Math 25a F24 CA Notes

yay

Fall 2024

1 9/4/2024

1.1 Welcome to Math 25a!

Course Head: Wes Cain
jcain2@math.harvard.edu
Science Center 515
Teaching Fellow: Jit Wu
Course Assistants: Anthony, Bright, Ely, Ethan, and Kieran (hi!)

Office hours for Wes start Friday 2:00-3:30. CA/TF office hours start Sunday. Schedule to appear on Canvas.
Homework: PS01 due next Wed 5:00pm by Canvas file upload.
Books: Hammock, Axler, Abbott, Treil

Norms:

- Judgement-free zone!
- Don't make assumptions about what your classmates know or don't know.
- Collaborate and help your math family :)

1.2 Motivation

Why study linear algebra?

$\mathbb{R}$ = set of all real numbers.

Basic functions: $f : \mathbb{R} \to \mathbb{R}$ (domain $\to$ codomain). $f(x) = mx + b$.

Studying linear transformations $T : \mathbb{R}^n \to \mathbb{R}^m$. (flat things) Will allow us to learn about all types of transformations by approximating non flat things with flat things.

Proposition 1. Let $a, b, c, d, f, s \in \mathbb{R}$. The system

$$\begin{aligned} ax + by &= r \\ cx + dy &= s \end{aligned}$$

has a unique solution (x, y both real) if and only if $ad - bc \neq 0$.

If and only if: P if and only if Q = if P then Q and if Q then P.

Proof. Step 1: Prove $\Leftarrow$ direction.

Assume $ad - bc \neq 0$. Goal: Show the system has a unique solution (x, y). Strategy: Find the solution!

$$\begin{aligned} acx + bcy &= cr \\ acx + ady &= as \\ acx &= cr - bcy \\ cr - bcy + ady &= as \\ y(ad - bc) &= as - cr. \end{aligned}$$

We can divide by $ad - bc$ to solve uniquely for y iff $ad - bc \neq 0$. Do something similar to prove for x.

Step 2: Prove $\Rightarrow$ direction (on the HW). Contrapositive: Prove that if $ad - bc = 0$, there are either 0 or multiple solutions to the system. $\square$

Contrapositive: Can sometimes be easier to prove.

- Statement: If it is Thursday, then I wear shoes.
- Converse: If I wear shoes, then it is Thursday. (Not equivalent)
- Contrapositive: If I don't wear shoes, then it is not Thursday. (Yes equivalent)

1.3 Intro to Sets

$S = \{1, 4, 9, 16\} = \{4, 1, 9, 16\} = \{1, 4, 4, 9, 16\}$. 4 elements.

$4 \in S$: 4 is in / is an element of S.

$\phi = \varnothing = \{\} =$ empty set. Not the same as $\{\varnothing\}$ which has one element!

1.3.1 Set builder notation

- $\mathbb{N} = \{1, 2, 3, \ldots\}$ natural numbers
- $\mathbb{Z} = \{\ldots, -1, 0, 1, \ldots\}$ integers
- $\mathbb{Q} = \left\{\frac{p}{q} : p \in \mathbb{Z}, q \in \mathbb{Z}, q \neq 0\right\}$ rational numbers

Why set builder notation is useful: Define $S_1 = \{x \in \mathbb{R} : x^2 < \pi\} = (-\sqrt{\pi}, \sqrt{\pi})$ open interval. Now define $S_2 = \{x \in \mathbb{Q} : x^2 < \pi\}$. Interval notation doesn't work because we're skipping a ton of real numbers. So set builder notation is helpful here.

1.3.2 Unions and intersections

If A and B are sets, the *intersection* $A \bigcap B = \{x : x \in A \text{ and } x \in B\}$. The *union* $A \bigcup B = \{x : x \in A \text{ or } x \in B\}$ (can be in both).

Sigma notation: $\bigcup_{j=1}^{n} A_j = \{x : x \in A_j \text{ for some } j = 1, \ldots, n\}$. Some = at least one. $\bigcap_{j=1}^{n} A_j = \{x : x \in A_j \text{ for all } j = 1, \ldots, n\}$.

1.3.3 Subsets

$A \subseteq B$ means that if $x \in A$ then $x \in B$. Examples: $\mathbb{Q} \subseteq \mathbb{R}$. $\{\sqrt{2}\} \nsubseteq \mathbb{Q}$ because $\sqrt{2} \notin \mathbb{Q}$.

Prove $A = B$ by proving $A \subseteq B$ and $B \subseteq A$.

Example 1. For each $n \in \mathbb{N}$, let $A_n = [-2 + 1/n, 2 - 1/n]$. Claim: $\bigcup_{n=1}^{\infty} A_n = (-2, 2)$.

Proof. Step 1: Prove $\bigcup_{n=1}^{\infty} A_n \subseteq (-2, 2)$.

First, assume $x \in \bigcup_{n=1}^{\infty} A_n$. This means for some $m \in \mathbb{N}$, $x \in A_m = [-2 + 1/m, 2 - 1/m]$. So $-2 < x < 2$, i.e., $x \in (-2, 2)$.

Step 2: Prove $(-2, 2) \subseteq \bigcup_{n=1}^{\infty} A_n$.

Now assume $x \in (-2, 2)$. Intuition: If n is really big, you can contain most of $(-2, 2)$ including x. Let $\epsilon = \min\{x - (-2), 2 - x\}$. (Invoking Archimedean Property: There is always a rational number between any two real numbers.) Choose $m \in \mathbb{N}$ large enough that $1/m < \epsilon$. Claim: $x \in A_m$.

$$\epsilon \leqslant x - (-2) = x + 2$$
$$x \geqslant -2 + \epsilon > -2 + 1/m.$$

Likewise,

$$\epsilon \leqslant 2 - x$$
$$x \leqslant 2 - \epsilon < 2 - 1/m.$$

So $x \in A_m$, so $x \in \bigcup_{n=1}^{\infty} A_n$. $\square$

1.3.4 Miscellaneous other set related things

Cartesian product: $A \times B = \{(x, y) : x \in A \text{ and } y \in B\}$. $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$ (the Cartesian plane).

If $A \subseteq S$, $S - A = \{x \in S : x \notin A\}$. $\mathbb{R} - \mathbb{Q}$ = set of irrational numbers.

2 9/6/2024

A **set** is a collection of objects, e.g, $\{1, \pi, -5, 3\}$ where order doesn't matter. A **list** of length $n (n \in \mathbb{N})$ is an ordered collection of objects $(x_1, x_2, \ldots, x_n)$ where order DOES matter and duplication may occur.

$\mathbb{R}^n = \mathbb{R} \times \mathbb{R} \times \cdots \times \mathbb{R}$ n times. $= \{(x_1, x_2, \ldots, x_n) : x_j \in \mathbb{R} \text{ for each } j = 1, \ldots, n\}$.

2.1 Intro to Logic

A **statement** is a sentence or math expression that is either definitely true or false (ie. $\sqrt{2} \in \mathbb{N}$ is false)

If P and Q statements, $\neg P$ is the negation of P (not P). $P \vee Q$ is P or Q and $P \wedge Q$ is P and Q.

$\neg(P \vee Q)$ is logically equivalent to $(\neg P) \wedge (\neg Q)$. Nota Bene: you can form truth tables.

Conditional Statements are of the form if P then Q, denoted $P \Rightarrow Q$ (P implies Q, Q if P, etc.). This means that Q is true *under the condition that* P is true. The **converse** of $P \Rightarrow Q$ is $Q \Rightarrow P$. The **contrapositive**, which is $\neg Q \Rightarrow \neg P$ is logically equivalent. Note Bene: proof by contraposition is a useful technique.

Suppose $n \in \mathbb{N}$. Prove that n^2 is odd $\Rightarrow n$ is odd. The contrapositive is n is even $\Rightarrow n^2$ is even. If n is even, we write $n = 2k$, where $k \in \mathbb{N}$. Then, $n^2 = 4k^2 = 2(2k^2)$. Since $2k^2 \in \mathbb{N}$, we have showed that n^2 is even, as desired.

For proof by **contradiction**, $P \wedge (\neg P)$ is false, regardless of the truth of P. $P \vee (\neg P)$ is true, regardless of the truth of P (ie. a tautology). For a conditional $P \Rightarrow Q$ is logically equivalent to $P \wedge (\neg Q) \Rightarrow R \wedge (\neg R)$ is a contradiction.

For instance, let $x > 0$. If $x^2 = 2$, then $x \notin \mathbb{Q}$ (ie. $\sqrt{2}$ irrational). Assume $P \wedge (\neg Q)$ – that is, $x^2 = 2$ but $x \in \mathbb{Q}$. We can write $x = \frac{a}{b}$, where $a, b \in \mathbb{Z}, b \neq 0$. Assume $a, b > 0$ and $\frac{a}{b}$ is a fraction in lowest terms. Squaring both sides, we obtain that $x = \frac{a}{b} \Rightarrow x^2 = 2 = \frac{a^2}{b^2} \Rightarrow 2b^2 = a^2$. Hence, a^2 is even, which implies that a is even. Write $a = 2k$, where $k \in \mathbb{N}$. Hence, $a^2 = 4k^2 = 2b^2 \Rightarrow b^2 = 2k^2$ so b^2 is even which implies that b is even. Since we have found that a is even and b is even, $\frac{a}{b}$ is *not* in lowest terms, arriving at a contradiction. We therefore conclude that $x \notin \mathbb{Q}$.

Biconditional statements, such as $P \Leftrightarrow Q$ means $P \Rightarrow Q$ and $Q \Rightarrow P$ means P and Q are equivalent, usually written P if and only if Q (abbreviated **iff**). For instance, let $x, y \in \mathbb{R}$. We seek to prove that $x = y$ iff $\forall \epsilon > 0, |x - y| < \epsilon$. We first prove the forward direction directly. If $x = y$. Then, $|x - y| = 0 < \epsilon$ for every $\epsilon > 0$. We prove the converse using contraposition. Assume $x \neq y$. Then $|x - y| \geqslant 0$. Take $\epsilon_1 = |x - y|$, which is positive.

2.2 Logical Quantifiers

$\forall$ means for all. $\exists$ means there exists (at least one). For instance, $\forall x \in \mathbb{R}$ with $x > 0, \exists y \in \mathbb{R}$ with $0 < y < x$ is true (ie. take $y = \frac{x}{2}$).

Negation of Statements with Quantifiers: If $P : \forall \epsilon \exists \delta > 0$ such that $\delta < \epsilon$. $\neg P : \exists \epsilon > 0$ such that $\forall \delta > 0, \delta \geqslant \epsilon$.

Proof by **induction** is useful for things that can be indexed by the natural numbers. We have a base case of P_1. We then suppose P_k true and work to prove that P_{k+1} is true for each $k \in \mathbb{N}$.

For example, if S is a nonempty finite set with n elements, then S has

2^n elements. For our base case, let $S = \{a\}$, which has two subsets. Assume P_k holds for some $k \in \mathbb{N}$.Let S have $k+1$ elements. Pick any $x \in S$. Let $S^- = S - \{x\}$. By hypothesis, S^- has 2^k subsets. S subsets are subsets of S^{-1} or $x \cup \{\text{subset of } S^-\}$.

3 9/9/2024

3.1 Gaussian Elimination

Types of elementary row operations:

- Interchange two rows
- Multiply one equation by a nonzero constant
- Multiply a row by nonzero constant and add result to some other row

Reading each useful left-to-right, the first variable remaining is a leading variable. The other variables are called free variables. Often, do Gaussian elimination using augmented matrices for bookkeeping.

Let A be a $m \times n$ (m rows, n columns) coefficient matrix. We say that A is in echelon form is both:

- if a row consists only of zeros, so do any rows below it
- if 1st nonzero entry of row i is in column j, then all entries below row i in columns 1 through j are 0.

Further, we say that A is in reduced row echelon form (RREF) if A is in echelon form and a) 1st nonzero entry in each row (if any) is 1 and) all entries above a leading 1 are zeros. **Proposition:** The RREF is unique (see article for proof).

If we have the following system

$$x + 2y = 3$$

$$2x + 4y = 7$$

we arrive at $0x + 0y = 1$, which is false $\forall (x, y) \in \mathbb{R}^2$. Hence, $\neg\exists (x, y) \in \mathbb{R}^2$ satisfying the system. There is no solutions and we can it *inconsistent*.

Here is an example RREF:

$$\left[\begin{array}{ccccc|c} 1 & 2 & 0 & 3 & 0 & 7 \\ 0 & 0 & 1 & -4 & 0 & 5 \\ 0 & 0 & 0 & 0 & 1 & -2 \end{array}\right]$$

We have x_1, x_3, x_5 leading variables. We have x_2 and x_4 as free variables. We know that $x_5 = -2$. x_3 must equal $5 + 4x_4$. $x_1 = 7 - 2x_2 - 3x_4$. If we wanted to present a solution, we would have

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 7 - 2x_2 - 3x_4 \\ x_2 \\ 5 + 4x_4 \\ x_4 \\ -2 \end{bmatrix} = \begin{bmatrix} 7 \\ 0 \\ 5 \\ 0 \\ -2 \end{bmatrix} + x_2 \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -3 \\ 0 \\ 4 \\ 1 \\ 0 \end{bmatrix}$$

We can this an *under-determined* system since we have more unknowns (5) that equations (3).

In $\mathbb{R}^3$, consider $\{(x, y, z) : y = 2x - 5\}$. We have $x = \frac{y}{2} + \frac{5}{2}$, we have y, z free. So if we were to describe this plane we have

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \frac{5}{2} \\ 0 \\ 0 \end{bmatrix} + y \begin{bmatrix} \frac{1}{2} \\ 1 \\ 0 \end{bmatrix} + z \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

3.2 Vectors

Think of a vector in $\mathbb{R}^n$ as a column $x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ where $x_j \in \mathbb{R}$. If $x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ and $y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$, we define the sum $x + y$ component-wise: that is, $x + y = \begin{bmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{bmatrix}$

If $\alpha \in \mathbb{R}$ is a scalar (ie. a number, **not** a vector) and $v \in \mathbb{R}^n$, then $\alpha v := \begin{bmatrix} \alpha v_1 \\ \alpha v_2 \\ \vdots \\ \alpha v_n \end{bmatrix}$. α just 'scales' the vector. The zero vector is the vector whose components are all equal to 0. $-v$ means $(-1) \cdot v$. We also have $v - w = v + (-1)w$

4 9/16/2024

4.1 Fields

A *field* is a set $\mathbb{F}$ with two binary operations $\mathbb{F} \times \mathbb{F} \to \mathbb{F}$, addition (+) and multiplication ($\times$ or $\cdot$) such that:

1. Addition and multiplication are both commutative and associative, e.g., $a + b = b + a$ and $a + (b + c) = (a + b) + c$ for all $a, b, c \in \mathbb{F}$.

2. Additive and multiplicative identities: $\exists$ two different elements of $\mathbb{F}$, usually denoted 0 and 1, such that $a + 0 = a$ and $1 \cdot a = a$ for all $a \in \mathbb{F}$.

3. Additive and multiplicative inverses: $\forall a \in \mathbb{F}$, $\exists b \in \mathbb{F}$ such that $a+b = 0$, and $\forall a \neq 0$, $\exists c \in \mathbb{F}$ such that $ac = 1$. We write $b = -a$ and $c = a^{-1}$.

4. Distributivity: $a(b + c) = ab + ac$.

The smallest field you can have is $\mathbb{F}_2 = \{0, 1\}$.
Examples:

- $\mathbb{R}$ and $\mathbb{Q}$ are fields
- $\mathbb{Z}$ is not a field; for example, 2 has no multiplicative inverse
- $\mathbb{C}$ is a field which looks like $\mathbb{R}^2$ (but has more fun properties)

Example: $\mathbb{F}_3 = \{0, 1, 2\}$ with addition and multiplication taken mod 3.

4.2 Complex numbers

Complex number $z = a + bi$. $\overline{z} = a - bi$ is the conjugate of z. Note that $z\overline{z} = a^2 + b^2$ (useful fact!). The *modulus* of z is $|z| = \sqrt{z\overline{z}}$.

Note that $\mathbb{C}$ is not ordered, i.e., you can't compare all complex values as being greater than or less than each other.

4.3 Vector Spaces

Review: Previously, we defined vectors in $\mathbb{R}^n$: $\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ where $x_1, x_2, \ldots, x_n \in \mathbb{R}$.
Operations are addition and scalar multiplication.

Magnitude of $\vec{v}$ is $\|v\| = \sqrt{x_1^2 + \cdots + x_n^2}$.

A *vector space* V over a field $\mathbb{F}$: Let V be a set $\subseteq \mathbb{F}^n$ for which addition and scalar multiplication are defined. V is a vector space over $\mathbb{F}$ if:

1. V is closed under addition: $\forall u, v \in V$, $u + v \in V$.

2. Commutativity and associativity of addition: If $u, v, w \in V$, then $u + (v + w) = (u + v) + w$ and $u + v = v + u$.

3. $\exists 0 \in V$ such that $v + 0 = v$ for all $v \in V$. Equivalent to the condition: If $v \in V$, $\exists w \in V$ such that $v + w = 0$. (Equivalence follows from closure under addition.) It also follows that $0 \cdot v = 0$. 0 and the inverses are unique!

4. V is closed under scalar multiplication: $\forall v \in V$ and $\alpha \in \mathbb{R}$, $\alpha v \in V$.

5. Associativity of scalar multiplication: $\alpha(\beta v) = (\alpha\beta)v$.

6. Distributivity in two ways: $\alpha(u+v) = \alpha u + \alpha v$, AND $(\alpha+\beta)v = \alpha v + \beta v$ for $\alpha, \beta \in \mathbb{R}$ and $u, v \in V$.

Examples:

- $V = \mathbb{R}^n$ and $\mathbb{F} = \mathbb{R}$.

- $V = \mathbb{C}^n$ and $\mathbb{F} = \mathbb{C}$.

- $V = \mathcal{C}[a, b]$ the set of continuous functions from $[a, b] \to \mathbb{R}$. $(f+g)(x) = f(x) + g(x)$, $(\alpha f)(x) = \alpha \cdot f(x)$.
- $S = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2 : y = \alpha x \right\}$ where α is a fixed scalar (a line through the origin! very important object yay)

Let V be a vector space over field $\mathbb{F}$. A nonempty subset $S \subseteq V$ is called a *subspace* if S is a vector space over $\mathbb{F}$.

Knowing V is a vector space and $S \subseteq V$ nonempty, it's sufficient to show closure under addition and scalar multiplication!

5 9/18/2024

Review from last time about the definition of subspace!

Example 2. $\mathcal{C}(\mathbb{R}) = \{f : \mathbb{R} \to \mathbb{R} : f \text{ is continuous}\}$. $\mathcal{C}^1(\mathbb{R}) = \{f : \mathbb{R} \to \mathbb{R} : f$ has a continuous derivative$\}$. To check if $\mathcal{C}^1(\mathbb{R})$ is a subspace of $\mathcal{C}(\mathbb{R})$, we must check:

1. $\mathcal{C}^1(\mathbb{R})$ is a nonempty subset of $\mathcal{C}(\mathbb{R})$
2. $\mathcal{C}^1(\mathbb{R})$ is closed under addition (if $f, g \in \mathcal{C}^1(\mathbb{R})$ then $f + g \in \mathcal{C}^1(\mathbb{R})$)
3. $\mathcal{C}^1(\mathbb{R})$ is closed under scalar multiplication (if $f \in \mathcal{C}^1(\mathbb{R})$ and $\alpha \in \mathbb{R}$, then $\alpha f \in \mathcal{C}^1(\mathbb{R})$)

What are the subspaces of $V = \mathbb{R}^2$ (vector space over $\mathbb{R}^2$ with usual addition and multiplication)?

V and $\{0\}$ are always subspaces of V. The other subspaces are of the form

$$\left\{ \begin{bmatrix} t \\ \alpha t \end{bmatrix} : t \in \mathbb{R} \right\}$$

for some fixed α (lines through the origin)! linear algebra woohoo

Let V be a vector space over $\mathbb{F}$. Let $U_1, U_2, \ldots, U_m$ be subsets of V. The *sum* $U_1 + U_2 + \cdots + U_m$ is the set of all possible sums $\vec{u_1} + \vec{u_2} + \cdots + \vec{u_m}$ where $\vec{u_j} \in U_j$ for all $j = 1, 2, \ldots, m$.

Proposition 2. If $U_1, U_2, \ldots, U_m$ are subspaces of V, then $U_1+U_2+\cdots+U_m$ is the smallest subspace of V containing every U_j for $j = 1, 2, \ldots, m$. (Smallest = any other subspace of V must contain this subspace.)

To prove this, (1) check for closure under addition and scalar multiplication and (2) prove any other subspace must contain this sum. Proof of (2): Any subspace of V containing $U_1, U_2, \ldots, U_m$ must be closed under addition. So any sum of elements in $U_1, U_2, \ldots, U_m$ must be in the subspace.

$U_1+U_2+\cdots+U_m$ is called a *direct sum* if each element of $U_1+U_2+\cdots+U_m$ can be represented UNIQUELY as a sum $\vec{u_1} + \vec{u_2} + \cdots + \vec{u_m}$.

To prove a sum is direct, [direct proof] take an arbitrary vector in the sum and calculate the unique $\vec{u_1}, \vec{u_2}, \ldots, \vec{u_m}$ whose sum is that vector. Alternately, [indirect proof] suppose there are two ways $\vec{u_1}+\vec{u_2}+\cdots+\vec{u_m} = \vec{u_1'}+\vec{u_2'}+\cdots+\vec{u_m'}$ to write the sum and prove that $\vec{u_1} = \vec{u_1'}$, $\vec{u_2} = \vec{u_2'}$, etc.

Proposition 3. Suppose V is a vector space over $\mathbb{F}$ and $U_1 + U_2 + \cdots U_m$ is a sum of subspaces. The sum is direct if and only if the only way to write the zero vector as $\vec{u_1} + \vec{u_2} + \cdots + \vec{u_m}$ is if $\vec{u_1} = \vec{u_2} = \cdots = \vec{u_m} = 0$. (This will become exactly the definition of linear independence.)

x

Proof. $\Rightarrow$: Assume the sum is direct. Then there's only one way to write the zero vector by the definition of direct sum. We know $0 + 0 + \cdots + 0$ works, so there can't be any other way.

$\Leftarrow$: Assume the only way to write 0 is $0 + 0 + \cdots + 0$. Suppose for the sake of contradiction that $\vec{v} = \vec{u_1} + \vec{u_2} + \cdots + \vec{u_m} = \vec{u_1'} + \vec{u_2'} + \cdots + \vec{u_m'}$. Then $0 = \vec{v} - \vec{v} = (\vec{u_1} - \vec{u_1'}) + (\vec{u_2} - \vec{u_2'}) + \cdots + (\vec{u_m} - \vec{u_m'})$, where $(\vec{u_1} - \vec{u_1'}) \in U_1$, $(\vec{u_2} - \vec{u_2'}) \in U_2$, etc. Since $0+0+\cdots+0$ is the only way to write 0, $\vec{u_1} - \vec{u_1'} = \vec{u_2} - \vec{u_2'} = \vec{u_m} - \vec{u_m'} = 0 \Rightarrow \vec{u_1} = \vec{u_1'}$, $\vec{u_2} = \vec{u_2'}$, etc. So there must be only one way to write each vector in the sum. □

Proposition 4. Let V a v.s. over $\mathbb{F}$ and U_1, U_2 subspaces. Then $V = U_1 \oplus U_2$ if and only if $U_1 \cap U_2 = \{0\}$. (Note: ONLY works for two subspaces, not more!)

Proof. $\Rightarrow$: By contraposition. Assume $x \in U_1 \cap U_2$, $x \neq 0$. Then by closure under multiplication, $x, -x \in U_1$ and $x, -x \in U_2$. So we can write $0 = x + (-x)$ where $x \in U_1$, $-x \in U_2$. Then $U_1 + U_2$ is not direct.

$\Leftarrow$: Assume $U_1 \cap U_2 = \{0\}$. Let $0 = u_1 + u_2$ where $u_1 \in U_1$, $u_2 \in U_2$. Inverses are unique, so $u_2 = -u_1$. Then by closure under addition, $u_1, u_2 \in U_1 \cap U_2$. The only way this is possible is if $u_1 = u_2 = 0$. So $V = U_1 \oplus U_2$. $\square$

6 9/20/2024

6.1 Span and Linear Combinations

Let V be a vector space over $\mathbb{F}$. If $\vec{v_1}, \vec{v_2}, \dots, \vec{v_m} \in V$, then a vector of the form

$$\vec{v} = \sum_{j=1}^{m} \alpha_j \vec{v}_j$$

for each $\alpha_j \in \mathbb{F}$ is called a *linear combination* of $\vec{v_1}, \vec{v_2}, \dots, \vec{v_m}$. The set of all linear combinations of vectors in $S = \{\vec{v_1}, \vec{v_2}, \dots, \vec{v_m}\}$ is called the *span* of S,

$$\operatorname{span}(S) = \left\{ \sum_{j=1}^{m} \alpha_j \vec{v}_j : \alpha_j \in \mathbb{F} \text{ for } j = 1, 2, \dots, m \right\}.$$

Proposition 5. $\operatorname{span}(S)$ is the smallest subspace of V containing $\vec{v_1}, \vec{v_2}, \dots,$ and $\vec{v_m}$.

Some definitions:

1. If $\operatorname{span}\{\vec{v_1}, \vec{v_2}, \dots, \vec{v_m}\} = V$, we say that $\vec{v_1}, \vec{v_2}, \dots, \vec{v_m}$ *spans* V.
2. If there is a finite set of vectors that spans V, we say V is *finite dimensional.*

Example 3. A spanning set for $\mathbb{R}^3$ is $\{(1,0,0), (0,1,0), (0,0,1)\}$. (Called the standard basis!) Technically, $\{(1,0,0), (0,1,0), (0,0,1), (1,1,1)\}$ also works, but there's an unnecessary number of vectors (not all linearly independent).

IMPORTANT DEFINITION! A set of vectors $\vec{v_1}, \vec{v_2}, \dots, \vec{v_m}$ is *linearly dependent* if $\exists \alpha_1, \alpha_2, \dots, \alpha_m$, not all zero, such that

$$\alpha_1 \vec{v_1} + \alpha_2 \vec{v_2} + \cdots + \alpha_m \vec{v_m} = 0.$$

Then, the set of vectors is *linearly independent* if it is not linearly dependent. Or the more usual definition: If the only way for $\alpha_1 \vec{v_1} + \alpha_2 \vec{v_2} + \cdots + \alpha_m \vec{v_m} = 0$ is if $\alpha_1 = \alpha_2 = \cdots = \alpha_m = 0$.

Observe: If we have a linearly dependent list, it is possible to remove some vector from the list without affecting the span.

Example 4. Consider $V = \mathcal{C}(\mathbb{R})$, continuous functions $f : \mathbb{R} \to \mathbb{R}$. This is infinite dimensional!

7 9/23/2024

f.d.v.s. = finite dimensional vector space (over $\mathbb{F}$) which we will be talking about today!

A *basis* for a f.d.v.s. V is a set of vectors $v_1, v_2, \ldots, v_m$ that's linearly independent and spans V.

Proposition 6. $B = \{v_1, v_2, \ldots, v_m\}$ is a basis for V iff each $v \in V$ has a unique representation as a linear combination $\sum_{j=1}^m \alpha_j v_j$ with each $\alpha_j \in \mathbb{F}$.

Proof. $\Rightarrow$: Assume B is a basis for V. Since B is a basis, it spans V, so we can write any $v \in V$ as a linear combination of $\alpha_j v_j$'s. Suppose

$$v = \sum_{j=1}^m \alpha_j v_j = \sum_{j=1}^m \beta_j v_j$$

for $v_j \in B$. We will prove $\alpha_j = \beta_j$ for each j.

$$0 = \sum_{j=1}^m \alpha_j v_j - \sum_{j=1}^m \beta_j v_j = \sum_{j=1}^m (\alpha_j - \beta_j) v_j.$$

But by linear independence of B, $\alpha_j - \beta_j = 0$ for all j.

$\Leftarrow$: Assume each $v \in V$ has a unique representation $\sum_{j=1}^m \alpha_j v_j$. This means $v \in \operatorname{span}(B)$ for all $v \in V$, so B spans V.

If any $v \in V$ has a UNIQUE representation as a linear combination of vectors in B, none of the vectors in B can be in the span of the other vectors, or else there would be multiple representations (i.e., one where the extra vector has coefficient zero and many where it doesn't). So, B must be linearly independent. $\square$

Observe: If V has spanning set $\{w_1, \ldots, w_n\}$ and $\{v_1, \ldots, v_n\} \subseteq V$ is linearly independent, then $m \leqslant n$. Based on the fact that given a set $\{v, w_1, \ldots, w_m\}$ where $v \in \operatorname{span}(\{w_1, \ldots, w_m\})$, we can delete one of the w vectors without affecting the span.

Proposition 7. If $B_1 = \{v_1, \ldots, v_m\}$ and $B_2 = \{w_1, \ldots, w_n\}$ are bases of V, then $m = n$.

The *dimension* of V, $\dim(V)$, is the number of vectors in any basis of V.

Standard basis for $\mathbb{R}^n = \{e_1, \ldots, e_n\}$ where e_j has all zeroes except for a one in the jth space. I.e., $e_1 = \langle 1, 0, \ldots, 0 \rangle$, $e_2 = \langle 0, 1, 0, \ldots, 0 \rangle$, etc.

Proposition 8. Let S_1, S_2 be subspaces of V. Then $\dim(S_1+S_2) = \dim(S_1)+\dim(S_2) - \dim(S_1 \cap S_2)$.

8 9/25/2024

So far, we gave discussed vector space algebra. subspaces, direct sum decomposition, bases, and dimension. From now on, we will focus on functions between vector spaces.

Let V and W be vector spaces over some field $\mathbb{F}$. A function $T : V \to W$ is called a linear transformation if $T(v_1 + v_2) = T(v_1) + T(v_2)$ and $T(\lambda v_1) = \lambda T(v_1)$.

How T acts on basis vectors $\{e_1, e2\}$ gives you a lot of information.

Let $\mathbb{F}$ be a field. $\mathbb{F}^{m \times n}$ denotes the set of all $m \times n$ matrices with entries in $\mathbb{F}$. $\mathbb{F}^n$ is the set of all $n \times 1$ vectors.

Proposition: If $T : \mathbb{F}^n \to \mathbb{F}^m$ is a linear transformation (here $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$), then there is a matrix $A \in \mathbb{F}^{m \times n}$ such that

$$T(v) = Av \forall v \in \mathbb{F}^n$$

Example: $T : \mathbb{R}^2 \to \mathbb{R}^2$ defined by orthogonal projection onto line $y = 3x$. Is T linear? Yes. Try drawing some pictures for geometric intuition.

Definition: Let V and W be vector spaces over $\mathbb{F}$. Then, $\mathcal{L}(V, W)$ denotes the set of all linear transformations from V into W. If $T_1, T_2 \in \mathcal{L}(V, W)$ ($T_1 : V \to W, T_2 : V \to W$), define $T_1 + T_2$ as

$$(T_1 + T_2)(x) = T_1(x) + T_2(x) \forall x \in V$$

$$(\lambda T_1)(x) = \lambda(T_1(x)) \forall x \in V$$

$\mathcal{L}(V, W)$ is a vector space over $\mathbb{F}$.

9 9/27/2024

Some definitions!

Recall: $\mathcal{L}(V, W)$ denotes the set of all linear transformations from V into W, where V And W are v.s.es over field $\mathbb{F}$. $\mathcal{L}(V, W)$ is a vector space over $\mathbb{F}$.

If U, W, W are vector spaces over $\mathbb{F}$ and $T \in \mathcal{L}(U, V)$, $S \in \mathcal{L}(V, W)$, then $ST \in \mathcal{L}(U, W)$ is defined by $(ST)(\vec{u}) = S(T(\vec{u}))$ for all $\vec{u} \in U$.

Let V, W be v.s. over $\mathbb{F}$ and $T \in \mathcal{L}(V, W)$. The *kernel* (nullspace) of T is $\text{Ker}(T) = \{\vec{v} \in V : T\vec{v} = \vec{0}\}$. The *range* (image) of T is $\text{range}(T) = \{T\vec{V} : \vec{V} \in V\}$. Note that Ker$(T)$ is a subspace of V and range(T) is a subspace of W. (You can check all the conditions.)

Let V and W be f.d.v.s. over $\mathbb{F}$, and $T \in \mathcal{L}(V, W)$. The *rank* of T is the dimension of the range. The *nullity* of T is the dimension of the kernel.

10 10/2/2024

Happy October! (Test review)

11 10/4/2024

11.1 Matrix of a transformation

If $T \in \mathcal{L}(V, W)$, where V, W are f.d.v.s./$\mathbb{F}$, we can write the matrix for T *with respect to* bases $\mathcal{B} = \{v_1, v_2, \dots, v_n\}$ for V and $\mathcal{C} = \{w_1, w_2, \dots, w_m\}$ for W:

$$_{\mathcal{C}}[T]_{\mathcal{B}} = \begin{bmatrix} \vdots & \vdots & & \vdots \\ Tv_1 & Tv_2 & \cdots & Tv_n \\ \vdots & \vdots & & \vdots \end{bmatrix}.$$

Remember to express the matrix entries in terms of the codomain W: If $Tv_k = A_{1k}w_1 + A_{2k}w_2 + \cdots + A_{mk}w_m$, then

$$_{\mathcal{C}}[T]_{\mathcal{B}} = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{m1} & A_{m2} & \cdots & A_{mn} \end{bmatrix}$$

an $m \times n$ matrix. (The notation $_{\mathcal{C}}[T]_{\mathcal{B}}$ means the matrix of T from V w.r.t. basis $\mathcal{B}$ into W w.r.t. basis $\mathcal{C}$. It was not discussed in class but it personally helps me understand... no need to think about that unless it helps you!)

11.2 Matrix multiplication

Let $T : U \to V$ and $S : V \to W$ be linear transformations, with U, V, W f.d.v.s./$\mathbb{F}$, and bases $\mathcal{B} = \{u_1, \dots, u_m\}$, $\mathcal{C} = \{v_1, \dots, v_n\}$, and $\mathcal{D} = \{w_1, \dots, w_p\}$ respectively. Let $B = {}_{\mathcal{C}}[T]_{\mathcal{B}}$ be the matrix for T and $A = {}_{\mathcal{D}}[S]_{\mathcal{C}}$ be the matrix for S. How do we define matrix multiplication so that AB is the matrix for ST? I.e., ${}_{\mathcal{D}}[S]_{\mathcal{C}}{}_{\mathcal{C}}[T]_{\mathcal{B}} = {}_{\mathcal{D}}[ST]_{\mathcal{B}}$. We want:

$${}_{\mathcal{D}}[ST]_{\mathcal{B}} = \begin{bmatrix} STu_1 & STu_2 & \cdots & STu_m \end{bmatrix}.$$

Observe:

$$\begin{aligned} STu_k &= S(Tu_k) \\ &= S\left(\sum_{j=1}^{n} B_{jk} v_j\right) \\ &= \sum_{j=1}^{n} B_{jk} S v_j \\ &= \sum_{j=1}^{n} B_{jk} \sum_{i=1}^{p} A_{ij} w_i \\ &= \sum_{i=1}^{p} \left(\sum_{j=1}^{n} A_{ij} B_{jk}\right) w_i. \end{aligned}$$

So we'll define matrix multiplication like this: If $A \in \mathbb{F}^{p \times n}$ and $B \in \mathbb{F}^{n \times m}$, then $AB \in \mathbb{F}^{p \times m}$ is the matrix whose ik-th entry is $\sum_{j=1}^{n} A_{ij} B_{jk}$.

Thankfully, this is the definition of matrix multiplication we already know! (if it wasn't, that would be bad)

11.3 Invertible matrices

$T \in \mathcal{L}(V, W)$ is *invertible* if $\exists S \in \mathcal{L}(W, V)$ such that $ST = \mathrm{id}_V$ and $TS = \mathrm{id}_W$.

Proposition 9. If T is invertible, it has a *unique* inverse. Denote it as T^{-1}.

Proof. Proof structure: If S_1, S_2 are both inverses of T, prove that $S_1 = S_2$.

Assume S_1, S_2 are both inverses of T.

$$S_1 = S_1 \circ \mathrm{id}_W = S_1 \circ (T \circ S_2) = (S_1 \circ T) \circ S_2 = \mathrm{id}_V \circ S_2 = S_2,$$

so $S_1 = S_2$. $\square$

Proposition 10. T is invertible if and only if it is bijective.

Proof. $\Rightarrow$: Assume T invertible.

Injective: If $Tv = Tw$, then $T^{-1}Tv = T^{-1}Tw \Rightarrow v = w$.

Surjective: For any $w \in W$, let $v = T^{-1}w$. Then $Tv = TT^{-1}w = w$.

$\Leftarrow$: Assume T is bijective. For each $w \in W$, there exists exactly one $v \in V$ such that $Tv = w$. Define $S : W \to V$ by the rule $Sw = v$ where $Tv = w$ (this is well-defined by the statement we just said). S is exactly the T^{-1} that we want! □

12 10/7/2024

The set $GL)n(\mathbb{F})$ consists of all invertible $n \times n$ matrices with entries in $\mathbb{F}$. (Called the *general linear group*.)

Proposition 11. $GL_n(\mathbb{F})$ is a non-abelian group with respect to matrix multiplication.

Proof. Check criteria for a group: Identity, closure, associativity, and inverse.

Closure: Suppose $A, B \in GL_n(\mathbb{F})$. Claim: AB is invertible and $(AB)^{-1} = B^{-1}A^{-1}$. (Shoes and socks theorem.)

Confirm this by calculating: $(B^{-1}A^{-1})(AB) = B^{-1}B = I$.

Check non-abelian: Find two elements that don't commute. □

12.1 Change of Basis

Consider $T \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$. Given a basis $\mathcal{B} = \{v_1, \ldots, v_n\}$ for $\mathbb{R}^n$, how can we convert from a standard matrix for T to a matrix with respect to $\mathcal{B}$?

Let $x \in \mathbb{R}^n$ be given w.r.t the standard basis E,

$$x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}.$$

Convert to $\mathcal{B}$-coordinates: $x = c_1v_1 + \cdots + c_nv_n$. We want

$$[x]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix},$$

so that for $S = \begin{bmatrix} \vdots & & \vdots \\ v_1 & \cdots & v_n \\ \vdots & & \vdots \end{bmatrix}$,

$$S[x]_{\mathcal{B}} = \begin{bmatrix} \vdots & & \vdots \\ v_1 & \cdots & v_n \\ \vdots & & \vdots \end{bmatrix} \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = [x]_E.$$

We eventually find out: If A is the standard matrix for T, then $A = SBS^{-1}$ where $B = {}_{\mathcal{B}}[T]_{\mathcal{B}}$. We usually notate this $[T]_{\mathcal{B}}$.

12.2 Determinants

Defining determinant on 2×2 matrices!

13 10/9/2024

13.1 More about Determinants

Example: Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be defined by $Te_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$, $Te_2 = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$, so that $[T]_E = \begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix}$.

$$\det \begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix} = 2 \cdot 3 - 1 \cdot 1 = 5.$$

If instead $Te_1 = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$, $Te_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$, then

$$\det \begin{bmatrix} 1 & 2 \\ 3 & 1 \end{bmatrix} = 1 \cdot 1 - 2 \cdot 3 = -5.$$

Conclusion: The order of the basis vectors matters (but only up to sign).

For a set of vectors A, $\det A$ is the signed volume of the parallelopiped formed by the vectors in A. Either positive or negative depending on the orders (e.g., for two vectors, the signed volume of the parallelogram depends on the direction of the cross product).

Properties of $\det : \mathbb{R}^{n \times n} \to \mathbb{R}$:

- If one of the rows or columns is multiplied by a scalar α, det is multiplied by α.
- det is multilinear, i.e., linear in each coordinate.
- If two of the rows or columns are switched, the sign of det is switched.
- If one of the basis vectors v_2 is replaced by $v_2 + \alpha v_1$, det is unchanged.
- If the rows or columns are linearly dependent, $\det = 0$.

Def: $A \in \mathbb{R}^{n\times n}$ is *diagonal* if $a_{ij} = 0$ whenever $i \neq j$. A is *upper triangular* if $a_{ij} = 0$ whenever $i > j$.

If $A = \begin{bmatrix} a_1 & & \\ & \ddots & \\ & & a_n \end{bmatrix}$ any triangular matrix, then $\det A = \prod_{i=1}^{n} a_i$. (Doesn't have to be diagonal!)

14 10/11/2024

The *transpose* of $A \in \mathbb{R}^{n\times n}$ is the matrix $A^T \in \mathbb{R}^{n\times n}$ such that $A^T_{ij} = A_{ji}$ for all i, j. (Columns and rows are switched)

Proposition 12. If $A \in \mathbb{R}^{n\times n}$ and $E \in \mathbb{R}^{n\times n}$ is elementary, then $\det(AE) = (\det A)(\det E)$.

Proposition 13. If $A \in \mathbb{R}^{n\times n}$ is invertible, then A can be written as a product of elementary matrices.

Proof. $I = E_m \cdots E_1 A$. All E_i are invertible, so $A = E_1^{-1} \cdots E_m^{-1}$ (with each E_i^{-1} also invertible). The product of invertible matrices is invertible, so A is invertible (specifically, $A^{-1} = E_m \cdots E_1$).

Method using the following corollary: If A is invertible, its RREF is the identity. $\det I = 1$. We have done elementary row operations to get to RREF. So for E_1 through E_m these elementary row operations, $1 = \det I = (\det A)(\det E_1)\cdots(\det E_m)$. So $\det A$ cannot possibly be 0. $\square$

Proposition 14 (Corollary)**.** If A is invertible then $\det A \neq 0$.

Proposition 15. If $A \in \mathbb{R}^{n\times n}$, then $\det A = \det A^T$.

15 10/16/2024

Recursive definition of the determinant (cofactor expansion)

15.1 Eigenvalues and Eigenvectors

Let V be a f.d.v.s. over $\mathbb{F}$. A nonzero $v \in V$ is called an *eigenvector* if $\exists \lambda \in \mathbb{F}$ such that $Tv = \lambda v$. λ is called the *eigenvalue* for v.

Example: $T : \mathbb{R}^2 \to \mathbb{R}^2$, projection onto the line $y = 3x$. An eigenvector for T is $v = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$, since $Tv = \begin{bmatrix} 1 \\ 3 \end{bmatrix} = 1 \cdot v$. v has corresponding eigenvalue 1. Another eigenvector is $w = \begin{bmatrix} -3 \\ 1 \end{bmatrix}$, since $Tv = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, and the corresponding eigenvalue for w is 0.

We will show: if you have an *eigenbasis* (basis of all eigenvectors), the matrix w.r.t. that eigenbasis is diagonal. This is super useful!

Proposition 16. For $T \in \mathcal{L}(V)$, λ is an eigenvalue if and only if $T - \lambda I$ is not invertible.

Proof. $\Rightarrow$: Suppose λ is an eigenvalue, so that $\exists v \in V$ such that $Tv = \lambda v$ and $v \neq 0$. Then $(T - \lambda I)(v) = Tv - \lambda v = 0$. So $v \in \ker(T - \lambda I)$, but $v \neq 0$, so the kernel is not trivial; therefore, $T - \lambda I$ is not injective.

$\Leftarrow$: Suppose $T - \lambda I$ is not invertible. Since surjective = injective for finite dimensional linear operators, $T - \lambda I$ is not injective. Then $\exists v \in \ker(T - \lambda I)$, not 0, such that $(T - \lambda I)(v) = 0$. Then $Tv - \lambda v = 0$ so $Tv = \lambda v$, so λ is an eigenvalue for v. $\square$

Observe that this means $\det(T - \lambda I) = 0$. This is how to calculate eigenvalues!

Example: Let $A = \begin{bmatrix} -1 & -2 \\ -2 & -1 \end{bmatrix}$. $A - \lambda I = \begin{bmatrix} -1 - \lambda & -2 \\ -2 & -1 - \lambda \end{bmatrix}$. Take the determinant: $\det(A - \lambda I) = (-1-\lambda)^2 - (-2)^2 = \lambda^2 + 2\lambda - 3 = (\lambda + 3)(\lambda - 1) = 0$. Solutions of this equation are 1 and -3, so the eigenvalues are 1 and -3.

$\lambda^2 + 2\lambda - 3 = 0$ is known as the *characteristic equation* for A.

To find the corresponding eigenvectors, plug in the eigenvalues and find the kernel of the matrix by row reduction / other matrix solving techniques. For example, for $\lambda = -3$: $A - (-3)I = \begin{bmatrix} 2 & -2 \\ -2 & 2 \end{bmatrix}$. The kernel of this matrix is $\left\{ \begin{bmatrix} x \\ x \end{bmatrix} : x \neq 0 \right\}$. Observe that if v is an eigenvector, all scalar multiples of v are also eigenvectors with the same eigenvalue. $Tv = \lambda v \Rightarrow T(av) = aTv = a\lambda v = \lambda(av)$.

16 10/18/2024

a

16.1 AMGM for Linear Algebra

Recall that the *geometric multiplicity* of an eigenvalue λ is the number of linearly independent eigenvectors (the dimension of the eigenspace), while the *algebraic multiplicity* is the multiplicity as a root of the characteristic polynomial.

Remark: If we write the characteristic equation as $(\lambda - \lambda *)^m q(\lambda) = 0$ where $q(\lambda *) \neq 0$ (q represents the rest of the polynomial), then m is the algebraic multiplicity of $\lambda *$. The geometric multiplicity is $\dim(\ker(T - \lambda * I))$.

Proposition 17. If V f.d.v.s. over $\mathbb{F}$ and $\lambda *$ an eigenvalue of $T \in \mathcal{L}(\mathcal{V})$, then the geometric multiplicity of $\lambda * \leqslant$ algebraic multiplicity of $\lambda *$.

Lemma 1. If $M = \left[\begin{array}{c|c} A & B \\ \hline 0 & C \end{array}\right]$ is a block matrix and A and C are square blocks, then $\det(M) = \det(A)\det(C)$. Same if M is lower triangular.

Lemma 2 (Corollary)**.** If $M = \left[\begin{array}{c|c} A & B \\ \hline 0 & C \end{array}\right]$ is a block matrix and A and C are square blocks, then the characteristic equation for M is $\det(M - \lambda I) = 0 \Rightarrow \det(A - \lambda I)\det(B - \lambda I) = 0$. This makes it easier to calculate eigenvalues. Warning: The I may not be the same size.

Lemma 3. If A and B are similar (i.e., $\exists S$ invertible s.t. $A = SBS^{-1}$, then A and B have the same eigenvalues.

In other words: If we change the basis of a transformation, it doesn't change the eigenvalues. This makes sense because scalars aren't under the basis of the vector space (they're just floating in the field), so they should be unchanged by change of basis.

Proof of lemma 3. Suppose $A = SBS^{-1}$.

The characteristic equation for A is $0 = \det(A - \lambda I) = \det(SBS^{-1} - \lambda I) = \det(SBS^{-1} - S(\lambda I)S^{-1})$ since λI is unchanged by change of basis. $= \det(S(B - \lambda I)S^{-1}) = (\det S)\det(B - \lambda I)(\det S^{-1})$ by linearity of the determinant. $= (\det S)(\det S^{-1})\det(B - \lambda I)$ by commutativity of scalar multiplication. $= (\det SS^{-1})\det(B - \lambda I)$ by linearity again. $= \det(B - \lambda I)$.

Conclusion: $\det(A - \lambda I) = \det(B - \lambda I)$, so the characteristic polynomials for A and B are the same, i.e., they have the same eigenvalues (with the same multiplicity!). □

Proof of AMGM for Linear Algebra. [Note that $\lambda{*}I_n$ means $\lambda{*}$ times I_n and not λ times I_n! :)]

Write the characteristic equation as $(\lambda - \lambda{*})^m q(\lambda) = 0$ where q is a polynomial and $q(\lambda{*}) \neq 0$. That is to say, m is as big as possible. Let the geometric multiplicity be $k = \dim(\ker(T - \lambda{*}I_n))$. Choose a basis $v_1, \ldots, v_k$ for $\ker(T - \lambda{*}I_n)$ and extend to a basis $v_1, \ldots, v_k, w_{k+1}, \ldots, w_n$ for V.

All $v_1, \ldots, v_k$ are eigenvectors w.r.t. $\lambda{*}$, so $Tv_i = \lambda{*}v_i$ for all $i = 1, \ldots, k$. So the first k columns of $[T]$ are very nice looking:

$$[T] = \left[\begin{array}{ccc|c} \lambda{*} & & & \\ & \ddots & & * \\ & & \lambda{*} & \\ \hline & 0 & & B \end{array}\right]$$

for B a $(n-k) \times (n-k)$ block matrix, $*$ some random stuff.

The characteristic equation of $[T]$ is:

$$\begin{aligned} 0 &= \det([T] - \lambda I_n) \\ &= \det(\lambda{*}I_k)\det(B - \lambda I_{n-k}) \\ &= (\lambda{*} - \lambda)^k \det(B - \lambda I_{n-k}) \\ &= (\lambda - \lambda{*})^k (-1)^k \det(B - \lambda I_{n-k}). \end{aligned}$$

This looks very similar to the characteristic equation! Assume for the sake of contradiction that $k > m$ (GM > AM). Then

$$0 = (\lambda - \lambda{*})^m \left[(\lambda - \lambda{*})^{k-m}(-1)^k \det(B - \lambda I_{n-k})\right].$$

So based on the char. eq., it must be that $(\lambda - \lambda{*})^{k-m}(-1)^k \det(B - \lambda I_{n-k}) = q(\lambda)$. But this contradicts the fact that $q(\lambda{*}) \neq 0$. So it can't be true that $k > m$.

So, GM $\leqslant$ AM. □

16.2 Diagonalizability

Proposition 18. If $\lambda_1, \ldots, \lambda_m$ are distinct eigenvalues of $T \in \mathcal{L}(V)$, and they have corresponding eigenvectors $v_1, \ldots, v_m$, then $v_1, \ldots, v_m$ are linearly independent.

THIS IS SUPER IMPORTANT AND USEFUL. FOR EXAMPLE:

Proposition 19 (Corollary)**.** Suppose $T \in \mathcal{L}(V)$ and $\dim(V) = n$. If T has n distinct eigenvalues, it has an *eigenbasis*, i.e., we can make a basis for V out of one eigenvector from each eigenspace. And, the matrix of T w.r.t. that basis is

$$[T]_B = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$$

for $\lambda_1, \ldots, \lambda_n$ the eigenvalues of T.

SO BEAUTIFUL!!!

If there exists a basis for V w.r.t. which $[T]$ is diagonal, then we call T *diagonalizable*.

Note: Every square matrix is diagonalizable in $\mathbb{C}$, but not every square matrix is diagonalizable in $\mathbb{R}$. This is why we use Jordan canonical form.

16.3 Invariance

If $T \in \mathcal{L}(V)$, a subspace $U \subseteq V$ is *invariant* under T if $T(U) \subseteq U$. Note that this works for infinite dimensional V as well!

Eigenspaces of T are invariant under T!

17 10/23/2024

17.1 Differentiating vector-valued functions

If $x : \mathbb{R} \to \mathbb{R}^n$ is defined by $x(t) = \begin{bmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{bmatrix}$, then we define $x'(t) = \begin{bmatrix} x_1'(t) \\ \vdots \\ x_n'(t) \end{bmatrix}$ where $x_1'(t)$ is the derivative of $x_1(t)$ defined the usual way.

As with single variable derivatives, $x'(t_0)$ describes a tangent vector to the curve at the point $x(t_0)$. The tangent line would be defined as $\ell(t) = x(t_0) + x'(t_0)t$. But, note that there are many possible tangent lines! The derivative describes the specific line which is both tangent to the curve AND points in the direction of the "velocity."

Consider the linear transformation $A : \mathbb{R}^n \to \mathbb{R}^n$ which maps $Ax = x'$. It maps each point in $\mathbb{R}^n$ to a velocity vector x'. (Warning: x is NOT a function anymore, now it's just a point!) A defines a *vector field*.

If $A = D$ some diagonal matrix such that $x' = Dx$, then $x_j' = \lambda_j x_j$ has a unique solution $x_j(t) = x_j(0)e^{\lambda_j t}$. If $A = SDS^{-1}$ a diagonalizable matrix, let $y = S^{-1}x$, then $y' = Dy$.

If $x \in \mathbb{R}^n$, the Euclidean norm of x is $\|x\| = \sqrt{x_1^2 + \cdots + x_n^2}$.

Proposition 20. Suppose $x' = Ax$ and A is diagonalizable over $\mathbb{R}$. Then all solutions of $x' = Ax$ satisfy $\|x(t)\| \to 0$ as $t \to \infty$ if and only if *every* eigenvalue of A is negative.

18 10/25/2024

18.1 Norms

A norm on V is a nonnegative, real-valued function $\|\cdot\| : V \to [0, \infty)$ satisfying the following conditiongs:

1. $\|v\| \geqslant 0 \ \forall v \in V$, and $\|v\| = 0$ iff $v = 0$.
2. If $c \in \mathbb{F}$, then $\|cv\| = |c|\|v\|$.
3. (Triangle Inequality) $\forall v, u \in V$, $\|v + u\| \leqslant \|v\| + \|u\|$.

Applications of the norm:

Recall that we defined $v \cdot w = v_1 w_1 + \cdots + v_n w_n$ for v, w of dimension n. We now have that $v \cdot w = \|v\|\|w\| \cos\theta$, where θ is the minor angle between the vectors. We also call this higher-dimensional dot product the *inner product*, $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{F}$. Inner product satisfies:

1. Positivity: $\langle v, v \rangle \geqslant 0 \ \forall v \in V$.
2. Definiteness: $\langle v, v \rangle = 0$ iff $v = 0$. [these two are usually combined into positive definiteness]

3. Bilinearity: $\langle cu + v, w\rangle = c\langle u, w\rangle + \langle v, w\rangle$ and $\langle v, cw + u\rangle = c\langle v, w\rangle + \langle v, u\rangle$.

4. Conjugate symmetry: $\langle w, v\rangle = \overline{\langle v, w\rangle}$. If in $\mathbb{R}$, then $\langle w, v\rangle = \langle v, w\rangle$.

Given an inner product space, we have an induced norm $\|v\| = \sqrt{\langle v, v\rangle}$.

19 11/1/2024

- Happy All Saints' Day!
- Fittingly, I (Kieran) have returned from illness hell (yay)

19.1 Adjoint of an operator

Recap:

1. When an operator $T \in \mathcal{L}(V)$ is diagonalizable, this is VERY useful. Can solve differential equations, other applications.

2. If V has an inner product, very useful to work with an orthonormal basis. Makes it easier to write coordinates. Some examples: Fourier coefficients, project onto subspaces, set up optimization problems, etc.

Question: When can we have an *orthonormal eigenbasis*? Then we can do a lot of stuff because our structure is nice.

Let $T \in \mathcal{L}(V, W)$ where V, W are finite dimensional inner product spaces over $\mathbb{F}$. Fix $w \in W$. Consider the following linear functional: $\phi \in \mathcal{L}(V, \mathbb{F})$ such that $\phi(v) = \langle Tv, w\rangle$. By Riesz Theorem, $\exists! x \in V$ such that $\phi(v) = \langle v, x\rangle$ for all $v \in V$. Note: x depends on w, so write $x = T^*w$.

The function $T^* : W \to V$ defined by $\langle Tv, w\rangle = \langle v, T^*w\rangle$ for all $v \in V, w \in W$ is called the *adjoint* of T.

Properties of the adjoint:

Proposition 21. Let V, W be FDIPS. If $T \in \mathcal{L}(V, W)$ then $T^* \in \mathcal{L}(W, V)$ (T^* is linear).

Lemma 4. If $x, y \in V$ are such that $\langle v, x\rangle = \langle v, y\rangle$ for all $v \in V$, then $x = y$.

More properties:

- If $S, T \in \mathcal{L}(V, W)$ and $\lambda \in \mathbb{F}$, then $(S+T)^* = S^* + T^*$.
- $(\lambda S)^* = \overline{\lambda} S^*$.
- $(T^*)^* = T$.
- $I^* = I$. [note that they might not be I over the same space]

If $A \in \mathbb{F}^{m\times n}$ ($\mathbb{F}$ is $\mathbb{R}$ or $\mathbb{C}$), then the *conjugate transpose* of A is the matrix $B \in \mathbb{F}^{n\times m}$ obtained by taking the complex conjugate of each entry in A^*. Note that if $\mathbb{F} = \mathbb{R}$, $B = A^*$.

Proposition 22. Let $T \in \mathcal{L}(V, W)$ where V, W are FDIPS over $\mathbb{F}$. Let $\mathcal{B} = v_1, \ldots, v_n$ and $\mathcal{C} = w_1, \ldots, w_m$ be orthonormal bases for V and W respectively. If $A = {}_{\mathcal{C}}[T]_{\mathcal{B}}$ (matrix of T w.r.t. these bases), then $B = {}_{\mathcal{C}}[T^*]_{\mathcal{B}}$ is the conjugate transpose of A.

T is *self-adjoint* if $T = T^*$. That is, $\langle Tu, v\rangle = \langle u, Tv\rangle$.

Proposition 23. If T is self-adjoint, then every eigenvalue of T is real.

Proposition 24 (Corollary)**.** If T is self-adjoint, then T has an eigenvalue (even if $\mathbb{F} = \mathbb{R}$).

20 11/4/2024

Assume V is FDIPS over $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$.

Some observations:

- T is self-adjoint iff the standard matrix A is symmetric ($A = A^T$).
- We say T is *normal* if $T^*T = TT^*$.

Example 5. Let T be such that $[T]_E = A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$. The conjugate transpose is $B = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$. $B = [T^*]_E$ by Prop 22 which we showed last time. We calculate that $AB = I$ and $BA = I$, which means $TT^* = T^*T$. So T is normal. But, since A is not symmetric, T is not self-adjoint.

We observe: All self-adjoint operators are normal, but not all normal operators are self-adjoint!

Proposition 25. $T \in \mathcal{L}(V)$ is normal if and only if $\|Tv\| = \|T^*v\|$ for all $v \in V$.

Proposition 26 (Corollary)**.** Suppose T is normal. Then if T has eigenvector v with eigenvalue λ, then v is also an eigenvector of T^* with eigenvalue $\overline{\lambda}$.

Proposition 27. If T is normal, then eigenvectors corresponding to distinct eigenvalues are orthogonal.

20.1 Spectral Theorem

Motivation: When does an operator have an orthonormal eigenbasis?

If $\mathbb{F} = \mathbb{R}$, this occurs iff the operator is self-adjoint. If $\mathbb{F} = \mathbb{C}$, it occurs iff the operator is normal.

Proposition 28. If $M \in \mathbb{R}^{n\times n}$ has orthonormal columns, then $M^T = M^{-1}$.

Proposition 29. If $A \in \mathbb{R}^{n\times n}$ can be written as $A = SDS^{-1}$ where S has orthonormal columns and D is diagonal, then $A = A^T$.

21 11/6/2024

21.1 Proof of Real Spectral Theorem

Review of three facts:

1. If V a FDIPS and $T \in \mathcal{L}(V)$ is self-adjoint, then the eigenvalues of T (if they exist) are real. (corollary of the fact that T has an orthonormal eigenbasis)

2. If V a FDIPS and $T \in \mathcal{L}(V)$ is self-adjoint, then at least one (real) eigenvalue does exist.

3. If V a FDIPS and $T \in \mathcal{L}(V)$ and W is a T-invariant subspace of V, then $W^\perp$ is invariant under T^*.

Also recall: If $A = [T]_{\mathcal{B}}$, and $\mathcal{B}$ is an orthonormal basis, and B is the conjugate transpose of A, then $B = [T^*]_{\mathcal{B}}$. (Prop 22)

Proof of Spectral Theorem if $\mathbb{F} = \mathbb{R}$:

Proof. $\Rightarrow$: Assume $\exists$ an orthonormal eigenbasis $\mathcal{B}$.

$$A = [T]_{\mathcal{B}} = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}.$$

Since $\mathbb{F} = \mathbb{R}$, all the eigenvalues are real, so A is its own conjugate transpose. Since the basis is orthonormal, the conjugate transpose of $A = [T^*]_{\mathcal{B}}$. So, $[T]_{\mathcal{B}} = [T^*]_{\mathcal{B}} \Rightarrow T = T^*$.

$\Leftarrow$: Assume T is self-adjoint. Will induct on $\dim V$.

Base case: Assume $\dim V = 1$. By Facts 1 and 2, there is at least one real eigenvalue. Let v be a corresponding eigenvector where $\|v\| = 1$ (since we can always normalize). Let $\mathcal{B} = \{v\}$. This is an orthonormal eigenbasis for V, so we're done.

Inductive step: Assume $\exists n > 1$ such that if $\dim V < n$, the claim holds. Let $\dim V = n$. By Facts 1 and 2, there is at least one real eigenvalue with eigenvector v_1 where $\|v_1\| = 1$. Let $W = \operatorname{span}\{v_1\}$, which is T-invariant since it's an eigenspace. Then by Fact 3, $W^\perp$ is invariant under T^*. We know $\dim W^\perp = n - 1$.

If $T = T^*$, then $T|_{W^\perp} = T|^*_{W^\perp}$ (if T is self-adjoint, then restriction of T to an invariant subspace is also self-adjoint). To prove, just use the definition $\langle Tv, w\rangle = \langle v, Tv\rangle$ and the restriction works out very nicely from the fact that $W^\perp$ is T-invariant.

So we can now use the inductive hypothesis: If T is self-adjoint, then there is an orthonormal eigenbasis for $W^\perp$, $\mathcal{B} = \{v_2, \ldots, v_n\}$. Add v_1 to get an orthonormal eigenbasis $\mathcal{C} = \{v_1, v_n, \ldots, v_n\}$ for $V = W \oplus W^\perp$. $\square$

21.2 Intro to positive operators

We sometimes want to decompose complicated operators into compositions of simpler operators. In particular: Singular Value Decompositions!

A very useful type of operator for these applications is Positive Operators.

A linear operator T is positive if:

1. T is self-adjoint

2. $\langle Tv, v\rangle \geqslant 0 \ \forall v \in V$.

22 11/8/2024

22.1 Positive operators ctd.

Example: Let U be a f.d. subspace of V. Then the orthogonal projection P_U of V onto U, where $P_U(v) = v - \langle v, w\rangle w$ for any unit vector $w \in U^\perp$, is a positive operator on V.

Proposition 30. $T \in \mathcal{L}(V)$ is positive iff T is self-adjoint and every eigenvalue is nonnegative and real. [$\langle Tv, v\rangle \geqslant 0$ $\forall v \in V$ iff every eigenvalue is nonnegative and real.]

22.2 Isometries (Wes's least favorite subject)

If $T \in \mathcal{L}(V)$ is such that $\|Tv\| = \|v\|$ for all $v \in V$, T is called an isometry.

Proposition 31. If $S \in \mathcal{L}(V)$, then S is an isometry iff $\langle Su, Sv\rangle = \langle u, v\rangle$ for all $u, v \in V$.

Other facts: S is an isometry iff:

- $S^*S = SS^* = I$.
- S^* is an isometry.
- S is invertible and $S^* = S^{-1}$.

Proposition 32. If V is a FDIPS over $\mathbb{C}$, then S is an isometry iff $\exists$ an orthonormal eigenbasis whose eigenvalue all satisfy $|\lambda| = 1$.

23 11/11/2024

- Happy 11/11! Make a wish :)

23.1 Preliminaries to SVD

If $T \in \mathcal{L}(V)$, an operator $S \in \mathcal{L}(V)$ is called a *square root* of T if $S^2 = T$.

Proposition 33. $T \in \mathcal{L}(V)$ is positive iff T has a positive square root. Moreover, a positive operator has a *unique* positive square root, denoted $\sqrt{T}$.

Observation: T^*T is a positive operator with a unique positive square root $\sqrt{T^*T}$ (which will appear a lot).

Proposition 34 (Polar Decomposition)**.** If $T \in \mathcal{L}(V)$, V a FDIPS, then there exists a unique isometry $S \in \mathcal{L}(V)$ such that $T = S\sqrt{T^*T}$.

This shows we can decompose any operator into two operators we understand well: An isometry S and a positive operator $\sqrt{T^*T}$. This isn't enormously useful to us on its own, but we will use it to prove SVD.

23.2 Singular Value Decomposition

The singular values of T are the eigenvalues of $\sqrt{T^*T}$.

Example: $T : \mathbb{R}^2 \to \mathbb{R}^2$ with standard matrix $A = \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix}$. Find singular values. Since the (standard) basis is orthonormal, the matrix for T^* is $A^T = \begin{bmatrix} 1 & 0 \\ 2 & 3 \end{bmatrix}$. Then the matrix for T^*T is $A^TA = \begin{bmatrix} 1 & 2 \\ 2 & 13 \end{bmatrix}$ (note this is symmetric, which is expected since T^*T is supposed to be self-adjoint). Compute the eigenvalues: $(1 - \lambda)(13 - \lambda) - 4 = 0 \Rightarrow \lambda = 7 \pm 2\sqrt{10}$. This means the singular values are $\sqrt{7 + 2\sqrt{10}}$ and $\sqrt{7 - 2\sqrt{10}}$.

Theorem 1 (Singular Value Decomposition)**.** Suppose T has singular values $\sigma_1, \ldots, \sigma_n$ (possibly repeated). Then there exist orthonormal bases $v_1, \ldots, v_n$ and $w_1, \ldots, w_n$ for V with the following property:

$$Tx = \sigma_1\langle x, v_1\rangle w_1 + \cdots + \sigma_n\langle x, v_n\rangle w_n$$

for every $x \in V$.

"A Funny Result" - Wes: This means that we can diagonalize ANY linear operator as long as we're allowed to use a different basis for the domain and the codomain!

Proof. Let $T : V \to V$, and consider $\sqrt{T^*T} : V \to V$. Since $\sqrt{T^*T}$ is a positive operator, it is self-adjoint. Then by the Spectral Theorem, there exists an orthonormal eigenbasis $v_1, \ldots, v_n$ for the domain (which is V, but specifically the first V).

For any $x \in V$, use the Fourier expansion:

$$x = \langle x, v_1 \rangle v_1 + \cdots + \langle x, v_n \rangle v_n$$
$$\sqrt{T^*T}(x) = \sigma_1 \langle x, v_1 \rangle v_1 + \cdots + \sigma_n \langle x, v_n \rangle v_n$$

where $\sigma_1, \ldots, \sigma_n$ are the eigenvalues of $\sqrt{T^*T}$ corresponding to the eigenvectors $v_1, \ldots, v_n$.

By Polar Decomposition, there exists a unique isometry S such that $T = S\sqrt{T^*T}$. Apply S to both sides:

$$\begin{aligned} Tx &= S(\sqrt{T^*T}x) \\ &= \sigma_1 \langle x, v_1 \rangle Sv_1 + \cdots \sigma_n \langle x, v_n \rangle Sv_n. \end{aligned}$$

Let $w_k = Sv_k$. Claim: $w_1, \ldots, w_n$ are an orthonormal basis for the codomain V.

Proof of Claim: We know $\|w_k\| = \|Sv_k\| = \|v_k\| = 1$ since S is an isometry and $v_1, \ldots, v_n$ are orthonormal. And, $\langle w_j, w_k \rangle = \langle Sv_j, Sv_k \rangle = \langle v_j, S^*Sv_k \rangle$. $S^*S = I$ since S is an isometry, so $\langle w_j, w_k \rangle = \langle v_j, v_k \rangle = 1$ if $j = k$, 0 else. So, $w_1, \ldots, w_n$ are orthonormal. □

Proposition 35 (Corollary)**.** Given $A \in \mathbb{R}^{n \times n}$, there exist $n \times n$ matrices U and V with orthonormal columns such that $A = U\Sigma V^T$ where $\Sigma = \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \end{bmatrix}$ and $\sigma_n \geqslant \cdots \geqslant \sigma_1 \geqslant 0$.

This is how you will actually apply SVD: By writing any operator A as a product of matrices with orthonormal columns + the diagonal matrix Σ!

Example: Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be

$$T \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2x_1 + 2x_2 \\ -x_1 + x_2 \end{bmatrix}.$$

The standard matrix is $A = \begin{bmatrix} 2 & 2 \\ -1 & 1 \end{bmatrix}$ which is not nice. But, T does not have an orthonormal eigenbasis, so we can't directly find an orthonormal basis under which T is diagonal.

Observe: If $A = U\Sigma V^T$ where U And V have orthonormal columns, notice that

$$\begin{aligned}
A^TA &= (U\Sigma V^T)^T(U\Sigma V^T)\\
&= (V\Sigma^T U^T)(U\Sigma V^T)\\
&= V\Sigma^T(U^TU)\Sigma V^T\\
&= V\Sigma^T\Sigma V^T \text{ since } U \text{ has o.n. cols}\\
&= V\Sigma^2V^T \text{ since } S \text{ is diagonal}\\
&= V\Sigma^2V^{-1} \text{ since } V \text{ has o.n. cols.}
\end{aligned}$$

So,

$$A^TA = V\begin{bmatrix} \sigma_1^2 & & \\ & \ddots & \\ & & \sigma_n^2 \end{bmatrix}V^{-1}.$$

This means the columns of V are eigenvectors of A^TA!

Apply it: $A^TA = \begin{bmatrix} 5 & 3 \\ 3 & 5 \end{bmatrix}$, which has eigenvalues 2 and 8. By the Spectral Theorem, A^TA has an orthonormal eigenbasis. I.e., the corresponding eigenvectors of 2 and 8, if normalized, will necessarily form an orthonormal eigenbasis. So all we have to do is find the corresponding eigenvectors and normalize them. For this problem, $v_1 = \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$ and $v_2 = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$. So,

$$V = \begin{bmatrix} -1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}.$$

We can find the singular values as in the example above: They are $\sqrt{2}, 2\sqrt{2}$. So

$$\Sigma = \begin{bmatrix} \sqrt{2} & 0 \\ 0 & 2\sqrt{2} \end{bmatrix}.$$

Then we can just solve for U using $A = U\Sigma V^{-1}$, and we're done.

24 11/20/2024

I'm back! (Sorry guys I have no excuse this time)

Recall: If V a FDVS over $\mathbb{F}$ and p is a polynomial with coefficients in $\mathbb{F}$, then $\ker p(T)$ and range $p(T)$ are invariant under T.

Corollary: If λ is an eigenvalue of T, then $G(\lambda, T)$ is invariant. Remember that $G(\lambda, T) = \ker(T - \lambda I)^n$.

Theorem 2 (Generalized Eigenspace Decomposition)**.** Let V be a FDVS over $\mathbb{C}$. Let $T \in \mathcal{V}$ have eigenvalues $\lambda_1, \dots, \lambda_m$ (no repeats, m may be less than $\dim V$). Then $V = \bigoplus_{j=1}^{n} G(\lambda_j, T)$. So V has a basis of generalized eigenvectors.

Two remarks:

1. Soon we will see how to construct a useful generalized eigenbasis: Jordan form!

2. If λ is an eigenvalue for T, then the restriction $(T - \lambda I)|_{G(\lambda,T)}$ is nilpotent.

Theorem 3 (Weak Jordan canonical form)**.** Let V be a complex FDVS and $T \in \mathcal{L}(V)$. If T has distinct eigenvalues $\lambda_1, \dots, \lambda_m$ with algebraic multiplicities $d_1, \dots, d_m$ respectively. Then there exists a basis for V in which the matrix of T is block diagonal

$$A = \begin{bmatrix} A_1 & & \\ & \ddots & \\ & & A_m \end{bmatrix}$$

where each block A_j is upper triangular

$$A_j = \begin{bmatrix} \lambda_j & * & \cdots & * \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \cdots & 0 & \lambda_j \end{bmatrix}.$$

A polynomial with leading coefficient 1 is called *monic.*

If V is a complex FDVS and $T \in \mathcal{L}(V)$ where $\lambda_1, \dots, \lambda_m$ are the distinct eigenvalues with algebraic multiplicities $d_1, \dots, d_m$ respectively, then the *characteristic polynomial* is $p(\lambda) = (\lambda - \lambda_1)^{d_1} + \cdots + (\lambda - \lambda_m)^{d_m}$.

Theorem 4 (Cayley-Hamilton, pt 1)**.** If V is a complex FDVS and T has characteristic polynomial q, then q annihilates T: $q(T) = 0$.