is in $\operatorname{im}(T)$, so we know that $\exists\ u_1, \dots, u_e \in V$ such that $T(u_i) = w_i$. Claim: $\{v_1, \dots, v_d, u_1, \dots, u_e\} \subseteq V$ is a basis for V. To see that the set spans, let $v \in V$ and we will show that v is a linear combination of $\{v_1, \dots, v_d, u_1, \dots, u_d\}$.
Remark: How can we solve $v = \alpha_1 v_1 + \cdots + \alpha_d v_d + \beta_1 u_1 + \cdots + \beta_e u_e$. Notice if we apply T to both sides:

$$\begin{aligned}
T(v) &= T\left(\alpha_1 v_1 + \cdots + \alpha_d v_d + \beta_1 u_1 + \cdots + \beta_e u_e\right) \\
T(v) &= \alpha_1 T\left(v_1\right) + \cdots + \alpha_d T\left(v_d\right) + \beta_1 T\left(u_1\right) + \cdots + \beta_e T\left(u_e\right) \\
T(v) &= \beta_1 w_1 + \cdots + \beta_e w_e \leftarrow \text{ basis for } \operatorname{im}(T)
\end{aligned}$$

Look at $v - \beta_1 u_1 - \cdots - \beta_e u_e$. Notice $T\left(v - \beta_1 u_1 - \cdots - \beta_e u_e\right) = T(v) - \beta w_1 - \cdots - \beta_e w_e = 0$, so $v - \beta_1 u_1 - \cdots - \beta_e u_e$ is in $\ker(T)$. So $\exists\ \alpha_1, \dots, \alpha_d \in F$ such that $v - \beta_1 u_1 - \cdots - \beta_e u_e = \alpha_1 v_1 + \cdots + \alpha_d v_1$.

Back to the proof

Since $w_1, \dots, w_e$ spans $\operatorname{im}(T)$, there exists $\beta_1, \dots, \beta_e \in F$ such that $T(v) = \beta_1 w_1 + \cdots + \beta_e w_e$. Then this means $v - \beta_1 u_1 - \dots \beta_e u_e$ is in $\ker(T)$. Hence $\exists \alpha_1, \dots, \alpha_d \in F$ such that $v - \beta_1 u_1 - \cdots - \beta_e u_e = \alpha_1 v_1 + \cdots + \alpha_d v_d \implies v \in \operatorname{span}\{v_1, \dots, v_d, u_1, \dots, u_e\}$. To show that $v_1, \dots, v_d, u_1, \dots, u_e$ is linearly independent. Suppose that $\alpha_1 v_1 + \cdots + \alpha_d v_a + \beta_1 u_1 + \cdots + \beta_e u_e = 0$. Applying T gives

$$\begin{aligned}
T\left(\alpha_1 v_1 + \cdots + \alpha_d v_d + \beta_1 u_1 + \dots \beta_e u_e\right) = T(0) \\
\implies \alpha_1 T\left(v_1\right) + \cdots + \alpha_d T\left(v_d\right) + \beta_1 w_1 + \cdots + \beta_e w_e = 0
\end{aligned}$$

since $w_1, \dots, w_e$ are linearly independent $\implies \beta_1 = \cdots = \beta_e = 0$. So now, $\alpha_1 v_1 + \cdots + \alpha_d v_d = 0 \implies \alpha_1 = \cdots = \alpha_d = 0$ since $\{v_1, \dots, v_d\}$ is linearly independent. $\square$

Theorem 0.8

If V is n-dimensional, $n \in \mathbb{N}$, then $V \cong F^n$

Proof: V has a basis $\{b_1, \ldots, b_n\}$, F^n has a basis $\{e_1, \ldots, e_n\}$, where $e_i = (0, \ldots, 0, 1, 0, \ldots, 0)$. We have a bijection between the bases, which extends to an isomorphism from V to F^n. $\square$

Matricies

Let F be a field. Given $m, n \in \mathbb{N}$, we let $M_{m,n}(F)$ denote the set of rectagular $m \times n$ arrays,

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{pmatrix}$$

where the a_{ij} 's are in F.

Terminologies

We call a_{ij} the (i, j)-entry of A and we call A an $m \times n$ matrix with entries in F and we write $A(i, j) = a_{ij}$.

Example 0.22

$$\begin{pmatrix} 2 & \pi & i \\ e & 3.1 & -6 \end{pmatrix} \in M_{2,3}(\mathbb{C})$$

Notice that $M_{m,n}(F)$ is an F-vector space.

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{pmatrix} + \begin{pmatrix} b_{11} & b_{12} & b_{13} & \cdots & b_{1n} \\ b_{21} & b_{22} & b_{23} & \cdots & b_{2n} \\ \vdots & \vdots & \vdots & \ldots & \vdots \\ b_{m1} & b_{m2} & b_{m3} & \cdots & b_{mn} \end{pmatrix}$$

$$= \begin{pmatrix} a_{11} + b_{11} & a_{12} + b_{12} & a_{13} + b_{13} & \cdots & a_{1n} + b_{1n} \\ a_{21} + b_{21} & a_{22} + b_{22} & a_{23} + b_{23} & \cdots & a_{2n} + b_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} + b_{m1} & a_{m2} + b_{m2} & a_{m3} + b_{m3} & \cdots & a_{mn} + b_{mn} \end{pmatrix}$$

Result 0.4

Matrices preserve closedness under addition and scalar multiplication

$$(A+B)(i,j) = A(i,j) + B(i,j)$$
$$(\lambda A)(i,j) = \lambda \cdot (A(i,j))$$

The 0 element is

$$0_{m,n} = \begin{pmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{pmatrix}$$

For $1 \leq i \leq m, 1 \leq j \leq n$, we let

$$E_{i,j} = \begin{pmatrix} 0 & 0 & \cdots & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & 0 \end{pmatrix}$$

where $a_{ij} = 0$ (i^{th} row, j^{th} column)

$$E_{i,j}(k, \updownarrow) = \delta_{i,k}\delta_{j,\updownarrow} \text{ for } 1 \leq k \leq m, 1 \leq j \leq n.$$

$$\delta_{a,b} = \begin{cases} 1 & \text{if } a = b \\ 0 & \text{otherwise} \end{cases}$$

$$E_{a,b} = \begin{cases} 1 & \text{if } i = k, j = \mathcal{L} \\ 0 & \text{otherwise} \end{cases}$$

Claim: $\{E_{ij} : 1 \leq i \leq m, 1 \leq j \leq n\}$ is a basis for $M_{m,n}(F)$

Proof: of spanning

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \in M_{m,n}(F)$$

Thus we have

$$\begin{aligned} A &= a_{11}E_{11} + a_{12}E_{12} + \cdots + a_{1n}E_{1n} \\ &+ a_{11}E_{21} + a_{12}E_{22} + \cdots + a_{1n}E_{2n} \\ &+ \ \vdots \\ &+ a_{11}E_{m1} + a_{12}E_{m2} + \cdots + a_{1n}E_{mn} \end{aligned}$$

as desired. $\square$

Proof: of independence

$$\text{If } \sum_{i=1}^{m}\sum_{j=1}^{n} c_{i,j}\cdot E_{i,j} = 0 \Longrightarrow c_{i,j} = 0,\ \forall\, i,j \text{ by looking at the } (i,j)\text{-entry.} \ \square$$

Corollary 0.3

$M_{m,n}(F) \cong F^{mn}$ as F-vector space.

Proof: We know that $M_{m,n}(F)$ has a basis of size $m \cdot n$. $\square$

Result 0.5

Let $m, n, p \in \mathbb{N}$. We have a multiplication $\cdot$:

$$M_{m,n}(F) \times M_{n,p}(F) \to M_{m,p}(F)$$

Remark: Notice that the columns of the first matrix and the number of rows of the second matrix must be the same.

Example 0.23

Let $m = 2, n = 3, p = 4$.

$$\begin{pmatrix} 1 & -2 & 6 \\ 4 & 0 & 5 \end{pmatrix} \begin{pmatrix} 1 & 0 & 2 & 0 \\ 0 & 0 & 1 & 3 \\ 1 & 0 & -1 & 2 \end{pmatrix} = \begin{pmatrix} \boxed{7} & \boxed{0} & \boxed{-6} & \boxed{6} \\ \boxed{9} & \boxed{0} & \boxed{3} & \boxed{10} \end{pmatrix}$$

Exercise: For instance, row 1 , column 4 $(a_{1,4})$ would be $(1\ -2\ 6)\begin{pmatrix} 0 \\ 3 \\ 2 \end{pmatrix} = 1\cdot 0 + (-2)\cdot 3 + 6 \cdot 2 = 6.$

Conjecture

$\forall\, \epsilon > 0$, we can find an algorithm for matrix multiplication $n \times u$ matrices runs in $\mathcal{O}\left(n^{2+\epsilon}\right)$ time.

Lecture 10 - Mon - Jan 29 - 2024

Let F be a field, and $m, n \in \mathbb{N}$, then $M_{m,n}(F)$ is the set of $m \times n$ matrices with entries in F.

Discovery 0.2: Matrix Multiplication is Associative

Let $A \in M_{m,n}(F)$, $B \in M_{n,p}(F)$, and $C \in M_{p,q}(F)$, and also let $i \in \{1, \ldots, m\}$ and $j \in \{1, \ldots, q\}$.

$$\Big((A \cdot B) \cdot C\Big)(i,j) = \sum_{k=1}^{p} \left(\sum_{l=1}^{n} A(i,l)B(l,k) \right) C(k,j)$$
$$\Big(A \cdot (B \cdot C)\Big)(i,j) = \sum_{s=1}^{p} A(i,s) \left(\sum_{t=1}^{n} B(s,t) \right) C(t,j)$$

Notice that $\Big((A \cdot B) \cdot C\Big)(i,j) = \Big(A \cdot (B \cdot C)\Big)(i,j)$, thus matrix multiplication is Associative.

Exercise: Show that matrix multiplication is also distributive. In particular,

$$A\left(B + \lambda C\right) = AB + \lambda AC$$

Definition 0.13: $M_n(F)$

We use $M_n(F)$ to denote $M_{n,n}(F)$.

Theorem 0.9

$M_n(F)$ is a ring.

Proof: We saw that $M_n(F)$ has $+$, and the addition axioms are satisfied since $\Big(M_n(F), +\Big)$ is a F-vector space (and hence an abelian group). Recall we also have multiplication:

$$M_n(F) \times M_n(F) \to M_n(F)$$

Remark:

$$\begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}^n = \begin{bmatrix} F_{n-1} & F_n \\ F_n & F_{n+1} \end{bmatrix} \qquad \forall\, n \geq 1$$

where F_i is the i^{th} Fibonacci number.

and the matrix multiplication is associative. Moreover, the multiplicative identity is $(i,j) = \delta_{ij}$. Last but not the least,

Exercise: Distributivity (relation between addition and multiplication).

Therefore, $M_n(F)$ is a ring. $\square$

Example 0.24: Idempotent in $M_2(\mathbb{R})$

Elements in $M_2(\mathbb{R})$ that are idempotent:

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \qquad \begin{bmatrix} 7 & e \\ -\frac{42}{e} & -6 \end{bmatrix} \qquad \cdots$$

Result 0.6: Idempotent in $M_2(\mathbb{R})$

Exercise: Show that for $a, b \in F \backslash \{0\}$,

$$\begin{pmatrix} a & b \\ x & 1-a \end{pmatrix}$$

is idempotent when $a(1-a) = xb$.

Proof: We can get that

$$\begin{pmatrix} a & b \\ x & 1-a \end{pmatrix}^n = \begin{pmatrix} a^2 + bx & ab + b - ba \\ ax + x - ax & bx + (1-a)^2 \end{pmatrix}$$

Therefore, if $a(1-a) = xb$, then we will have $a^2 + bx = a$ and $bx + (1-a)^2 = 1 - a$. □

Definition 0.14: Diagonal

A matrix $D \in M_n(F)$ is called **diaginal** if $D(i, j) = 0$ whenever $i \neq j$.
Remark: The set $D_n(F)$ of $n \times n$ diagonal matrices is a subring of $M_n(F)$ and <u>is also commutative</u>.

Definition 0.15: Upper Triangular

A matrix U is upper triangular if it is of the form

$$U = \begin{pmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ 0 & u_{22} & \cdots & u_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & u_{nn} \end{pmatrix}$$

in particular, $U(i, j) := 0$ when $i > j$.

Lecture 11 - Wed - Jan 31 - 2024

Theorem 0.10

For $E_{a,b}, E_{c,d} \in M_n(F)$, then we have

$$E_{a,b} \cdot E_{c,d} = \begin{cases} 0 & \text{if } b \neq c \\ E_{a,d} & \text{if } b = c \end{cases}$$

Proof: Def'n check. □

Discovery 0.3

What is $E_{i,j}^2$?
Answer: 0 if $i \neq j$ or $E_{i,j}$ if $i = j$.

Proposition 0.6

Upper-Δ matrices form a subring of $M_n(F)$.

It is easy to check that it has 0 and ± 1, and it is closed under addition. Here, we will prove that it is closed under multiplication:

Proof: For two matrices $A, B \in M_n(F)$, we can write them as in the form

$$A = \sum_{i \leq j} a_{ij} E_{ij} \qquad B = \sum_{k \leq l} b_{kl} E_{kl}$$

Therefore

$$\begin{aligned} A \cdot B &= \sum_{i \leq j} \sum_{k \leq l} a_{ij} E_{ij} b_{kl} E_{kl} \\ &= \sum_{i \leq j} \sum_{k \leq l} a_{ij} b_{kl} \underbrace{E_{ij} E_{kl}}_{\neq 0 \text{ @ } j=k} \\ &= \sum_{i \leq j \leq l} a_{ij} b_{jl} E_{il} \end{aligned}$$

so the (i, l) entry is $\sum_{j=i}^{l} a_{ij} b_{jl}$, which implies that it is an empty sum when $i > j$, which then implies that $A \cdot B$ is upper-Δ. $\square$

$GL_n(F)$ – General Linear group of F

Discovery 0.4

$M_n(F)$ is a ring, so its units form a group. We let $GL_n(F)$ denote its set of units.

Definition 0.16: Invertible

A matrix A is called **invertible** (or non-singular) if it is in $GL_n(F)$ and it is called **non-invertible** or singular otherwise.

Theorem 0.11

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in GL_2(F) \quad \text{if and only if} \quad ad - bc \neq 0.$$

Proof:

1. ($\Longrightarrow$)
   Suppose $ad = bc$ and A is invertible, then we know that there exists $B \in GL_2(F)$ such that

   $$B \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} B = I$$

   Consider $\begin{bmatrix} b \\ -a \end{bmatrix} \in M_{2,1}(F)$, so we have

   $$B\left(\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{bmatrix} b \\ -a \end{bmatrix}\right) = B\begin{bmatrix} 0 \\ bd - ad \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$
   $$\left(B\begin{pmatrix} a & b \\ c & d \end{pmatrix}\right)\begin{bmatrix} b \\ -a \end{bmatrix} = I\begin{bmatrix} b \\ -a \end{bmatrix} = \begin{bmatrix} b \\ -a \end{bmatrix}$$

   which implies that $a = b = 0$. Similarly, we have $c = d = 0$. However, this contradicts the fact that $B\begin{pmatrix} a & b \\ c & d \end{pmatrix} = I$. Hence we have that $ad \neq bc$.

2. ($\Longleftarrow$)
   Since we have $ad - bc \neq 0$, thus we consider $B = \dfrac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = A^{-1}$

woo hoo. $\square$

For what follows, we will think of F^n as $M_{n,1}(F)$

Example 0.25

We have

$$F^n = \Big\{(a_1, a_2, \dots, a_n), a_1, \dots, a_n \in F\Big\}$$
$$M_{n,1}(F) = \left\{\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}, a_1, \dots, a_n \in F\right\}$$

Remark: If $A \in M_{m,n}(F)$ and $v \in M_{n,1}(F)$, so $A \cdot v \in M_{m,1}(F) = F^m$.

Result 0.7

If A is a $m \times n$ matrix, then it can be viewed as a map

$$T_A : F^n \to F^m$$

given by

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

Example: Notice that T_A is linear.

Universal Property

Theorem 0.12: Universal Property

] Let V and W be vector spaces and let $\mathcal{B}$ be a basis for V, Our goal is to understand the linear maps from V to W (hard). **Universal Property tells us that**:

> Let $f : \mathcal{B} \to W$ be a set map, then there $\exists!$ (exists a unique) linear map $T : V \to W$ such that $T\big|_{\mathcal{B}} = f$ (with restriction to $\mathcal{B}$).

Lecture 12 - Fri - Feb 2 - 2024

Definition 0.17: Linear Transformation

Let V, W be F-vector space, and $T : V \to W$ linear, we'll say T is a linear transformation. When $V = W$, we'll say that T is a linear operator.

Universal Property

We can make linear transformation

$$T : V \to W$$

as follows:

> We fix a basis $\mathcal{B}$ for V and specify what T does to $\mathcal{B}$ (no constraints) and then T is uniquely determined.

Example 0.26

Suppose we have

$$T : \mathbb{R}^2 \to \mathbb{R}^3$$
$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} \mapsto \begin{bmatrix} 5 \\ -\pi \\ \sqrt{e} \end{bmatrix}$$
$$\begin{bmatrix} 0 \\ 1 \end{bmatrix} \mapsto \begin{bmatrix} 5 \\ -\pi \\ \sqrt{e} \end{bmatrix}$$

Thus we have

$$T\left(\begin{bmatrix} a \\ b \end{bmatrix}\right) = T\left(a\begin{bmatrix} 1 \\ 0 \end{bmatrix} + b\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) \cdots$$

Theorem 0.13

Let V and W be vector spaces. Let $\mathcal{B}$ be a basis for V. If $f : \mathcal{B} \to W$ (set map) then there exists linear transformation

$$T : V \to W$$

such that $T(b) = f(b)$ for all $b \in \mathcal{B}$. Moreover, if $S : V \to W$ such that $S(b) = f(b)$ for all $b \in \mathcal{B}$, then $S = T$.

Proof: We recall that if $\mathcal{B}$ is a basis for V, then if $v \in V$, there is a unique linear combination

$$v = \sum_{b \in \mathcal{B}} \lambda_b \cdot b \qquad \lambda_b \in F \text{ only finitely many are non-zero}$$

Then if $T : V \to W$ is linear and $T(b) = f(b)$ for all $b \in \mathcal{B}$, we must have

$$\begin{aligned} T\left(\sum_{b \in \mathcal{B}} \lambda_b \cdot b\right) &= \sum_{b \in \mathcal{B}} T(\lambda_b \cdot b) \\ &= \sum_{b \in \mathcal{B}} \lambda_b \cdot f(b) \end{aligned}$$

so we see that if this T is linear, so it is unique. Let's check if T is linear:
Remark: T is linear iff $T(v + \lambda w) = T(v) + \lambda T(w)$ for all $v, w \in V$ and $\lambda \in F$.

Hence let $v, w \in V$ and $\lambda \in F$, then we can write

$$v = \sum \alpha_b \cdot b \qquad w = \sum \beta_b \cdot b$$

so

$$\begin{aligned} v + \lambda w &= \sum (\alpha_b + \lambda \beta_b) \cdot b \\ \Rightarrow \quad T(v + \lambda w) &= T(v) + \lambda T(w) \end{aligned}$$

Thus completing the proof. □

Result 0.8

In what follows, we identify

$$F^n \longleftrightarrow M_{n,1}(F) = \left\{ \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} : a_1, \ldots, a_n \in F \right\}$$

$$F^m = \longleftrightarrow M_{m,1}(F)$$

Theorem 0.14

Every linear transformation

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

is of the form $T(v) = Av \in F^m$, $v \in M_{n,1}(F)$ for some $A \in M_{m,n}(F)$.

Remark: Notice if A is matrix, then $v \to Av$ is linear.

Proof: Take the standard basis for F^n. Let

$$A = \begin{pmatrix} | & | & & | \\ T(e_1) & T(e_2) & \cdots & T(e_n) \\ | & | & & | \end{pmatrix}$$

then $A(e_j) = T(e_j)$. Since A and T does the same on a basis, then they are the same by Universal Property.
$\square$

Example 0.27

Find the matrix of the linear operator, $T : \mathbb{R}^2 \to \mathbb{R}^3$, where T rotates points counterclockwise $45°$ with respect to the origin:

$$T\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} \qquad T\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$$

then the matrix for T is

$$A = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}$$

Example 0.28

What does the linear transformation

$$\begin{bmatrix} x \\ y \end{bmatrix} \mapsto \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

do to the unit disc?

Exercise: Ellipse.

Definition 0.18

We say that the matrix

$$A = \begin{pmatrix} | & | & & | \\ T(e_1) & T(e_2) & \cdots & T(e_n) \\ | & | & & | \end{pmatrix}$$

is the matrix of T with respect to the standard basis.

We'd like to extend this idea to understand linear transformation between abstract finite dimensional vector space.

Example 0.29

Let

$$V = \mathbb{R}[x]_{\leq 2} \qquad W = \mathbb{R}[x]_{\leq 3}$$

then

$$T(p(x)) = \int_0^x p(t)\, dt$$

is a linear transformation.

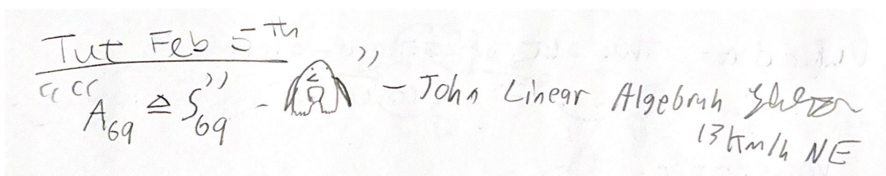
Let the following be ordered set

$$\mathcal{B} = (1, x, x^2) \quad \text{for } V$$
$$\mathcal{C} = (1, x, x^2, x^3) \quad \text{for } W$$

so we can use $\mathcal{B}$ and $\mathcal{C}$ to give us elements of V and W coordinates:

$$\left[a + bx + cx^2\right]_{\mathcal{B}} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$$

$$\left[\alpha + \beta x + \gamma x^2 + \delta x^3\right]_{\mathcal{C}} = \begin{bmatrix} \alpha \\ \beta \\ \gamma \\ \delta \end{bmatrix}$$



Regular Language

Definition 0.19

$\sum$ be finite alphabet.

e.g. $\{a, b\}$

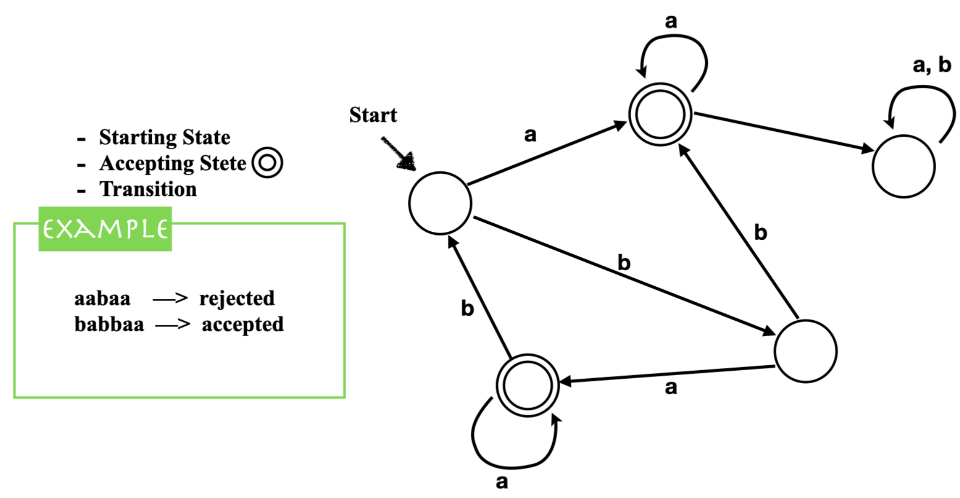
Definition 0.20: Semigroup

$\sum^*$ = set of finite strings (words) on $\sum$

e.g.

$$\sum = \{a, b\} \quad \Rightarrow \quad \sum^* \ni a, ab, baba, abba, aaaaa, \ldots$$

Example 0.30



Definition 0.21: DFA

A deterministic finite-state automaton (DFA)

1. input alphabet $\sum$
2. has finite set of states

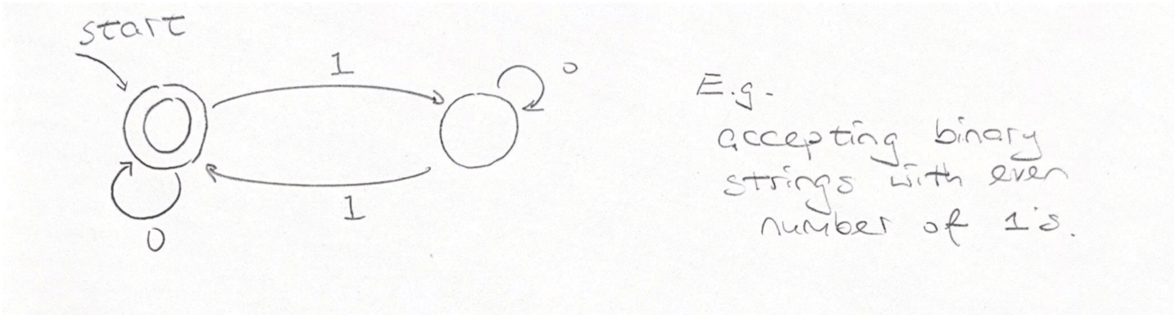
Definition 0.22: Regular Language

The set of all strings accepted by a DFA is called a regular Language.

Theorem 0.15: Church-Turing Thesis

Every algorithm can be similated by what's called a Turing machine, whuch is simulated with finite number of states and two stacks.

Example 0.31

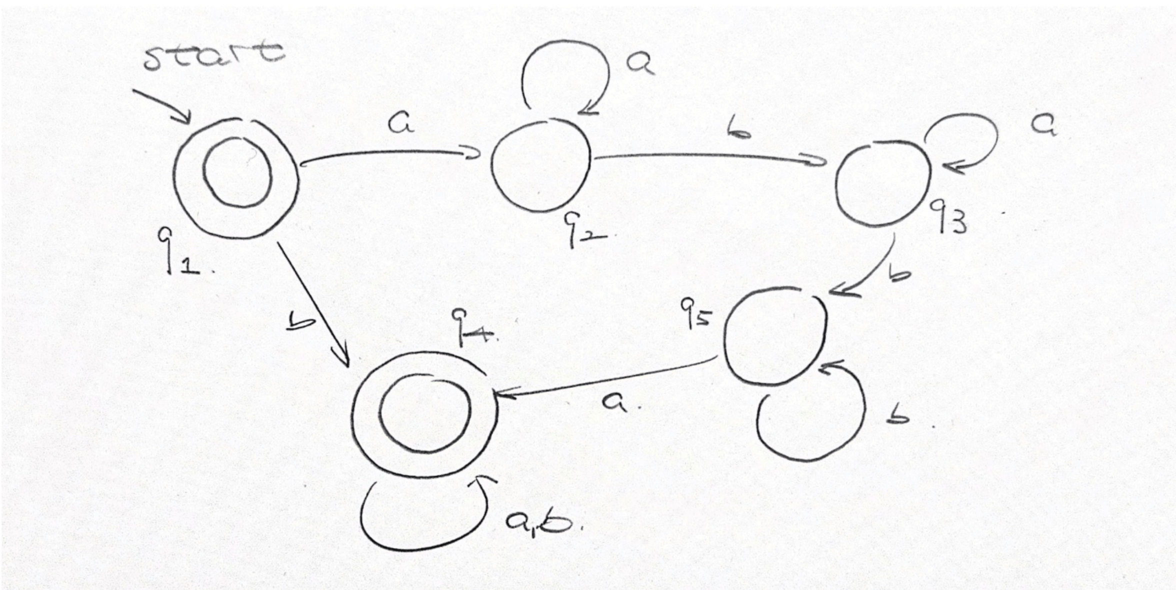


Definition 0.23: Adjacency Matrix

See the example below:

Example 0.32

Given a DFA, we associate an **adjacency matrix** as follows:



We define matrix A whose $(i, j)^{th}$ entry is the number of paths from q_i to q_j:

$$A := \begin{pmatrix} 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{pmatrix}$$

Theorem 0.16

Let A be the adjacency matrix of DFA with states $q_1, \dots, q_s$, then $A^n(i,j)$ = number of paths of length n from $q_i \to q_j$.

Proof: By induction.

1. Base case: definition
2. Hypothesis: Yes for length k
3. Proof Step: Yes for length $k+1$

orz. □

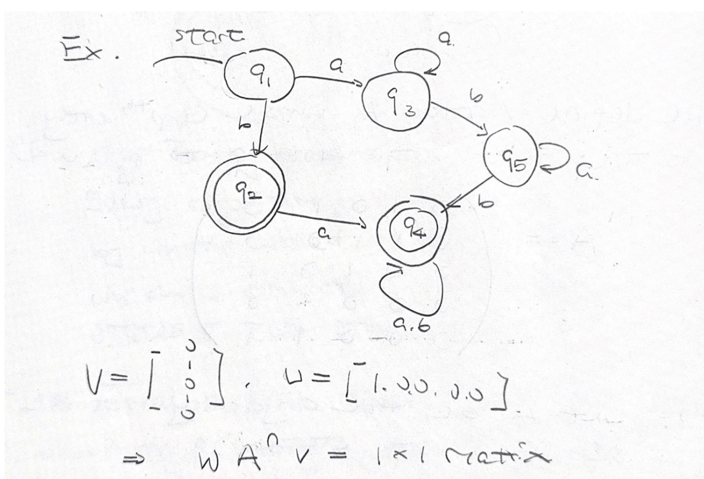
Now we can compute the strings accepted by DFA's. For an $s \times s$ adjacency matrix A, we have

$$v = \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_s \end{bmatrix} \qquad \text{where } \varepsilon_i = \begin{cases} 1 & q_i \text{ is accepted} \\ 0 & q_i \text{ is not accepted} \end{cases}$$

and

$$w = \begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix}$$

Example 0.33

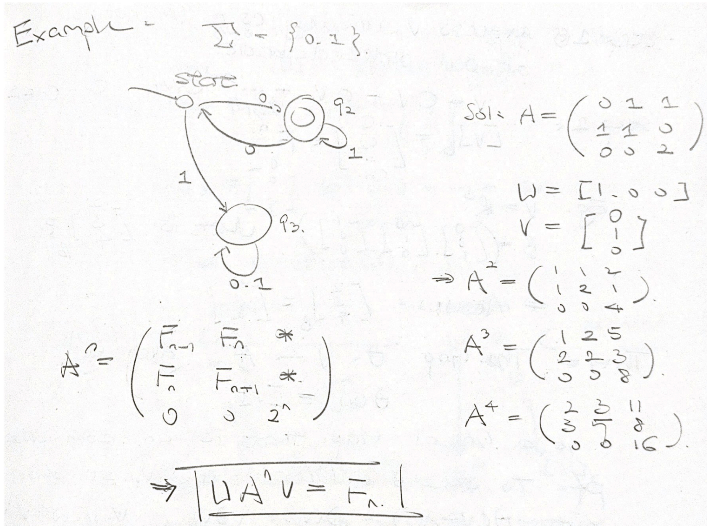


Theorem 0.17

Have wA^nV is the number of strings of length n accepted by the DFA.

Proof: Number of strings accepted by DFA is the number of paths accepted of length n started from q_1. □

Example 0.34



Lecture 13 - Mon - Feb 5 - 2024

Definition 0.24: Ordered Basis

An ordered basis $\mathcal{B}$ for V is a list $v_1, v_2, \ldots, v_n$ of distinct vectors such that $\{v_1, \ldots, v_n\}$ is a basis for V.

Discovery 0.5

Significance:
WE can use ordered bases to endow vector spaces with coordinates.

But how do we do this?
Given $v \in V$,

1. Step 1:

   Express v uniquely as a linear combination of our ordered basis

2. Step 2:

$$[v]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} \in F^n$$

Example 0.35

$V = \mathbb{R}^3$ and $\mathcal{B} = \left(\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \right)$, what is $\begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix}_{\mathcal{B}}$?

Solution: The answer is $\begin{bmatrix} 4 \\ 3 \\ 2 \end{bmatrix}$

Theorem 0.18

$\dim V = n$, the map

$$\theta : V \to F^n$$
$$\theta(v) = [v]_{\mathcal{B}}$$

is a linear map that is an isomorphism.

To prove the theorem, we first prove the following proposition.

Proposition 0.7

If V and U are two n-dimensional vector spaces ($n < \infty$) and $T : V \to U$ is linear then

$$T \text{ is 1-to-1 } \Leftrightarrow T \text{ is onto } \Leftrightarrow T \text{ is isomorphism}$$

Proof: We use Rank-Nullity Theorem,

$$\ker(T) = (0) \Rightarrow \dim V = \dim(\text{im } T) \Rightarrow \text{ onto}$$

Similar proof follow. $\square$

Proof: of theorem To see that θ is linear, we must show that

$$\theta(v + \lambda w) = \theta(v) + \lambda\theta(w) \qquad \forall\, v, w \in V,\ \lambda \in F$$

Recall $\mathcal{B} = (v_1, \ldots, v_n)$, let $[v]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$ and $[w]_{\mathcal{B}} = \begin{bmatrix} d_1 \\ \vdots \\ d_n \end{bmatrix}$, then we have

$$v = c_1 v_1 + \cdots + c_n v_n$$
$$w = d_1 v_1 + \cdots + d_n v_n$$

so

$$v + \lambda w = (c_1 + \lambda d_1)v_1 + \cdots + (c_n + \lambda d_n)v_n$$

which gives us

$$[v + \lambda w]_{\mathcal{B}} = [v]_{\mathcal{B}} + \lambda[w]_{\mathcal{B}}$$

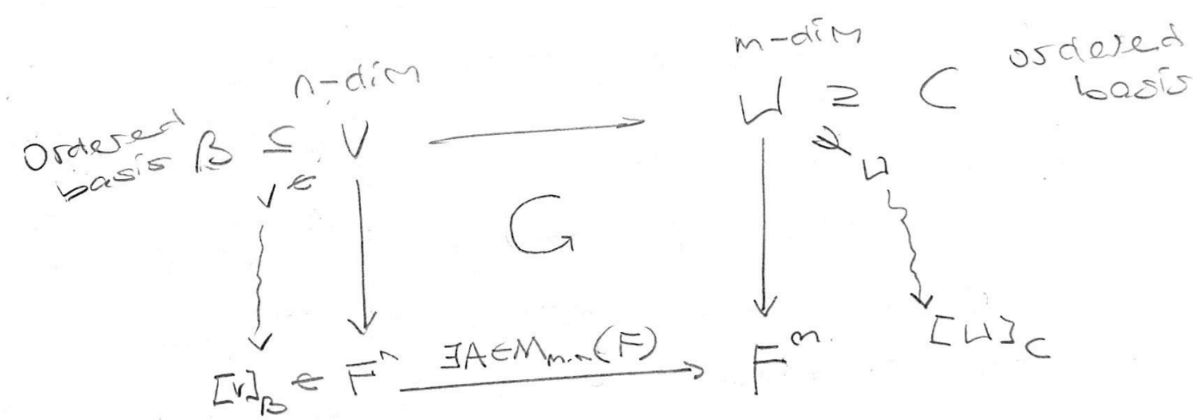
as desired.

Therefore, to show that θ is isomorphism, STP θ is one-to-one, so STP $\ker(\theta) = (0)$. Suppose that $\theta(v) = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$, so we have

$$\begin{aligned} [v]_{\mathcal{B}} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} = 0 \cdot v_1 + \cdots + 0 \cdot v_n = 0 \\ \Rightarrow \quad \ker(\theta) = (0) \end{aligned}$$

as desried. $\square$

Commutative Diagram



$\exists\,!$ a matrix A that makes the diagram commutative. I.e.

$$\forall\, v \in V, \quad [T(v)]_{\mathcal{C}} = \underbrace{A}_{m\times n} \underbrace{[v]_{\mathcal{B}}}_{n\times 1} \in F^m$$

Example 0.36

Let $V = \mathbb{R}[x]_{\leq 2}$ has basis $\mathcal{B} = (1, x, x^2)$ and $W = \mathbb{R}[x]_{\leq 3}$ has basis $\mathcal{C} = (1, x, x^2, x^3)$. Let $T : V \to W$ such that $T(p(x)) = \displaystyle\int_0^x p(t)\,dt$. Find $A \in M_{4,3}(\mathbb{R})$ that makes the diagram commutative.

Look at the standard basis for $\mathbb{R}^3$:

$$e_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad e_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad e_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Thus we have

$$e_1 \to \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \quad e_2 = \begin{bmatrix} 0 \\ 0 \\ \frac{1}{2} \\ 0 \end{bmatrix}, \quad e_3 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \frac{1}{3} \end{bmatrix}$$

which implies that

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

Definition 0.25

In general, the matrix A with

$$\begin{tikzcd} \mathcal{B} \subseteq & V \arrow[r] \arrow[dd] & W \arrow[dd] & \supseteq \mathcal{C} \\ v \arrow[d, mapsto] & & & w \arrow[d, mapsto] \\ {[v]_{\mathcal{B}}} & \mathbb{F}^n \arrow[r, "A"] & \mathbb{F}^m & {[w]_{\mathcal{C}}} \end{tikzcd}$$

is called the matrix of T with respect to the ordered basis $\mathcal{B}$ and $\mathcal{C}$ and we write

$$A = [T]_{\mathcal{B},\mathcal{C}}$$

(denoting convert $\mathcal{B}$ to $\mathcal{C}$).

Discovery 0.6

How do we find $[T]_{\mathcal{B},\mathcal{C}}$?

Suppose we have

$$\mathcal{B} = (b_1, \ldots, b_n)$$
$$\mathcal{C} = (c_1, \ldots, c_m)$$

then we would have

$$A = [T]_{\mathcal{B},\mathcal{C}} = \begin{pmatrix} | & & | \\ [T(b_1)]_{\mathcal{C}} & \cdots & [T(b_n)]_{\mathcal{C}} \\ | & & | \end{pmatrix}$$

a $m \times n$ matrix.

Example 0.37

Let $V = \mathbb{R}[x]_{\leq 2}$ and $W = \{p(x) : \deg(p) \leq 3, p(1) = 0\}$. Let $T : V \to W$ such that $T(p(x)) = p(x)(x-1)$, which is a linear map. Find $[T]_{\mathcal{B},\mathcal{C}}$ with $\mathcal{B} = (x^2, x, 1)$ and $\mathcal{C} = (x-1, x^2-1, x^3-1)$.

Proof: Notice that $T(x^2) = x^3 - x^2 = (x^3-1) - (x^2-1)$, and with similar process, we can find that

$$\left.\begin{aligned} [T(x^2)]_{\mathcal{C}} &= \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} \\ [T(x)]_{\mathcal{C}} &= \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \\ [T(1)]_{\mathcal{C}} &= \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \end{aligned}\right\} \quad [T]_{\mathcal{B},\mathcal{C}} = \begin{bmatrix} 0 & -1 & 1 \\ -1 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

as desired. $\square$

Lecture 14 - Wed - Feb 7 - 2024

Recall the commutative diagram,

$$\begin{tikzcd} V \arrow[r, "T"] \arrow[d] \arrow[dr, phantom, "\circlearrowleft"] & W \arrow[d] \\ F^n \arrow[r, "{[T]_{\mathcal{B},\mathcal{C}}}"] & F^m \end{tikzcd}$$

Theorem 0.19

For all $v \in V$,

$$[T]_{\mathcal{B},\mathcal{C}}[v]_{\mathcal{B}} = [T(v)]_{\mathcal{C}}$$

Proof: Consider the LHS, for $v = b_j$, we have

$$\begin{aligned} LHS &= [T]_{\mathcal{B},\mathcal{C}}[b_j]_{\mathcal{B}} \\ &= [T]_{\mathcal{B},\mathcal{C}} \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} \\ &= j^{th} \text{ column of } [T]_{\mathcal{B},\mathcal{C}} \\ &= [T(b_j)]_{\mathcal{C}} \end{aligned}$$

so we have $[T]_{\mathcal{B},\mathcal{C}}[b_j]_{\mathcal{B}} = [T(b_j)]_{\mathcal{C}}$ for $j = 1, \dots, n$. Hence the linear transformation for

$$\begin{array}{ll} V \to F^m & \text{given by} \\ v \mapsto [T]_{\mathcal{B},\mathcal{C}}[v]_{\mathcal{B}} & \text{and the ones given by} \\ v \mapsto [T(v)]_{\mathcal{B}} & \text{agree on a basis for } V \end{array}$$

so by Universal Property, they are the same. □

Composition

Suppose we have

V	W	U
n-dim	m-dim	p-dim
$\mathcal{B} = (b_1, \dots, b_n)$	$\mathcal{C} = (c_1, \dots, c_m)$	$\mathcal{D} = (d_1, \dots, d_p)$

Therefore

$$\left.\begin{array}{r} T : V \to W \\ S : W \to U \end{array}\right\} \Rightarrow S \circ T : V \to U$$

Now we are interested in what the relationship is between

$$[S \circ T]_{\mathcal{B},\mathcal{D}} \quad \& \quad [T]_{\mathcal{B},\mathcal{C}} \quad \& \quad [S]_{\mathcal{C},\mathcal{D}}$$

Theorem 0.20

We have

$$\underbrace{[S \circ T]_{\mathcal{B},\mathcal{D}}}_{p\times n} = \underbrace{[S]_{\mathcal{C},\mathcal{D}}}_{p\times m} \cdot \underbrace{[T]_{\mathcal{B},\mathcal{C}}}_{m\times n}$$

Proof: We have the commutative diagrams

$$\begin{tikzcd} V \arrow[r, "T"] \arrow[d] \arrow[dr, phantom, "\circlearrowright"] & W \arrow[d] \\ F^n \arrow[r, "{[T]_{\mathcal{B},\mathcal{C}}}"] & F^m \end{tikzcd} \qquad \begin{tikzcd} W \arrow[r, "T"] \arrow[d] \arrow[dr, phantom, "\circlearrowright"] & U \arrow[d] \\ F^m \arrow[r, "{[S]_{\mathcal{C},\mathcal{D}}}"] & F^p \end{tikzcd}$$

Thus concatenating the two diagrams gives us

$$\begin{tikzcd} V \arrow[r, "T"] \arrow[d] \arrow[dr, phantom, "\circlearrowright"] & W \arrow[r, "S"] \arrow[d] & U \arrow[d] \\ F^n \arrow[r, "{[T]}"] & F^m \arrow[r] & F^p \end{tikzcd}$$

Therefore we have

$$\begin{array}{lcl} \to\to\,⤵ & := & [S(T(v))]_{\mathcal{D}} \\ ↳\,\to\to & := & [S]_{\mathcal{C},\mathcal{D}}[T]_{\mathcal{B},\mathcal{C}}[v]_{\mathcal{B}} \end{array}$$

hence

$$[S \circ T]_{\mathcal{B},\mathcal{D}}[v]_{\mathcal{B}} = [S(T(v))]_{\mathcal{D}} = [S]_{\mathcal{C},\mathcal{D}}[T]_{\mathcal{B},\mathcal{C}}[v]_{\mathcal{B}}$$
$$\Rightarrow \quad [S \circ T]_{\mathcal{B},\mathcal{D}} = [S]_{\mathcal{C},\mathcal{D}}[T]_{\mathcal{B},\mathcal{C}}$$

thus we complete the proof. □

Remark: If $A_1, A_2 \in M_{m,n}(F)$ such that $A_1 v = A_2 v \ \forall\, v \in F^n$, then $A_1 = A_2$.

Midterm

Definition 0.26: Midterm

The midterm is 1 hour and 50 minutes long consisting four parts.

(a) 5 multiple choices, (2 marks each, 10 marks in total)

(b) 5 true or falses, (2 marks each, 10 marks in total)

(c) 3 parts, linear independece, dependence, and span (10 marks)

(d) 3 parts, matrices and linear transformation (10 marks)

Lecture 15 - Fri - Feb 9 - 2024

Here we introduce a second proof for the Theorem introduced above (instead of depicting diagrams). We first introduce a lemma:

Lemma 0.5

Let $A, B \in M_{m,n}(F)$, if $Av = Bv$ for all $v \in F^n$, then $A = B$.

Proof: of the lemma
This is easy to see if we pass in the standard basis vectors to show that each entry of A is equal to the corresponding entry in B, thus A and B are equal to each other. □

Proof: of the theorem
Let $[v]_{\mathcal{B}} \in F^n$, consider $[S]_{\mathcal{C},\mathcal{D}}[T]_{\mathcal{B},\mathcal{C}}[v]_{\mathcal{B}}$. Hence by associativity, we have

$$\begin{aligned}[S]_{\mathcal{C},\mathcal{D}}[T]_{\mathcal{B},\mathcal{C}}[v]_{\mathcal{B}} &= [S]_{\mathcal{C},\mathcal{D}}\Big([T]_{\mathcal{B},\mathcal{C}}[v]_{\mathcal{B}}\Big) \\ &= [S]_{\mathcal{C},\mathcal{D}}\Big([T(v)]_{\mathcal{C}}\Big) \\ &= [S(T(v))]_{\mathcal{D}}\end{aligned}$$

Notice that $[S(T(v))]_{\mathcal{D}} = [(S \circ T)(v)_{\mathcal{C}}]$. On the other hand, $[S \circ T]_{\mathcal{B},\mathcal{C}}[v]_{\mathcal{B}} = [(S \circ T)(v)]_{\mathcal{D}}$. □

Result 0.9

Key properties:

1. $[T]_{\mathcal{B},\mathcal{C}}[v]_{\mathcal{B}} = [T(v)]_{\mathcal{C}}$
2. $[S]_{\mathcal{C},\mathcal{D}}[w]_{\mathcal{C}} = [S(w)]_{\mathcal{D}}$
3. $[S \circ T]_{\mathcal{B},\mathcal{D}} = [S]_{\mathcal{C},\mathcal{D}}[T]_{\mathcal{B},\mathcal{C}}$

Example for Commutative Diagram Composition

Suppose we have a vector space V with ordered basis $\mathcal{B} = (b_1, \ldots, b_n)$, and a linear transformation $T : V \to V$, (more precisely, a linear operator).

Remark: Often it is the case that we have another ordered basis $\mathcal{C}$ for V and we would like to know that relation between $[T]_{\mathcal{B}}$ and $[T]_{\mathcal{C}}$.

Definition 0.27

We let $[T]_{\mathcal{B}}$ denote $[T]_{\mathcal{B},\mathcal{B}}$.

Example 0.38

Suppose for a linear transformation

$$T : \mathbb{R}^3 \to \mathbb{R}^2$$
$$T\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} x + 2y + 3z \\ x + y + z \\ 2x + y + z \end{bmatrix}$$

and suppose we have two ordered basis $\mathcal{B}$ and $\mathcal{C}$ such that

$$\mathcal{B} = \left(\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\right)$$
$$\mathcal{C} = \left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}\right)$$

Discovery 0.7

We want to find $[T]_{\mathcal{B}}$ and $[T]_{\mathcal{C}}$.

We first find $[T]_\mathcal{B}$:

$$1^{st} \text{ column} : T(b_1) = \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 3 \end{bmatrix}_\mathcal{B}$$

$$2^{nd} \text{ column} : T(b_2) = \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}_\mathcal{B}$$

$$3^{rd} \text{ column} : T(b_3) = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}_\mathcal{B}$$

Thus we have $[T]_\mathcal{B} = \begin{pmatrix} 1 & 1 & 2 \\ 1 & 1 & 1 \\ 3 & 2 & 1 \end{pmatrix}$ as desired. Similarly we would find $[T]_\mathcal{C} = \begin{pmatrix} 0 & 1 & 3 \\ -1 & -1 & -1 \\ 2 & 3 & 4 \end{pmatrix}$.

Discovery 0.8: What is the relationship?

$$\begin{tikzcd}
{[v]_\mathcal{B}} \arrow[r, phantom, "\in"] \arrow[d] \arrow[rrr, bend left] & \mathbb{R}^3 \arrow[r, "{[T]_\mathcal{B}}"] \arrow[d] \arrow[dr, phantom, "\circlearrowright"] & \mathbb{R}^3 \arrow[r, phantom, "\ni"] \arrow[d] & {[T(v)]_\mathcal{B}} \arrow[d] \\
{[v]_\mathcal{C}} \arrow[r, phantom, "\in"] \arrow[rrr, bend right] & \mathbb{R}^3 \arrow[r, "{[T]_\mathcal{C}}"] & \mathbb{R}^3 \arrow[r, phantom, "\ni"] & {[T(v)]_\mathcal{C}}
\end{tikzcd}$$

Let $S : \mathbb{R}^3 \to \mathbb{R}^3$ be the linear map with the property that $S[v]_\mathcal{B} = [v]_\mathcal{C}$, so S is a matrix. Therefore for matrix S:

$$1^{st} \text{ column} : S[b_1]_\mathcal{B} = [b_1]_\mathcal{C} \;\Rightarrow\; \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}_\mathcal{C} = \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}$$

$$2^{nd} \text{ column} : S[b_2]_\mathcal{B} = [b_2]_\mathcal{C} \;\Rightarrow\; \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}_\mathcal{C} = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$$

$$3^{rd} \text{ column} : S[b_3]_\mathcal{B} = [b_3]_\mathcal{C} \;\Rightarrow\; \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}_\mathcal{C} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

Thus we have $[S]_{\mathcal{B}\to\mathcal{C}} = \begin{pmatrix} 0 & -1 & 1 \\ -1 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$ as desired. Similarly we would find $[S]_{\mathcal{C}\to\mathcal{B}} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$.

Result 0.10

We have

$$[S]_{\mathcal{B}\to\mathcal{C}}[S]_{\mathcal{C}\to\mathcal{B}} = I$$
$$[S]_{\mathcal{C}\to\mathcal{B}}[S]_{\mathcal{B}\to\mathcal{C}} = I$$
$$[S]_{\mathcal{C}\to\mathcal{B}}[T]_{\mathcal{B}}[S]_{\mathcal{B}\to\mathcal{C}} = [T]_{\mathcal{C}}$$

Lecture 16 - Mon - Feb 12 - 2024

Recall last lecture's example. In general, if for vector space V with ordered basis $\mathcal{B} = (b_1, \ldots, b_n)$, suppose $T : V \to V$ is linear, then

$$[T]_{\mathcal{B}} = \begin{pmatrix} | & & | \\ [T(b_1)]_{\mathcal{B}} & \cdots & [T(b_n)]_{\mathcal{B}} \\ | & & | \end{pmatrix}$$

Definition 0.28: Similarity

Let A and B be $n \times n$ matrices. We way that $\mathcal{B}$ is similar to A if there exists $S \in \mathrm{GL_n(F)}$ (units of $M_n(F)$) such that $B = S^{-1}AS$.

Proposition 0.8: Similarity is equivalence relation

Similarity is equivalence relation.

Proof: We need to prove reflexivity, symmetry and trnasitivity:

1. Reflexive
   We have $A = I^{-1}AI$, so A is similar to A.

2. Symmetric
   If B is similar to A, then there exists $S \in \mathrm{GL_n(F)}$ such that $B = S^{-1}AS$, thus we have

   $$A = SBS^{-1} = (S^{-1})^{-1}B(S^{-1})$$

3. Transitive
   If A is similar to B which is similar to C, then we know

   $$A = S^{-1}BS$$
   $$B = T^{-1}CT$$

   which yields us that

   $$A = (TS)^{-1}C(TS)$$

:3 □

Definition 0.29: Notation

If A and B are similar, we write $A \sim B$.

Definition 0.30: Trace

Given an $n \times n$ matrix A, we define the trace of the matrix as

$$\operatorname{tr}(A) = \text{ sum of entries on the main diagonal}$$

Theorem 0.21

For $A, B \in M_n(F)$, we have

$$\operatorname{tr}(AB) = \operatorname{tr}(BA)$$

Proof: we have

$$\begin{aligned} \operatorname{tr}(AB) &= \sum_{i=1}^{n} (AB)(ii) \\ &= \sum_{i=1}^{n} \left(\sum_{k=1}^{n} A(ik)B(ki) \right) \\ \operatorname{tr}(BA) &= \sum_{i=1}^{n} \left(\sum_{k=1}^{n} B(ik)A(ki) \right) \end{aligned}$$

same thing. □

Corollary 0.4

If A and B are similar, then $\operatorname{tr}(A) = \operatorname{tr}(B)$.

Proof: We know that for some $S \in \mathrm{Gl_n(F)}$ we have

$$B = S^{-1}AS$$

Thus we let $X = S$ and $Y = AS^{-1}$, which gives us that

$$\operatorname{tr}(B) = \operatorname{tr}(XY) = \operatorname{tr}(YX) = \operatorname{tr}(A)$$

as desired. □

Result 0.11

Since $[T]_{\mathcal{B}}$ and $[T]_{\mathcal{C}}$ are similar, so

$$\operatorname{tr}([T]_{\mathcal{B}}) = \operatorname{tr}([T]_{\mathcal{C}})$$

Definition 0.31: Notation

If V and W are F-vector spaces, then we write either $\mathcal{L}(V, W)$ or $\mathrm{Hom_F(V, W)}$ for the set of linear maps from V to W.

Example 0.39

$\mathrm{Hom_F(F^n, F^m) = M_{m,n}(F)}$.

Theorem 0.22

Let V and W be F-vector spaces, suppose that $\dim V = n$ and $\dim W = m$, then we have

$$\mathrm{Hom_F(V, W) \cong M_{m,n}(F)} \text{ as } F\text{-vector spaces}$$

Discovery 0.9: $\mathrm{Hom_F(V, W)}$ **is an** F**-Vector Space**

Suppose we have $S, T \in \mathrm{Hom_F(V, W)}$, we know that for $v \in V$, we have $(T+S)(v) \in W$. Moreover,

$$\begin{aligned}(T+S)(v_1 + \lambda v_2) &= T(v_1 + \lambda v_2) + S(v_1 + \lambda v_2)\\ &= T(v_1) + T(\lambda v_2) + S(v_1) + S(\lambda v_2)\\ \mathrm{Hom_F(V, W)} \ni 0 &:= 0_{V,W}(v) = 0\\ -T &\text{ is also linear}\end{aligned}$$

Thus we have that $\mathrm{Hom_F(V, W)}$ is an abelian group under $+$.

Lecture 17 - Wed - Feb 14 - 2024

Proof: of the above theorem: Define

$$\begin{aligned}\psi : \mathrm{Hom_F(V, W)} &\to M_{m,n}(F)\\ \psi(T) &= [T]_{\mathcal{B},\mathcal{C}}\\ T &: V \to W\end{aligned}$$

We need to show that ψ is linear and one-to-one and onto.

1. Linear
   We want to show that
   $$\psi(T + \lambda S) = \psi(T) + \lambda\psi(S)$$
   for $T, S : V \to W$ and $\lambda \in F$. i.e.
   $$[T + \lambda S]_{\mathcal{B},\mathcal{C}} = [T]_{\mathcal{B},\mathcal{C}} + \lambda[S]_{\mathcal{B},\mathcal{C}}$$

Recall that j^{th} column of $y : V \to W$, $[y]_{\mathcal{B},\mathcal{C}}$, is $[y(b_j)]_{\mathcal{C}}$. STP the j^{th} columns are the same. We know that

$$\begin{aligned} j^{th} \text{ column of } [T+\lambda S]_{\mathcal{B},\mathcal{C}} &= [(T+\lambda S)(b_j)]_{\mathcal{C}} \\ &= [T(b_j) + \lambda S(b_j)]_{\mathcal{C}} \\ &= [T(b_j)]_{\mathcal{C}} + [\lambda S(b_j)]_{\mathcal{C}} \\ &= j^{th} \text{ column of } [T]_{\mathcal{B},\mathcal{C}} + \lambda \cdot j^{th} \text{ column of } [S]_{\mathcal{B},\mathcal{C}} \\ &= j^{th} \text{ column of } \left([T]_{\mathcal{B},\mathcal{C}} + \lambda \cdot [S]_{\mathcal{B},\mathcal{C}}\right) \end{aligned}$$

which implies linearity.

2. One-to-one
   STP $\ker(\psi) = (0)$, suppose $T \in \ker(\psi)$,

$$\begin{aligned} \psi(T) &= 0 \\ \Rightarrow \quad [T]_{\mathcal{B},\mathcal{C}} &= [0] \\ \Rightarrow \quad j^{th} \text{ column of } [T]_{\mathcal{B},\mathcal{C}} &= 0 \quad \text{ for } j = 1, 2, \ldots, n \\ \Rightarrow \quad [T(b_j)]_{\mathcal{C}} &= \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} \quad \text{ for } j = 1, 2, \ldots, n \\ \Rightarrow \quad T(b_j) &= 0 \quad \text{ for } j = 1, 2, \ldots, n \\ \Rightarrow \quad T &= 0 \end{aligned}$$

3. Onto
   Consider $\psi : \mathrm{Hom_F(V, W)} \to \mathrm{M_{m,n}(F)}$, we let

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

   we must find $T : V \to W$ such that $\psi(T) = [T]_{\mathcal{B},\mathcal{C}} = A$. We need to find the j^{th} column of $[T]_{\mathcal{B},\mathcal{C}} = \begin{bmatrix} a_{ij} \\ \vdots \\ a_{nj} \end{bmatrix} \Leftrightarrow a_{1j}c_1 + a_{2j}c_2 + \cdots + a_{mj}c_m = T(b_j)$. By Universal Property, there exists $T : V \to W$ linear such that $T(b_j) = a_{1j}c_1 + a_{2j}c_2 + \cdots + a_{mj}c_m$, $\forall\, j$, which implies that

$$[T]_{\mathcal{B},\mathcal{C}} = A = \begin{pmatrix} \ddots \end{pmatrix}$$

thus we complete the proof. $\square$

Rank, Nullity, Transpose

We know that both $M_{m,n}(F)$ and $M_{n,m}(F)$ both have dimension $m \times n$. In particular, they are isomorphic. We define the transpose map

$$^T : M_{m,n}(F) \to M_{n,m}(F)$$
$$A \rightsquigarrow A^T$$

Flipping along the main diagonal

Example 0.40

For $A, B \in M_{m,n}(F)$, we have $(A + \lambda B)^T = A^T + \lambda B^T$, in particular, ^{T} is a linear map.

Theorem 0.23

If $A \in M_{m,n}(F)$, $B \in M_{n,p}(F)$, then

$$(AB)^T = B^T A^T$$

Proof: Consider $(AB)^T(ij) = (AB)(ji) = \displaystyle\sum_{k=1}^{n} A(jk)B(ki)$.

Similarly $B^T A^T(ij) = \displaystyle\sum_{k=1}^{n} B^T(ik)A^T(kj) = \sum_{k=1}^{n} B(ki)A(jk)$.

Thus they are the same. □

Corollary 0.5

$(A_1 A_2 \cdots A_d)^T = {A_d}^T {A_{d-1}}^T \cdots {A_1}^T$ whenever the product makes sense.

Example 0.41

Consider $D : \mathbb{R}[x]_{\leq 2} \to \mathbb{R}[x]_{\leq 1}$ be the differentiation map, with basis $\mathcal{B} = (1, x, x^2), \mathcal{C} = (1, x)$ respectively. Suppose we want to find a linear map $T : \mathbb{R}[x]_{\leq 1} \to \mathbb{R}[x]_{\leq 2}$ such that $[T]_{\mathcal{C},\mathcal{B}} = [D]^T_{\mathcal{B},\mathcal{C}}$.

$$[D]_{\mathcal{B},\mathcal{C}} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} \quad \Rightarrow \quad D^T_{\mathcal{B},\mathcal{C}} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 2 \end{pmatrix}$$

Hence we have

$$[T]_{\mathcal{C},\mathcal{B}} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 2 \end{pmatrix} \implies [T(1)]_{\mathcal{B}} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \text{ and } [T(x)]_{\mathcal{B}} = \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix} \Rightarrow T(1) = x, T(x) = 2x^2$$

thus $T(a + bx) = ax + 2bx^2$.

Lecture 18 - Fri - Feb 16 - 2024

Let A be a $m \times n$ matrix in $M_{m,n}(F)$, then we know that A indces a linear map:

$$
\begin{array}{ccc}
T_A : F^n & \xrightarrow{T} & F^m \\
\downarrow & & \downarrow \\
M_{n,1}(F) & \xrightarrow{[T]_{\mathcal{B},\mathcal{C}}} & M_{m,1}(F)
\end{array}
$$

$$T(v) = A \cdot v$$

We define

1. $\operatorname{rank}(A) := \operatorname{rank}(T_A) = \dim(\operatorname{im}(T_A))$
2. $\operatorname{nullity}(A) := \operatorname{nullity}(T_A) = \dim(\ker(T_A))$
3. null space of $A := \ker(T_A)$

In general,

$$
\begin{aligned}
&(x_1, \ldots, x_m) \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \\
=& x_1[a_{11}, a_{12}, \ldots, a_{1n}] + x_2[a_{21}, a_{22}, \ldots, a_{2n}] + \cdots + x_m[a_{m1}, a_{m2}, \ldots, a_{mn}]
\end{aligned}
$$

Corollary 0.6

Let $A \in M_{m,n}(F)$, then $\operatorname{rank}(A) = \dim(\operatorname{span}(\text{col of } A))$, i.e.

$$A = \begin{pmatrix} | & | & \cdots & | \\ \vec{c_1} & \vec{c_2} & \cdots & \vec{c_n} \\ | & | & \cdots & | \end{pmatrix} \qquad \vec{c_i} \in M_{m,1}(F)$$

$$\Rightarrow \quad \operatorname{rank}(A) = \operatorname{rank}(\operatorname{span}(\{\vec{c_1}, \vec{c_2}, \ldots, \vec{c_n}\}))$$

Proof: We know that $\operatorname{rank}(A) = \dim(\operatorname{im}(T_A))$, where

$$F^n \to F^m \qquad T_A(v) = A \cdot v$$

Notice that $T_A \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = A \cdot \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = x_1\vec{c_1} + \cdots + x_n\vec{c_n}$, so $\operatorname{im}(T_A) \subseteq \operatorname{span}(\{\vec{c_1}, \vec{c_2}, \ldots \vec{c_n}\})$. Notice $\vec{c_j} \in \operatorname{im}(T_A)$ since $\vec{c_j} = A \cdot e_j$, so $\operatorname{span}(\{\vec{c_1}, \vec{c_2}, \ldots \vec{c_n}\}) \subseteq \operatorname{im}(T_A)$, so $\operatorname{rank}(A) = \dim(\operatorname{im}(T_A)) = \dim(\operatorname{span}(\{\vec{c_1}, \vec{c_2}, \ldots \vec{c_n}\}))$.
$\square$

Definition 0.32: Column Rank & Row Rank

If $A = \begin{pmatrix} | & | & \cdots & | \\ \vec{c_1} & \vec{c_2} & \cdots & \vec{c_n} \\ | & | & \cdots & | \end{pmatrix}$, we call $\dim(\text{span}(\{\vec{c_1}, \vec{c_2}, \dots \vec{c_n}\}))$ **the column rank of** A, if $A = \begin{pmatrix} - & \vec{r_1} & - \\ - & \vec{r_2} & - \\ \vdots & \vdots & \vdots \\ - & \vec{r_m} & - \end{pmatrix}$, then we call $\dim(\text{span}(\{\vec{r_1}, \vec{r_2}, \dots \vec{r_m}\}))$ the **row rank of** A.

Theorem 0.24

Let A be $m \times n$, then row rank $A =$ column rank A

Lemma 0.6

Let A be an $m \times n$ matrix. Let $\vec{c_1}, \vec{c_2}, \dots, \vec{c_p} \in M_{m,1}(F)$ be a basis of the span of columns of A, then there exists a $p \times n$ matrix R such that

$$A = \begin{pmatrix} | & | & \cdots & | \\ \vec{c_1} & \vec{c_2} & \cdots & \vec{c_p} \\ | & | & \cdots & | \end{pmatrix} R, \qquad \text{let } C = \begin{pmatrix} | & & | \\ \vec{c_1} & \cdots & \vec{c_p} \\ | & & | \end{pmatrix}$$

Proof: Let U_j denote the j^{th} column of A so

$$A = \begin{pmatrix} | & | & \cdots & | \\ \vec{u_1} & \vec{u_2} & \cdots & \vec{u_n} \\ | & | & \cdots & | \end{pmatrix}$$

By assumption, each U_j is in the span of $\vec{c_1}, \vec{c_2}, \dots, \vec{c_p}$. In particular, there exists $r_{1j}, r_{2j}, \dots, r_{pj} \in F$ such that

$$\vec{U_j} = r_{1j}\vec{c_1} + \cdots + r_{pj}\vec{c_p}$$

Let $R(ij) = r_{ij}$, then we have

$$C \cdot R = \begin{pmatrix} | & & | \\ C \begin{bmatrix} r_{11} \\ \vdots \\ r_{p1} \end{bmatrix} & \cdots & C \begin{bmatrix} r_{1n} \\ \vdots \\ r_{pn} \end{bmatrix} \\ | & & | \end{pmatrix} = \begin{pmatrix} | & & | \\ \vec{U_1} & \cdots & \vec{U_n} \\ | & & | \end{pmatrix} = A$$

□

Proof: of Theorem

Let $p =$ column rank of A and let $\vec{c_1}, \vec{c_2}, \ldots, \vec{c_p}$ be a basis for the column space of A. Let $C = \begin{pmatrix} | & & | \\ \vec{c_1} & \cdots & \vec{c_p} \\ | & & | \end{pmatrix}$. Then there exists a $p \times n$ matrix R such that $A = C \cdot R$. Write $R = \begin{pmatrix} - & \vec{r_1} & - \\ - & \vec{r_2} & - \\ \vdots & \vdots & \vdots \\ - & \vec{r_m} & - \end{pmatrix}$, then notice every row of $C \cdot R$ is a linear combination of $\vec{r_1}, \ldots, \vec{r_p}$, which implies that every row of A is in the span$(\{\vec{r_1}, \ldots, \vec{r_p}\})$. $\square$

Lecture 19 - Mon - Feb 26 - 2024

(1) How does one find the null space of a matrix?

(2) How does one find the image of a matrix?

(3) How does one find the inverse of an invertible matrix

(4) How does one find the rank of a matrix?

(5) How does one find the $[x]_{\mathcal{B}}$ for $x \in V$, $\mathcal{B}$ an ordered basis?

(6) How to check if a set is linearly independent?

(7) How do we check if a vector is in the span of a set?

Remark: For 1 and 2, we typically want to find a basis for these spaces.

Vector Equations

For $A \in M_{m,n}(F)$, $\vec{x} \in F^n$ and $\vec{b} \in F^m$, a vector equation is in the form of

$$A\vec{x} = \vec{b}$$

Moreover, when $\vec{b} = 0$ we call the equation homogenous, otherwise we say it is non-homogenous.

Elementary Row Operations

Ler A be an $m \times n$ matrix, we will say $A' \in M_{m,n}(F)$ can be obtained from A via a elementary row operations if one of the following these holds:

(a) The i^{th} row of A' is c times i^{th} row of A for $c \in F$ and all other rows are the same.

(b) We swap any two rows of A

(c) A' i^{th} row is obtained by replacing i^{th} row of A by i^{th} row of A and $c \times (j^{th})$ row of A for $i \neq j$ and $c \in F$.

Remark: If we can obtain A' from A via elementary row operations, then we can obtain A from A' from elementary row operations.

Definition 0.33: Row-Equivalent

We will write $A \to A'$ if A' can be ontained by an elementary row operations, we will say $A, B \in M_{m,n}(F)$ are row equivalent if there exists $d \geq 0$ and $A = A_0, \ldots, A_d = B \in M_{m,n}(F)$ such that

$$A = A_0 \to \cdots \to A_d = B$$

Proposition 0.9

Row equivalence is an equivalence relation on $M_{m,n}(F)$.

Proof: Definition check. □

Theorem 0.25

Let $A, B \in M_{m,n}(F)$, if A and B are row equivalent, then $A\vec{x} = \vec{0}$ and $B\vec{x} = \vec{0}$ have the same set of solutions.

Lecture 20 - Wed - Feb 28 - 2024

Discovery 0.10: Fact 1

$(I + c \cdot E_{ij})A$, $i \neq j$
performs an elementary row operation to A, in which we take the i^{th} row of A and add $c \cdot j^{th}$ row of A to it.

Remark: Notice that if $i \neq j$, then we have

$$(I + c \cdot E_{ij})(I - c \cdot E_{ij}) = I$$

Important to note that difference of squares does not generally hold for matrices.

Discovery 0.11: Fact 2

$$\left(\sum_{r \neq i} E_{rr} + c \cdot E_{ii}\right) A, \; c \neq 0$$

performs an elementary row operation to A, in which we take the i^{th} row of A and scale it by $c \neq 0$.

Discovery 0.12: Fact 3

$$\left(\sum_{r \neq i,j} E_{rr} + E_{ij} + E_{ji}\right) A, \; c \neq 0$$

performs an elementary row operation to A, in which we interchange row i and j.

Corollary 0.7

If A and B are row equivalent $m \times n$ matrices, then there exists $U \in \mathrm{GL}_m(F)$ such that $B = UA$.

Proof: Since B and A are row equivalent, there exist $A = A_0, A_1, \dots, A_d = B$ such that

$$A = A_0 \to A_1 \to \cdots \to A_d = B$$

Thus by Fact 1-3, for $i = 0, 1, \dots, d-1$, we have $A_{i+1} = U_i A_i$ for $U_i \in \mathrm{GL}_m(F)$, so we have

$$B = A_d = U_{d-1} U_d \cdots U_0 A_0$$

Because for the fact that $U_{d-1}, U_d, \dots, U_0 \in \mathrm{GL}_m(F)$ which is a group, so we can take their product to be U, which then yields us that $B = UA$. $\square$

Proof: of theorem 0.5
By Corollary, $B = UA$ for $U \in \mathrm{GL}_m(F)$, so we have

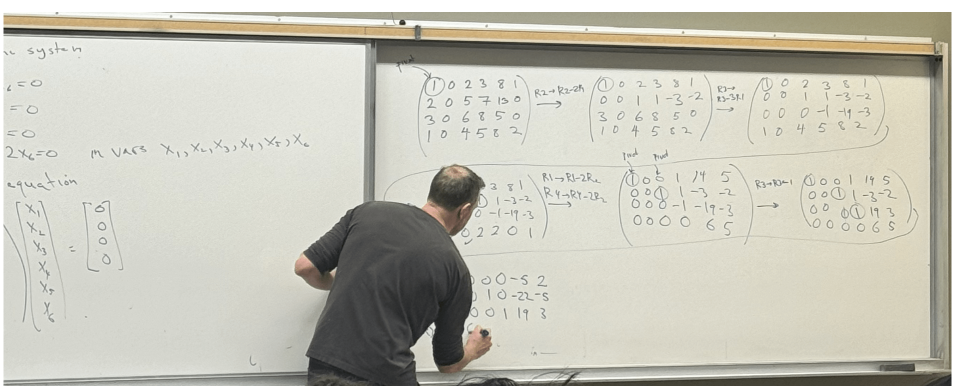
$$B\vec{x} = \vec{0} \;\Leftrightarrow\; UA\vec{x} = \vec{0} \;\Leftrightarrow\; U^{-1}UA\vec{x} = U^{-1}\vec{0} \;\Leftrightarrow\; A\vec{x} = \vec{0}$$

thus completing the proof. $\square$

Example 0.42

Find all solution to the system:

$$\begin{aligned} x_1 + 2x_3 + 3x_4 + 8x_5 + x_6 &= 0 \\ 2x_1 + 5x_3 + 7x_4 + 3x_5 &= 0 \\ 3x_1 + 6x_3 + 8x_4 + 5x_5 &= 0 \\ x_1 + 4x_3 + 5x_4 + 8x_5 + 2x_6 &= 0 \end{aligned}$$



Bell cooking.

Definition 0.34: RREF

A matrix A is in row-reduced echelon form (RREF) if the following hold:

- the first nonzero entry in each nonzero row of A is equal to 1 (we call these the pivots of A and their corresponding columns the pivot columns of A);
- each column containing the leading 1 of some non-zero row has all of its other entries equal to zero;
- all zero rows of A are below all nonzero rows;
- if $\vec{r_1}, \dots, \vec{r_r}$ are the nonzero rows of A and the leading nonzero 1 of $\vec{r_i}$ occurs in position k_i (i.e., in column k_i) then $k_1 < k_2 < \cdots < k_r$.

We'll see that if A and B are in RREF and are row equivalent, then $A = B$. In other words, RREF of a matrix is unique.

Theorem 0.26

Every $m \times n$ matrix A is equivalent to a matrix in RREF.

Proof: We do this by induction on the number of rows (m)

1. Base Case, $m = 1$:
   Then either A is the zero row and we are done, or A is not zero, then there exists some i such that the i^{th} column of A is not zero and the column before that is 0, where we can simply scale the row by $1/c$ and thus obtain the matrix in RREF.

2. Induction Hypothesis, let $k \geq 2$ and the result holds whenever $m < k$:

3. Induction Step, consider the case when $m = k$:
   Let $\vec{c_1}, \dots, \vec{c_n}$ denote the columns of A and let i be the smallest index for which $\vec{c_j} \neq 0$.

   After performing a row swap, we can arrange it so that the first coordinate of $\vec{c_i} \neq 0$. After scaling the first row, we can then assume that

   $$\vec{c_i} = \begin{bmatrix} 1 \\ \vdots \\ * \end{bmatrix}$$

   Now by performing row operations, we can ensure that everything below the 1 in $\vec{c_i}$ is 0. Then by the induction hypothesis, we can use the elementary row operations to put the small section in the southeast corner into a RREF.

   Then we use those pivots to clear all the non-zero entry above them, so we are done.

□

Solving Equations

$$A\vec{x} = \vec{b}$$

Input A which is $m \times n$ and $b \in F^m$, where $x_1, \ldots, x_n$ are unknowns, and we want to solve for $\vec{x}$.

Algorithm 0.1

1. Step 1:
   Make an $m \times (n+1)$ matrix $\left(A \mid \vec{b}\right)$
2. Step 2:
   We use elementary row operations to put $\left(A \mid \vec{b}\right)$ into RREF
3. Step 3:
   Let $\left(B \mid \vec{c}\right)$ denote the RREF matrix obtained in Step 2, solve $B\vec{x} = \vec{c}$ directly gives us the solution, which are precisely the solutions to $A\vec{x} = \vec{b}$.

We saw that if $\left(A \mid \vec{b}\right)$ and $\left(B \mid \vec{c}\right)$ are row equivalent $m \times (n+1)$ matrices, then there exists $U \in \mathrm{GL}_m(F)$ such that

$$\left(B \mid \vec{c}\right) = U \cdot \left(A \mid \vec{b}\right)$$

Remark:

$$U \cdot \begin{pmatrix} | & | & & | \\ \vec{c_1} & \vec{c_2} & \cdots & \vec{c_r} \\ | & | & & | \end{pmatrix} = \begin{pmatrix} | & | & & | \\ U \cdot \vec{c_1} & U \cdot \vec{c_2} & \cdots & U \cdot \vec{c_r} \\ | & | & & | \end{pmatrix}$$

so $U \cdot \left(A \mid \vec{b}\right) = \left(U \cdot A \mid U \cdot \vec{b}\right) = \left(B \mid \vec{c}\right)$, so $B = U \cdot A$, $\vec{c} = U \cdot \vec{b}$. So now

$$B\vec{x} = \vec{c} \Leftrightarrow UA\vec{x} = U\vec{b} \Leftrightarrow A\vec{x} = \vec{b} \quad \because U \text{ invertible}$$

Remark: If $\left(A \mid \vec{b}\right)$ is in RREF, then $A\vec{x} = \vec{b}$ has no solution iff the last column is a pivot column.

Lecture 22 - Mon - Mar 4 - 2024

Example 0.43

Suppose we have $V = \mathbb{R}[x]_{\leq 2}$ and $\mathcal{B} = \{1 + x + x^2, 1 + 2x + 4x^2, 1 + 3x + 9x^2\}$. How would you find $[1 + x]_{\mathcal{B}}$?

Proof: We write

$$[1 + x]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} \in \mathbb{R}^3$$

This means that

$$1 + x = c_1(1 + x + x^2) + c_2(1 + 2x + 4x^2) + c_3(1 + 3x + 9x^2)$$

Comparing the coefficients, we have

$$\begin{aligned} x^0 &: 1 = c_1 + c_2 + c_3 \\ x^1 &: 1 = c_1 + 2c_2 + 3c_3 \\ x^2 &: 0 = c_1 + 4c_2 + 9x_3 \end{aligned}$$

which corroesponds to

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 4 & 9 \end{pmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \quad \rightsquigarrow \quad \text{RREF} \to \text{solution}$$

□

Theorem 0.27

If $(A \mid \vec{b})$ is in RREF, then $A\vec{x} = \vec{b}$ has a solution if and only if the last column of it is not a pivot column.

Proof: 1. ($\Longrightarrow$)
Last time we proved this using contrapositive.
2. ($\Longleftarrow$)
Recall that if A and B are row equivalent, then

$$\{\vec{x} : A\vec{x} = \vec{0}\} = \{\vec{x} : B\vec{x} = \vec{0}\}$$

Remark: In general, if $j_1 < j_2 < \cdots < j_k$ are the pivot columns and we let $u_1, \ldots, u_d$ denote the free variables, then the equation $B\vec{x} = \vec{0}$ gives rise to k non-trivial linear equations for the form

$$x_{ji} + \lambda_{i1}u_1 + \lambda_{i2}u_2 + \cdots + \lambda_{id}u_d = 0$$

for $i = 1, \ldots, k$, and $\lambda_{ij} \in F$.

In particular, we see how to give all solutions,

1. we can easily assign any value in out field to the free vars
2. the bound variables are uniquely determined by this assignment

We continue the proof:

If $(A \mid \vec{b})$ is in RREF and the last column is not a pivot column, then $A\vec{x} = \vec{b}$ has a solution. We have

$$\left(\begin{array}{c|c} A & b_1 \\ & b_2 \\ & \vdots \\ & b_m \end{array}\right) \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n+1} \end{bmatrix} = \vec{0}$$

If the last column is not a pivot column, then x_{n+1} is a free variable, so we can give it any value and there will still be a solution.

Example 0.44

$$\begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 3 \end{pmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

In general, we have

$$(A \mid \vec{b}) \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \\ -1 \end{bmatrix} = \vec{0} \quad \Leftrightarrow \quad A \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} - \vec{b} = 0 \quad \Leftrightarrow \quad A \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} = \vec{b}$$

as desired. □

We now show that RREF of a matrix is unique.

Theorem 0.28

If B and C are $m \times n$ matrices in RREF that are row equivalent, then $B = C$.

Corollary 0.8

RREF of a $m \times n$ matrix A is unique.

Proof: If $A \to B$ and $A \to C$, B, C in RREF, then because row equivalence is transitive, then B and C are equivalent, so $B = C$. □

Proof: of Theorem
We proved this by induction on $n =$ number of columns of B and C

1. Base Case: $n = 1$
   A $m \times 1$ matrix in RREF is either

$$\begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

2. Now we assume the result holds whenever $n \leq k$

3. Consider the case when $n = k + 1$

□

Lecture 23 - Wed - Mar 6 - 2024

Recall $A\vec{x} = \vec{0}$, we can row reduce A to a B in RREF such that B has certain columns that are pivot columns and certain columns that are not. The free variables are those indexed by the non-pivot columns while bounded variables are indexed by pivot columns $x_{j0}, \ldots, x_{jk}$. Thus $A\vec{x} = \vec{0}$ is equivalent to the system

$$\begin{aligned} x_{j1} &= \lambda_{11}u_1 + \cdots + \lambda_{1d}u_1 \\ x_{j2} &= \lambda_{21}u_1 + \cdots + \lambda_{2d}u_1 \\ &\vdots \\ x_{jk} &= \lambda_{k1}u_1 + \cdots + \lambda_{kd}u_1 \end{aligned}$$

Corollary 0.9

Nullity of A = number of free variables = number of non-pivot columns.

Proof: Let $u_1, \ldots, u_d$ denote the free variables and $x_{j1}, \ldots, x_{jk}$ denote the $\mathfrak{B}$ounded variables. The nullspace of A is

$$\{\vec{x} : A\vec{x} = \vec{0}\}$$

and the nullity is the dimension of the nullspace. After relabelling,

$$\{\vec{x} : A\vec{x} = \vec{0}\} = \{(u_1, \ldots, u_d, x_{j1}, \ldots, x_{jk}) : \textit{system above}\}$$

We create a map $T : F^d \to \text{nullspace}(A)$ such that

$$T(c_1, \ldots, c_d) = (c_1, \ldots, c_d, \lambda_{11}c_1 + \cdots + \lambda_{1d}c_d, \ldots, \lambda_{k1}c_1 + \cdots + \lambda_{kd}c_d)$$

Notice T is linear, and the kernel of T is $(0, \ldots, 0)$ because

$$T(a_1, \ldots, a_d) = (\underbrace{0, \ldots, 0}_{d}, \underbrace{0, \ldots, 0}_{k})$$

thus $a_1 = \cdots = a_d = 0$, which implies that T is one-to-one. Notice T is also onto since every elemetn of the nullspace is uniquely determined by an assignment of the free variable. Therefore, T is an isomorphism, so the nullspace of $A \cong F^d$, which then gives us that nullity of A is the same as the number of free variables. $\square$

Example 0.45

If A is $m \times n$ with columns $c_1, \ldots, c_n$, thus

$$A \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = x_1\vec{c_1} + \cdots + x_n\vec{c_n}$$

If A is $m \times n$ with

$$A = \begin{pmatrix} | & | & & | \\ \vec{c_1} & \vec{c_2} & \cdots & \vec{c_n} \\ | & | & & | \end{pmatrix}$$

then

$$
\begin{aligned}
T_A &: F^n \to F^m \\
T_A(\vec{x}) &= A\vec{x} \\
\operatorname{im}(T_A) &= \{A \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} : x_1, \dots, x_n \in F\} \\
&= \{x_1\vec{c_1} + \cdots + x_n\vec{c_n} : x_1, \dots, x_n \in F\} \\
&= \operatorname{span}(\text{cols of } A) \\
&= \operatorname{Col}(A)
\end{aligned}
$$

How do we find a basis for the column space of A? or the basis for $\operatorname{im}(T_A)$

Result 0.12

We put A to B in RREF. If $j_1, \dots, j_k$ are the pivot columns, then the corresponding columns of A is the basis of the column space of A .

Lemma 0.7

If B is in RREF, then the pivot columns of B form a basis for the column space of B

Proof: Let $b_1, \dots, b_k$ denote the pivot columns of B and let $j_1 < j_2 < \cdots < j_k$ denote the coordinate where 1 appears in $\vec{b_1}, \dots, \vec{b_k}$. Then

$$
\begin{aligned}
\vec{b_1} &= \vec{e_{j_1}} \\
\vec{b_2} &= \vec{e_{j_2}} \\
&\vdots \\
\vec{b_k} &= \vec{e_{j_k}}
\end{aligned}
$$

Notice that $\{e_1, \dots, e_m\}$ is linearly independent, so that $\{\vec{e_{j_1}}, \dots, \vec{e_{j_k}}\} \subseteq \{e_1, \dots, e_m\}$ is also linearly independent. Therefore, $\{\vec{b_1}, \dots, \vec{b_k}\}$ is also linearly independent.

Now to show that $\vec{b_1}, \dots, \vec{b_k}$ span the column space, it suffices to show that if $\vec{b}$ is another column of B, then $\vec{b} \in \{\vec{b_1}, \dots, \vec{b_k}\}$.

Now let $\vec{b}$ be a not pivot column of B, our claim is that if the i^{th} coordinate of $\vec{b}$ is non-zero, then $i \in \{j_1, j_2, \dots, j_k\}$.

Proof: of the claim:

If i^{th} coordinate is non-zero, then the i^{th} row is also non-zero, so the first column with a non-zero entry in the i^{th} row is a pivot column, which implies that $i \in \{j_1, j_2, \dots, j_k\}$. $\square$

Therefore,

$$
\begin{aligned}
\vec{b} &= \lambda_1 \vec{e_{j_1}} + \cdots + \lambda_k \vec{e_{j_k}} \\
&= \lambda_1 \vec{b_1} + \cdots + \lambda_k \vec{b_k}
\end{aligned}
$$

so $\vec{b} \in \operatorname{span}\{\text{pivot cols of } B\}$ $\square$

Theorem 0.29

If A is row equivalent to B in RREF and B has pivot columns $j_1, \dots, j_k$, which implies that the $j_1, \dots, j_k$ columns of A form a basis for the column space of A and $\text{rank}(A) = k =$ number of pivot columns.

Proof: Write

$$A = \begin{pmatrix} | & | & & | \\ \vec{c_1} & \vec{c_2} & \cdots & \vec{c_n} \\ | & | & & | \end{pmatrix} \qquad B = \begin{pmatrix} | & | & & | \\ \vec{u_1} & \vec{u_2} & \cdots & \vec{u_n} \\ | & | & & | \end{pmatrix}$$

If A and B are row equivalent, then there exists an invertible $n \times n$ matrix U such that $A = UB$. If $\vec{u_{j_1}}, \dots, \vec{u_{j_k}}$ are the pivot columns of B. Notice

$$\begin{aligned} UB &= U \begin{pmatrix} | & | & & | \\ \vec{u_1} & \vec{u_2} & \cdots & \vec{u_n} \\ | & | & & | \end{pmatrix} \\ &= \begin{pmatrix} | & | & & | \\ U \cdot \vec{u_1} & U \cdot \vec{u_2} & \cdots & U \cdot \vec{u_n} \\ | & | & & | \end{pmatrix} \end{aligned}$$

Our claim is that the columns above form a basis for the column space of A:
This is because U is invertible so a unique linear combination is still unique. $\square$

Corollary 0.10

If A is $m \times n$, then

$$\begin{aligned} n = & \text{ number of columns} \\ = & \text{ number of pivot columns } + \text{ number of non-pivot columns} \\ = & \text{ rank} + \text{nullity} \end{aligned}$$

Lecture 24 - Fri - Mar 8 - 2024

Result 0.13

For $A \in M_n(F)$, the following are equivalent:

1. A is invertible
2. $T_A : F^n \to F^m$, $T_A(\vec{x}) = A \cdot \vec{x}$ is bijective
3. Nullspace for A is (0)
4. $\text{Null}(A) = 0$
5. $\text{rank}(A) = n$

6. The columns of A form a basis for F^n

Proof:

1. $1 \to 2$
   If A is invertible, then there exists B such that $BA = AB = I$, so $T_A(B\vec{x}) = A \cdot (B\vec{x}) = (AB) \cdot \vec{x} = \vec{x}$, so T_A is onto. By rank-nullity theorem, it is also one to one.

2. $2 \to 3$
   $\ker(T) = \{\vec{x} : A\vec{x} = \vec{0}\}$, but T_A is one to one, thus $\ker(T) = \{\vec{0}\}$, which implies that null space of A is $(\vec{0})$.

3. $3 \to 4$
   Immediate

4. $4 \to 5$
   Follows from rank-nullity theorem

5. $5 \to 6$
   If rank $= n$, then column rank of $A = n$, then if $\vec{v_1}, \ldots, \vec{v_n}$ are the columns of A, then $\operatorname{span}\{\vec{v_1}, \ldots, \vec{v_n}\} = F^n$, which implies that $\vec{v_1}, \ldots, \vec{v_n}$ is a basis for F^n

6. $6 \to 1$
   If the columns of A form a basis, then let $\vec{v_1}, \ldots, \vec{v_n}$ denote these columns, and since they are basis, then for every $j \in \{1, \ldots, n\}$, we have $e_j = b_{1j}\vec{v_1} + \cdots + b_{nj}\vec{v_n} : b_{1j}, \ldots, b_{nj} \in F$, so this means

$$
\begin{aligned}
A \begin{bmatrix} b_{1j} \\ \vdots \\ b_{nj} \end{bmatrix} &= \begin{pmatrix} | & | & & | \\ \vec{v_1} & \vec{v_2} & \cdots & \vec{v_n} \\ | & | & & | \end{pmatrix} \begin{bmatrix} b_{1j} \\ \vdots \\ b_{nj} \end{bmatrix} \\
&= b_{1j}\vec{v_1} + \cdots + b_{nj}\vec{v_n} \\
&= \vec{e_j}
\end{aligned}
$$

   Therefore, we know that there exists matrix B such that $A \cdot B = I$

We formed a loop. □

We recall that for 2×2 matrices $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$, we had a very simple test for invertibility:

$$
\begin{pmatrix} a & b \\ c & d \end{pmatrix} \text{ is invertible} \quad \Leftrightarrow \quad ad - bc \neq 0
$$

We'd like to extend this for larger matrices.

Determinants

For what follows,

$$A = \begin{pmatrix} — & \vec{r_1} & — \\ — & \vdots & — \\ — & \vec{r_n} & — \end{pmatrix}$$

and we will sometimes write $A = (\vec{r_1}, \dots, \vec{r_n})$

n-linear

Definition 0.35: n-linear

We say that a function

$$D : M_n(F) \to F$$

is n-linear if, when we fix all rows other than the i^{th} row and let the i^{th} rows vary, we obtain a linear function of the i^{th} row, for $i = 1, 2, \dots, n$

Exercise: Show that $D : M_2(F) \to F$ that sends $aE_{1,1} + bE_{1,2} + cE_{2,1} + dE_{2,2}$ to $ad - bc$ is 2-linear.

Example 0.46

If

$$D : M_2(F) \to F$$
$$D\left(\begin{pmatrix} a & b \\ c & d \end{pmatrix}\right) = ac$$

Then D is 2-linear.

Example 0.47

Let

$$D\left(\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}\right) = aei \qquad E\left(\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}\right) = a + e + i$$

Then D is 3-linear, and E is not 3-linear. E is not linear because you would send 0 to non-zero element.

Definition 0.36: Alternating

Let $D : M_n(F) \to F$ be n-linear, then we say D is **alternating** if the following hold:

> If B is obtained from A by interchanging rows $i, j, i \neq j$, then $D(B) = -D(A)$ and whenever A has two equal rows, $D(A) = 0$.

Result 0.14

Let F be a field of characteristic not equal to two. Then if D is n-linear then $D(A) = -D(B)$ whenever B is obtained by interchanging two rows of some matrix A if and only if $D(C) = 0$ whenever C has two identical rows. Moreover, the converse holds without any restriction on the characteristic of the field.

Proof: Suppose that $D(A) = -D(B)$ whenever B is obtained by interchanging two rows of some matrix A. Then if C has two identical rows, then if we switch these rows we see $D(C) = -D(C)$ and so $2D(C) = 0$ which gives that $D(C) = 0$. Conversely, now assume that $D(C) = 0$ whenever C has two identical rows and let A be a matrix with rows $r_1, \dots, r_n$. Then create a new matrix A' with rows $s_1, \dots, s_n$ where $s_i = r_i$ for all $i \neq j, k, j < k$, and $s_j = r_j + r_k, s_k = r_j + r_k$. Then $D(A') = 0$ since A' has two equal rows. But notice since D is a linear function of the j^{th} row, we see that

$$0 = D(A') = D(s_1, \dots, sn) = D(s1, \dots, s_{i-1}, r_j, s_{i+1}, \dots, s_n) + D(s1, \dots, s_{i-1}, r_k, s_{i+1}, \dots, s_n)$$

Next we use linearity of the k^{th} row to get that the RHS is the sum of four terms

$$\begin{gathered}
D(s_1, \dots, s_{i-1}, r_k, s_{i+1}, \dots, sk-1, r_j, s_{k+1}, \dots, s_n), \\
D(s_1, \dots, s_{i-1}, r_k, s_{i+1}, \dots, sk-1, r_k, s_{k+1}, \dots, s_n), \\
D(s_1, \dots, s_{i-1}, r_j, s_{i+1}, \dots, sk-1, r_j, s_{k+1}, \dots, s_n), \\
D(s_1, \dots, s_{i-1}, r_j, s_{i+1}, \dots, sk-1, r_k, s_{k+1}, \dots, s_n).
\end{gathered}$$

Notice the second and third of these terms is zero, since the matrices have two equal rows. Since the four terms sum to zero, we see that $D(A) = -D(B)$ where A is the matrix with rows $r_1, \dots, r_n$ and B is obtained by switching the j^{th} and k^{th} rows of A. $\square$

Remark: Notice that, however, if $D(A) = -D(B)$ when we interchange two rows of A to obtain B and D is n-linear, this **does not** imply $D(C) = 0$ whenever C has two equal rows. What if we are working in field with characteristic 2?

Lemma 0.8

If $A \to B$ through RREF, then $D(A) = D(B)$ if it is $R_i \to R_i + cR_j$; $D(A) = -D(B)$ if it is $R_i \to R_j$ & $R_j \to R_i$; and $D(B) = cD(A)$ if it is $R_i = cR_i, c \neq 0$.

Proof: If we scale the i^{th} row of A by a nonzero scalar c, then $D(A') = cD(A)$ since D is linear as a function of the i^{th} row; if A' is obtained by interchanging two rows of A then $D(A') = -D(A)$. Finally, we can check if A has rows $r_1, \dots, r_n$ and we replace row i by $r_i + cr_j$ then by n-linearity $D(r_1, \dots, r_{i-1}, r_i + cr_j, r_{i+1}, \dots, r_n) = D(A)$. It follows that if A and A' are row equivalent then $D(A) = \alpha D(A')$ with α a nonzero element of F. $\square$

Now consider an n-linear alternating function

$$D : M_n(F) \to F$$

Result 0.15

Our claim is that if $D(I) \neq 0$, then $D(A) = 0 \iff A$ is not invertible.

Proof: We row reduce A to a matrix B in RREF. By the above, $D(A) \neq 0 \iff D(B) \neq 0$. If B is the identity matrix then A is invertible since it has full rank and since $D(I) \neq 0$ we see that $D(A) \neq 0$. If B is not the identity matrix, then we do not have a pivot in every column and since the number of rows is equal to the number of columns, we must have a zero row. Thus $D(B) = 0$ and since A is row equivalent to B we see that $D(A) = 0$. $\square$

Lecture 25 - Mon - Mar 11 - 2024

Definition 0.37: Determinant

We say a map

$$D : M_n(F) \to F$$

is a **determinant** function if

1. It is n-linear
2. It is alternating
3. $D(I) = 1$

Remark: We will see that for all $n \geq 1$, there exists a unique function, which we will call the determinant.

Example 0.48

For $n = 1$, we would have $D((a)) = a$, which is indeed n-linear, alternating, and $D((1)) = 1$.

Example 0.49

For $n = 2$, we would have

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

is 2-linear, alternating, and $D((I)) = 1$.

Discovery 0.13

We will now show that the determinant function exist for every n by induction on n.

Proof:

1. Base cass, $n = 1, 2$:
2. Suppose that $d \geq 3$, there exists a determinant function whenever $n < d$.
3. We will show that there exists a determinant function

$$D : M_d(F) \to F$$

Definition 0.38: $A(i \mid j)$

Given an $m \times m$ matrix A, we let $A\,(i \mid j)$ denote the $(m-1) \times (m-1)$ matrix obtained by deleting the i^{th} row and the j^{th} column.

Proposition 0.10

Let $\det : M_{l-1}(F) \to F$ be a determinant function. For $j = 1, 2, \ldots, d$ we define

$$E_j : M_l(F) \to F$$
$$E_j(A) = \sum_{i=1}^{d} a_{ij}(-1)^{i+j} \det(A\,(i \mid j))$$

then each E_j is a determinant function; i.e. they are n-linear, alternating and $E_j(I) = 1$.

Proof: We first compute $E_i(I_n)$. Since in this case the (i,j)-entry of I is zero unless $j = i$ and is 1 if $i = j$, we see that $E_j(A) = (-1)^{2j} D(I(j \mid j)) = D(I_{n-1}) = 1$.
We now show E_j is alternating. Notice that it suffices to show that $D(A) = 0$ whenever A has two identical rows and it suffices to consider the case when the two rows are adjacent, so suppose that the p^{th} and $(p+1)^{st}$ row of A are the same. Then $A(i|j)$ has two identical rows if $i \notin \{p, p+1\}$ and so $D(A(i|j)) = 0$ for all i except when $i \in \{p, p+1\}$, and so $E_j(A) = a_{i,j}(-1)^{i+j} D(A(i|j)) + a_{i+1,j}(-1)^{i+j+1} D(A(i+1|j))$. One can now easily check these two terms cancel. Now let us consider the case where A has rows r_k for $k \neq i$ and row i is $r_i + cs_i$. We must show that $E_j(A) = E_j(A_1) + cE_j(A_2)$, where A_1 is the matrix in which the k^{th} row is r_k for all k, and A_2 is the matrix in which the k^{th} row is r_k if $k \neq i$ and the i^{th} row is s_i. Then by $(n-1)$-linearity of D, we see that $D(A(p|j)) = D(A_1(p|j)) + cD(A_2(p|j))$ if $p \neq i$ and $D(A(i|j)) = D(A_1(i|j)) = D(A_2(i|j))$ since all rows other than the i^{th} are the same in A, A_1, and A_2. So from these fact is follows that $E_j(A) - E_j(A_1) - cE_j(A_2) = (A(i,j) - A_1(i,j) - cA_2(i,j))(-1)^{i+j} D(A(i|j))$. Notice $A(i,j) - A_1(i,j) - cA_2(i,j) = 0$ so we obtain linearity in each row, so E_j is a determinant function.
□ □

We were tying to show that determinant functions exists and they are unique.

Recall that to check an n-linear map is alternating, it suffices to check that $D(C) = 0$ whenever C has two consecutive equal rows, and we were showing that the determinant functions exist by induction on n.

Discovery 0.14

We now show that there exist exactly one determinant function.

Proof: For the proof, let e_j denote a row of n zeros with j^{th} spot being 1. Let

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \qquad \vec{r_i} = (a_{i1}, \dots, a_{in})$$

Therefore we can write the first row of A as $a_{11}\vec{e_1} + \cdots + a_{1n}\vec{e_n}$. If we fix all other rows and look at D as a linear function of the first row, then

$$\begin{aligned} D(A) &= a_{11} D \begin{pmatrix} — & \vec{e_1} & — \\ — & \vdots & — \\ — & \vec{r_n} & — \end{pmatrix} + \cdots + a_{n1} D \begin{pmatrix} — & \vec{e_n} & — \\ — & \vdots & — \\ — & \vec{r_n} & — \end{pmatrix} \\ &= \sum_{j=1}^{n} a_{1j} D \begin{pmatrix} — & \vec{e_j} & — \\ — & \vdots & — \\ — & \vec{r_n} & — \end{pmatrix} \end{aligned}$$

Similarly, we can proceed the same process for the second row of A. Therefore we can rewrite this as

$$D(A) = \sum_{j_1=1}^{n} \sum_{j_2=1}^{n} a_{1j_1} a_{2j_2} D \begin{pmatrix} — & \vec{e_1} & — \\ — & \vec{e_2} & — \\ — & \vdots & — \\ — & \vec{r_n} & — \end{pmatrix}$$

Continuiting in this manner, we see that

$$D(A) = \sum_{j_1=1}^{n} \cdots \sum_{j_n=1}^{n} \left[a_{1j_1} \cdots a_{nj_n} \cdot \underbrace{D \begin{pmatrix} — & \vec{e_1} & — \\ — & \vec{e_2} & — \\ — & \vdots & — \\ — & \vec{e_n} & — \end{pmatrix}}_{\text{scalar}} \right]$$

Example 0.50

For 2×2, and D is 2-linear with $D(I) = 1$, then

$$\begin{aligned} D\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} &= \cancel{a_{11}a_{21}D\begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}} + a_{11}a_{22}D\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + a_{12}a_{21}D\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + \cancel{a_{12}a_{22}D\begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}} \\ &= a_{11}a_{22}D\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - a_{12}a_{21}D\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ &= a_{11}a_{22} - a_{12}a_{21} \end{aligned}$$

Suppose that D is n-linear and alternating, $D : M_n(F) \to F$, thus

$$D(A) = \sum_{j_1=1}^{n} \cdots \sum_{j_n=1}^{n} \left[\prod_{i=1}^{n} a_{ij_i} \cdot D\begin{pmatrix} — & \vec{e_1} & — \\ — & \vec{e_2} & — \\ — & \vdots & — \\ — & \vec{e_n} & — \end{pmatrix} \right]$$

Notice if there exist $p \neq q$ such that $j_p = j_q$ implies that $\vec{e_{j_p}} = \vec{e_{j_q}}$, which impies that $D(matrix) = 0$. Thus we may assume that $j_1, \ldots, j_n$ are pairwise distinct. Since they all take values in $\{1, 2, \ldots, n\}$ we see that $\{1, 2, \ldots, n\}$ must be a rearrangement of $1, 2, \ldots, n$. That is, there must be a one-to-one and onto map $\sigma : \{1, \ldots, n\} \to \{1, \ldots, n\}$ such that $j_i = \sigma(i)$ for $i = 1, 2, \ldots, n$. We recall that the set of bijective set maps is called the set of permutations of $\{1, 2, \ldots, n\}$ and it forms a group under composition; this group is called the n-th symmetric group and we let S_n denote this group:

$$S_n = \{\sigma : \{1, 2, \ldots, n\} \to \{1, 2, \ldots, n\}\}$$

Then we can rewrite the sum as

$$D(A) = \sum_{\sigma \in S_n} \prod_{i=1}^{n} a_{i\sigma(i)} \cdot D\begin{pmatrix} — & \vec{e_1} & — \\ — & \vec{e_2} & — \\ — & \vdots & — \\ — & \vec{e_n} & — \end{pmatrix}$$

For $\sigma \in S_n$, we define $\operatorname{sgn}(\sigma) := D(matrix)$. We call this quantity the *sign* of the permutation σ. We will show that this sign is always 1 or -1. $\square$

Lemma 0.9

Let $\sigma \in S_n$, then we can perform a series of row interchanges to transform $\begin{pmatrix} — & \vec{e}_{\sigma(1)} & — \\ — & \vdots & — \\ — & \vec{e}_{\sigma(n)} & — \end{pmatrix}$ to the identity matrix. If $d, e \in \mathbb{N}$ and we can row reduce $\begin{pmatrix} — & \vec{e}_{\sigma(1)} & — \\ — & \vdots & — \\ — & \vec{e}_{\sigma(n)} & — \end{pmatrix} \to I$ using d row interchanges and using e interchanges, then $d \equiv e \pmod 2$.

Proof:
Proof: of first statement:
We induct on n, notice that it is true for $n = 1$.
Induction Hypothesis: Assume true for $n < d$, $d \geq 2$.
Consider the case when $n = d$:

$$\begin{pmatrix} — & \vec{e}_{\sigma(1)} & — \\ — & \vdots & — \\ — & \vec{e}_{\sigma(d)} & — \end{pmatrix} \rightsquigarrow \begin{pmatrix} — & \vec{e}_{\tau(1)} & — \\ — & \vdots & — \\ — & \vec{e}_{\tau(d-1)} & — \\ 0 & \cdots & 1 \end{pmatrix} = \begin{pmatrix} — & \vec{e}_{\tau(1)} & 0 \\ — & \vdots & 0 \\ — & \vec{e}_{\tau(d-1)} & 0 \\ 0 & \cdots & 1 \end{pmatrix}$$

$\sigma(1), \ldots, \sigma(d)$ is the rearrangements of $1, \ldots, d$, so there exists i such that $\sigma(i) = d$. If $i = d$, we do nothing, otherwise, we interchange rows i and d as shown above. Notice that $\tau(1), \ldots, \tau(d-1)$ is a permutation of $1, 2, \ldots, d-1$. By induction hypothesis, we can perform row interchanges that turn

$$\begin{pmatrix} — & \vec{e}_{\tau(1)} \\ — & \vdots \\ — & \vec{e}_{\tau(d-1)} \end{pmatrix} \to I_{d-1}$$

and if we do these interchanges on

$$\begin{pmatrix} — & \vec{e}_{\tau(1)} & — \\ — & \vdots & — \\ — & \vec{e}_{\tau(d-1)} & — \\ 0 & \cdots & 1 \end{pmatrix} \to I_d$$

as desired. $\square$

Proof: of statement 2:
For this, we let D be a determinant function $D : M_n(\mathbb{C}) \to \mathbb{C}$. Consider a matrix

$$A = \begin{pmatrix} — & \vec{e}_{\sigma(1)} & — \\ — & \vdots & — \\ — & \vec{e}_{\sigma(n)} & — \end{pmatrix}$$

Suppose that we have two ways of transforming A to the identity cia row interchanges:

$$A \to A_1 \to A_2 \to \cdots \to A_{d-1} \to I$$
$$A' \to A'_1 \to A'_2 \to \cdots \to A'_{e-1} \to I$$

If we go by the top path, $D(A) = (-1)^d$, and if we go by the bottom path, we have $D(A) = (-1)^e$, thus we can conclude that d and e have the same parity. $\square$ $\square$

We define for $\sigma \in S_n$

$$\operatorname{sgn}(\sigma) = (-1)^d$$

where we can row reduce $\begin{pmatrix} — & \vec{e}_{\sigma(1)} & — \\ — & \vdots & — \\ — & \vec{e}_{\sigma(n)} & — \end{pmatrix} \to I$ with d row interchanges. This is well-defined because if we can row reduce with d_1 row interchanges and d_2 interchanges, then $d_1 \equiv d_2 \pmod 2$ and thus $(-1)^{d_1} = (-1)^{d_2}$.

Corollary 0.11

If $D : M_n(F) \to F$ is a determinant function, then we have

$$D \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} = \sum_{\sigma \in S_n} \left(\prod_{i=1}^{n} a_{i,\sigma(i)} \right) \operatorname{sgn}(\sigma)$$

Theorem 0.30

For $A, B \in M_n(F)$, we have

$$\det(AB) = \det(A)\det(B)$$

Proof: Notice if $\det(C) = 0$ if and only if C is not invertible. Hence AB is not invertivle if and only if A is not invertible or B is not invertible. Thus we have

$$\begin{aligned} 0 = \det(AB) &\iff \det(A) = 0 \qquad \text{or} \det(B) = 0 \\ &\iff \det(A)\det(B) = 0 \end{aligned}$$

So STP the case when $\det(A) \neq 0$ and $\det(B) \neq 0$. i.e., A and B are both invertible.

Define a map $D : M_n(F) \to F$ via the rule $D(A) = \dfrac{\det(AB)}{\det(B)}$. CLAIM: D is n-linear, alternating, and $D(I)$ is equal to 1. Notice $D(I) = 1$. We then show it is n-linear, we first write

$$A = \begin{pmatrix} — & \vec{r_1} & — \\ — & \vdots & — \\ — & \vec{r_n} & — \end{pmatrix} \qquad \implies \qquad AB = \begin{pmatrix} — & \vec{r_1} \cdot B & — \\ — & \vdots & — \\ — & \vec{r_n} \cdot B & — \end{pmatrix}$$

So to show D is n-linear, we fix all rows but i^{th} and let it vary. Thus we know that it is n-linear because both matrix multiplication and determinant function are linear. Alternating is easy to see because two same rows multiply by B are still the same, and because determinant function is alternating, thus D is also alternating. $\square$

Result 0.16

This gives us a nice result:
If $\sigma, \tau \in S_n$, then $\operatorname{sgn}(\sigma \circ \tau) = \operatorname{sgn}(\sigma)\operatorname{sgn}(\tau)$.

Proof: Def check. $\square$

Lecture 28 - Mon - Mar 18 - 2024

Recall from last lecture, we have

$$\det(A) = \sum_{\sigma \in S_n} \left[\operatorname{sgn}(\sigma) \cdot \prod_{i=1}^{n} a_{i,\sigma(i)} \right]$$

Discovery 0.15

S_n is a group under $\circ$ (`\circ`), and $\{\pm 1\}$ is a group under $\cdot$ (`\cdot`). Moreover,

$$\begin{aligned} \operatorname{sgn} : S_n &\to \{\pm 1\} \\ \operatorname{sgn}(\sigma \circ \tau) &= \operatorname{sgn}(\sigma)\operatorname{sgn}(\tau) \qquad \forall\, \sigma, \tau \in S_n \end{aligned}$$

sgn is a group homomorphism. And

$$A_n = \ker(\operatorname{sgn}) = \{\sigma : \operatorname{sgn}(\sigma) = 1\} = \text{ kernel of sgn}$$

notice that A_n is the alternating group, a subgroup of S_n. It is also closed under $\circ$, taking inverses, it has the identity:

$$\begin{aligned} \operatorname{sgn}(\sigma) = \operatorname{sgn}(\tau) \quad &\Rightarrow \quad \operatorname{sgn}(\sigma \circ \tau) = \operatorname{sgn}(\sigma)\operatorname{sgn}(\tau) = 1 \\ \operatorname{sgn}(\sigma) = 1 \quad &\Rightarrow \quad \operatorname{sgn}(\sigma^{-1}) = 1 \end{aligned}$$

Theorem 0.31

Let $A \in M_n(F)$, then $\det(A^T) = \det(A)$.

Proof: We have

$$\begin{aligned} \det(A) &= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} a_{i,\sigma(i)} \\ \det(A^T) &= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} a_{\sigma(i),i} \end{aligned}$$

Notice that $\sigma(1), \dots, \sigma(n)$ is a permutation of $1, \dots, n$, so the following follows the definition. Basically, we can just rearrange them, in other words, take their inverses:

$$\det(A^T) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma^{-1}) \prod_{j=1}^{n} a_{j,\sigma^{-1}(j)}$$

gg $\square$

Recall we showed that

$$E_j(A) = \sum_{i=1}^{n} a_{ij}(-1)^{i+j} \det(A(i \mid j))$$

is a determinant function. Since determinant function is unique, we have

Definition 0.39: Cofactor Expansion Along j^{th} Column

$$\det(A) = \sum_{i=1}^{n} a_{ij}(-1)^{i+j} \det(A(i \mid j))$$

Example 0.51

Use cofactor expansion to find

$$\det \begin{pmatrix} 2 & 6 & 18 \\ 0 & 5 & 906 \\ 0 & 0 & 3 \end{pmatrix} = 2(-1)^{1+1} \det \begin{pmatrix} 5 & 906 \\ 0 & 3 \end{pmatrix} = 30$$

The Classical Adjoint

Theorem 0.32

If A, $n \times n$, and $\det(A) \neq 0$, then

$$A^{-1} = \frac{1}{\det(A)} \Big((-1)^{i+j} \det(A(j \mid i)) \Big)_{1 \leq i,j \leq n}$$

Theorem 0.33

If $\dim V = n$, and $T : V \to V$ is defined as $\det(T) := \det([T]_{\mathcal{B}})$ where $\mathcal{B}$ is an ordered basis, then $[T]_{\mathcal{C}} = S^{-1}[T]_{\mathcal{B}}S$, and we have

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

Lecture 29 - Wed - Mar 20 - 2024

Eigenvalues and Eigenvectors

Consider

$$T : \mathbb{R}^2 \to \mathbb{R}^2$$
$$T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \left(\begin{bmatrix} 3x \\ 2y \end{bmatrix}\right)$$

The linear transformation sends $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ to $\begin{bmatrix} 3 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ to $\begin{bmatrix} 0 \\ 2 \end{bmatrix}$, and we know that $T\left(\begin{bmatrix} 5 \\ 7 \end{bmatrix}\right) = T\left(\begin{bmatrix} 5 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 7 \end{bmatrix}\right)$.

Definition 0.40: Eigenvalue & Eigenvector

Let $T : V \to V$ be a linear operator. We say that the scalar $c \in F$ is an **eigenvalue** of T if there exists $0 \neq v \in V$ such that $T(v) = cv$, and we call v the **eigenvector** of T.

Example 0.52

Consider the example above, we notice that $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ is an eigenvector of T since we have $T\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) = 3 \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, where 3 would be the eigenvalue.

Remark: If $\vec{v}$ is an eigenvector of T with corresponding eigenvalue c, so is $\lambda \cdot v$ for $\lambda \neq 0$, $\lambda \in F$.

$$T(\lambda \cdot v) = \lambda \cdot T(v) = \lambda \cdot c \cdot v = c \cdot (\lambda \cdot v)$$

Remark: For $A \in M_n(F)$, $c \in F$ is an eigenvalue of A if there exists $\vec{v} \neq 0$ such that $A\vec{v} = c \cdot \vec{v}$.
Exercise: $V = \mathbb{R}[x]$, we define two linear operator:

$$T : V \to V \qquad T(p(x)) = \int_0^x p(t)\, dt$$
$$S : V \to V \qquad S(p(x)) = \frac{d}{dx} p(x)$$

What are the eigenvalues / eigenvectors of T and S?
Proof: T does not have eigenvector. If there exists $0 \neq p(x) \in \mathbb{R}[x]$ such that

$$T(p(x)) = c \cdot p(x), \quad c \in \mathbb{R}$$

which implies that

$$\int_0^x p(t)\, dt = c \cdot p(x)$$

which then implies that $p(x) = c \cdot p'(x)$, which clearly has no solution with $0 \neq p(x) \in \mathbb{R}[x], c \in \mathbb{R}$ because degree of $p'(x)$ is always less than degree of $p(x)$ for $p(x) \neq 0$.

How about S? If $0 \neq p(x)$ is an eigenvector, we need $S(p(x)) = c \cdot p(x)$ for some $c \in \mathbb{R} \quad \Rightarrow$ $p'(x) = c \cdot p(x)$. Case 1: $c \neq 0$, so we have no solution, because $\deg(p(x)) > \deg(p'(x))$. Case 2: $c = 0$, $p'(x) = 0$, so we can conclude that $p(x)$ is a constant. Thus eigenvalue is 0, and eigenvector is nonzero constant polynomials. □

Example 0.53

Show that $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \in M_2(\mathbb{R})$ has no real eigenvalue.

Proof: Suppose we have a solution, so we have $c \in \mathbb{R}$ and $\begin{bmatrix} a \\ b \end{bmatrix} \neq \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ such that $A \begin{bmatrix} a \\ b \end{bmatrix} = c \cdot \begin{bmatrix} a \\ b \end{bmatrix}$, which would then give us $\begin{bmatrix} b \\ -a \end{bmatrix} = \begin{bmatrix} ca \\ cb \end{bmatrix}$, since a, b not both zero, which implies that $c \neq 0, a \neq 0, b \neq 0$. Now, we have $-1 = c^2$. If we work over $\mathbb{C}$, then i is the eigenvalue, and the eigenvector would be $\begin{bmatrix} 1 \\ i \end{bmatrix}$ $\square$

Example 0.54

Let $A = I_n \in M_n(F)$, then the eigenvalue is simply 1, and everything is its eigenvector.

Definition 0.41: Eigenspace

Let $T : V \to V$ be an operator (includes $A \in M_n(F)$). Then if $c \in F$, we let $W_c = \{\vec{v} \in V : T(\vec{v}) = c \cdot \vec{v}\}$, which is called the **eigenspace of T associated to** c.

Theorem 0.34

For all $c \in F$, W_c is a subspace of V.

Proof: If suffices to show that $\vec{0} \in W_c$ and W_c is closed under addition and scalar Multiplication.

$$T(\vec{0}) = \vec{0} = c \cdot \vec{0} \;\Rightarrow\; \vec{0} \in W_c$$

and more definition check. $\square$

Remark: c is an eigenvalue of T if and only if $W_c \neq (\vec{0})$ if and only if $\dim(W_c) \geq 1$.

Theorem 0.35

Let $T : V \to V$ be an linear operator for $\dim(V) \ll \infty$, and let $c \in F$, then TFAE:

(1) c is an eigenvalue of T

(2) $W_c \neq (\vec{0})$

(3) $T - cI$ is not invertible.

(4) $\det(T - cI) = 0$

Proof: We saw $1 \Leftrightarrow 2$

$2 \Rightarrow 3$. If $W_c \neq (\vec{0})$, then there exists $0 \neq v$ such that $T(v) = cv$, which is equivalent to $(T - cI)(v) = 0$, which means that $\ker(T - cI) \neq (\vec{0})$, which implies that $T - cI$ is not 1-1, hence not invertible.

$3 \to 4$, $T - cI$ not invertible $\Rightarrow \det(T - cI) = 0$.

$4 \Rightarrow 1$, $\det(T - cI) = 0$, implies that $\ker(T - cI) \neq (\vec{0})$, so there exists $0 \neq \vec{v}$ such that $(T - cI)(v) = \vec{0}$, thus $T(v) = c \cdot v$. $\square$

Characteristic Polynomial

Definition 0.42

In general, if we let x be an indeterminate, then $xI - A \in M_n(F[x])$. We will see that $\det(xI - A)$ is a monic degree n polynomial in $F[x]$ whose roots are the eigenvalues of A.

Lecture 30 - Fri - Mar 22 - 2024

Discovery 0.16

Let x be a variable, let $p_A(x) = \det(xI - A)$, where $xI - A \in M_n(F[x])$, then c is a root of $p_A(x) \quad \Leftrightarrow \quad c$ is an eigenvalue.

Proof: Let $B = xI - A$ and we write $b = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nn} \end{pmatrix}$ Therefore, we have $b_{ij} = x\delta_{ij} - a_{ij}$ where $\delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$, following that, we have

$$\begin{aligned} \det(xI - A) = \det(B) &= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) b_{1\sigma(1)} \cdots b_{n\sigma(n)} \\ &= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma)(x\delta_{1\sigma(1)} - a_{1\sigma(1)}) \cdots (x\delta_{n\sigma(n)} - a_{n\sigma(n)}) \\ &= \sum_{\sigma \in S_n} \Big(\operatorname{sgn}(\sigma) x^n \delta_{1\sigma(1)} \cdots \delta_{n\sigma(n)} + \text{ lower degree terms}\Big) \in F[x] \end{aligned}$$

Notice that the coefficient for x^n in $p_A(x)$ is $\sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \delta_{1\sigma(1)} \cdots \delta_{n\sigma(n)} = \det \begin{pmatrix} \delta_{11} & \delta_{12} & \cdots & \delta_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ \delta_{n1} & \delta_{n2} & \cdots & \delta_{nn} \end{pmatrix} = 1$

Therefore, $p_A(x)$ is a monic polynomial of x with degree n, and thus

$$c \text{ is a root for } p_A(x) \quad \Leftrightarrow \quad \det(cI - A) = 0 \quad \Leftrightarrow \quad c \text{ is an eigenvalue of } A$$

$\square$

Example 0.55

Find $p_A(x)$ when $A = \begin{pmatrix} 0 & 2 & 0 \\ 3 & 1 & -1 \\ 3 & 7 & -1 \end{pmatrix}$.

We have

$$\begin{aligned} p_A(x) &= \det(xI - A) \\ &= x \det \begin{pmatrix} x-1 & 1 \\ -7 & x+1 \end{pmatrix} - (-2) \det \begin{pmatrix} -3 & 1 \\ -3 & x+1 \end{pmatrix} \\ &= x\left[(x-1)(x+1) + 7\right] + 2\left[-3(x+1) + 3\right] \\ &= x^3 \end{aligned}$$

which means that the only eigenvalue is 0, and to find the corresponding eigenvectors, we solve for $(A - oI)\vec{x} = \vec{0}$.

$$\begin{pmatrix} 0 & 2 & 0 \\ 3 & 1 & -1 \\ 3 & 7 & -1 \end{pmatrix} \vec{x} = \vec{0} \quad \rightsquigarrow \quad \left(\begin{array}{ccc|c} 0 & 2 & 0 & 0 \\ 3 & 1 & -1 & 0 \\ 3 & 7 & -1 & 0 \end{array}\right)$$

$$\left(\begin{array}{ccc|c} 1 & 0 & -\frac{1}{3} & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

where we can find that the eigenvector is $\begin{bmatrix} x_3/3 \\ 0 \\ x_3 \end{bmatrix}$ for $x_3 \neq 0$.

Theorem 0.36

If A and B are similar, (i.e., there exists $S \in \mathrm{GL}_n(F)$ such that $B = S^{-1}AS$) then $p_A(x) = p_B(x)$. In particular, they have the same eigenvalues.

Proof: We have B is $S^{-1}AS$ for some $S \in \mathrm{GL}_n(F)$, thus

$$\begin{aligned} p_B(x) &= \det(xI - B) \\ &= \det(xI - S^{-1}AS) \\ &= \det(S^{-1}(xI - A)S) \\ &= \det(S^{-1}) \det(xI - A) \det(S) \\ &= \det(xI - A) = p_A(x) \end{aligned}$$

gg $\square$

Remark: So they have the same eigenvalues.

Theorem 0.37

If A is upper triangular, then $p_A(x) = (x - a_{11})(x - a_{22}) \cdots (x - a_{nn})$, which implies that a_{ii} are the eigenvalues.

Definition 0.43: Multiplicity

If c is an eigenvalue of A, then $p_A(x) = (x - c)^n q(x)$ for $q(c) \neq 0$. We call n the **multiplicity** of the eigenvalue c.

n is not the dimension of the eigenspace in general, but we so have $\dim W_c \leq n$.

Result 0.17

If $p(x) = x^n + c_{n-1}x^{n-1} + \cdots c_1 x + c_0$, we let $C = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & -c_0 \\ 1 & 0 & 0 & \cdots & 0 & -c_1 \\ 0 & 1 & 0 & \cdots & 0 & -c_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & -c_n \end{pmatrix}$ is called the companion matrix of $p(x)$

Theorem 0.38

$p_C(x) = x^n + c_{n-1}x^{n-1} + \cdots c_1 x + c_0 \in F[x]$, we prove this by induction on n. We assume true for $n < d$, $d \geq 2$. Consider the case when $n = d$, so

$$C = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & -c_0 \\ 1 & 0 & 0 & \cdots & 0 & -c_1 \\ 0 & 1 & 0 & \cdots & 0 & -c_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & -c_n \end{pmatrix} \rightarrow xI - C = \begin{pmatrix} x & 0 & 0 & \cdots & 0 & c_0 \\ -1 & x & 0 & \cdots & 0 & c_1 \\ 0 & -1 & x & \cdots & 0 & c_2 \\ \vdots & \vdots & \vdots & \ddots & x & \vdots \\ 0 & 0 & 0 & \cdots & -1 & c_n \end{pmatrix}$$

$\ldots$

Lecture 31 - Mon - Mar 25 - 2024

Quotients

Let V be a vector space over the field F and let W and W' be subspaces of V. We say that the sum of W and W' is the subspace of V given by $\{w + w' : w \in W, w' \in W'\}$. We say that this sum is direct if $W \cap W' = \{0\}$. Notice that if the sum is direct that if $v \in W + W'$ then there is unique way to write it in the form $w + w'$ with $w \in W$ and $w' \in W'$, since if we have two decompositions

$$v = w_1 + w_1' = w_2 + w_2'$$

then subtracting yields $w_1 - w_2 = w_2' - w_1'$ and so $w_1 - w_2 \in W \cap W' = \{0\}$. Thus $w_1 = w_2$ and $w_1' = w_2'$. When a sum is direct, we write $W \oplus W'$ to denote the sum $W + W'$. We say that a subspace W' of V is complementary to W if $V = W + W'$ and this sum is direct; that is, $V = W \oplus W'$. Notice that if W is a subspace of V then a complementary subspace exists: take a basis $\mathcal{B}$ for W; extend it to a basis $\mathcal{C}$ for V; now let W' denote the span of $\mathcal{C}\backslash\mathcal{B}$; then one can check $W + W'$ is direct and equal to V. Now let W be a subspace of a vector space V. Just as we formed $\mathbb{Z}/n\mathbb{Z}$ in MATH 145 using an equivalence relation, we can similarly define a quotient space V /W by putting an equivalence relation on V as follows. We write $v \sim_W v'$ if $v - v' \in W$. Notice that $\sim_W$ is an equivalence relation: it's transitive since $0 \in W$; it's symmetric, since if $w \in W$ then $-w \in W$; finally, if $v \sim_W v'$ and $v' \sim_W v''$ then $v - v'$ and $v' - v''$ are in W so their sum, $v - v''$, is in W, which gives $v \sim_W v''$. Notice that if $v \in V$ then every element of V that is equivalent to v with respect to $\sim_W$ is of the form $v + w$ for some $w \in W$ and all elements of this form are equivalent to v. For this reason, we write $v + W$ to denote the equivalence class of a vector $v \in V$ and we call an equivalence class a coset of W. We should think of a coset of W as translating the subspace W by a vector in V. We let V/W denote the set of cosets of W.

Example 0.56

Let $V = \mathbb{R}^2$ and let W denote the span of $(1, 1)$, which is the "diagonal" line through the origin. Then the cosets of W (i.e., the equivalence classes) are the sets $\{(x, x + c) : x \in \mathbb{R}\}$ with $c \in \mathbb{R}$.

Proof: Every coset is a set of the form $(a, b) + W$. Notice that $W = \{(x, x) : x \in \mathbb{R}\}$ so the coset $(a, b) + W = \{(a + x, b + x) : x \in \mathbb{R}\} = \{(y, c + y) : y \in \mathbb{R}\}$, where $c = b - a$. □

Proposition 0.11

Let W be a subspace of V. Then V/W is a vector space with addition given by $(v_1 + W) + (v_2 + W) = v_1 + v_2 + W$ and scalar multiplication $c \cdot (v + W) = cv + W$.

Theorem 0.39: Universal Property

Let V be a vector space and let W be a subspace of V and let $\pi : V \to V/W$ be the quotient map. If $T : V \to U$ is a linear map to another vector space U and $W \subseteq \ker(T)$ then there is a unique linear map $\tilde{T} : V/W \to U$ such that the following diagram commutes:

$$\begin{tikzcd} V \arrow[r] \arrow[d] & U \\ V/W \arrow[ur] & \end{tikzcd}$$

where the top arrow is the map T, the downward arrow is the quotient map π and the diagonal map is $\tilde{T}$. In other words, there is a unique linear map $\tilde{T}$ such that $T = \tilde{T} \circ \pi$.

Theorem 0.40

Let W be a subspace of V and let W' be a complementary subspace to W. Then $W' \cong V/W$.

Corollary 0.12

If V is n-dimensional and W is a d-dimensional subspace, then V/W has dimension $n - d$.

Lecture 32 - Wed - Mar 27 - 2024

Compute $p_A(x) = \det(xI - A)$ and find its roots. For each eigenvalue c, we compute the nullspace of $A - cI$ whose nonzero elements are eigenvectors for eigenvalue c.

Definition 0.44: Diagonalization

Let $A \in M_n(F)$, we say that A is diagonalizable if there exists $S \in \mathrm{GL}_n(F)$ and a diagonal matrix D such that $D = S^{-1}AS$ (i.e., A is similar to a diagonal matrix).

Note that $\phi : M_n(F) \to M_n$, $\phi(X) = S^{-1}XS$ is an isomorphism.

Theorem 0.41

If $A \in M_n(F)$ has the property that F^n has a basis $\vec{v_1}, \dots, \vec{v_n}$ consisting of eigenvalues of A, then A is diagonalizable and if $A\vec{v_i} = c_i\vec{v_i}$, then

$$\begin{pmatrix} c_1 & & & 0 \\ & c_2 & & \\ & & \ddots & \\ 0 & & & c_n \end{pmatrix} = S^{-1}AS$$

where $S = \begin{pmatrix} | & | & & | \\ \vec{v_1} & \vec{v_2} & \cdots & \vec{v_n} \\ | & | & & | \end{pmatrix}$.

Proof: Let $S = \begin{pmatrix} | & | & & | \\ \vec{v_1} & \vec{v_2} & \cdots & \vec{v_n} \\ | & | & & | \end{pmatrix}$, then S is invertible because Colrank of $S = n$, which implies that the determinant of S is non-zero. Consider

$$A \cdot S = A \cdot \begin{pmatrix} | & | & & | \\ \vec{v_1} & \vec{v_2} & \cdots & \vec{v_n} \\ | & | & & | \end{pmatrix} = \begin{pmatrix} | & | & & | \\ A\vec{v_1} & A\vec{v_2} & \cdots & A\vec{v_n} \\ | & | & & | \end{pmatrix} = \begin{pmatrix} | & | & & | \\ c_1\vec{v_1} & c_2\vec{v_2} & \cdots & c_n\vec{v_n} \\ | & | & & | \end{pmatrix}$$

so we have $S^{-1}(AS) = S^{-1} \begin{pmatrix} | & | & & | \\ c_1\vec{v_1} & c_2\vec{v_2} & \cdots & c_n\vec{v_n} \\ | & | & & | \end{pmatrix} = \begin{pmatrix} | & | & & | \\ c_1S^{-1}\vec{v_1} & c_2S^{-1}\vec{v_2} & \cdots & c_nS^{-1}\vec{v_n} \\ | & | & & | \end{pmatrix}$ Notice that the identity is simply $\begin{pmatrix} | & | & & | \\ \vec{e_1} & \vec{e_2} & \cdots & \vec{e_n} \\ | & | & & | \end{pmatrix} = I = S^{-1}S$. Hence we know that $S^{-1}AS = \begin{pmatrix} | & | & & | \\ c_1 & c_2 & \cdots & c_n \\ | & | & & | \end{pmatrix}$, which is the diagonal matrix $\begin{pmatrix} c_1 & 0 & & 0 \\ 0 & c_2 & & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & & c_n \end{pmatrix}$ $\square$

Example 0.57

Let $A = \begin{pmatrix} 1 & 1 \\ -6 & 6 \end{pmatrix}$, find a diagonal matrix D and a matrix invertible matrix S such that $S^{-1}AS = D$.

1. Step 1:
   Compute the characteristic polynomial we have
   $$\begin{aligned} p_A(x) = \det(xI - A) &= \det \begin{pmatrix} x-1 & -1 \\ 6 & x-6 \end{pmatrix} \\ &= (x-3)(x-4) \end{aligned}$$
   Hence 3 and 4 are our eigenvalues
2. Step 2: Find corresponding eigenvectors
   When $c = 3$, we have $A - 3I = \begin{pmatrix} -2 & 1 \\ -6 & 3 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & -\frac{1}{2} \\ 0 & 0 \end{pmatrix}$, which implies that $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ is an eigenvector.
   Similarly, when $c = 4$, we have $A - 4I = \begin{pmatrix} -3 & 1 \\ -6 & 2 \end{pmatrix}$, which implies that $\begin{bmatrix} 1 \\ 3 \end{bmatrix}$ is another eigenvector.
   Therefore, they are the basis for F^2, $S = \begin{pmatrix} 1 & 1 \\ 2 & 3 \end{pmatrix}$
3. Step 3: Find the inverse for S
   We can find that $S^{-1} = \begin{pmatrix} 3 & -1 \\ -2 & 1 \end{pmatrix}$
4. Calculate
   Therefore
   $$D = S^{-1}AS = \begin{pmatrix} 3 & -1 \\ -2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -6 & 6 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 2 & 3 \end{pmatrix} = \begin{pmatrix} 3 & 0 \\ 0 & 4 \end{pmatrix}$$

Discovery 0.17: Not all matrices are diagonalizable

Consider the matrix $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ because the only eigenvalues are 0.

Example 0.58

Find A^{100}.

Proof: Since we have $A = S\begin{pmatrix} 3 & 0 \\ 0 & 4 \end{pmatrix} S^{-1} = SDS^{-1}$, therefore

$$\begin{aligned} A &= \underbrace{(SDS^{-1})(SDS^{-1})\cdots(SDS^{-1})}_{100} && (1) \\ &= SD^{100}S^{-}1 && (2) \\ &= S\begin{pmatrix} 3^{100} & 0 \\ 0 & 4^{100} \end{pmatrix} S^{-1} && (3) \end{aligned}$$

□

Theorem 0.42

In fact, An $n \times n$ matrix A is diagonalizable over F if and only if there is a basis for F^n consisting of eigenvectors of A.

Proof: Backward is trivial, conversely, if A is diagonalizable, then there exist $c_1, \ldots, c_n$ and $S \in \mathrm{GL}_n(F)$ such that

$$\begin{pmatrix} c_1 & 0 & & 0 \\ 0 & c_2 & & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & & c_n \end{pmatrix} = S^{-1}AS$$

so $S\begin{pmatrix} c_1 & 0 & & 0 \\ 0 & c_2 & & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & & c_n \end{pmatrix} = AS$. If $S = \begin{pmatrix} | & | & & | \\ \vec{v_1} & \vec{v_2} & \cdots & \vec{v_n} \\ | & | & & | \end{pmatrix}$, then $S\begin{pmatrix} c_1 & 0 & & 0 \\ 0 & c_2 & & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & & c_n \end{pmatrix} = A\begin{pmatrix} | & | & & | \\ \vec{v_1} & \vec{v_2} & \cdots & \vec{v_n} \\ | & | & & | \end{pmatrix}$, then $\begin{pmatrix} | & | & & | \\ Sc_1\vec{e_1} & Sc_2\vec{e_2} & \cdots & Sc_n\vec{e_n} \\ | & | & & | \end{pmatrix}$, then $\begin{pmatrix} | & | & & | \\ c_1\vec{v_1} & c_2\vec{v_2} & \cdots & c_n\vec{v_n} \\ | & | & & | \end{pmatrix} = \begin{pmatrix} | & | & & | \\ A\vec{v_1} & A\vec{v_2} & \cdots & A\vec{v_n} \\ | & | & & | \end{pmatrix}$. So the vectors are eigenvectors of A and they form a basis because S is invertible. □

Theorem 0.43

Let $A \in M_n(F)$ and suppose that $p_A(x)$ has n distinct roots in F. Then A is diagonalizable.

Proof: Let $c_1, \ldots, c_n$ be the distinct roots of the characteristic polynomial of A and let $v_1, \ldots, v_n$ be corresponding eigenvectors. We claim that $v_1, \ldots, v_n$ forms a basis for F^n. To see this, it suffices to show that the vectors are linearly independent. So SFAC that $\{v_1, \ldots, v_n\}$ is dependent. Then there is a minimal dependent subset: $\{v_{i_1}, \ldots, v_{i_k}\}$. Notice $k \geq 2$ since the vectors are nonzero. So we have a non-trivial dependence

$$\sum_{j=1}^{k} \lambda_j v_{i_j} = 0$$

Left multiplying by A gives us another relation

$$\sum_{j=1}^{k} \lambda_j c_{i_j} v_{i_j} = 0$$

Then multiplying our first relation by c_{i_k} and subtracting our second relation we get

$$\sum_{j=1}^{k-1} \lambda_j (c_{i_k} - c_{i_j}) v_{i_j} = 0$$

By minimality of our dependent subset that $\lambda_j(c_{i_k} - c_{i_j}) = 0$ for $j = 1, \ldots, k-1$; since $c_1, \ldots, c_n$ are distinct, we then see $\lambda_1 = \cdots = \lambda_{k-1} = 0$, so our original relation becomes $\lambda_k v_{i_k} = 0$ so $\lambda_k = 0$ too, contradicting the fact that our dependence was non-trivial. $\square$

Lecture 33 - Mon - Apr 1 - 2024

Definition 0.45: Linear Recurrence

A sequence $f_0, f_1, f_2, \ldots$ taking values in a field F satisfies a **linear recurrence** is there exists $d \geq 1$ and constants $c_0, c_1, \ldots, c_d \in F$ such that

$$f_n = c_1 f_{n-1} + c_2 f_{n-2} + \cdots + c_d f_{n-d} \qquad \forall\, n \geq d$$

Example 0.59

Fibonacci sequence is a classical example, so does $f_n = 2^n$ and $g_n = n^2$.

Solving Linear Recurrences with Linear Algebra

Suppose we have a function $f(n)$ for $n = 0, 1, 2, \ldots$ and a $d \geq 1$ such that $f(0), f(1), \ldots, f(d-1)$ are our initial values, and we have $f(n) = c_1 f(n-1) + c_2 f(n-2) + \cdots + c_d f(n-d)$. We wonder how we can solve this using linear algebra, consider

$$F^d \ni \vec{v_0} = \begin{bmatrix} f(0) \\ f(1) \\ \vdots \\ f(d-1) \end{bmatrix}$$

Let $A = \in M_d(F)$ be

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 \\ & & & & \ddots & \\ 0 & 0 & 0 & 0 & \cdots & 1 \\ c_d & c_{d-1} & c_{d-2} & c_{d-3} & \cdots & c_1 \end{pmatrix}$$

Let $\vec{v_n} = A^n \vec{v_0}$ for all $n \geq 0$,

Theorem 0.44

We will have

$$\vec{v_n} = \begin{bmatrix} f(n) \\ f(n+1) \\ \vdots \\ f(n+d-1) \end{bmatrix} \qquad \forall\, n \geq 0$$

Proof: The proof involves induction, which is a lot typing, just check definitions. □

Now, consider

$$\underbrace{\begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix}}_{d} (A^n \cdot \vec{v_0}) = \begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix} \cdot \vec{v_n}$$

$$= \begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} f(n) \\ f(n+1) \\ \vdots \\ f(n+d-1) \end{bmatrix} = f(n)$$

Suppose that A is diagonalizable and $D = s^{-1}AS$, then $A^n = S^{-1}D^nS$. Therefore,

$$f(n) = \begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix} A^n \cdot \vec{v_0} \tag{4}$$

$$= \left(\begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix} S \right) D^n (S^{-1}\vec{v_0}) \tag{5}$$

Example 0.60

Consider $F_0 = 0, F_1 = 1$, and $F_n = F_{n-1} + F_{n-2}$.

Therefore, $\vec{v_0} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ and $A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$, thus $A^n \cdot \vec{v_0} = \begin{bmatrix} F_n \\ F_{n+1} \end{bmatrix}$. Diagonalizes A. Step 1, we need to first find the characteristic polynomial, which simply is $x^2 - x - 1$, which has distinct roots: $(1+\sqrt{5})/2$ and $(1-\sqrt{5})/2$. Step 2, Then we want to find the eigenvectors: If λ is ab eigenvalue, of $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$, then we need to solve

$$\left(\begin{array}{cc|c} a-\lambda & b & 0 \\ c & d-\lambda & 0 \end{array} \right)$$

so if $(a-\lambda, b) \neq (0,0)$. then $\begin{bmatrix} b \\ \lambda - a \end{bmatrix}$ is an eigenvector. The matrix S has columns u_1, u_2 where u_1, u_2 are eigenvectors for these eigenvalues. We compute and find

$$S = \begin{pmatrix} 1 & 1 \\ \frac{1+\sqrt{5}}{2} & \frac{1-\sqrt{5}}{2} \end{pmatrix}$$

$$\Rightarrow \quad S^{-1}AS = \begin{pmatrix} \frac{1+\sqrt{5}}{2} & 0 \\ 0 & \frac{1-\sqrt{5}}{2} \end{pmatrix}$$

Hence solving for A we have

$$A = S\begin{pmatrix} \frac{1+\sqrt{5}}{2} & 0 \\ 0 & \frac{1-\sqrt{5}}{2} \end{pmatrix} S^{-1} \Rightarrow \quad A^n = S\begin{pmatrix} \left(\frac{1+\sqrt{5}}{2}\right)^n & 0 \\ 0 & \left(\frac{1-\sqrt{5}}{2}\right)^n \end{pmatrix} S^{-1}$$

Therefore,

$$\begin{aligned} F_n = \begin{bmatrix} 1 & 0 \end{bmatrix} A^n \begin{bmatrix} 0 \\ 1 \end{bmatrix} &= \begin{bmatrix} 1 & 0 \end{bmatrix} S \begin{pmatrix} \left(\frac{1+\sqrt{5}}{2}\right)^n & 0 \\ 0 & \left(\frac{1-\sqrt{5}}{2}\right)^n \end{pmatrix} S^{-1} \begin{bmatrix} 0 \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{pmatrix} \left(\frac{1+\sqrt{5}}{2}\right)^n & 0 \\ 0 & \left(\frac{1-\sqrt{5}}{2}\right)^n \end{pmatrix} \begin{bmatrix} \frac{1}{\sqrt{5}} \\ -\frac{1}{\sqrt{5}} \end{bmatrix} \\ &= \frac{1}{\sqrt{5}} \left\{ \left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n \right\} \end{aligned}$$

Remark: Notice that F_n is the closest integer to $\frac{1}{\sqrt{5}} \cdot \rho^n$, where ρ is the golden ratio.

Result 0.18

Not all matrices are diagonalizable, if, however, we work over $\mathbb{C}$, then all matrices are triangularizable: (i.e., there exists $S \in \mathrm{GL}_n(\mathbb{C})$ and an upper-triangular matrix U such that $U = S^{-1}AS$).

Theorem 0.45

Let F be a field, and let $A \in M_n(F)(\subseteq M_n(K))$, then there exists a finite extension K of F and there exists $S \in \mathrm{GL}_n(K)$ such that

$$S^{-1}AS \quad \text{is upper-triangular}$$

Proof:

$\square$

Lecture 34 - Wed - Apr 3 - 2024

Review for Math 145

Theorem 0.46

If $p(x) \in F[x]$ (e.g. $x^2 + 1 \in \mathbb{R}[x]$), there exists a **field extension** K of F (i.e. $F \subseteq K$) such that $p(x)$ factors into linear factors:

$$p(x) = C(x - \lambda_1) \cdots (x - \lambda_d) \qquad \lambda_1, \ldots, \lambda_d \in K$$

Definition 0.46

f K is an extension of F, K is an F-vector space. So K has a dimension. If $\dim_F K < \infty$, we say that K is a finite extension of F.

Corollary 0.13

If $A \in M_n(F)$, then there exists a finite extension K of F such that A has an eigenvalue as a matrix in $M_n(K)$.

Proof: Let $p_A(x) \in F[x]$ be the characteristic polynomial of A, then there exists a finite extension K of F such that

$$p_A(x) = (x - \lambda_1) \cdots (x - \lambda_d) \qquad \lambda_1, \ldots, \lambda_d \in K$$

so A has eigenvalues as a matrix in $M_n(K)$. (In fact n eigenvalues when we count by multiplicity). □

Recall that not all matrices are diagonalizable,

Example 0.61

For instance, $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$

Definition 0.47

A matrix $A \in M_n(F)$ is triangularizable over F if there exists $S \in \mathrm{GL}_n(F)$ such that $S^{-1}AS$ is upper-triangular.

Theorem 0.47

Let $A \in M_n(F)$, then there exists a finite extension K of F such that $A \in M_n(F) \subseteq M_n(K)$ is triangularizable over K and if $F = \mathbb{C}$, we can take $K = \mathbb{C}$.

Proof: Induction over n:

1. Base case: $n = 1$
   Trivial case

2. Assume true

3. We know there exists a finite extension K of F such that

   $$p_A(x) = (x - \lambda_1) \cdots (x - \lambda_d) \qquad \lambda_1, \ldots, \lambda_d \in K$$

   and when $F = \mathbb{C}$, we can take $K = \mathbb{C}$. Now there exists an eigenvalue $\vec{v_1} \in K^d$ such that

   $$A\vec{v_1} = \lambda_1 \vec{v_1}$$

We can extend $\vec{v_1}$ to an ordered basis $\vec{v_1},\dots,\vec{v_d}$ for K^d. Let $S = \begin{pmatrix} | & | & & | \\ \vec{v_1} & \vec{v_2} & \cdots & \vec{v_d} \\ | & | & & | \end{pmatrix} \in \mathrm{GL}_d(K)$.

Notice that

$$A \begin{pmatrix} | & | & & | \\ \vec{v_1} & \vec{v_2} & \cdots & \vec{v_d} \\ | & | & & | \end{pmatrix} = \begin{pmatrix} | & | & & | \\ A\vec{v_1} & A\vec{v_2} & \cdots & A\vec{v_d} \\ | & | & & | \end{pmatrix} = \begin{pmatrix} | & | & & | \\ \lambda_1\vec{v_1} & A\vec{v_2} & \cdots & A\vec{v_d} \\ | & | & & | \end{pmatrix}$$

so

$$S^{-1}AS = S^{-1} \begin{pmatrix} | & | & & | \\ \lambda_1\vec{v_1} & A\vec{v_2} & \cdots & A\vec{v_d} \\ | & | & & | \end{pmatrix} = \begin{pmatrix} | & | & & | \\ S^{-1}\lambda_1\vec{v_1} & S^{-1}A\vec{v_2} & \cdots & S^{-1}A\vec{v_d} \\ | & | & & | \end{pmatrix}$$

Recall $I = \begin{pmatrix} | & | & & | \\ \vec{e_1} & \vec{e_2} & \cdots & \vec{e_d} \\ | & | & & | \end{pmatrix} = S^{-1}S$, hence

$$S^{-1}AS = \begin{pmatrix} \lambda_1 & c_1 & c_2 & \cdots & c_{d-1} \\ 0 & & & & \\ 0 & & & & \\ \vdots & & & & \\ 0 & & & & A' \end{pmatrix} \qquad A' \in M_{d-1}(K), c_1 \dots, c_d \in K$$

By induction hypothesis, we know that there exists $T \in \mathrm{GL}_{d-1}(K)$ such that $T^{-1}A'T$ is upper-triangular, thus we consider

$$\begin{aligned}
\begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & & \\ 0 & & & T^{-1} \end{pmatrix} \left(S^{-1}AS\right) \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & & \\ 0 & & & T \end{pmatrix} &= \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & & \\ 0 & & & T^{-1} \end{pmatrix} \left[\begin{pmatrix} \lambda_1 & c_2 & \cdots & c_{d-1} \\ 0 & & & \\ \vdots & & & \\ 0 & & & A' \end{pmatrix} \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & & \\ 0 & & & T \end{pmatrix} \right] \\
&= \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & & \\ 0 & & & T^{-1} \end{pmatrix} \begin{pmatrix} \lambda_1 & * & \cdots & * \\ 0 & & & \\ \vdots & & & \\ 0 & & & A'T \end{pmatrix} \\
&= \begin{pmatrix} \lambda_1 & * & \cdots & * \\ 0 & & & \\ \vdots & & & \\ 0 & & & T^{-1}A'T \end{pmatrix} \\
&= \begin{pmatrix} \lambda_1 & * & \cdots & * \\ 0 & & & \\ \vdots & & & \\ 0 & & & U \end{pmatrix}
\end{aligned}$$

so if we let $T_1 = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & & \\ 0 & & & T \end{pmatrix}$, then $T_1^{-1}S^{-1}AST_1$ is upper-Triangular. $\square$

For simplicity, we will work in $M_n(\mathbb{C})$, so

$$p_A(x) = (x - \lambda_1)\cdots(x - \lambda_n)$$

and there exists $S \in \mathrm{GL}_n(\mathbb{C})$ such that $S^{-1}AS = \begin{pmatrix} \gamma_1 & & & * \\ & \gamma_2 & & \\ & & \ddots & \\ 0 & & & \gamma_n \end{pmatrix}$. Because $S^{-1}AS$ is similar to A,

$$P_{S^{-1}AS}(x) = p_A(x) = (x - \lambda_1)\cdots(x - \lambda_n)$$

so $\det(xI - S^{-1}AS) = (x - \gamma_1)\cdots(x - \gamma_n)$, which implies that γ_i's are rearrangements of λ_i's.

Discovery 0.18

Recall the following facts:

1. Similar matrices have the same trace
2. Similar matrices have the same determinant

So $\mathrm{tr}(A) = \mathrm{tr}(S^{-1}AS) = \gamma_1 + \gamma_2 + \cdots + \gamma_n = \lambda_1 + \lambda_2 + \cdots + \lambda_n$, and
$\det(A) = \det(S^{-1}AS) = \gamma_1\gamma_2\cdots\gamma_n = \lambda_1\lambda_2\cdots\lambda_n$.

Definition 0.48: Vandermonde Matrices

For $\lambda_1, \ldots, \lambda_n \in F$, we define the matrix

$$V_n(\lambda_1, \ldots, \lambda_n) = \begin{pmatrix} 1 & \lambda_1 & \cdots & \lambda_1^{n-1} \\ 1 & \lambda_2 & \cdots & \lambda_2^{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \lambda_n & \cdots & \lambda_n^{n-1} \end{pmatrix}$$

to be called the Vandermonde matrices.

Lecture 35 - Fri - Apr 5 - 2024

Recall the Vandermonde Matrix.

Theorem 0.48

We have

$$\det(V_n(\lambda_1, \dots, \lambda_n)) = \prod_{j>i} (\lambda_j - \lambda_i)$$

In particular, if $\lambda_1, \dots, \lambda_n$ are distinct, then $V_n(\lambda_1, \dots, \lambda_n)$ is invertible. We prove that if $P_0, P_1, \dots, P_n$ are polynomials with P_i monic and degree i, then the determinant still holds for the matrix defined as

$$\begin{pmatrix} P_0(\lambda_1) & P_1(\lambda_1) & \cdots & P_{n-1}(\lambda_1) \\ P_0(\lambda_2) & P_1(\lambda_2) & \cdots & P_{n-1}(\lambda_2) \\ \vdots & \vdots & \ddots & \vdots \\ P_0(\lambda_n) & P_1(\lambda_n) & \cdots & P_{n-1}(\lambda_n) \end{pmatrix}$$

Lemma 0.10

Let $p(x) \in F[x]$ be monic of degree n and let $c \in F$, then there exists a monic polynomial $q(x)$ of degree $n-1$ such that

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

Proof: we know that

$$x^i - c^i = (x-c)(x^{i-1} + x^{i-2}c + \cdots + c^{i-1})$$

so we have

$$\begin{aligned} p(x) - p(c) &= \left(\sum_{j=0}^{n-1} a_j x^j + x^n \right) - \left(\sum_{j=0}^{n-1} a_j c^j + c^n \right) \\ &= (x^n - c^n) + \sum_{j=0}^{n-1} a_j (x^j - c^j) \\ &= (x-c) \left[x^{n-1} + x^{n-2}c + \cdots + c^{n-1} + \sum_{j=0}^{n-1} a_j (x^{j-1} + x^{j-2}c + \cdots + c^{j-1}) \right] \end{aligned}$$

and $q(x)$ is monic of degree $n-1$. $\square$

Proof: of the theorem:

We prove the theorem by induction on n.

Base case, $n = 2$: we simply have

$$\det \begin{pmatrix} 1 & \lambda_1 + a \\ 1 & \lambda_2 + a \end{pmatrix} = (\lambda_2 - \lambda_1)$$

Suppose the statement is true for $n = d$ where $d \geq 2$.

Consider the case when $n = d + 1$:

$$A = \begin{pmatrix} 1 & P_1(\lambda_1) & \cdots & P_{n-1}(\lambda_1) \\ 1 & P_1(\lambda_2) & \cdots & P_{n-1}(\lambda_2) \\ \vdots & \vdots & \ddots & \vdots \\ 1 & P_1(\lambda_n) & \cdots & P_{n-1}(\lambda_n) \end{pmatrix}$$

Perform row operations, $R_i \longrightarrow R_i - R_1$ for $i = 2, \ldots, d+1$, we get A becomes

$$\begin{pmatrix} 1 & P_1(\lambda_1) & \cdots & P_{n-1}(\lambda_1) \\ 0 & P_1(\lambda_2) - P_1(\lambda_1) & \cdots & P_{n-1}(\lambda_2) - P_{n-1}(\lambda_1) \\ \vdots & \vdots & \ddots & \vdots \\ 0 & P_1(\lambda_n) - P_1(\lambda_1) & \cdots & P_{n-1}(\lambda_n) - P_{n-1}(\lambda_1) \end{pmatrix}$$

Let $c = \lambda_1$, by our lemma, for all $i \geq 1$, there exists a polynomial $Q_{i-1}(x)$ of degree $i - 1$ such that

$$P_i(x) - P_i(c) = (x - c)Q_{i-1}(x)$$

We know that

$$\begin{aligned} P_i(\lambda_j) - P_i(\lambda_i) &= P_i(\lambda_j) - P_i(c) \\ &= (\lambda_j - c)Q_{i-1}(\lambda_j) \qquad \text{plug in } x = \lambda_j \end{aligned}$$

Hence we can rewrite our matrix as

$$\begin{pmatrix} 1 & P_1(\lambda_1) & \cdots & P_{n-1}(\lambda_1) \\ 0 & (\lambda_2 - \lambda_1)Q_0(\lambda_2) & \cdots & (\lambda_2 - \lambda_1)Q_{n-2}(\lambda_2) \\ \vdots & \vdots & \ddots & \vdots \\ 0 & (\lambda_n - \lambda_1)Q_0(\lambda_n) & \cdots & (\lambda_n - \lambda_1)Q_{n-2}(\lambda_n) \end{pmatrix}$$

We use cofactor expansion along the first column to obtain

$$\begin{aligned} \det &= (-1)^{1+1} \det \begin{pmatrix} (\lambda_2 - \lambda_1)Q_0(\lambda_2) & \cdots & (\lambda_2 - \lambda_1)Q_{n-2}(\lambda_2) \\ \vdots & \ddots & \vdots \\ (\lambda_n - \lambda_1)Q_0(\lambda_n) & \cdots & (\lambda_n - \lambda_1)Q_{n-2}(\lambda_n) \end{pmatrix} \\ &= (\lambda_2 - \lambda_1)(\lambda_3 - \lambda_1) \cdots (\lambda_n - \lambda_1) \det \begin{pmatrix} Q_0(\lambda_2) & \cdots & Q_{n-2}(\lambda_2) & & \\ Q_0(\lambda_3) & \cdots & Q_{n-2}(\lambda_3) & & \\ \vdots & \ddots & \vdots Q_0(\lambda_n) & \cdots & Q_{n-2}(\lambda_n) \end{pmatrix} \end{aligned}$$

Because Q_i is monic, we use the induction hypothesis to say this is

$$= \prod_{j>i\geq 2} (\lambda_j - \lambda_i) = \prod_{j>i} (\lambda_j - \lambda_i)$$

QED. $\square$

Lagrange Interpolation

Set-up

We have distinct values $\lambda_1, \ldots, \lambda_n \in F$ and $a_1, \ldots, a_n \in F$ (which are not necessarily distinct), we want to find a deg $\leq n-1$ polynomial $p(x)$ such that

$$p(\lambda_1) = a_1 \qquad \ldots \qquad p(\lambda_n) = a_n$$

We solve

$$p(x) = c_0 + c_1 x + c_2 x^2 + \cdots + c_{n-1} x^{n-1}$$

using the above knowns.

Notice that

$$p(\lambda_1) = a_1 \qquad \ldots \qquad p(\lambda_n) = a_n$$

is equivalent to saying

$$\begin{pmatrix} 1 & \lambda_1 & \cdots & \lambda_1^{n-1} \\ 1 & \lambda_2 & \cdots & \lambda_2^{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \lambda_n & \cdots & \lambda_n^{n-1} \end{pmatrix} \begin{bmatrix} c_0 \\ c_1 \\ \vdots \\ c_{n-1} \end{bmatrix} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$$

Since we know that all λ_i's are distinct, so the matrix $V(\lambda_1, \ldots, \lambda_n)$ is invertible. That is,

$$\begin{bmatrix} c_0 \\ c_1 \\ \vdots \\ c_{n-1} \end{bmatrix} = V^{-1}(\lambda_1, \ldots, \lambda_n) \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$$

Final

Definition 0.49: Final

The final is 2 hours and 30 minutes long consisting five parts.

(a) 5 true or falses, 2 points each

(b) 5 short answer questions, 2 points each

(c) 3 long answer questions, on determinants

(d) 4 long answer questions, on diagonalizability and characteristic polynomial

(e) 3 long answer questions, on eigenspaces and similarity