MATH 247: Honours Applied Linear Algebra

Anna Brandenberger, Daniela Breitman

February 22, 2021

This is a transcript of the lectures given by Prof. Axel Hundemer during the winter semester of the 2017-2018 academic year (01-04 2018) for the Honours Applied Linear Algebra class (MATH 247). **Subjects covered** are: Matrix algebra, determinants, systems of linear equations; Abstract vector spaces, inner product spaces, Fourier series; Linear transformations and their matrix representations; Eigenvalues and eigenvectors, diagonalizable and defective matrices, positive definite and semidefinite matrices; Quadratic and Hermitian forms, generalized eigenvalue problems, simultaneous reduction of quadratic forms; and Applications.

1 Abstract Vector Spaces

1.1 Group Theory

LECTURE 01/09

Definition 1.1. Group: Let G be a set with a binary operation

$$\begin{cases} G \times G \to G \\ (a,b) \mapsto a+b \end{cases} \quad \text{such that:}$$

- **associativity** $a+(b+c) = (a+b)+c, \forall a,b,c \in G$.
- there exists a **neutral element**
  $\exists 0 \in G$ s.t. $a+0 = 0+a = a, \forall a \in G$.
- there exists an (additive) **inverse element**
  $\exists -a \in G$ s.t. $a+(-a) = (-a)+a = 0, \forall a \in G$.

Definition 1.2. A group $(G,+)$[1] is called **Abelian** if **commutativity** is satisfied:

- $\forall a,b \in G, a+b = b+a$.

[1] A group G with operation $+$.

Examples 1.1. of Abelian groups:

- $(\mathbb{R},+)$, $(\mathbb{C},+)$, $(\mathbb{Q},+)$, $(\mathbb{Z},+)$ standard addition.
- $(\mathbb{R}\backslash\{0\},\cdot)$, $(\mathbb{C}\backslash\{0\},\cdot)$ standard multiplication.
- $\mathbb{R}^n$, $\mathbb{C}^n$ with component-wise addition:
  $(x_1,x_2,\ldots,x_n)+(y_1,y_2,\ldots,y_n) = (x_1+y_1,\ldots,x_n+y_n)$
- $\mathbf{Mat}(m\times n,\mathbb{R})$, $\mathbf{Mat}(m\times n,\mathbb{C})$: $m \times n$ matrices with coefficients in $\mathbb{R}$ or $\mathbb{C}$ with component-wise addition.

However, $\mathbb{N} = \{1,2,3\ldots\}$ with standard addition is not a group because there is no inverse element in $\mathbb{N}$ for any $n \in \mathbb{N}$.

- the set $F(\mathbb{R})$ of all real-valued functions with domain $\mathbb{R}$, with:
  - addition: $(f+g)(x) \equiv f(x)+g(x)$
  - neutral element: zero function $0(x) \equiv 0$
  - inverse element: $(-f)(x) \equiv -f(x), \forall x \in \mathbb{R}$

Example 1.2. of a non-Abelian group:

- $\mathbf{GL}(n,\mathbb{R})$, $n \geq 2$: "General Linear Group" i.e. the set of all invertible $n \times n$ matrices with coefficients in $\mathbb{R}$, with matrix multiplication.
  - neutral element: $I_n = \begin{bmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{bmatrix}$ s.t. $I_n \cdot A = A \cdot I_n = A$, $\forall A \in \mathbf{GL}(n,\mathbb{R})$.
  - inverse of A is A^{-1} since $A \cdot A^{-1} = A^{-1} \cdot A = I_n$.

Theorem 1.1. ***Cancellation Law***: *Let $(G,+)$ be a group and $a,b,c \in G$:*

(a) $a+b=a+c \implies b=c$

(b) $b+a=c+a \implies b=c$

Proof of (a):

$$\begin{aligned}
& a+b=a+c && \\
\implies & -a+(a+b) = -a+(a+c) && \text{existence of an inverse} \\
\implies & (-a+a)+b = (-a+a)+c && \text{associative} \\
\implies & 0+b = 0+c && \text{definition of neutral element} \\
\implies & b = c && \text{neutral element}
\end{aligned}$$

Exercise 1.3. Prove (b).

1.2 *Vector Spaces*

LECTURE 01/11

Definition 1.3. Let V be a set, and let K be either $\mathbb{R}$ or $\mathbb{C}$ together with two operations: $\begin{cases} + : V \times V \to V & \text{vector addition} \\ \cdot\ : K \times V \to V & \text{scalar multiplication} \end{cases}$, such that $(V,+)$ is an **abelian group**. If the following axioms for scalar multiplication hold $\forall k,l \in K$ and $u,v \in V$:

$$\begin{aligned}
(kl)v &= k(lv) && \text{associativity} \\
(k+l)v &= kv+lv && \text{distributivity of scalars over vectors} \\
k(u+v) &= ku+kv && \text{distributivity of vectors over scalars} \\
1 \cdot v &= v && \text{neutral element of scalar multiplication}
\end{aligned}$$

then $(V,+,\cdot)$ is called a **vector space over K**.

The neutral element is called the **zero vector**.

K stands for Körper ≈ field in German.

Examples 1.4. of vector spaces:

- $(\mathbb{R}^n, +, \cdot)$ where $+$ and $\cdot$ are standard addition of tuples and standard scalar multiplication.
- $\mathbf{Mat}(n \times m, K)$ with $+$ standard addition of matrices, and $\cdot$ standard multiplication by scalars.[2]
- Let s be the set of all **real sequences**, i.e. $(a_1, a_2, \dots) : a_i \in \mathbb{R}$ together with component-wise addition and scalar multiplication. The axioms of VS follow immediately from the fact that the operations are component-wise.
- Let c be the set of all **convergent real sequences** together with component-wise $+$ and $\cdot$. The only thing one must check is that c is closed under $+$ and $\cdot$[3], which is proven in ANALYSIS 1.
- Let c_0 be the set of all **real sequences converging to** 0, together with component-wise $+$ and $\cdot$. This set is closed under $+$ and $\cdot$ since the sum of two sequences with limit 0 has limit 0, same for scalar multiplication.
- Let c_{00} be the set of all **real sequences that are eventually 0**[4], with component-wise $+$ and $\cdot$.
- Let $I \subseteq \mathbb{R}$ be an interval. Let $F(I)$ be the set of all **real-valued functions** on the interval I, together with $+$ and $\cdot$ defined by:

$$\begin{aligned}(f+g)(x) &\equiv f(x) + g(x)\\ (kf)(x) &\equiv kf(x)\end{aligned}$$

- The set of all solutions to $y'' + y = 0$ or, more generally, the set of **solutions to any linear differential equation** $a_n y^n + a_{n-1} y^{n-1} + \dots + a_1 y^1 + a_0 y^0 = 0, a_i \in \mathbb{R}$. Just need to verify that the solutions are closed under $+$ and $\cdot$. [5]

SOME CONSEQUENCES of the axioms of a vector space:

Theorem 1.2. *Let $(V, +, \cdot)$ be a VS over K. Then*

(a) $k \cdot 0 = 0 \quad \forall k \in K$

(b) $0 \cdot v = 0 \quad \forall v \in V$

(c) $k \cdot v = 0 \iff k = 0 \vee v = 0$

[2] Checking axioms:
$(\mathbf{Mat}(n \times m, K), +)$ is an abelian group: obvious.
axioms for scalar multiplication also satisfied.
$\implies$ vector space.

[3] i.e. that the sum of two convergent sequences is convergent and the scalar multiple of a convergent sequence is convergent.

[4] i.e. all but finitely many terms of these are different from 0.

[5] Let y_1, y_2 be two solutions of $y'' + y = 0$, i.e. $y_1'' + y_1 = 0$ and $y_2'' + y_2 = 0$, and k an arbitrary constant.

$$\begin{aligned}(y_1+y_2)'' + (y_1+y_2) &= y_1'' + y_2'' + y_1 + y_2\\ &= (y_1'' + y_1) + (y_2'' + y_2)\\ &= 0\\ (ky_1'') + (ky_1) &= k(y_1'' + y_1)\\ &= 0\end{aligned}$$

$\implies y_1 + y_2$ and ky_1 are solutions of $y'' + y = 0$.

Proof. of (a)

$$\begin{aligned} k \cdot 0 &= k \cdot (0+0) && \text{neutral el of } + \\ &= k \cdot 0 + k \cdot 0 && \text{distribution of vectors} \\ \implies 0 + \cancel{k \cdot 0} &= \cancel{k \cdot 0} + k \cdot 0 && \text{neutral element} \\ \implies 0 &= k \cdot 0 && \text{cancellation law} \end{aligned}$$ □

Proof. of (b)

$$\begin{aligned} 0 \cdot v &= (0+0) \cdot v && \text{neutral element} \\ &= 0 \cdot v + 0 \cdot v && \text{distribution of scalars} \\ \implies 0 + \cancel{0 \cdot v} &= \cancel{0 \cdot v} + 0 \cdot V && \text{neutral element} \\ \implies 0 &= 0 \cdot v && \text{cancellation law} \end{aligned}$$ □

Proof. of (c): $k \cdot v = 0$.
"$\Leftarrow$" follows from (a) and (b).
"$\Rightarrow$": if $k = 0$, we are done. So let $k \neq 0$: must show that $v = 0$.

$$\begin{aligned} kv = 0 \implies \frac{1}{k}(kv) &= \frac{1}{k} \cdot 0 = 0 \\ &= (\frac{1}{k}k)v = 1 \cdot v = v && \text{associative} \\ \implies v &= 0 \end{aligned}$$ □

Theorem 1.3. *Let $(V, +, \cdot)$ be a VS over K. Then:*

(a) $n \cdot v = v + \cdots + v$ *(n times)*, $\forall n \in \mathbb{N}$

(b) $0 \cdot v = 0$

(c) $(-1) \cdot v = -v$, $\forall v \in V$

Proof. of (a): by Induction.
Base case $n = 1$: $1 \cdot v = v$ is an axiom.
Inductive step $n \to n+1$: assume $nv = v + \cdots + v$ for some $n \in \mathbb{N}$.

$$(n+1)v = nv + 1 \cdot v = \underbrace{(v + \cdots + v)}_{n \text{ times}} + v = \underbrace{v + \cdots + v}_{n+1 \text{ times}}$$ □

Proof. of (b) has already been done. □

Proof. of (c)

$$\begin{aligned} v + (-1)v &= 1 \cdot v + (-1) \cdot v \\ &= (1 + (-1))v && \text{associative} \\ &= 0 \cdot v = 0 && \text{prev. proven} \end{aligned}$$

$\implies (-1)v$ is the additive inverse of v, i.e. $(-1)v = -v$. □

1.3 Subspaces

LECTURE 01/16

Definition 1.4. Let $(V, +, \cdot)$ be a VS over K and let $W \subseteq V$. W is called a **subspace** of V, in symbols $W \leq V$, if $(W, +, \cdot)$ is a VS over K.

Examples:

1. all subspaces of $\mathbb{R}^2$ ($\{0\}$, all lines through the origin and $\mathbb{R}^2$ ($\{0\}$).
2. $V \equiv \mathbf{Mat}(n \times n, K)$
   $U \leq V$, where U is the set of all upper triangular $n \times n$ matrices with coefficients in K:
   $U \equiv \{A = (a_{ij}) \in \mathbf{Mat}(n \times n, K) : a_{ij} = 0, \forall i > j, 1 \leq i, j \leq n\}$
3. sequences $c_{00} \leq c_0 \leq c \leq s$
4. let $P_n(K)$ be the set of all polynomials of degree at most n. Let $P(K)$ be the set of all polynomials with standard $+$ and $\cdot$. Then:
   $P_0(K) \leq P_1(K) \leq\leq \cdots \leq P_n(K) \leq \cdots \leq P(K)$

Theorem 1.4. *Let $(V, +, \cdot)$ be a VS over K and let $W \subseteq V$. Then $W \leq V$ iff W is closed under $+$ and $\cdot$ and $0 \in W$.*

Proof.
("$\Rightarrow$") Since $(W, +, \cdot)$ is a VS over K, W is closed under $+$ and $\cdot$.
Furthermore, $0 \in W$ since $(W, +)$ is an (abelian) group, therefore contains a neutral element.
("$\Leftarrow$") Must check the axioms:

- addition: associativity and commutativity hold on W because they hold on V.
- neutral element: $0 \in W$ by condition.
- additive inverse: Let $u \in W$. Since W closed under $\cdot$, we have $(-1)u = -u \in W$.

All axioms of addition hold.
All axioms of multiplication hold on W since they hold on V.
$\implies$ W is a VS w.r.t $+$ and $\cdot$, i.e. $W \leq V$. □

Theorem 1.5.

(a) Let $U, W \leq V$ VS over K. Then $U \cap W \leq V$.

(b) Let $U_1, \ldots, U_n \leq V$. Then $U_1 \cap \cdots \cap U_n = \bigcap_{i=1}^{n} U_i \leq V$.

(c) I an arbitrary index set $U_i \leq V$ $\forall i \in I$. Then $\bigcap_{i \in I} \leq V$.

Unions are not closed under $+$! For example, take $V = \mathbb{R}^2$ and let U and W be respectively vectors on the x and y axes. Their union is not closed under $+$: adding vectors from U and W can yield any vector in $\mathbb{R}^2$.
$U \cup W \leq V \iff U \subseteq W \vee W \subseteq U$

Proof. of (a) *[(b) and (c) left as exercises]*

- $0 \in U, 0 \in W \implies 0 \in U \cap W$
- Let $v_1, v_2 \in U \cap W$, especially $v_1, v_2 \in U \implies v_1 + v_2 \in U$. Similarly, $v_1 + v_2 \in W$. So, $v_1 + v_2 \in U \cap W \implies U \cap W$ is closed under $+$.
- $\begin{cases} v_1 \in U \implies kv_1 \in U & \forall k \in K \\ v_1 \in W \implies kv_1 \in W & \forall k \in K \end{cases}$
  $\implies kv_1 \in U \cap W$ So $U \cap W$ is closed under $\cdot$.

$\implies U \cap W \leq V$ □

Definition 1.5. Union. Let V be a VS over K, $U, W \leq V$, then $U + W \equiv \{u + w : u \in U, w \in W\}$.

Theorem 1.6. *Let V be a VS over K, $U, W \leq V$, then*

(a) $U + W \leq V$

(b) $U + W$ *is the* smallest *subspace of V containing both $U \cup W$ i.e if $\tilde{V}$ is* any *subspace of V, then $U + W \subseteq \tilde{V}$*

Proof. of a)

- $0 \in U, 0 \in W \implies 0 \in U \cup W \because 0 + 0 = 0$
- Let $v_1, v_2 \in U + W$, then $\exists u_1, u_2 \in U, w_1, w_2 \in W$ such that $(v_1 = u_1 + w_1) + (v_2 = u_2 + w_2) = v_1 + v_2 = (u_1 + u_2) \in U + (w_1 + w_2) \in W \in U + W$.
- $kv_1 = k(u_1 + w_1) = ku_1(\in U) + kw_1(\in W) \in U + W \Rightarrow U + W \leq V$

□

Proof. of b) Let $\tilde{V}$ be any subspace of $V \subset U \cup W$. Let $u \in U$, $w \in W$ be arbitrary, then $u + w \in U + W$.
$u + w \in \tilde{V} \implies U + W \subseteq \tilde{V}$. So $U + W$ is the smallest subspace of V that contains $U \cup W$. □

In a similar way we define $u_1 + \cdots + u_n = \{u_1 + \cdots + u_n : u_k \in U_k \forall 1 \leq k \leq n\}$, where $u_1, \cdots, u_n \leq V$, then $u_1 + \cdots + u_n \leq V$.

1.4 Span and Linear Independence

Definition 1.6. Linear Combination. Let V be a VS over K, $v_1, \ldots, v_n \in V$, $k_1, \ldots, k_n \in K$. Then an expression $k_1 v_1 + \cdots + k_n v_n$ is a **linear combination** of $v_1, \ldots, v_n$.

This notion can be extended to infinitely many vectors: $k_1 v_1 + k_2 v_2 + \ldots$ where all but FINITELY MANY of the k_i are 0, i.e. their sum is actually finite.[6]

[6] Since infinitely many additions are not defined.

Definition 1.7. Span. Let V be a VS over K, $v_1, \ldots, v_n \in V$. Then $span\{v_1, \ldots, v_n\} = \{k_1 v_1 + \cdots + k_n v_n : k_i \in K, 1 \leq i \leq n\}$.

Note that:

$$\begin{aligned} & span\{v_1, \ldots, v_n\} \\ &= span\{v_1\} + \cdots + span\{v_n\} \\ &= \{k_1 v_1 : k_1 \in K\} + \cdots + \{k_n v_1 : k_1 \in K\} \\ &= \{k_1 v_1 + \cdots + k_n v_n : k_i \in K\} \leq V \end{aligned}$$

For infinitely many vectors $v_1, v_2, \cdots \in V$, $span\{v_1, v_2, \ldots\} = \{k_1 v_1 + k_2 v_2 + \cdots : k_i \in K :$ all but finitely many are 0$\}$.

Note that $span\{v_1, \ldots, v_n\}$ is the smallest subspace of V that contains all vectors $v_1, \ldots, v_k$.

Exercise 1.5. Prove directly that $span\{v_1, \ldots, v_n\} \leq V$

Examples 1.6.

1. $V = \mathbb{R}^2$, $\begin{cases} e_1 = (1, 0, 0, \cdots, 0) \\ e_2 = (0, 1, 0, \cdots, 0) \\ \cdots \\ e_n = (0, 0, 0, \cdots, 0) \end{cases}$ then $\mathbb{R}^2 = span\{e_1, \ldots, e_n\}$.

2. Consider VS $P_n(K) = \{1, x, x^2, \ldots, x^n\}$, then $span\{1, x, x^2, \ldots, x^n\} = \{k_0 1 + x_1 x + \cdots + x_n x^n : k_i \in x\prime 0 \leq k \leq n\} = P_n(K)$.

3. Consider VS s of all sequences with coefficients in $\mathbb{R}$.
   Define $\begin{cases} s_1 = (1, 0, 0, \ldots) \\ s_2 = (0, 1, 0, \ldots) \\ s_3 = (0, 0, 1, \ldots) \\ \cdots \end{cases} \implies span\{s_1, s_2, \ldots\} = c_{00}$.

Note that $\{s_1, s_2, \ldots\}$ does NOT span s: every LC of s results in a sequences that eventually constantly equals 0.

Definition 1.8. Linear independence and Dependence. Let V be a VS over K, $v_1, \ldots, v_n \in V$. We say that $v_1, \ldots, v_n$ are **linearly independent** if $k_1 v_1 + \cdots + k_n v_n = 0 \Leftrightarrow k_1 = k_2 = \cdots = k_n = 0$. We say that $v_1, \ldots, v_n$ are **linearly dependent** if they are not linearly independent.

Theorem 1.7. *Let V be a VS over K.*

(a) any (finite) subset of V that contains 0 is LD (linearly dependent).

(b) If $v_1, \ldots, v_n$ are LD, there exists at least one j, $1 \leq j \leq n$ such that v_j is a LC of $v_1, \ldots, v_{j-1}, v_{j+1}, \ldots, v_n$. In that case, $span\{v_1, \ldots, v_n\} = span\{v_1, \ldots, v_{j-1}, v_{j+1}, \ldots, v_n\} = span\{v_1, \ldots, \hat{v}_j, \ldots, v_n\}$.

LECTURE 01/18

Proof. (a) Consider $0, v_1, \ldots, v_n$. Then $1 \cdot 0 + 0 \cdot 1 + \cdots + 0 \cdot v_n = 0$. Non trivial LC of $0, v_1, \ldots, v_n$ since the coefficient of the zero vector is non-zero.[7]

[7] **Remark**: Even $\{0\}$ is LD since $1 \cdot 0 = 0$ which is a non-trivial LC of the zero-vector resulting in the zero-vector.

(b) Let $v_1, \ldots, v_n$ be LD. Thus $\exists k_1, \ldots, k_n \in K$, where not **all** of them 0, such that $k_1 v_1 + \cdots + k_j v_j + \cdots + k_n v_n = 0$. Let $k_j \neq 0$.

$$\implies k_j v_j = -k_1 v_1 - \cdots - k_{j-1} v_{j-1} - k_{j+1} v_{j+1} - \cdots - k_n v_n \quad (1)$$

$$\implies v_j = -\frac{k_1}{k_j} v_1 - \cdots - \frac{\hat{k_j}}{k_j} v_j - \cdots - \frac{k_n}{k_j} v_n \quad (2)$$

Thus v_j is a LC of $\{v_1, \hat{v}_j, v_n\}$.
Regarding spans: obviously $span\{v_1, \ldots, \hat{v}_j, \ldots, v_n\} \subseteq span\{v_1, \ldots, v_n\}$.
We still need to prove that $span\{v_1, \ldots, v_n\} \subseteq span\{v_1, \ldots, \hat{v}_j, \ldots, v_n\}$.
Let $v = a_1 v_1 + \cdots + a_j v_j + \cdots + a_n v_n \in span\{v_1, \ldots, v_n\}$.

Then by (3):

$$v = a_1v_1 + \cdots + a_j(-\frac{k_1}{k_j}v_1 - \cdots - \frac{\hat{k_j}}{k_j}v_j - \cdots - \frac{k_n}{k_j}v_n) + \cdots + a_nv_n$$
$$\implies v = (a_1 - \frac{k_1}{k_j}v_1)v_1 + \cdots + (a_{j-1} - \frac{k_{j-1}}{k_{j-1}})v_{j-1} + (a_{j+1} -$$
$$\frac{k_{j+1}}{k_{j+1}})v_{j+1} + \cdots + (a_n - \frac{k_n}{k_j})v_n$$

which is a LC of $v_1, \ldots, \hat{v}_j, \ldots v_n$. □

Theorem 1.8. *Let $v_1, \ldots, v_n \in V$, linearly independent. Let $v_{n+1} \notin span\{v_1, \ldots, v_n\}$. Then $v_1, \ldots, v_n, v_{n+1}$ are LI.*

Proof. Let $k_1v_1 + \cdots + k_jv_j + \cdots + k_nv_n = 0$.
We need to show that $k_1 = \cdots = k_n = k_{n+1} = 0$.
Assume that $k_{n+1} \neq 0$. Then,

$$k_{n+1}v_{n+1} = -k_1v_1 - \cdots - k_nv_n$$
$$\implies v_{n+1} = -\frac{k_1}{k_{n+1}}v_1 - \cdots - \frac{k_n}{k_{n+1}}v_n$$
$$\implies v_{n+1} \in span\{v_1, \ldots, v_n\}$$

Thus the assumption was wrong. $\implies k_{n+1} = 0$.
$\implies k_1 = \cdots = k_n = k_{n+1} = 0$ since $v_1, \ldots, v_n$ are LI.
$\implies v_1, \ldots, v_n, v_{n+1}$ are LI. □

Remark: Let V be a VS over K, $W \subseteq V$. Assume that V is finite dimensional. We CURRENTLY cannot conclude that W is finite dimensional. This will be proven later.

1.5 *Bases and Dimension*

Definition 1.9. Let V be a VS over K. V is called **finite dimensional** if V has a **finite spanning set**, i.e. there are finitely many $v_1, \ldots, v_n \in V$ s.t. $V = span\{v_1, \ldots, v_n\}$. V is called **infinite dimensional** if it is NOT finite dimensional.

Caution: it is NOT enough to construct an infinitely spanning set for V to prove that it is inf. dim! rather, we need to show that NO finite spanning set can possibly exist.

Examples 1.7. of finite and infinite dimensional VSes:

1. $\mathbb{R}^n$ is finite dim.
   $\mathbb{R}^n = span\{e_1, \ldots, e_n\}$

2. c_{00} is inf. dim.

   Proof. Assume that c_{00} is finite dim; let $s_1, \ldots, s_n \in c_{00}$ s.t. $c_{00} = span\{s_1, \ldots, s_n\}$.
   $s_j, 1 \leq j \leq n$, is in c_{00}. Thus $\exists N_j$ s.t. $s_{j_k} = 0\ \forall k \geq N_j$.
   Let $N \equiv max\{N_1, \ldots, N_k\}$ and $s_j = (s_{j_1}, s_{j_2}, \ldots)$. Then:

$s_{j_k} = 0 \ \forall 1 \leq j \leq n, \forall k \geq N$.
The same holds for every $s \in span\{s_1, \dots, s_n\}$ i.e. $s = (s_1, s_2, \dots, s_{n-1}, s_n)$, then $s_k = 0 \ \forall k \geq N$.
Thus e.g. $\tilde{s} = (0, 0, \dots, 0, 1, 0, \dots, 0) \in c_{00}$ (1 at the N^{th} position) is not in the span of $s_1, \dots, s_n$. □

Theorem 1.9. ***Steinitz Exchange Lemma.***
*Let V be a **finite dimensional** VS over K. Let $u_1, \dots, u_m \in V$ be LI and let $v_1, \dots, v_n$ be spanning, i.e. $V = span\{v_1, \dots, v_n\}$. Then, $m \leq n$ and there exist finitely many indices $k_1, \dots, k_{n-m}$ s.t. $u_1, \dots, u_m, v_{k_1}, \dots, v_{k_{n-m}}$ is spanning.*

i.e. we EXCHANGED m of the vectors $v_1, \dots, v_n$ by $u_1, \dots, u_m$ without changing the span.

Proof. This is proven iteratively, in m steps, replacing one of the v_j by one of the u-s at each step.

1. STEP: Since $span\{v_1, \dots, v_n\} = V$, $\exists a_1, \dots, a_n \in K$ s.t. $u_1 = a_1 v_1 + \cdots + a_n v_n$.
Note that $u_1 \neq 0$ since $u_1, \dots, u_m$ are LI.
Thus at least one of the coefficients $a_1, \dots, a_n \neq 0$.
By potentially reordering $v_1, \dots, v_n$, we may w.l.o.g. assume that $a_1 \neq 0$.

$$\begin{aligned} \implies a_1 v_1 &= +u_1 - a_2 v_2 - \cdots - a_n v_n \\ v_1 &= +\frac{1}{a_1} u_1 - \frac{a_2}{a_1} v_2 - \cdots - \frac{a_n}{a_1} v_n \end{aligned}$$

$\implies v_1$ is a LC of $u_1, v_2, \dots, v_n$.
$\implies span\{v_1, v_2, \dots, v_n\} = V = span\{u_1, v_2, \dots, v_n\}$.

2. STEP: Since $u_1, v_2, \dots, v_n$ is spanning, $\exists b_1, a_2, \dots, a_n \in K$ s.t. $u_2 = b_1 u_1 + a_2 v_2 + \cdots + a_n v_n$.
We want to show that at least one of the a_j is non-zero.
Assume $a_2 = \cdots = a_n = 0$.
Then, $u_2 = b_1 u_1 \implies b_1 u_1 - 1 \cdot u_2 = 0$
which is a non-trivial combination of u_1, u_2 resulting in the zero vector. But u_1, u_2 are LI, thus at least one $a_j \neq 0$. w.l.o.g assume that $a_2 \neq 0$.

$$\begin{aligned} a_2 v_2 &= u_2 - b_1 u_1 - a_3 v_3 - \cdots - a_n v_n \\ v_2 &= \frac{1}{a_2} u_2 - \frac{b_1}{a_2} u_1 - \frac{a_3}{a_2} u_3 - \cdots - \frac{a_n}{a_2} u_n \end{aligned}$$

$\implies v_2$ is a LC of $u_1, u_2, v_3, \dots, v_n$. Thus $V = span\{u_1, v_2, v_3, \dots, v_n\} = span\{u_1, u_2, v_3, \dots, v_n\}$.

...
After the m^{th} step, we obtain a spanning set $\{u_1, \ldots, u_m, v_{m+1}, \ldots, v_n\}$ of V. Especially $m \leq n$. □

LECTURE 01/23

Theoretical Applications of the Steinitz Exchange Lemma

Theorem 1.10. *Every subspace of a finite dimensional VS over K is finite dimensional.*

Proof. Let V be a finite dimensional VS, $U \leq V.V$ is finite dimensional, thus it has a finite spanning set $\{v_1, \cdots, v_n\}$. We need to show that U has a finite spanning set. We proceed recursively.

Assume this is not the case:

Then $U \neq \{\vec{0}\}$. Let $u_1 \neq 0, u_1 \in U$. Then $\{u_1\}$ is linearly independent but not spanning. Thus $\exists u_2 \in U$ with $u_2 \notin span\{u_1\}$. By the previous theorem, $\{u_1, u_2\}$ is linearly independent; however, it is not spanning, thus $\exists u_3 \notin span\{u_1, u_2\}$, etc.
After n+1 iterations, we've obtained a linearly independent set $\{u_1, \cdots, u_{n+1}\} \subseteq U \subseteq V$. But V has a spanning set $\{u_1, \cdots, u_n\}$ containing n elements! $\implies$ CONTRADICTION to Steiniz! Our assumption was wrong. Thus U is finite dimensional. □

Examples 1.8.

1. The VSs of all sequences with coefficients in K is infinite dimensional.
2. We already know that c_{00} is infinite dimensional $c_{00} \leq s$. If s were finite dimensional then so would be c_{00} by previous Thm but c_{00} is infinite dimensional $\implies$ s is infinite dimensional.
3. Similarly, c_0 is also infinite dimensional since $c_{00} \leq c_0$.

Definition 1.10. **Basis**. Let V be a VS over K. A subset B of V is called a **basis** of V if:

(i) B is linearly independent.

(ii) B is spanning.

Exercise 1.9. Prove that any superspace of an infinite dimensional VS over K is infinite dimensional.

Examples 1.10.

1. $V = \mathbb{R}^n$, $B = \{e_1, \cdots, e_n\}$
2. $V = \mathbb{C}^n$, $B = \{e_1, \cdots, e_n\}$
3. $V = P_n$, Polynomials of degree $\leq n$, $B = \{1, x, x^2, \cdots, x^n\}$ [8]
4. $V = \mathbf{mat}(mxn, K)$. Let $M_{ij} = (m_{ij})$, where $m_{kl} \begin{cases} 0 \text{ if } (k,l) \neq (i,j) \\ 1 \text{ if } (k,l) = (i,j) \end{cases}$

   then $\{M_{ij}\}, 1 \leq i \leq m$, and $1 \leq j \leq n$ is a basis for V.[9]
5. $m = n = 2$: {4'identity' matrices} is a basis for $\mathbf{Mat}(2 \times 2, K)$.

[8] Showing that B is a basis for V:

Proof. B is spanning.
It is also LI: Let $a_0, a_1, \cdots, a_n \in K$ such that $a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n = 0$ ($\implies$ zero polynomial) $= 0 + 0 \cdot x + 0 \cdot x^2 + \cdots + 0 \cdot x^n \; \forall x \in K$
Two polynomials are identically equal iff they have the same coefficients. □

[9] See assignment 3.

Q: Does every VS have a basis?
A: Yes but we'll prove this only in the finite dimensional case.

Theorem 1.11. *Every finite dimensional VS has a basis.*

Proof. By *Casting Out Algorithm* (Steiniz proof), let V be finite dimensional, by definition of finite dimensional, V has a finite spanning set $\{u_1, \cdots, u_n\}$. If $\{u_1, \cdots, u_n\}$ is linearly independent, it is a basis and DONE.

If $\{v_i, \cdots, v_n\}$ is linearly dependent $\exists j : 1 \leq j \leq n$ such that v_j is a linear combination of $v_1, \cdots, \hat{v}_j, \cdots, v_n$. W.l.o.g., assume that $j = n$. As proven in previous theorem, we then have that $span\{v_1, \cdots, v_n\} = span\{v_1, \cdots, v_{n-1}\} \implies \{v_1, \cdots, v_{n-1}\}$ spans V. After repeating this arguments finitely many times, we obtain $v_1, \cdots, v_k, \ \ k \leq n$ such that $\{v_1, \cdots, v_k\}$ is spanning and linearly independent and thus a basis.

Q: What is a basis for $\{0\}$?
A: $B = \{\}$, since $span\{\} = \{0\}$.

Another version of the proof: *Bottom Up Algorithm*:

Let V be finite dimensional VS over K. Say V has a spanning set of n elements. If $V = \{0\}, \ \ B = \{\}$ and were done. If not, pick any $v_1 \in V, \ \ v_1 \neq 0$, then $\{v_1\}$ is linearly independent. It if is also spanning, we're done. If not, let $v_2 \in V$ be any vector with $v_2 \notin span\{v_1\}$, then by previous theorem, $\{v_1, v_2\}$ is linearly independent if $\{v_1, v_2\}$ is spanning, we're done. If not, $\exists v_3 \in V : v_3 \notin span\{v_1, v_2\}$, etc... until spanning. So by Steiniz, this algorithm has to terminate after at most n steps. We obtain $\{v_1, \cdots, v_k\}, \ \ k \leq n$, linearly independent and spanning, i.e. a basis. □

Example 1.11. Polynomials of degree ≤ 2, coefficients $\in K$. We know that $\{1, x, x^2\}$ is a basis and thus spanning. Finding another basis.

Proof. Let $v_0 \equiv 2$, then $span\{v_0\} \neq P_2$. Let $v_1 \equiv 3 + 5x$, then $v_1 \notin span\{v_0\}$
$\implies \{v_0, v_1\}$ is linearly independent, not spanning (no polynomial of degree 2 is in the span).
Let $v_2 \equiv 7 - 5x + 11x^2$, then $v_2 \notin span\{v_0, v_1\} \implies \{v_0, v_1, v_3\}$ is linearly independent. It needs to be spanning as well since otherwise $\exists v_3 \notin span\{v_0, v_1, v_2\} \implies \{v_0, v_1, v_2, v_3\}$ linearly independent $\implies$ CONTRADICTION to Steinitz. $\implies \{v_0, v_1, v_2\}$ is linearly independent and spanning, $\implies \{v_0, v_1, v_2\}$ is a basis for P_2. □

Theorem 1.12. *Let V be a* finite dimensional *VS over K. Let $\{u_1, \ldots, u_m\}$ and $\{u_1, \ldots, u_n\}$ be bases for V. Then $m = n$, i.e. any 2 bases for a finite dimensional vector space contain the* same *number of elements.*

Proof. Especially $\{u_1, \cdots, u_m\}$ is linearly independent and $\{u_1, \cdots, u_n\}$ is spanning. By Steiniz $\implies m \leq n$. But we also have that $\{u_1, \cdots, u_n\}$ is linearly independent and $\{u_1, \cdots, u_m\}$ is spanning. By Steiniz $\implies n \leq m. \therefore n = m.$ □

Remark: the notion of dimension is well-defined by previous theorem.

Definition 1.11. Dimension. Let V be finite dimensional VS over K. Let $\{v_1, \cdots, v_n\}$ be any basis for V. Then n is called the **dimension** of V.

LECTURE 01/25

Examples 1.12. *of dimension:*

1. $\mathbb{R}^n$ as VS over $\mathbb{R}$ has dimension n since $\{e_1, \ldots, e_n\}$ is a basis.

2. $\mathbb{C}^n$ as VS over $\mathbb{C}$ has dimension n since $\{e_1, \ldots, e_n\}$ is a basis.

3. $P_n(K)$ VS over K has dimension $n+1$ since $\{1, x, x^2, \ldots, x^n\}$ is a basis with $n+1$ elements.

4. $\mathbf{Mat}(m \times n, K)$ VS over K has dimension $m \cdot n$ since
$$\{E_{ij} = (m_{kl}) \text{ where } m_{kl} = \begin{cases} 1 \text{ if } (k,l) = (i,j) \\ 0 \text{ if } (k,l) \neq (i,j) \end{cases} \quad \}$$
is a basis with $m \cdot n$ elements.

Some Theorems on Bases and Dimension

Theorem 1.13. *Let V be a finite dimensional VS over K, $B = \{v_1, \ldots, v_n\}$ a basis for V. Let $u \in V$. Then there exist **uniquely determined** $k_1, \ldots, k_n \in K$ s.t. $v = k_1 v_1 + \cdots + k_n v_n$.*

Proof. B is spanning, thus v can be written as a LC of $v_1, \ldots, v_n$.
Let $v = k_1 v_1 + \cdots + k_n v_n$ and $v = l_1 v_1 + \cdots + l_n v_n$.

$$\begin{aligned} \implies 0 &= (k_1 v_1 + \cdots + k_n v_n) - (l_1 v_1 + \cdots + l_n v_n) \\ &= (k_1 v_1 - l_1 v_1) + \cdots + (k_n v_n - l_n v_n) \\ &= (k_1 - l_1) v_1 + \cdots + (k_n - l_n) v_n \end{aligned}$$

Since $v_1, \ldots, v_n$ are LI we have that $k_1 - l_1 = 0, \ldots, k_n - l_n = 0$.
$\implies k_1 = l_1, \ldots, k_n = l_n$.
$\implies v$ can be expressed in exactly one way as a LC of $v_1, \ldots, v_n$. □

Theorem 1.14. *Let V be a finite dimensional VS over K, let $U \leq V$. Then $\dim U \leq \dim V$.*

Remark: Note that U is finite dim since V is finite dim. Thus, $\dim U$ is defined.

Proof. Let $\{u_1, \ldots, u_m\}$ be a basis for U, $\{v_1, \ldots, v_n\}$ be a basis for V.
Note that $\{u_1, \ldots, u_m\}$ is LI and $\{v_1, \ldots, v_n\}$ is spanning. By Steinitz (Theorem 1.9), $m \leq n \implies \dim U \leq \dim V$. □

Theorem 1.15. *Let V be a finite dimensional VS over K with $\dim V = n$. Let $B = \{v_1, \ldots, v_n\}$ be LI. Then B is a basis for V.*

Proof. B is LI. We need to show it is spanning. Assume it is not.

Then $\exists\, v_{n+1} \in V, v_{n+1} \notin span\{v_1, \ldots, v_n\} \implies \{v_1, \ldots, v_n, v_{n+1}\}$ is LI. But since $\dim V = n$, V has a spanning set of length n. CONTRADICTION with Steinitz. Thus B is also spanning and $\therefore$ a basis. □

Theorem 1.16. *Let V be a finite dimensional VS over K with $dimV = n$. Let $B = \{v_1, \ldots, v_n\}$ be spanning. Then B is a basis for V.*

Proof. B is spanning. We need to show that B is LI. Assume that B is LD. Then $\exists j : 1 \leq j \leq n$ s.t. v_j is a LC of $v_1, \cdots, \hat{v}_j, \cdots, v_n$ and $V = span\{v_1, \cdots, v_n\} = span\{v_1, \cdots, \hat{v}_j, \cdots, v_n\}$.

i.e. V has a spanning set of $n-1$ vectors. But since $\dim V = n$, it has a LI set of length n, CONTRADICTION with Steinitz.

$\implies B$ is also LI and thus a basis. □

Example 1.13. $V = P_2(K)$, $\dim V = 3.$, $B = \{3, 5-9x, 7+3x+13x^2\}$.

B is LI for degree reasons, thus B is also spanning.

Theorem 1.17. *Let V be a finite dimensional VS over K, $U \leq V$. Let $B_u = \{u_1, \ldots, u_m\}$ be a basis for U. Then B_u can be extended to a basis B_v for V, i.e. $\exists\, u_{m+1}, \ldots, u_n \in V$ s.t. $B_v \equiv \{u_1, \ldots, u_m, u_{m+1}, \ldots, u_n\}$ is a basis for V.*

Proof. This is the same argument we used for proving the existence of a basis via the *Bottom Up* algorithm: we start with B_u in the initial step.[10] □

[10] **Exercise**: Work out the details!

Theorem 1.18. *Let V be a finite dimensional VS over K, $U \leq V$, $W \leq V$. Then $dim(U+W) = dimU + dimW - dim(U \cap W)$.*

Proof. Let $\{v_1, \cdots, v_k\}$ be a basis for $U \cap W$. $U \cap W \leq U$. $\therefore$ by the previous theorem, we can extend this to a basis $\{v_1, \cdots, v_k, u_{k+1}, \cdots, u_m\}$ for U. Similarly, we can find $w_{k+1}, \cdots, w_n$ s.t. $\{v_1, \cdots, v_k, w_{k+1}, \cdots, w_n\}$ is a basis for W.

Now consider the set $S \equiv \{v_1, \cdots, v_k, u_{k+1}, \cdots, u_m, w_{k+1}, \cdots, w_n\}$. We will show that S is a basis for $U+W$.

1. STEP: S spans $U+W$, i.e. $span\{S\} = U+W$.

First show "$\subseteq$". Let $v \in spanS$ be ARBITRARY. Then:

$$\begin{aligned} v = a_1v_1 + \cdots + a_kv_k & \qquad \in U \cap W \\ + b_{k+1}u_{k+1} + \cdots + b_mu_m & \qquad \in U \\ + c_{k+1}w_{k+1} + \cdots + c_nw_n & \qquad \in W \end{aligned}$$

where (a_i, b_j, c_l) are scalars.

$\implies v \in U+W \therefore span\{S\} \subseteq U+W$

Next show "$\supseteq$": Let $v \in U+W$ be arbitrary.

Then $\exists u \in U, w \in W$ s.t. $v = u + w \in span\{v_1, \cdots, v_k, w_{k+1},$

$\cdots, w_n\} \in span\{v_1, \cdots, v_k, u_{k+1}, \cdots, u_n\} \in spanS.$
$\implies U + W \subseteq spanS$
$\implies U + W = span\{S\}$
i.e. S spans $U + W$.

2. STEP: S is LI. Let the as, bs and cs $\in K$ be such that:

$$\begin{aligned}
0 &= a_1 v_1 + \cdots + a_k v_k \\
&\quad + b_{k+1} u_{k+1} + \cdots + b_m u_m \\
&\quad + c_{k+1} w_{k+1} + \cdots + c_n w_n && \text{(*)} \\
\implies & a_1 v_1 + \cdots + a_k v_k && \in U \cap W \\
&\quad + b_{k+1} u_{k+1} + \cdots + b_m u_m && \in U \\
&\quad = -c_{k+1} w_{k+1} - \cdots - c_n w_n && \in U \cap W \quad \text{(since =)}
\end{aligned}$$

Since $\{v_1, \ldots, v_k\}$ is a basis for $U \cap W$, $\exists a'_1, \cdots, a'_k \in K$ s.t.

$$\begin{aligned}
-c_{k+1} w_{k+1} - \cdots - c_n w_n &= a'_1 v_1 + \cdots + a'_k v_k \\
\implies 0 \cdot v_1 + \cdots 0 \cdot v_k - c_{k+1} w_{k+1} - \cdots - c_n w_n & \\
= a'_1 v_1 + \cdots + a'_k v_k + 0 \cdot w_{k+1} + \cdots + 0 \cdot w_n &
\end{aligned}$$

Note that $\{v_1, \cdots, v_k, w_{k+1}, \cdots, w_n\}$ is a basis for W and thus especially LI. Thus, the coefficients in the previous equations are uniquely determined.
$\implies [a'_1 = \cdots = a'_k = 0]$ and $c_{k+1} = \cdots = c_n = 0$.
Plugging this into (*):

$$0 = a_1 v_1 + \cdots + a_k v_k + b_{k+1} u_{k+1} + \cdots + b_m u_m$$

But $\{v_1, \cdots, v_k, u_{k+1}, \cdots, u_m\}$ is a basis for U. Thus, $a_1 = \cdots = a_k = 0$ and $b_{k+1} = \cdots = b_m = 0$. This proves that S is LI. Hence, S is a basis for $U + W$.

$$\begin{aligned}
dim(U + W) &= |S| && \text{(size of } S\text{)} \\
&= k + (m - k) + (n - k) && (vs, us, ws) \\
&= m + n - k \\
&= dimU + dimW - dim(U \cap W) && \square
\end{aligned}$$

Example 1.14. Any two planes through the origin in $\mathbb{R}^3$ have non trivial intersection, i.e. they intersect in at least a line through the origin. [11]

[11] *Proof.* Planes through the origin are the 2-dim subspaces of $\mathbb{R}^3$. Let U, W be 2 such planes. Then:

$$\begin{aligned}
dim(U + W) &= dimU + dimW - dim(U \cap W) \\
dim(U \cap W) &= dimU + dimW - dim(U + W) \\
&= 2 + 2 - dim(U + W) \\
&= 4 - dim(U + W) \\
&\quad (U + W \leq \mathbb{R}^3) \\
&\geq 4 - 3 = 1
\end{aligned}$$

$\implies$ the intersection of the planes is a subspace of $\mathbb{R}^3$ of dim ≥ 1 and is thus at least a line.

1.6 Direct Sums

LECTURE 01/30

Recall: V VS over K, and $U \leq V$, $W \leq V$, then $U + W \equiv \{u + w : u \in U, w \in W\}$ and $U + W \leq V$

Definition 1.12. A sum $U + W$ of subspaces U, W of V is called **direct** - in symbols, $U \oplus V$) if $\forall v \in U + W$, the representation $v = u + w$, with $u \in U,\ w \in W$ is **uniquely determined**.

Similarly a sum $U_1 + \cdots + U_n$ of subspaces of V is called **direct** if $\forall v \in U_1 + \cdots + U_n$, the representation $v = u_1 + \cdots + u_n, \quad u_j \in U_j\ \forall 1 \leq j \leq n$ is uniquely defined.

Theorem 1.19. *$U + W$ of subspaces of V is direct iff $U \cap W = \{0\}$.*

Proof. " $\Rightarrow$ " Let $U + W$ be direct. Let $v \in U \cap W$, then:

$$\underbrace{0}_{\in U+W} = \underbrace{0}_{\in U} + \underbrace{0}_{\in W} = \underbrace{v}_{\in U} + \underbrace{-v}_{\in W}$$

Since the representation is unique, $v = 0 \implies U \cap W = \{0\}$.

" $\Leftarrow$ " Let $v \in U + W$ be arbitrary. Let $v = u_1 + w_1$ and $v = u_2 + w_2$, for $u_1, u_2 \in U,\ \ w_1, w_2 \in W$.

$$\begin{aligned}
&v = u_1 + w_1 \\
&v = u_2 + w_2 \\
&0 = (u_1 - u_2) + (w_1 - w_2) \\
\implies &\underbrace{\underbrace{u_1 - u_2}_{\in U}}_{\in U \cap W} = \underbrace{\underbrace{w_2 - w_1}_{\in W}}_{\in U \cap W} \\
\implies &u_1 - u_2 = 0 = w_2 - w_1 \\
\implies &u_1 = u_2 \wedge w_1 = w_2
\end{aligned}$$

$\implies v$ is uniquely determined $\implies U + W$ is direct. □

Exercise 1.15. Let $v_1, \cdots, v_n \in V$ be LI, then $span\{v_1\} + \cdots + span\{v_n\}$ is direct.

2 Linear Maps

2.1 Mappings and Associated Subspaces

LECTURE 01/30

Definition 2.1. Let V and W be VS over K. A function $L : V \to W$ is called **linear** if:

(Same K for both V and W.)

1. $L(v_1 + v_2) = L(v_1) + L(v_2)$
2. $L(kv_1) = kL(v_1)$

$v_1, v_2 \in V$
$L(v_1), L(v_2) \in W$

Theorem 2.1. *Let $L : V \to W$ be linear, then $L(0 \in V) = 0 \in W$.*

Proof.

– *version 1.* $L(0) = L(0+0) = L(0) + L(0)$ definition (1.)

$\implies 0 + \cancel{L(0)} = L(0) + \cancel{L(0)}$ cancellation

$\implies L(0) = 0$

– *version 2.* $L(0) = L(0 \cdot 0) = 0 \cdot L(0)$ definition (2.) □

$= 0$ theorem

Definition 2.2. Subspaces associated with linear maps. Let $L : V \to W$ be linear. Define:

Warning 1. In calculus, the so-called "linear" maps $f : \mathbb{R} \to \mathbb{R}, x \mapsto ax + b$ are only linear if $b = 0$!

(a) The **kernel**, $\ker L$ is defined as $\ker L \equiv \{v \in V : L(v) = 0\} \subseteq V$.

(b) The **image** *im* L is defined as $im\ L \equiv \{L(v) : v \in V\} \subseteq W$

Remark 2. Axler calls *ker* L the *nullspace* of L and *im* L the *range* of L.

Theorem 2.2. *Let $L : V \to W$ be linear, then:*

1. $ker\ L \leq V$
2. $im\ L \leq W$

Proof.

1. $0 \in ker\ L$ since $L(0) = 0$. Let $v_1, v_2 \in ker\ L$, then $L(v_1 + v_2) = L(v_1) + L(v_2) = 0 + 0 = 0 \implies v_1 + v_2 \in ker\ L$ and $L(kv_1) = kL(v_1) = k \cdot 0 = 0 \implies kv_1 \in ker\ L$
   $ker\ L \leq V$
2. Since $L(0(\in V)) = 0(\in W) \implies 0(\in W) \in im\ L$. Let $w_1, w_2 \in im\ L$. Thus $\exists v_1, v_2 \in V$ with $L(v_1) = w_1$ and $L(v_2) = w_2$, then $L(v_1 + v_2) = L(v_1) + L(v_2) = w_1 + w_2 \implies w_1 + w_2 \in im\ L$ and $L(kv_1) = kL(v_1) = kw_1 \implies kw_1 \in im\ L \implies im\ L \leq W$. □

Examples 2.1.

1. L defined as: $L : K^n \to k^m$ is linear:
$$x \mapsto \underbrace{\underbrace{A}_{m\times n} \cdot \underbrace{x}_{n\times 1}}_{m\times 1}$$
$A(x+\tilde{x}) = Ax + A\tilde{x} \implies L(x+\tilde{x}) = L(x) + L(\tilde{x})$
$A(kx) = k(Ax) \implies L(kx) = kL(x)$
$\implies$ L is linear.[12]

2. D defined as: $D : P(\mathbb{R} \to P(\mathbb{R})$ is linear:
$$p \mapsto p'(\text{derivative})$$
$D(n+q) = (n+q)' = n' + q' = D(n) + D(q)$
$D(kn) = (kn)' = kn' = kD(n)$
$\implies D$ is linear.[13]

3. (a) D defined as: $D : P_n(\mathbb{R} \to P_{n-1}(\mathbb{R})$ is linear as shown
$$p \mapsto p'$$
above.[14]

   (b) Same function defined on $D : P_n(\mathbb{R}) \to P_n(\mathbb{R})$ is not surjective.[15]

4. Integration: $V = P(\mathbb{R})$, for $a, b \in \mathbb{R}$

   (a) L defined as: $L : P(\mathbb{R}) \to \mathbb{R}$ is linear:
$$p \mapsto \int_a^b p dx$$
$$L(p+q) = \int_a^b (p+q)dx = \int_a^b pdx + \int_a^b qdx = L(p) + L(q)$$
$$L(kp) = \int_a^b kpdx = k\int_a^b pdx = kL(p)$$
$\implies$ L linear.

   (b) $a \in \mathbb{R}$: $L : P(\mathbb{R}) \to P(\mathbb{R})$
$$p \mapsto \int_a^x p(t)dt$$
L is linear as shown above.

   (c) L defined as follows is linear: $L : P(\mathbb{R}) \to P_{n+1}(\mathbb{R})$
$$r \mapsto \int_a^x rdt$$

Assignment 4: $V = s,\ s = \{(x_1, x_2, x_3, \cdots) : x_i \in \mathbb{R}, \forall i\}$
Left-shift: $(x_1, x_2, x_3, \cdots) \mapsto (x_2, x_3, x_4, \cdots)$
Right-shift:$(x_1, x_2, x_3, \cdots) \mapsto (0, x_1, x_2, \cdots)$.

[12]
- $ker\ L = nullspace(A)$
- $im\ L = colspace(A)$ (c.f. assign. 2)

[13]
- $ker\ D = P_0(\mathbb{R})$
- $im\ D = P(\mathbb{R})$

[14]
- $ker\ D = P_0(\mathbb{R})$
- $im\ D = P_{n-1}(\mathbb{R})$

D is surjective.

[15]
- $ker\ L = P_0(\mathbb{R})$
- $im\ L = P_{n-1}(\mathbb{R}) \subsetneqq P_n(\mathbb{R})$

$\ker L = \{0\}$
$im\ L =$ Set of all polynomials with a root at $x = a$.

Definition 2.3. $I \subseteq \mathbb{R}$ interval.
$c^0(I) \equiv$ Set of all continuously functions on I.
$c^1(I) \equiv$ Set of all continuously differentiable functions $\cdots$
$\cdots$
$c^n(I) \equiv$ Set of all n-times continuously differentiable functions on I.
$\cdots$
$c^\infty(I) \equiv$ Set of all functions on I with derivatives of all orders.

Exercise 2.2. $c^0(I), c^n(I), c^\infty(I)$ are VS over $\mathbb{R}$.
Show $c^\infty(I) \leq \cdots \leq c^n(I) \leq \cdots \leq c^1(I) \leq c^0(I)$.

$D : c^n(I) \to c^{n-1}(I)$ or $D : c^\infty(I) \to c^\infty(I)$ are all linear.

LECTURE 02/01

Theorem 2.3. *Let V, W be VS over K, $L : V \to W$ linear. Then L is injective iff $ker\ L = \{0\}$.*

Proof.
"$\Rightarrow$" Let L be injective[16] and let $v \in ker\ L$. Then:
$L(v) = 0 = L(0) \implies v = 0 \implies ker\ L = \{0\}$.
"$\Leftarrow$" Let $ker\ L = \{0\}$. Let $v, \tilde{v} \in V$ s.t. $L(v) = L(\tilde{v})$.
$\implies L(v) - L(\tilde{v}) = 0 = L(v - \tilde{v})$
$\implies v - \tilde{v} \in ker\ L \implies v - \tilde{v} = 0 \implies v = \tilde{v}$
$\implies L$ is injective. □

[16] injective = one to one

Theorem 2.4. *Let V be a finite dimensional VS over K, W a VS over K and $\{v_1, \cdots, v_n\}$ a basis for V. Let $w_1, \cdots, w_n$ be arbitrary vectors in W. Then there exists a **uniquely determined** linear map $L : V \to W$ s.t. $L(v_1) = w_1, \cdots, L(v_n) = w_n$.*

Proof.
DEFINITION OF L:
Let $v \in V$. $\exists k_1, \cdots, k_n \in K$ s.t. $v = k_1 v_1 + \cdots + k_n v_n$. Define $L(v) \equiv k_1 w_1 + \cdots + k_n w_n$. Note that L is a function since $k_1, \cdots, k_n$ are uniquely determined.

LINEARITY OF L:
Let $v, \tilde{v} \in V$.

$$\begin{aligned} v &= k_1 v_1 + \cdots + k_n v_n \\ \tilde{v} &= l_1 v_1 + \cdots + l_n v_n \\ v + \tilde{v} &= (k_1 + l_1) v_1 + \cdots + (k_n + l_n) v_n \\ \implies L(v + \tilde{v}) &= (k_1 + l_1) w_1 + \cdots + (k_n + l_n) w_n \\ &= (k_1 w_1 + \cdots + k_n w_n) + (l_1 w_1 + \cdots + l_n w_n) \\ &= L(v) + L(\tilde{v}) \end{aligned}$$

and

$$\begin{aligned} kv &= (kk_1)v_1 + \cdots + (kk_n)v_n \\ \implies L(kv) &= (kk_1)w_1 + \cdots + (kk_n)w_n \\ &= k[k_1w_1 + \cdots k_nw_n] \\ &= kL(v) \end{aligned}$$

$\implies L$ is linear.

UNIQUENESS:

Let $\tilde{L} : V \to W$ be linear with $\tilde{L}(v_1) = w_1, \cdots, \tilde{L}(v_n) = w_n$.

$$\begin{aligned} &\implies \tilde{L}(k_1v_1) = k_1w_1, \cdots, \tilde{L}(k_nv_n) = k_nw_n \\ &\implies \tilde{L}(k_1v_1 + \cdots + k_nv_n) \\ &\qquad = \tilde{L}(k_1v_1) + \cdots + \tilde{L}(k_nv_n) \\ &\qquad = k_1w_1 + \cdots + k_nw_n \\ &\qquad = L(k_1w_1 + \cdots + k_nv_n) \\ &\implies \tilde{L}(v) = L(v)\ \forall v \in V \implies \tilde{L} = L \end{aligned}$$

$\implies L$ is uniquely determined. □

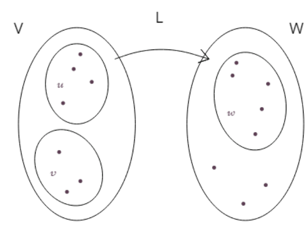

Figure 1: Representation of Thm 2.5

Note that the v_i are not necessarily uniquely determined!

Theorem 2.5. ***Dimension Formula****. Let V be a finite dimensional VS over K, W VS over K. Let $L : V \to W$ be linear. Then $\dim \ker L + \dim im\ L = \dim V$.*

Proof. Let $\{u_1, \cdots, u_k\}$ be a basis for $ker\ L$ and let $\{w_1, \cdots, w_n\}$ be a basis for $im\ L$. Since $w_1, \cdots, w_n \in im\ L\ \exists v_1, \cdots, v_n \in V$ with $L(v_1) = w_1, \cdots, L(v_n) = w_n$.

Let $B \equiv \{u_1, \cdots, u_k, v_1, \cdots, v_n\}$. CLAIM: B is a basis for V.

1. SHOW B IS LI:

   Let $a_1u_1 + \cdots + a_ku_k + b_1v_1 + \cdots + b_nv_n = 0 \qquad (\star)$.

   Applying L to both sides:

$$\begin{aligned} &\implies L(a_1u_1 + \cdots a_ku_k + b_1v_1 + \cdots + b_nv_n) = L(0) = 0 \\ &\implies \underbrace{a_1L(u_1)}_{=0} + \cdots + \underbrace{a_kL(u_k)}_{=0} + \underbrace{b_1L(u_1)}_{=w_1} + \cdots + \underbrace{b_nL(v_n)}_{=w_n} = 0 \\ &\implies b_1w_1 + \cdots + b_nw_n = 0 \end{aligned}$$

   since $w_1, \cdots, w_n$ are LI.

   It follows that $b_1 = \cdots = b_n = 0$ in $(\star)$: $a_1u_1 + \cdots + a_ku_k = 0$.
   Since $u_1, \cdots, u_k$ are LI, $\implies a_1 = \cdots = a_k = 0$, $\implies B$ is LI.

2. SHOW B IS SPANNING:

Let $v \in V$ be arbitrary. Let $w \equiv L(v)$. Then $w \in im\ L$. Thus $\exists b_1, \cdots, b_n \in K$ s.t. $w = b_1 w_1 + \cdots + b_n w_n$.

Consider $\tilde{v} \equiv b_1 v_1 + \cdots + b_n v_n$. Then:

$$\begin{aligned}
L(v - \tilde{v}) &= \underbrace{L(v)}_{=w} - \underbrace{L(\tilde{v})}_{=b_1 w_1 + \cdots + b_n w_n = w} = 0 \\
\implies & v - \tilde{v} \in ker\ L \implies \exists a_1, \cdots, a_k \text{s.t.} \\
& v - \tilde{v} = a_1 u_1 + \cdots + a_k u_k \\
\implies & v = a_1 u_1 + \cdots + a_k u_k + b_1 v_1 + \cdots + b_n v_n \\
\implies & B \text{ is spanning}
\end{aligned}$$

$\implies$ B is a basis for V.

$\implies$ We have: $dimV = k + n = dim\ ker\ L + dim\ im\ L$ □

Examples 2.3.

1. $D : P_n(K) \to P_n(K)$ DIFFERENTIATION.
   $ker\ D$ consists of all constant polynomials, i.e. $ker\ D = P_0(K)$.
   $dim\ ker\ D = 1$.
   $im\ D$ consists of all polynomials of degree at most $n-1$. i.e.
   $im\ D = P_{n-1}(K) \implies dim\ im\ D = n$.

$$\begin{aligned}
dim\ P_n(K) &\stackrel{?}{=} dim\ ker\ D + dim\ im\ L \\
n + 1 &= 1 + n \qquad \checkmark
\end{aligned}$$

2. Let $A \in \mathbf{Mat}(n \times m, \mathbb{R})$.
   Consider $Ax = 0$.
   Learned in previous linear algebra courses that:
   $dim\ nullspace(A) + dim\ colspace(A) = m$
   ($dim\ colspace(A) = dim\ rowspace(A) = rank(A)$)

ASSIGNMENT 2:
$L : x \mapsto Ax$, $L : K^m \to K^n$
$nullspace(A) = ker\ L$
$colspace(A) = im\ L$
since $dim\ ker\ L + dim\ im\ L = dim\ V$
$\implies \underbrace{dim\ nullspace(A)}_{nullity(A)} + rank(A) = m$

3. $tr : \mathbf{Mat}(n \times n, K) \to K$ "trace": $\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \cdots & & \cdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \mapsto a_{11} + \cdots a_{nn}$.
   tr is linear:

$$\begin{aligned}
tr(A+B) &= tr \begin{pmatrix} a_{11}+b_{11} & \cdots & a_{1n}+b_{1n} \\ \cdots & & \cdots \\ a_{n1}+b_{n1} & \cdots & a_{nn}+b_{nn} \end{pmatrix} \\
&= (a_{11} + b_{11}) + \cdots + (a_{nn} + b_{nn}) \\
&= (a_{11} + \cdots + a_{nn}) + (b_{11} + \cdots + b_{nn}) \\
&= trA + trB
\end{aligned}$$

$$\begin{aligned}
tr(kA) &= tr \begin{pmatrix} ka_{11} & \cdots & ka_{1n} \\ \cdots & & \cdots \\ ka_{n1} & \cdots & ka_{nn} \end{pmatrix} \\
&= ka_{11} + \cdots + ka_{nn} = k(a_{11} + \cdots + a_{nn}) \\
&= k \cdot trA
\end{aligned}$$

ker L is the set of all matrices where the sum of all elements along the main diagonal is 0.
$im\ L = K$, $dim\ im\ L = 1$
$dim\ \mathbf{Mat}(n \times n, K) = n^2$
$\implies dim\ ker\ L = n^2 - 1$.

4. Let $A \in \mathbf{Mat}(n \times m, K)$ with $rows = n < columns = n$. The lin. hom. system $Ax = 0$ has non-trivial solutions.
   Let $L : x \mapsto Ax : K^m \to K^n$
   Then $Ax = 0$ has non-trivial sol. iff L has a non-trivial kernel, i.e. iff $dim\ ker\ L > 0$.
   By dimensional Formula, $dim\ ker\ L + dim\ im\ L = dim K^m$

$$\implies dim\ ker\ L = dim\ K^m - dim\ im\ L = m - dim\ colspace(A)$$
$$= m - \underbrace{dim\ rowspace(A)}_{\leq n} = m - n > 0$$

   $\implies ker\ L$ is non-trivial.
   $\implies Ax = 0$ has non-trivial solutions.

LECTURE 02/06

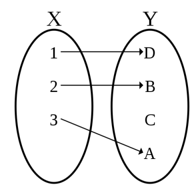

Figure 2: Injective (non surjective) map.

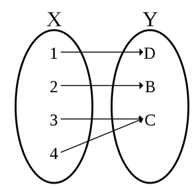

Figure 3: Surjective (non injective) map.

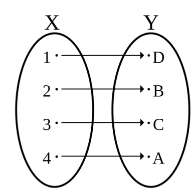

Figure 4: Bijective map.

2.2 *Endomorphisms and Isomorphisms*

Definition 2.4. An **endomorphism** is a linear map $L : V \to V$, V is VS over K.

Theorem 2.6. *Let V be finite dimensional VS over K, $L : V \to V$ be an endomorphism, then L is bijective iff L is surjective iff L is injective.*

Lemma 2.7. *Lemma: Let V be a finite dimensional VS over K, and let $U \leq V$ s.t* $\dim U = dim V$, *then* $U = V$.

Proof. Let $\{v_1, \cdots, v_n\}$ be a basis for $U \implies \{v_1, \cdots, v_n\}$ is LI, where $n = \dim U = \dim V \implies \{v_1, \cdots, v_n\}$ is a basis for V, especially, it spans $V \implies U = V$ $\square$

Proof. of Theorem 2.6.
There are 2 non-trivial implications:

1. injective $\implies$ surjective:
   By Dimension formula, $\dim \ker L + \dim imL = \dim V$. L injective $\implies kerL = \{0\}$.
   $\implies \dim kerL = 0 \implies \dim imL = \dim V$, where $imL \leq V$. By Lemma 2.7, $im\ L = V \implies L$ is surjective.

2. Surjective $\implies$ injective:
   Let L be surjective $\implies im\ L = V \implies \dim im\ L = \dim V$
   $\dim \ker L = 0 \implies \ker L = \{0\} \implies L$ is injective. $\square$

Remark 3. Comparable results are wrong in almost any other context, e.g.: $f : \mathbb{R} \to \mathbb{R}$ is injective but NOT surjective,
$x \mapsto e^x$
while $f : \mathbb{R} \to \mathbb{R}$ is surjective but not injective.
$x \mapsto x^3 - x$

Remark 4. The result is also wrong in general for endomorphisms between infinite dimensional VS (See Ass. 5).

Coordinate Vectors

All bases in this section are considered as **ordered**. $B = (v_1, v_2, \cdots, v_n)$ i.e bases will have a first, second, etc. element. Vectors in K^n will always be considered as **column vectors** in this section.

Definition 2.5. Let V be finite dimensional VS over K, $B = (v_1, v_2, \cdots, v_n)$ is its basis. Let $v \in V$; then $\exists$ **uniquely determined** $a_1, \cdots, a_n \in K$ s.t $v = a_1 v_1 + \cdots + a_n v_n$. The column vector $\begin{pmatrix} a_1 \\ \cdots \\ a_n \end{pmatrix} \in K^n$ uniquely represents v. It is called the coordinate vector of v with respect to B, written as $[v]_B$.

Many textbooks use $\{v_1, \cdots, v_n\}$ even if they mean **ordered** bases.

Theorem 2.8. *Let V be finite dimensional, B an ordered basis for V, then the coordinate map* $[\ ]_B : V \to K^n,\ v \mapsto [v]_B$ *is linear and bijective (this is an* ***isomorphism****).*

Proof. 1. Linearity of $[\ ]_B$:
Let $B = (v_1, v_2, \cdots, v_n)$.
$[v_1]_B = 1 \cdot v_1 + 0 \cdot v_2 + \cdots + 0 \cdot v_n = \begin{pmatrix} 1 \\ 0 \\ \cdots \\ 0 \end{pmatrix} = e_1$
$[v_2]_B = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \cdots \\ 0 \end{pmatrix} = e_2$
...
$[v_n]_B = \begin{pmatrix} 0 \\ 0 \\ \cdots \\ 1 \end{pmatrix} = e_n$
By the Existence Theorem (Theorem 2.4) for linear maps, $\exists$ a uniquely determined linear map $L : V \to K^n$ with:
$L(v_1) = e_1, \cdots, L(v_n) = e_n$, where
$L(a_1 v_1 + \cdots + a_n v_n) = a_1 e_1 + \cdots + a_n e_n = \begin{pmatrix} a_1 \\ a_2 \\ \cdots \\ a_n \end{pmatrix}$
$\implies L = [\ ]_B \implies [\ ]_B$ is linear.

Exercise: Prove this DIRECTLY, i.e without the use of the Existence Theorem.

2. Bijectivity of $[\ ]_B$:
Injectivity: $[v]_B = \begin{pmatrix} 0 \\ \cdots \\ 0 \end{pmatrix} \iff v = 0 \cdot v_a + \cdots + 0 \cdot v_n = 0$ i.e $\ker[\ ]_B = \{0\} \implies [\ ]_B$ is injective.
Surjectivity: Let $\begin{pmatrix} a_1 \\ \cdots \\ a_n \end{pmatrix} \in K^n$ be arbitrary. Let $v \equiv a_1 v_1 + \cdots + a_n v_n$, then $[v]_B = \begin{pmatrix} a_1 \\ \cdots \\ a_n \end{pmatrix} \implies [\ ]_B$ is surjective.
$\implies [\ ]_B$ is bijective $\implies$ isomorphism. $\square$

Theorem 2.9. *on isomorphisms: Let $L : V \to W$ be an isomorphism. Since L is bijective, it has an inverse map $L^{-1} : W \to V$, then L^{-1} is linear.*

Proof. Let $w_1,\ w_2 \in W$, $v_1,\ v_2 \in V$ with $L(v_1) = w_1$ and $L(v_2) = w_2$, then

$$\begin{aligned} L^{-1}(w_1 + w_2) &= L^{-1}(L(v_1) + L(v_2)) \\ &= L^{-1}(L(v_1 + v_2)) \\ &= v_1 + v_2 = L^{-1}(w_1) + L^{-1}(w_2) \\ &\implies L^{-1}(w_1 + w_2) = L^{-1}(w_1) + L^{-1}(w_2) \end{aligned}$$

Let $k \in K$, then

$$\begin{aligned} L^{-1}(kw_1) &= L^{-1}(kL(v_1)) \\ &\implies L^{-1}(L(kv_1)) = kL^{-1}(w_1) \end{aligned}$$

$\implies L^{-1}$ linear. □

Theorem 2.10. *Let V be VS over K of dimension n, then V is isomorphic to K^n.*

Proof. Let B be ordered basis for V, then $[\ ] : V \to K^n$ is an isomorphism by Theorem 2.8. □

2.3 Matrix Representation of Linear Maps

Q: Does there exist a $n \times m$ matrix A with coefficients in K that mimics the action of L, i.e. $\underbrace{[L(v)]_{B'} = \underbrace{A}_{n\times m} \cdot \underbrace{[v]_B}_{m\times 1}}_{n\times 1} \forall v \in V$

A: Such a matrix, if it exists, would rightfully be called the **matrix representation** of L wrt B and B'.

Idea: Let $V,\ W$ be finite dimensional VS over K, $L : V \to W$, let B be ordered basis for V, B' ordered basis for W. $\dim V = m,\ \ \dim W = n$. For $v \in V,\ \ w \in W$. $[v]_B \in K^m$, $[w]_B \in K^n$. Before we construct A in the general case, let's revesit the *classical* case of $V = K^m$, $W = K^n$, B standard basis on K^m, and B' standard basis on K^n.

$$A = \begin{pmatrix} a_{11} & \dots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nm} \end{pmatrix}$$

The linear maps $L : K^m \to K^n$ are of the form $L(x) = Ax$, where A is an $n \times m$ matrix. What can we say about the coefficients of A?

$$\underbrace{Ae_1}_{\in K^m} = L(e_1) = \begin{pmatrix} a_{11} \\ \vdots \\ a_{n1} \end{pmatrix} \text{ i.e the first column of } A.$$

$$\dots$$

$$\underbrace{Ae_m}_{\in K^m} = L(e_m) = \begin{pmatrix} a_{1m} \\ \vdots \\ a_{nm} \end{pmatrix} \text{ i.e the } m\text{th column of } A.$$

i.e.

$$A = (L(e_1) \mid L(e_2) \mid \cdots \mid L(e_m))$$

LECTURE 02/08

Definition 2.6. Let $L : V \to W$ linear, V, W vector space over K with $\dim V = m$, $\dim W = n$, $B = \{v_1, \cdots, v_m\}$ basis for V, $\{w_1, \cdots, w_n\}$ basis for W. Define $[L]_{B',B}$ the matrix representation of L with respect to B and B' by:

$$[L]_{B',B} = \underbrace{\overbrace{\left(\underbrace{[L(v_1)]_{B'}}_{n\times 1} \;\middle|\; \cdots \;\middle|\; \underbrace{[L(v_m)]_{B'}}_{n\times 1} \right)}^{m \text{ columns}}}_{n\times m \text{ matrix}}$$

Theorem 2.11. *$L : V \to W$ linear, dim $V = m$, dim $W = n$, $B = \{v_1, \cdots, v_m\}$ basis for V, $B' = \{w_1, \cdots, w_n\}$ basis for W. Then:*

$$\underbrace{[L(v)]_{B'}}_{n\times 1} = \underbrace{\underbrace{[L]_{B',B}}_{n\times m}\underbrace{[v]_B}_{m\times 1}}_{n\times 1} \qquad \forall v \in V$$

Proof. Let $v \in V$ be arbitrary,

$$v = a_1 v_1 + \cdots + a_m v_m \implies [v]_B = \begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix}. \text{ Then:}$$

$$\begin{aligned} [L]_{B',B}[v]_B &= ([L(v_1)]_{B'} \mid \cdots \mid [L(v_m)]_{B'}) \cdot \begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix} \\ &= a_1[L(v_1)]_{B'} + \cdots + a_m[L(v_m)]_{B'} \\ &= [a_1L(v_1)]_{B'} + \cdots + [a_mL(v_m)]_{B'} \\ &= [a_1L(v_1) + \cdots + a_mL(v_m)]_{B'} \\ &= [L(a_1v_1 + \cdots + a_mv_m)]_{B'} = [L(v)]_{B'} \end{aligned}$$ □

Note that $[\ ]_{B'} : W \to K^n$ and L are linear.

Theorem 2.12. *Let $L : V \to W$, V, W finite dimensional, L linear, B basis for V and B' basis for W, $A \equiv [L]_{B',B}$. Then ker L is isomorphic to $nullspace(A)$ and im L is isomorphic to $colspace(A)$. The isomorphism is given by:*

$$\begin{aligned} ker\ L &\to nullspace(A) \\ v &\mapsto [v]_B \\ \text{and} \quad im\ L &\to colspace(A) \\ w &\mapsto [w]_{B'} \end{aligned}$$

Proof. Let $v \in ker\ L \Leftrightarrow L(v) = 0 \Leftrightarrow [L(v)]_{B'} = 0 = [L]_{B',B}[v]_B \Leftrightarrow [v]_B \in nullspace(A)$

Let $w \in im\ L \Leftrightarrow \exists v \in V : L(v) = w \Leftrightarrow [L(v)]_{B'} = [w]_{B'} = [L]_{B',B}[v]_B \Leftrightarrow [w]_{B'} \in colspace(A)$ □

Example 2.4. $D : P_2(\mathbb{R}) \to P_2(\mathbb{R})$ ENDOMORPHISM

$B = B' = (1, x, x^2)$, $D(p) \equiv p'$

$[D]_{B,B} = ?$ $[D]_{B,B} = [D]_B$

$$\begin{aligned} D(1) &= 0 = 0 \cdot 1 + 0 \cdot x + 0 \cdot x^2 \\ D(x) &= 1 = 1 \cdot 1 + 0 \cdot x + 0 \cdot x^2 \quad \implies [D]_{B,B} = \left(\begin{smallmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{smallmatrix}\right) \\ D(x^2) &= 2x = 0 \cdot 1 + 2 \cdot x + 0 \cdot x^2 \end{aligned}$$

We can use $[D]_{B,B}$ to differentiate a polynomial via matrix multiplication: Let $p(x) = 3 - 2x + 5x^2$, $p'(x) = -2 + 10x$, $[p]_B = \left(\begin{smallmatrix} 3 \\ -2 \\ 5 \end{smallmatrix}\right)$

$$\left(\begin{smallmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{smallmatrix}\right)\left(\begin{smallmatrix} 3 \\ -2 \\ 5 \end{smallmatrix}\right) = \left(\begin{smallmatrix} -2 \\ 10 \\ 0 \end{smallmatrix}\right)$$

which is the coordinate vector of the polynomial $-2 \cdot 1 + 10 \cdot x = -2 + 10x = p'(x)$

ker D is the set of all constant polynomials, i.e. $P_0(\mathbb{R})$

im $D = P_1(\mathbb{R})$

$nullspace[D]_{B,B}$: $\left(\begin{smallmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{smallmatrix}\right) \sim \left(\begin{smallmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{smallmatrix}\right)$ with columns a_0, a_1, a_2

$\implies a_1 = a_2 = 0$, a_0 free parameter.

$\implies nullspace(A) = \left\{ \left(\begin{smallmatrix} a_0 \\ 0 \\ 0 \end{smallmatrix}\right) : a_0 \in \mathbb{R} \right\} = span \left\{ \left(\begin{smallmatrix} 1 \\ 0 \\ 0 \end{smallmatrix}\right) \right\}$

$\left(\begin{smallmatrix} a_0 \\ 0 \\ 0 \end{smallmatrix}\right)$ is exactly the coordinate vectors of constant polynomials.

$$\begin{aligned} colspace([D]_{B',B}) &= span \left\{ \left(\begin{smallmatrix} 1 \\ 0 \\ 0 \end{smallmatrix}\right) \left(\begin{smallmatrix} 0 \\ 2 \\ 0 \end{smallmatrix}\right) \right\} = span \left\{ \left(\begin{smallmatrix} 1 \\ 0 \\ 0 \end{smallmatrix}\right) \left(\begin{smallmatrix} 0 \\ 1 \\ 0 \end{smallmatrix}\right) \right\} \\ &= \left\{ \left(\begin{smallmatrix} a_0 \\ a_1 \\ 0 \end{smallmatrix}\right) : a_0, a_1 \in \mathbb{R} \right\} \end{aligned}$$

which is precisely the set of all coordinate vectors of linear polynomials.

Example 2.5. $L : \mathbf{Mat}(2 \times 2, \mathbb{R}) \to \mathbf{Mat}(2 \times 2, \mathbb{R})$

$A \mapsto A - A^t$

L is linear. Now let:

Note that L is linear since:

$$\begin{aligned} L(A+B) &= (A+B) - (A+B)^t \\ &= A + B - A^t - B^t \\ &= (A - A^t) + (B - B^t) \\ &= L(A) + L(B) \\ L(kA) &= kA - (kA)^t = kA - kA^t \\ &= k(A - A^t) = kL(A) \end{aligned}$$

$\implies$ L is linear.

$$B = B' = \{\underbrace{\left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right)}_{M_{11}}, \underbrace{\left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)}_{M_{12}}, \underbrace{\left(\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right)}_{M_{21}}, \underbrace{\left(\begin{smallmatrix} 0 & 0 \\ 0 & 1 \end{smallmatrix}\right)}_{M_{22}}\}$$

$$\begin{aligned} L\left(\left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right)\right) &= \left(\begin{smallmatrix} 0 & 0 \\ 0 & 0 \end{smallmatrix}\right) = 0 \cdot M_{11} + 0 \cdot M_{12} + 0 \cdot M_{21} + 0 \cdot M_{22} \\ L\left(\left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)\right) &= \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right) = 0 \cdot M_{11} + 1 \cdot M_{12} + (-1) \cdot M_{21} + 0 \cdot M_{22} \\ L\left(\left(\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right)\right) &= \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right) = 0 \cdot M_{11} + (-1) \cdot M_{12} + 1 \cdot M_{21} + 0 \cdot M_{22} \\ L\left(\left(\begin{smallmatrix} 0 & 0 \\ 0 & 1 \end{smallmatrix}\right)\right) &= \left(\begin{smallmatrix} 0 & 0 \\ 0 & 0 \end{smallmatrix}\right) = 0 \cdot M_{11} + 0 \cdot M_{12} + 0 \cdot M_{21} + 0 \cdot M_{22} \\ \implies [L]_{B,B} &= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \end{aligned}$$

Let $A \equiv \left(\begin{smallmatrix} 2 & 1 \\ 3 & -1 \end{smallmatrix}\right)$, $L(A) = \left(\begin{smallmatrix} 2 & 1 \\ 3 & -1 \end{smallmatrix}\right) - \left(\begin{smallmatrix} 2 & 3 \\ 1 & -1 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 0 & -2 \\ 2 & 0 \end{smallmatrix}\right)$

$$[L(A)]_B = \begin{pmatrix} 0 \\ -2 \\ 2 \\ 0 \end{pmatrix}, \; [A]_B = \begin{pmatrix} 2 \\ 1 \\ 3 \\ -1 \end{pmatrix}$$

$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \\ 3 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ -2 \\ 2 \\ 0 \end{pmatrix} = [L(A)]_B$$

ker L: compute $nullspace([L]_{B,B})$:

$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 & -1 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \sim \begin{pmatrix} 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

with columns $a_{11}, a_{12}, a_{21}, a_{22}$.

$a_{12} - a_{21} = 0 \Leftrightarrow a_{21} = a_{12} \Leftrightarrow a_{11}, a_{12}, a_{22}$ free parameters.

$$\implies nullspace \ = span\left\{\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}\right\} = \left\{\begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{pmatrix} : a_{11}, a_{12}, a_{22} \in \mathbb{R}\right\}$$

which are the coordinate vectors of the matrices $\left\{ \begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix} : a_{11}, a_{12}, a_{22} \in \mathbb{R} \right\}$ which is exactly the set of all 2×2 symmetric matrices.
Check:

$$L(A) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = 0 = A - A^t$$
$$\Leftrightarrow A = A^t \Leftrightarrow A \text{ is } 2 \times 2 \text{ symmetric.}$$

$$im\ L: \ colspace([L]_{B,B}) = span\{\begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ 1 \\ 0 \end{pmatrix}\}$$
$$= span\{\begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \end{pmatrix}\} = \{\begin{pmatrix} 0 \\ a_{12} \\ -a_{12} \\ 0 \end{pmatrix} : a_{12} \in \mathbb{R}\}$$

These are the coordinate vectors of the matrices:

$$\{\begin{pmatrix} 0 & a_{12} \\ -a_{12} & 0 \end{pmatrix} : a_{12} \in \mathbb{R}\} \quad A = -A^t$$

This is the set of all 2×2 skew-symmetric matrices.

LECTURE 02/13

We continue the above example: Let $L : \mathbf{Mat}(2 \times 2), \mapsto \mathbf{Mat}(2 \times 2, \mathbb{R})$
$A \mapsto A - A^t$

$$B = \{\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\} \qquad [L]_{B,B} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$B' = \{ \underbrace{\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}}_{\mapsto \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}}, \underbrace{\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}}_{\mapsto \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}}, \underbrace{\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}}_{\mapsto \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}}, \underbrace{\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}}_{\mapsto \begin{pmatrix} 0 & 2 \\ -2 & 0 \end{pmatrix}} \}$$

$$[L]_{B',B} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}$$

$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} - \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 2 \\ -2 & 0 \end{pmatrix}$
$= 0 \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + 0 \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$
$+ 0 \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} + 2 \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$

This matrix is diagonal, and simpler than the previous representation (See last example). Shows advantage of choosing different bases.

Compositions of Linear maps

Let U, V, W be finite dimensional VS over K with bases $B, B', B''.L_1 : U_{B'} \mapsto V_{B'}, L_2 : V_{B'} \mapsto W_{B''}, L_2 \circ L_1 : U_B \mapsto W_{B''}$. How are these three linear maps related?

Theorem 2.13. *Let U, V, W be finite dimensional VS over K with bases B, B', B''. $L_1 : U_B \mapsto V_{B'}, L_2 : V_{B'} \mapsto W_{B''}$. Then $L_2 \circ L_1 : U_B \mapsto W_{B''}$. If $\dim U = m$, $\dim V = k$, $\dim W = n$, then:*

$$\underbrace{[L_2 \circ L_1]_{B'',B}}_{n\times m} = \underbrace{[L_2]_{B'',B'}}_{n\times k} \cdot \underbrace{[L_1]_{B',B}}_{k\times m}$$

Proof. Let $u \in U$ be arbitrary vector, then

$$\begin{aligned}[L_2 \circ L_1]_{B'',B}[u]_B &= [(L_2 \circ L_1)(u)]_{B''} \\ &= [L_2(L_1(u))]_{B''} \\ &= [L_2]_{B'',B'}[L_1(u)]_{B'} \\ &= [L_2]_{B'',B'}[L_1]_{B',B}[u]_B \forall u \in U.\end{aligned}$$

$$\implies [L_2 \circ L_1]_{B'',B} = [L_2]_{B'',B'} \cdot [L_1]_{B',B}$$ ∎

Matrix Representation of Isomorphisms

Theorem 2.14. *V, W, finite dimensional VS over K. $L : V \mapsto W$ linear, B, B' bases for V, W respectively. L isomorphism $\iff$ $[L]_{B',B}$ invertible.*

Remark 5. Consider id: $V \mapsto V$ the identity map, $\dim V = n, B$ basis for V, then $[id]_{B,B} = I_n$ ($n \times n$ identity matrix). If $B = (v_1, \cdots, v_n)$, then

$$\begin{aligned}[id]_{B,B} &= \big([id(v_1)]_B \mid \cdots \mid [id(v_n)]_B\big) \\ &= \Big(\underbrace{[v_1]_B}_{=1\cdot v_1 + 0\cdot v_2 + \cdots + 0\cdot v_n} \mid \cdots \mid [v_n]_B\Big) \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = I_n\end{aligned}$$

Proof.

("⇒") Let $L : V_B \mapsto W_{B'}$ be an isomorphism. $\implies$ L is bijective $\implies$ L has an inverse $L^{-1} : W_{B'} \mapsto V_B$ which is linear. We have $L^{-1} \circ L : V_B \mapsto V_B$, $L^{-1} \circ L = id_V$ then:

$$[L^{-1} \circ L]_{B,B} = I = [L^{-1}]_{B,B'}[L]_{B',B} = I \implies [L]_{B',B} \text{ invertible.}$$
$$\text{Alternatively, } [L]_{B',B} \cdot [L^{-1}]_{B,B'} = I \implies [L]_{B',B} \text{ invertible.}$$

("⇐") Let $[L]_{B',B}$ be invertible $n \times n$ matrix.

We first prove that L is injective. Let $v \in \ker L$:

$$\implies L(v) = 0 \implies [L(v)]_{B'} = 0 \; \underbrace{[L]_{B',B}}_{invertible} [v]_B = [L(v)]_{B'} = 0$$

$$\implies [v]_B = 0 \implies v = 0 \implies \ker L = \{0\} \implies L \text{ is injective.}$$

Next, we prove L is surjective.

$$\begin{aligned} \dim V = \dim W = n \\ \dim \ker L + \dim imL = \dim V = n \\ \implies \dim imL = n = \dim W \\ \implies imL = W \implies L \ \ surjective. \end{aligned}$$

$\implies L$ is bijective $\implies L$ is an isomorphism. $\square$

2.4 Change of basis

Q_1: V finite dimensional VS over K, B, B' bases for V , let $v \in V$.
Given $[v]_B$, how can we compute $[v]_{B'}$?
Consider $id : V_B \mapsto V_{B'}$. $[id]_{B',B}[v]_B = [id(v)]_{B'} = [v]_{B'}$

Definition 2.7. So given $id : V_B \mapsto V_{B'}$, we have $[v]_{B'} = [id]_{B',B}[v]_B$. The matrix $[id]_{B',B}$ is called the **transition matrix** or **change-of-basis matrix** from B to B'.

CAUTION: Schaum's mistakenly calls this the transition matrix from B' to B.

Q_2: $L : V \mapsto V$ linear, B, B' bases for V. Given $[L]_{B,B}$, how can we compute $[L]_{B',B'}$?
Consider: $V_{B'} \xrightarrow{id} V_B \xrightarrow{L} V_B \xrightarrow{id} V_{B'}$. Then $[id \circ L \circ id] = [id]_{B',B}[L]_{B,B}[id]_{B,B'} = [L]_{B',B'}$. Let $P \equiv [id]_{B,B'}$.

$$\begin{aligned} & id : V_B \mapsto V'_B \\ & \text{then } [id]_{B',B} = \left([id]_{B,B'}\right)^{-1} = P^{-1} \\ & \implies [L]_{B',B'} = P^{-1}[L]_{B,B}P \end{aligned}$$

where P is the change-of-basis matrix from B' to B.

Definition 2.8. Two $n \times n$ matrices A, B with coefficients $\in K$ are **similar** if there exists an invertible $n \times n$ matrix P with $B = P^{-1}AP$, thus, any two matrix representations of the same linear map $L : V \mapsto V$ with respect to bases B, B' are similar. Note that converse also holds.

Special case: $V = K^n$ with two bases: standard basis S and arbitrary basis B. If we consider the vectors in K^n as row vectors, $[v]_S = v^t$, a column vector.

How can we compute $[v]_B$? Invert matrix: $[v]_B = [id]_{B,S}[v]_S$

$[id]_{B,S}$ is the transition matrix from S to B.

If $B = (v_1, \cdots, v_n)$ then $[id]_{S,B} = (v_1^t \mid \cdots \mid v_n^t)$. And we have $[id]_{B,S} = ([id]_{S,B})^{-1} \implies [v]_B = P^{-1} \underbrace{[v]_S}_{=v^t}$, where $P = (v_1^t \mid \cdots \mid v_n^t)$.

e.g. (added) in $\mathbb{R}^2$, $B = \{\vec{v}_1, \vec{v}_2\}$.
$\vec{a} = c_1\vec{v}_1 + c_2\vec{v}_2$
$\vec{a} = \underbrace{\begin{pmatrix} | & | \\ \vec{v}_1{}^t & \vec{v}_2{}^t \\ | & | \end{pmatrix}}_{P} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}$
i.e. $[\vec{a}]_S = P[\vec{a}]_B$
i.e. $P = [id]_{S,B}$ change-of-basis matrix.

LECTURE 02/15

Examples 2.6.

1. $V = \mathbb{R}^2$, S standard basis, $B = ((1,1),(3,2))$

$$[id]_{S,B} = \left(\begin{smallmatrix} 1 & 3 \\ 1 & 2 \end{smallmatrix}\right)$$
$$[id]_{B,S} = ([id]_{S,B})^{-1} = \left(\begin{smallmatrix} 1 & 3 \\ 1 & 2 \end{smallmatrix}\right)^{-1} = \frac{1}{-1}\left(\begin{smallmatrix} 2 & -3 \\ -1 & 1 \end{smallmatrix}\right) = \left(\begin{smallmatrix} -2 & 3 \\ 1 & -1 \end{smallmatrix}\right)$$
$$[v]_B = [id]_{B,S}[v]_S$$

Check:
$v = (1,0)$, $[v]_S = \left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right)$, $[v]_B = \left(\begin{smallmatrix} -2 \\ 1 \end{smallmatrix}\right)$; $-2(1,1) + 1(3,2) = (1,0)$
$v = (2,3)$, $[v]_S = \left(\begin{smallmatrix} 2 \\ 3 \end{smallmatrix}\right)$, $[v]_B = \left(\begin{smallmatrix} -2 & 3 \\ 1 & -1 \end{smallmatrix}\right)\left(\begin{smallmatrix} 2 \\ 3 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 5 \\ -1 \end{smallmatrix}\right)$

2. Consider $L : \mathbb{R}^2 \Rightarrow \mathbb{R}^2$, projection onto x-axis.
i.e. $L((1,0)) = (1,0), L((0,1)) = (0,0) \Rightarrow [L]_{S,S} = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right)$

$$\begin{aligned}[L]_{B,B} &= [id]_{B,S}[L]_{S,S}[id]_{S,B} \\ &= ([id]_{S,B})^{-1}[L]_{S,S}[id]_{S,B} \\ &= \left(\begin{smallmatrix} -2 & 3 \\ 1 & -1 \end{smallmatrix}\right)\left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right)\left(\begin{smallmatrix} 1 & 3 \\ 1 & 2 \end{smallmatrix}\right) \qquad = \left(\begin{smallmatrix} -2 & 3 \\ 1 & -1 \end{smallmatrix}\right)\left(\begin{smallmatrix} 1 & 3 \\ 0 & 0 \end{smallmatrix}\right) = \left(\begin{smallmatrix} -2 & -6 \\ 1 & 3 \end{smallmatrix}\right)\end{aligned}$$

Check: $v = (2,3) \overset{L}{\mapsto} w = (2,0)$
$[v]_B = \left(\begin{smallmatrix} 5 \\ -1 \end{smallmatrix}\right)$, $[w]_B = \left(\begin{smallmatrix} -2 & 3 \\ 1 & -1 \end{smallmatrix}\right)\left(\begin{smallmatrix} 2 \\ 0 \end{smallmatrix}\right) = \left(\begin{smallmatrix} -4 \\ 2 \end{smallmatrix}\right)$
$\left(\begin{smallmatrix} -2 & -6 \\ 1 & 3 \end{smallmatrix}\right)\left(\begin{smallmatrix} 5 \\ -1 \end{smallmatrix}\right) = \left(\begin{smallmatrix} -4 \\ 2 \end{smallmatrix}\right)$

Examples 2.7. $V = \mathbf{Mat}(2 \times 2, \mathbb{R})$

See last class.

$$L : A \mapsto A - A^t$$

$$B = \left(\left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right)\left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)\left(\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right)\left(\begin{smallmatrix} 0 & 0 \\ 0 & 1 \end{smallmatrix}\right)\right)$$
$$B' = \left(\left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right)\left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)\left(\begin{smallmatrix} 0 & 0 \\ 0 & 1 \end{smallmatrix}\right)\left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)\right)$$
$$[L]_{B,B} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, [L]_{B',B'} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}$$
$$[id]_{B,B'} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \quad [id]_{B',B} = ([id]_{B,B'})^{-1}$$

$$\left(\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \end{array}\right) \sim \left(\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -2 & 0 & -1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \end{array}\right)$$
$$\sim \left(\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & \frac{1}{2} & -\frac{1}{2} & 1 \end{array}\right) \sim \left(\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & \frac{1}{2} & -\frac{1}{2} & 1 \end{array}\right)$$

$$\implies [id]_{B',B} = \frac{1}{2}\begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 2 \\ 0 & 1 & -1 & 0 \end{pmatrix}$$

Check: $A = \left(\begin{smallmatrix} 3 & 3 \\ 1 & -1 \end{smallmatrix}\right)$, $\quad [A]_B = \begin{pmatrix} 3 \\ 3 \\ 1 \\ -1 \end{pmatrix}$

$$[A]_{B'} = \frac{1}{2}\begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 2 \\ 0 & 1 & -1 & 0 \end{pmatrix}\begin{pmatrix} 3 \\ 3 \\ 1 \\ -1 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 3 \\ 2 \\ -1 \\ 1 \end{pmatrix}$$

Check:

$$3\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + 2\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} - \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} 3 & 3 \\ 1 & -1 \end{pmatrix} = A$$

The Normal Form Problem

Let $A \in \mathbf{Mat}(n \times m, K).L_A : K^n \to K^n$, $v \mapsto Av$. S standard basis for K^n, $B = \{v_1, \cdots, v_n\}$ an arbitrary basis.

Then $[L_A]_{S,S} = A$

$[L_A]_{B,B} = \underbrace{[id]_{B,S}}_{=P^{-1}} \underbrace{[L_A]_{S,S}}_{=A} \underbrace{[id]_{S,B}}_{\equiv P} = P^{-1}AP$ where $P = (v_1^t \mid \cdots \mid v_n^t)$

The converse works as well, i.e. given an arbitrary $P \in \mathbf{GL}(n, K)$, where $\mathbf{GL}(n, K)$ is the "general linear" set of all **invertible** $n \times n$ matrices with coefficients in K, we can interpret $P^{-1}AP$ as a matrix representation of L_A w.r.t. some basis B of K^n.

$P = (v_1^t \mid \cdots \mid v_n^t)$

Since P is invertible, its columns are LI $\implies \{v_1, \cdots v_n\}$ are LI.

Since $dim\ K^n = n$, $\{v_1, \cdots v_n\}$ also span K^n, i.e. $\{v_1, \cdots v_n\} \equiv B$ is a basis for K^n. And $[L_A]_{B,B} = P^{-1}AP$.

There is thus a one-to-one correspondence between matrix representations of L_A and matrices similar to A. Especially, similar matrices can be interpreted as representing the *same* linear map (w.r.t different bases).

We are thus interested in determining the set of all matrices in $\mathbf{Mat}(n \times n, K)$ similar to A ("Similarity Problem").

2.5 *Similarity*

Definition 2.9. Let $A, B \in \mathbf{Mat}(n \times n, K)$. B is said to be **similar** to A, in symbols $B \sim A$ if $\exists P \in \mathbf{GL}(n, K)$ s.t. $B = P^{-1}AP$.

Definition 2.10. **Equivalence Relation**: Let S be a set. A relation, i.e. a subset of the cartesian product $S \times S$, denoted by $\sim$, with the following properties:

1. $\forall x \in S:\ x \sim x$ REFLEXIVE
2. $\forall x, y \in S:\ x \sim y \iff y \sim x$ SYMMETRIC
3. $\forall x, y, z \in S:\ x \sim y \wedge y \sim z \implies x \sim z$ TRANSITIVE

Example 2.8. $S = \mathbb{Z}$, $n \sim k :\iff n - k$ is even.

CLAIM: $\sim$ is an equivalence relation:

- $n \sim n$ since $n - n = 0$ is even.
- $n \sim k \iff n - k$ even $\iff k - n$ is even $\iff k \sim m$
- $n \sim k \wedge k \sim m \implies n - m = \underbrace{n-k}_{\text{even}} + \underbrace{k-m}_{\text{even}}$ is even $\implies n \sim m$

$\{n \in \mathbb{Z} : n \sim 0\} = \{\cdots, -2, 0, 2, 4, \cdots\}$ set of all even numbers.
$\{n \in \mathbb{Z} : n \sim 1\} = \{\cdots, -3, -1, 1, 3, \cdots\}$ set of all odd numbers.

Definition 2.11. Let S be a set, $s \in S$, $\sim$ an equivalence relation on S. The **equivalence class** $[s] = [s]_\sim \equiv \{t \in S : t \sim s\}$.

Note that $s \in [s]_\sim$, since $s \sim s$.

Example 2.9. $S = \mathbb{Z}, \sim$: $n - k$ is even.
$[0]$ is the set of all even numbers.
$[1]$ is the set of all odd numbers.
Note that these equivalence classes partition S.
This holds in general:

Theorem 2.15. *Let S be a set, $\sim$ an equivalence relation. Then the equivalence classes of $\sim$ partition S, i.e. every $s \in S$ is contained in an equivalence class and any two equivalence classes are either identical or disjoint.*

Exercise 2.10. Show that similarity is an equivalence relation.

Remark 6.

$$[I_n] = \{\underbrace{P^{-1}I_nP}_{P^{-1}P=I_n} : P \in \mathbf{GL}(n,K)\} \\ = \{I_n\}$$

LECTURE 02/20

Proof. Since $s \in [s]$, thus the equivalence classes cover S. It remains to show that any two equivalence classes are either identical or disjoint. Let $s,\ t \in S$ such that $[s] \cap [t] \neq \emptyset$. Let $u \in [s] \cap [t]$.

$$\implies u \sim s \wedge u \sim t$$
$$\implies \text{by symmetry } s \sim u \wedge u \sim t \implies \text{by transitivity } s \sim t \wedge t \sim s$$

Let $v \in$ [s] be arbitrary $\implies v \sim s$ and $s \sim t \implies$ by transitivity $\implies v \sim t \implies v \in$ [t] $\implies$ [s] $\subseteq$ [t]. ☐

Theorem 2.16. *Similarity, i.e. $B \sim A \iff \exists P \in \mathbf{GL}(n,K) : B = P^{-1}AP$ is an equivalence relation.*

Proof.

- **Reflexivity**: $AI_n^{-1}AI_n \implies A \sim A$.
- **Symmetry**: Let $B \sim A$, $B = P^{-1}AP \implies PB = AP \implies A = PBP^{-1} = (P^{-1})^{-1}BP^{-1} \implies A \sim B$.
- **Transitivity**: Let $C \sim B \wedge B \sim A$. $C = Q^{-1}BQ, B = P^{-1}AP$
$\implies C = Q^{-1}(P^{-1}AP)Q = (Q^{-1}P^{-1})APQ = (PQ)^{-1}A(PQ)$
$\implies C \sim A$

Thus, similarity is an equivalence relation. Thus the similarity classes $[s]_{\sim} \equiv \{B \in \mathbf{Mat}(n \times n, K) : \exists P \in \mathbf{GL}(n, K) : B = P^{-1}AP\}$. Partition $\mathbf{Mat}(n \times n, K)$. All matrices in a given similarity class can be considered to represent the same linear map wrt different bases. □

Remark 7. Useful trick for showing that two given matrices are not similar:

1. Similar matrices have the same determinant: $B = P^{-1}AP$

   Proof.
   $$\det B = \det(P^{-1}AP) = \det P^{-1} \det A \underbrace{\det P}_{\neq 0} = \frac{1}{\cancel{\det P}} \det A \cancel{\det P}$$
   $$= \det A$$
   $\implies$ if $\det A \neq \det B$ then A, B can't be similar. □

2. Similar matrices have the same trace.

2.6 *Diagonalization: Eigenvalues and Eigenvectors*

Given a linear map $L : V \to V$, V finite dimensional VS over K, we can represent L by its matrix representation $[L]_B$, B basis for V.The set of all such representation forms a similarity class of matrices in $\mathbf{Mat}(n \times n, K)$. We want to find the *simplest* representative within this similarity class.

Definition 2.12. Invariant subspace: Let $L : V \to V$, V VS over K, L linear subspace. A subspace $U \leq V$ is called **invariant under** L or L-**invariant** if $L(U) = \{L(u) : u \in U\} \subseteq U$.

Assume V finite dimensional VS over K, $L : V \to V$, linear, and U, W L-invariant subspaces of V with $V = U \oplus W$. Let $\{u_1, \cdots, u_k\}$ be basis for U, and $\{w_1, \cdots, w_m\}$ be basis for W. So $B \equiv \{u_1, \cdots, u_k, w_1, \cdots, w_m\}$ is a basis for V. Then $[L]_B$ is of the form $\begin{pmatrix} \overbrace{\star}^{k} & 0 \\ 0 & \underbrace{\star}_{m} \end{pmatrix}$ since $L(u_j)$ is an LC of $u_1, \cdots, u_k$ $\forall 1 \leq j \leq k$ and $L(w_i)$ is LC of $w_1, \cdots, w_m$ $\forall 1 \leq i \leq m$.

Similarly, if $v = u_1 \oplus u_2 \oplus \cdots \oplus u_k$, where all u_j are L-invariant with bases $\{u_1^1, \cdots, u_m^1\}$ for $u_1, \cdots, \{u_1^k, \cdots, u_m^k\}$ for u_k, and $B \equiv \{u_1^1, \cdots, u_{m_k}^k\}$ basis for V, then $[L]_B$ is of the form: $\begin{pmatrix} \overbrace{\star}^{m_1} & \cdots & 0 \\ \vdots & \star & 0 \\ 0 & \cdots & \underbrace{\star}_{m_k} \end{pmatrix}$

$m_1 + \cdots + m_k = n = \dim V$ We are thus interested in one-dimenisonal L-invariant subspaces of V. Let $v \neq 0$ such that $span\{v\}$ is L-invariant, then $L(v) \in span\{v\} \implies \exists \lambda \in K : L(v) = \lambda v$.

Note that the representation is simplest if all of the L-invariant subspaces have dimension 1. Is it possible to achieve this? In general, no, but sometimes yes.

Definition 2.13. V finite dimensional VS over K, $L : V \to V$ linear. $v \in V, v \neq 0$ such that $L(v) = \lambda v$ for some $\lambda \in K$, then v is called an **eigenvector** of L to the **eigenvalue** λ.

Assume that V has a basis of eigenvectors $v_1, \cdots, v_n$- to eigenvalues $\lambda_1, \cdots, \lambda_n$. $B \equiv \{v_1, \cdots, v_n\}; L(v_j) = \lambda_j v_j = 0 \cdot v_1 + \cdots + \lambda_j v_j + \cdots + 0 \cdot v_n$, then $[L]_B = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$ In this case, the normal form problem is solved.

We can also define eigenvalues and eigenvectors for matrices in $\mathbf{Mat}(n \times n, K)$:

Definition 2.14. Let $A \in \mathbf{Mat}(n \times n, K), v \in K^n, v \neq 0$ such that $Av = \lambda v$ for some $\lambda \in K$, then v is called an eigenvector of A to the eigenvalue λ.

Let $L : V \to V, L(v) = \lambda v$, B basis for V. Then $[L]_B[v]_B = [L(v)]_B = [\lambda v_B] = \lambda[v]_B$.

$\implies$ λ eigenvalue of $[L]_B$, $[v]_B$ is corresponding eigenvector. $[L]_B$ passes through a similarity class of matrices. All similar matrices have thus the **same** eigenvalues. This can be shown directly: Let $A \in \mathbf{Mat}(n \times n, K), v \neq 0, \lambda \in K$ s.t $Av = \lambda v$.

$$\begin{aligned} B &= P^{-1}AP \\ (P^{-1}AP)(P^{-1}v) &= P^{-1}A(PP^{-1}v = P^{-1}Av = P^{-1}\lambda v = \lambda(P^{-1}v) \\ \implies \underbrace{(P^{-1}v)}_{\neq 0} &= \lambda(P^{-1}v) \implies \lambda \text{is an eigenvalue of } B. \end{aligned}$$

Some Properties of Eigenvectors and Eigenvalues

Theorem 2.17. *Let $L : V \Rightarrow V, \lambda$ eigenvalue of L. $E_\lambda(L) \equiv \{v \in V : L(v) = \lambda v\}$, then $E_\lambda \leq V$ which is called the* ***eigenspace*** *of the eigenvalue λ. For matrices: $E_\lambda(A) \equiv \{v \in K^n : Av = \lambda v\}$*

LECTURE 02/22

Remark 8. E_λ consists of all eigenvectors of L w.r.t. λ plus the zero vector.

Proof.

$0 \in E_\lambda(L)$ since $L(0) = 0 = \lambda \cdot 0$ ✓.

Closedness under $+$: Let $v_1, v_2 \in E_\lambda(L)$.

Then $L(v_1) = \lambda v_1$, $L(v_2) = \lambda v_2$

$\implies L(v_1 + v_2) = L(v_1) + L(v_2) = \lambda v_1 + \lambda v_2 = \lambda(v_1 + v_2)$

$\implies v_1 + v_2 \in E_\lambda(L)$ ✓.

Closedness under scalar $\cdot$: Let $k \in K$.
Then $L(kv_1) = kL(v_1) = k(\lambda v_1) = \lambda(kv_1)$
$\implies kv_1 \in E_\lambda(L)$ ✓.

Thus $E_\lambda(L) \leq V$. □

Theorem 2.18. *$E_\lambda(L)$ is L-invariant.*

Proof. Let $v \in E_\lambda(L)$. Then $L(v) = \lambda v \in E_\lambda(L) \implies E_\lambda(L)$ is L-invariant. □

Theorem 2.19. *Eigenvectors to distinct eigenvalues are linearly independent, i.e. if $L : V \to V$, $v_1, \cdots, v_k$ eigenvectors of L to the eigenvalues $\lambda_1, \cdots, \lambda_k$ which are pairwise distinct, then $\{v_1, \cdots, v_k\}$ are LI.*

Proof. by Induction.

$k = 1$: Let v_1 eigenvector of L to evaluate λ_1. Then $v_1 \neq 0$ by definition of eigenvectors. Thus $\{v_1\}$ is LI.

$k \Rightarrow k+1$: We assume that the statement holds for some value of k and need to prove it for $k+1$: Let $v_1, \cdots, v_k, v_{k+1}$ be eigenvectors of L to pairwise distinct eigenvalues $\lambda_1, \cdots, \lambda_{k+1}$. Let:

$$\begin{aligned}
a_1v_1 + \cdots + a_kv_k + a_{k+1}v_{k+1} &= 0 \\
L(a_1v_1 + \cdots + a_kv_k + a_{k+1}v_{k+1}) &= L(0) = 0 \\
\implies a_1L(v_1) + \cdots + a_kL(v_k) + a_{k+1}L(v_{k+1}) &= 0 \\
\implies a_1\lambda_1v_1 + \cdots + a_k\lambda_kv_k + a_{k+1}\lambda_{k+1}v_{k+1} &= 0 \\
\text{Also } \lambda_{k+1} * \text{line 1: } a_1\lambda_{k+1}v_1 + \cdots + a_k\lambda_{k+1}v_k + a_{k+1}\lambda_{k+1}v_{k+1} &= 0 \\
\text{subtracting } \implies a_1(\lambda_1 - \lambda_{k+1})v_1 + \cdots + a_k(\lambda_k - \lambda_{k+1})v_k &= 0
\end{aligned}$$

Since $v_1, \cdots, v_k$ is LI by inductive hypothesis, it follows that:

$$\begin{aligned}
a_1(\underbrace{\lambda_1 - \lambda_{k+1}}_{\neq 0}) = 0, \cdots a_k(\underbrace{\lambda_k - \lambda_{k+1}}_{\neq 0}) &= 0 \\
\implies a_1 = 0, \cdots, a_k &= 0 \\
\text{Since } \underbrace{a_1v_1 + \cdots a_kv_k}_{=0} + a_{k+1}v_{k+1} = 0 \implies a_{k+1}\underbrace{v_{k+1}}_{\neq 0} &= 0 \\
\implies a_{k+1} = 0 \implies a_1 = \cdots = a_k = a_{k+1} &= 0
\end{aligned}$$

$\implies \{v_1, \cdots, v_{k+1}\}$ is LI. □

Theorem 2.20. *Let V be a finite dimensional VS over K. $L : V \to V$. Then L is diagonalizable, i.e. there exists a basis B for L s.t. $[L]_B$ is a diagonalizable matrix, iff V has a basis of eigenvectors.*

Proof.

("⇒") Let L be diagonalizable, let B be a basis for V such that $[L]_B$ is a diagonal matrix $D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$, $B = (v_1, \cdots, v_n)$.

$$[L(v_1)]_B = D \cdot e_1 = \begin{pmatrix} \lambda \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \lambda_1 e_1 = [\lambda_1 v_1]_B \implies L(v_1) = \lambda_1 v_1$$

$$\vdots$$

$$[L(v_n)]_B = D \cdot e_n = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \lambda_n \end{pmatrix} = \lambda_n e_n = [\lambda_n v_n]_B \implies L(v_n) = \lambda_n v_n$$

$\implies v_1, \cdots v_n$ are eigenvectors of L.

("⇐") Let $\{v_1, \cdots, v_n\}$ be a basis of eigenvectors of L, say to eigenvalues $\lambda_1, \cdots, \lambda_n$. Let $B = \{v_1, \cdots, v_n\}$. Then:

$$\begin{aligned} L(v_1) &= \lambda_1 v_1 + 0 \cdot v_2 + \cdots + 0 \cdot v_n \\ &\vdots \\ L(v_n) &= 0 \cdot v_1 + \cdots + \lambda_n v_n \end{aligned} \implies [L]_B = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}$$

so $[L]_B$ is a diagonal matrix. □

Theorem 2.21. *V VS over K, $\dim V = n$, $L : V \to V$ linear. Assume that L has n pairwise distinct eigenvalues. Then L is diagonalizable.*

Proof. We need to prove (by Theorem 2.20) that V has a basis of eigenvectors. Let $v_1, \cdots, v_n$ be eigenvectors to the n pairwise distinct eigenvalues $\lambda_1, \cdots, \lambda_n$. By Theorem 2.19, $\{v_1, \cdots, v_n\}$ is LI and thus a basis for eigenvectors for V, $\implies L$ is diagonalizable. □

NOT every linear map $L : V \to V$ (V finite dimensional) is diagonalizable.

Example 2.11. The differentiation map $D : P_n(\mathbb{R}) \to P_n(\mathbb{R})$ is NOT diagonalizable $\forall n \geq 2$.

Proof. We have to find all eigenvalues and eigenvectors of D. Let $p(x) = a_0 + a_1 x + \cdots + a_n x^n$ be arbitrary $\neq 0$.

$$D(p(x)) = a_1 + 2a_2 x + 3a_3 x^2 + \cdots + n a_n x^{n-1}$$

$p(x)$ is an eigenvector of D iff $\exists\lambda \in \mathbb{R}$ st. $D(p(x)) = \lambda p(x)$, i.e.:

$$\begin{aligned} a_1 &= \lambda a_0 \\ 2a_2 &= \lambda a_1 \\ 3a_3 &= \lambda a_2 \\ &\vdots \\ na_n &= \lambda a_{n-1} \\ 0 &= \lambda a_n \end{aligned}$$

1. Case: $\lambda \neq 0$

$$\begin{aligned} a_n &= 0 \\ a_{n-1} &= 0 \\ &\vdots \\ a_0 &= 0 \end{aligned}$$

$\implies p = 0$ ↯

$\implies$ no $\lambda \neq 0$ is an eval of D

2. Case: $\lambda = 0$

$$\begin{aligned} a_n &= 0 \\ a_{n-1} &= 0 \\ &\vdots \\ a_1 &= 0 \end{aligned}$$

But a_0 can be arbitrary!

Thus $\lambda = 0$ is an eval of D with:

$E_0(D) = \{c_0\} = span\{1\}$

$dim\ E_0(D) = 1 \neq dim\ P_n(\mathbb{R}) = n + 1$

Thus we cannot find a basis of eigenvectors for $P_n(\mathbb{R})$ if $n > 1$

$\implies$ D is not diagonalizable for any $n > 1$.

□

Remark 9. While $0 \in V$ is not considered an eigenvector, $0 \in K$ can very well be an eigenvalue.

Theorem 2.22. *Let $L : V \to V$. Then 0 is an eigenvalue of L iff L is* NOT *injective. In that case, $E_0(L) = ker\ L$.*

Proof.

("$\Rightarrow$") Let 0 be an eval of L. Then $\exists v \neq 0$ with $L(v) = 0 \cdot v = 0 \implies$ L has non-trivial kernel, i.e. is non-injective.

("$\Leftarrow$") Let L be non-injective $\implies$ L has non-trivial kernel $\implies$ $\exists v \neq 0$ with $L(v) = 0 = 0 \cdot v \implies v$ evec to eval $0 \implies 0$ is eval.

And: $E_0(L) = \{v \in V : L(v) = 0 \cdot v = 0\} = ker\ L$. □

LECTURE 02/27

2.7 *The Characteristic Polynomial and The Minimal Polynomial*

Goal: Let $A \in \mathbf{Mat}(n \times n, K)$: find all eigenvalues (and eigenvectors) of A.

λ is an eigenvalue of $A \Leftrightarrow \exists x \in K^n, x \neq 0 : Ax = \lambda x \Leftrightarrow \lambda x - Ax = 0 \Leftrightarrow \lambda \cdot I_n x - Ax = 0 \Leftrightarrow \exists x \neq 0 : \underbrace{(\lambda I - A)}_{n \times n\ matrix} x = 0 \Leftrightarrow$ the linear system $(\lambda I - A)x = 0$ has a non-trivial solution, $\Leftrightarrow \lambda I - A$ is non-invertible $\leftrightarrow det(\lambda I - A) = 0$.

Definition 2.15. $det(\lambda I - A) = det \begin{pmatrix} \lambda - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a21 & \lambda - a_{22} & \cdots & -a2n \\ \vdots & \vdots & \ddots & \vdots \\ -a_{n1} & \cdots & & \lambda - a_{nn} \end{pmatrix}$

is a polynomial in λ of degree n and leading term λ^n, and is called the **characteristic polynomial** of A, denoted $\chi_A(\lambda)$.

Some coefficients of the characteristic polynomial:

- Constant Coefficient: Let $\chi_A(\lambda) = \lambda^n + c_{n-1}\lambda^{n-1} + \cdots + c_1\lambda_1 + c_0$
  $\Rightarrow \chi_A(0) = c_0 = det(-A) = (-1)^n detA$
  $\Rightarrow c_0 = (-1)^n detA$
- Coefficient of λ^{n-1}: the only term in the determinant that possibly contributes a λ^{n-1} is the product of the elements on the main diagonal.

$$\begin{aligned}\chi_A(\lambda) &= (\lambda - a_{11}) \cdot \ldots \cdot (\lambda - a_{nn}) + \text{terms of order at most } n-2 \text{ in } \lambda \\ &= \lambda^n + \lambda^{n-1}(-a_{11} - a_{22} - \cdots - a_{nn}) + \cdots \\ &= \lambda^n - trA \cdot \lambda^{n-1} + \text{ lower order terms}\end{aligned}$$

i.e. $c_{n-1} = -trA$.

$$detA = \sum_{\substack{\sigma \text{ perm of} \\ \{1,\cdots,n\}}} \overbrace{sgn(\sigma)}^{\pm 1} \cdot a_{1\sigma(1)} \cdot \ldots \cdot a_{n\sigma(n)}$$

Theorem 2.23. *Similar matrices have the same characteristic polynomial.*

Proof. Let $B \sim A$, $A, B \in \mathbf{Mat}(n \times n, K)$.
$B = P^{-1}AP$. Then

$$\begin{aligned}\chi_B(\lambda) &= det(\lambda I - P^{-1}AP) \\ &= det(\lambda P^{-1}P - P^{-1}AP) \\ &= det(P^{-1}(\lambda I)P - P^{-1}AP) \\ &= det(P^{-1}(\lambda I - A)P) \\ &= det(P^{-1})det(\lambda I - A)detP \\ &= \frac{1}{\cancel{detP}} \cdot \chi_A(\lambda) \cdot \cancel{detP} = \chi_A(\lambda)\end{aligned}$$

□

Remark 10. Since $(-1)^n detA$ and $-trA$ are coefficients of $\chi_A(\lambda)$. We thus have that similar matrices have the same determinant (we already knew that) and trace.
Especially, if $A, B \in \mathbf{Mat}(n \times n, K)$ with $trA \neq trB$ then A, B are NOT similar.

Proof. Number of eigenvalues of $A \in \mathbf{Mat}(n \times n, K)$: λ_0 eigenvalue of $A \Leftrightarrow \chi_A(\lambda_0) = 0$. But $\chi_A(\lambda)$ is a polynomial of degree n and thus has at most n roots. Thus A has AT MOST n distinct eigenvals. □

Proof. without using $\chi_A(\lambda)$: Let $\lambda_1, \cdots, \lambda_k$ be the distinct eigenvalues of A, and let $v_1, \cdots, v_k$ be corresponding eigenvecs. Eigenvectors to distinct eigenvalues are LI $\implies \{v_1, \cdots, v_k\}$ is LI $\implies k \leq n \implies A$ has at most n eigenvals. □

Definition 2.16. Let $L : V \to V$, V finite dimensional vector space, L linear. Let B be any basis for V. The characteristic polynomial $\chi_L(\lambda) \equiv \chi_A(\lambda)$ where $A = [L]_B$.

Recall that λ eigenvalue of A iff $\chi_A(\lambda) = 0$, i.e. the eigenvalues of A (or L) are just the roots of $\chi_A(\lambda)$. (We can use numerical methods to find roots, but that is actually not efficient, there exist much better algorithms to approximate eigenvalues and eigenvectors - MATH 327 MATRIX NUMERICAL ANALYSIS.)

Remark 11. Since matrix representations to different bases are similar, it follows from previous theorem (Theorem 2.23) that $\chi_L(\lambda)$ is well-defined, i.e. does NOT depend on B.

Example 2.12. Let $A = \begin{pmatrix} 1 & 2 & 4 \\ 1 & 0 & 2 \\ -1 & -1 & -3 \end{pmatrix}$. Find all eigenvalues, eigenvectors; find $P \in \mathbf{GL}(n, \mathbb{R})$ s.t. $P^{-1}AP$ is diagonal.

$$\begin{aligned}\chi_A(\lambda) &= det(\lambda I - A) = det\begin{pmatrix}\lambda-1 & -2 & -4\\ -1 & \lambda & -2\\ 1 & 1 & \lambda+3\end{pmatrix}\\ &= det\begin{pmatrix}\lambda+1 & -2 & -4\\ -\lambda-1 & \lambda & -2\\ 0 & 1 & \lambda+3\end{pmatrix} = (\lambda+1)det\begin{pmatrix}1 & -2 & -4\\ -1 & \lambda & -2\\ 0 & 1 & \lambda+3\end{pmatrix}\\ &= (\lambda+1)det\begin{pmatrix}1 & -2 & -4\\ 0 & \lambda-2 & -6\\ 0 & 1 & \lambda+3\end{pmatrix} = (\lambda+1)det\begin{pmatrix}\lambda-2 & -6\\ 1 & \lambda+3\end{pmatrix}\\ &= (\lambda+1)(\lambda^2+3\lambda-2\lambda-\cancel{6}+\cancel{6}) = (\lambda+1)(\lambda^2+\lambda)\\ &= \lambda(\lambda+1)^2\end{aligned}$$

$\implies$ the eigenvalues are $\lambda = 0, \lambda = -1$.

Eigenspaces: x eigenvector s.t. $Ax = \lambda x \iff (\lambda I - A)x = 0$. So, eigenvectors are all nontrivial solutions of this equation (since 0 cannot be an eigenvector).

Can use any elementary row or column operation for determinant computations: here we computed col 2 - col 1, then row 1 + row 2.

CASE 1: $\lambda = -1$:

$$E_{-1} : \begin{pmatrix}-2 & -2 & -4\\ -1 & -1 & -2\\ 1 & 1 & 2\end{pmatrix} \sim \begin{pmatrix}1 & 1 & 2\\ 0 & 0 & 0\\ 0 & 0 & 0\end{pmatrix} \implies \begin{cases}x_1 + x_2 + 2x_3 = 0\\ x_2, x_3 \text{ free parameters}\end{cases}$$

$x_2 = 0, x_3 = 1 : x_1 = -2, \begin{pmatrix}-2\\0\\1\end{pmatrix}$; $x_2 = 1, x_3 = 0 : x_1 = -1, \begin{pmatrix}-1\\1\\0\end{pmatrix}$

$\Rightarrow E_{-1}(A) = span\{\begin{pmatrix}-2\\0\\1\end{pmatrix}, \begin{pmatrix}-1\\1\\0\end{pmatrix}\}$

CASE 2: $\lambda = 0$:

$$E_0 : \begin{pmatrix}1 & 2 & 4\\ 1 & 0 & 2\\ -1 & -1 & -3\end{pmatrix} \sim \begin{pmatrix}1 & 2 & 4\\ 0 & -2 & -2\\ 0 & 1 & 1\end{pmatrix} \sim \begin{pmatrix}1 & 2 & 4\\ 0 & 1 & 1\\ 0 & 0 & 0\end{pmatrix} \implies \begin{cases}x_1 + 2x_2 + 4x_3 = 0\\ x_2 + x_3 = 0\\ x_3 \text{ free parameter}\end{cases}$$

$x_3 = 1 : x_2 = -1, x_1 - 2 + 4 = 0, x_1 = -2$

$\Rightarrow E_0(A) = span\{\begin{pmatrix}-2\\-1\\1\end{pmatrix}\}$ $[= nullspace(A)]$

So $P \equiv \begin{pmatrix}-2 & -2 & -1\\ -1 & 0 & 1\\ 1 & 1 & 0\end{pmatrix}$ then $P^{-1}AP = \begin{pmatrix}0 & 0 & 0\\ 0 & -1 & 0\\ 0 & 0 & -1\end{pmatrix}$. [17]

[17] where the first column: eigenvector to $\lambda = 0$ and the second two columns are the eigenvectors to $\lambda = -1$.

Q: Does $A \in \mathbf{Mat}(n \times n, K)$ always have at least one eigenvalue?

1. NO if $K = \mathbb{R}$

2. YES if $K = \mathbb{C}$.

1. If $K = \mathbb{R}$, $\exists$ linear maps $L : V \to V$, V VS over $\mathbb{R}$ without eigenvalues (this occurs when both roots of the characteristic polynomial are complex). $A = \begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix}$, $L : x \mapsto Ax, L : \mathbb{R}^2 \to \mathbb{R}^2$. Since A is a rotation by $\frac{\pi}{2}$, we can already see that a rotation of $\vec{v}$ will never be a multiple of itself unless $\vec{v} = \vec{0} \neq$ eigenvalue. We define the characteristic polynomial:

$$\chi_A = \det(\lambda I - A) = \det\left(\begin{smallmatrix} \lambda & 1 \\ -1 & \lambda \end{smallmatrix}\right) = \lambda^2 + 1$$

This has no real roots, and hence $L : \mathbb{R}^2 \to \mathbb{R}^2$ has no eigenvalues in the reals. We note that if we take L as a map from $\mathbb{C}^2$ to $\mathbb{C}^2$, $x \mapsto Ax$, then for $\chi_A = \lambda^2 + 1$, we have roots $\lambda = \pm i$ and thus two eigenvalues $\implies$ A is diagonalizable over $\mathbb{C}$:

$$A = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right) \sim \left(\begin{smallmatrix} i & 0 \\ 0 & -i \end{smallmatrix}\right)$$

To find eigenvectors, we set $\lambda = i$:

$$(\lambda I - A) = \left(\begin{smallmatrix} i & 1 \\ -1 & i \end{smallmatrix}\right) = \left(\begin{smallmatrix} 1 & -i \\ i & 1 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 1 & -i \\ 0 & 0 \end{smallmatrix}\right) \qquad \implies \left(\begin{smallmatrix} i \\ 1 \end{smallmatrix}\right).$$

We then set $\lambda = -i$.

$$(\lambda I - A) = \begin{pmatrix} -i & 1 \\ -1 & -i \end{pmatrix} = \left(\begin{smallmatrix} 1 & i \\ -i & 1 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 1 & i \\ 0 & 0 \end{smallmatrix}\right) \qquad \implies \left(\begin{smallmatrix} -i \\ 1 \end{smallmatrix}\right)$$
$$\implies P \equiv \left(\begin{smallmatrix} i & -i \\ 1 & 1 \end{smallmatrix}\right), P^{-1}AP = \left(\begin{smallmatrix} i & 0 \\ 0 & -i \end{smallmatrix}\right)$$

Proof. 2. λ_0 eigenval of $A \Leftrightarrow \chi_A(\lambda_0) = 0$, i.e. A has at least one eigenval iff $\chi_A(\lambda)$ has at least one root. $\chi_A(\lambda)$ is a polynomial of degree n. By the Fundamental Theorem of Algebra, every polynomial of degree n factors over $\mathbb{C}$ into a product of n linear factors, i.e. every polynomial of degree n has n roots over $\mathbb{C}$ (counted with multiplicity) and thus has at least one root. $\square$

LECTURE 03/01

We can find eigenvalues and eigenvectors without the use of the characteristic polynomial.

Definition 2.17. Matrix Polynomials: let $p(x) \equiv a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$, $A \in Mat(n \times n, K)$, then $A^n \in \mathbf{Mat}(n \times n, K)$ and $\underbrace{p(A)}_{n\times n} = a_0I + a_1A + \cdots + a_nA^n$.

ALGORITHM FOR FINDING EIGENVALUE OF $A \in Mat(n \times n, \mathbb{C})$:

- Let $x_0 \in \mathbb{C}^n$ be arbitrary $\neq 0$.
- Consider $x_0, A^2x_0, A^3x_0, \cdots$ at most n of those can be LI. Let k be the smallest index s.t $\{x_0, Ax_0, \cdots, A^kx_0\}$ is LD.
- So if we let $a_0, \cdots, a_k \in \mathbb{C}$, not all 0 s.t $a_0x_0 + a_1Ax_0 + \cdots + a_kA^kx_0 = 0 \implies (a_0I + a_1A + \cdots + a_kA^k)x_0 = 0$. This is a homogeneous linear system with nontrivial solution since $x_0 \neq 0 \implies B \equiv a_0I + a_1A + \cdots + a_kA^k$ is not invertible.
- Let $p(x) \equiv a_0 + a_1x + \cdots + a_kx^k \implies B = p(A)$. By the Fundamental Theorem of Algebra, we know that p factors fully over $\mathbb{C} : p(x) = a_k(x - c_1) \cdot \cdots \cdot (x - c_k)$.

- Since the product of invertible matrices implies that each matrix involved is individually invertible, we have that $p(A) = a_k(A - c_1 I) \cdot \ldots \cdot (A - c_k I)$ is not invertible, and thus at least ONE of $(A - c_j I)$ is not invertible, say $A - c_j$, then $\implies \exists x \in \mathbb{C}^n, x \neq 0$ s.t $(A - c_j I)x = 0 \iff Ax - c_j Ix = Ax - c_j x = 0 \iff Ax = c_j x$ is an eigenvalue of A.

Caution: note that NOT ALL of the c_i's are necessarily eigenvalues, but AT LEAST ONE is.

Example 2.13. Let $A = \begin{pmatrix} 1 & 2 & 4 \\ 1 & 0 & 2 \\ -1 & -1 & -3 \end{pmatrix}$. We choose the simplest column as Ax_0, so $x_0 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$ (x_0 is just a column from Id. We can choose it to be anything, but want it as simple as possible).

We then have that

$$x_0 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \qquad Ax_0 = \begin{pmatrix} 2 \\ 0 \\ -1 \end{pmatrix} \qquad A^2 x_0 = A(Ax_0) = \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix}$$
$$\implies \{x_0, Ax_0, A^2x_0\} LD$$
$$\implies Ax_0 + A^2 x_0 = 0, (A + A^2)x_0 = 0$$
$$\implies p(\lambda) = \lambda + \lambda^2 = \lambda(\lambda + 1)$$
$$\implies \lambda = 0, \lambda = -1$$

We obtain the exact same result if we take $x_0 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$

2.8 Normal Forms for Non-diagonalizable Matrices

Definition 2.18. A matrix $A \in Mat(n \times n, K)$ of the form $\begin{pmatrix} \lambda_1 & \cdots & \star \\ & \ddots & \\ 0 & \cdots & \lambda_n \end{pmatrix}$ is **upper triangular**. Similarly, a transpose of an upper triangular matrix is lower triangular.

Theorem 2.24. *For upper triangular $A \in Mat(n \times n, \mathbb{C})$, the diagonal elements $\lambda_1, \cdots, \lambda_n$ are the eigenvalues of A.*

Proof. $\chi_A(\lambda) = \det(\lambda I - A) = \det \begin{pmatrix} \lambda - \lambda_1 & & \star \\ & \ddots & \\ 0 & & \lambda - \lambda_n \end{pmatrix}$. Since the determinant of a triangular matrix is the product of its diagonal entries, we have that $\chi_A(\lambda = (\lambda - \lambda_1) \cdot \cdots \cdot (\lambda - \lambda_n)$

$\implies \lambda_1, \cdots, \lambda_n$ are the eigenvalues of A. The same holds for lower triangular matrices. □

Theorem 2.25. *Every matrix $A \in Mat(n \times n, \mathbb{C})$, is similar to an upper triangular matrix.*

Proof. We proceed by recursive induction on n.

$n = 1$: Every 1×1 matrix is upper triangular.

$n-1 \to n$: (Inductive step) Assume the statement holds for $n-1$. Let $A \in Mat(n \times n, \mathbb{C})$. A has at least one eigenvalue $\lambda_1 \in \mathbb{C}$. Let v_1 be corresponding eigenvector. We extend to a basis $B \equiv (v_1, v_2, \cdots v_n)$ of $\mathbb{C}^n$.

$P \equiv (v_1|v_2|\cdots|v_n)$, then $P^{-1}AP \equiv B = \left(\begin{array}{c|ccc} \lambda_1 & \star & \cdots & \star \\ \hline 0 & & & \\ \vdots & & C & \\ 0 & & & \end{array}\right)$

$A \sim B$, show $B \sim$ upper triangular matrix $T \implies A \sim T$ by transitivity. By inductive hypothesis, C is an upper triangular matrix i.e $\exists \tilde{Q} \in \mathbf{GL}(n-1, \mathbb{C}), \tilde{Q}^{-1}C\tilde{Q} = \begin{pmatrix} \lambda_2 & \cdots & \star \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$.

Let $Q \equiv \left(\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & \tilde{Q} & \\ 0 & & & \end{array}\right)$. We then extend Q for $n \times n \in$

$\mathbf{Mat}(n \times n, \mathbb{C}) \implies Q$ is invertible with $Q^{-1} = \left(\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & \tilde{Q}^{-1} & \\ 0 & & & \end{array}\right)$

Since $QQ^{-1} = \left(\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & \tilde{Q} & \\ 0 & & & \end{array}\right) \left(\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & \tilde{Q}^{-1} & \\ 0 & & & \end{array}\right) =$

$\left(\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & \tilde{Q} \cdot \tilde{Q}^{-1} & \\ 0 & & & \end{array}\right) = \left(\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & I_n & \\ 0 & & & \end{array}\right)$

$\implies Q^{-1} = \left(\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & \tilde{Q}^{-1} & \\ 0 & & & \end{array}\right)$

$\implies Q^{-1}BQ = \left(\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & \tilde{Q}^{-1} & \\ 0 & & & \end{array}\right) \left(\begin{array}{c|ccc} \lambda_1 & \star & \cdots & \star \\ \hline 0 & & & \\ \vdots & & C & \\ 0 & & & \end{array}\right) \left(\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & \tilde{Q} & \\ 0 & & & \end{array}\right) =$

$\left(\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & \tilde{Q}^{-1} & \\ 0 & & & \end{array}\right) \left(\begin{array}{c|ccc} \lambda_1 & \star & \cdots & \star \\ \hline 0 & & & \\ \vdots & & C\tilde{Q} & \\ 0 & & & \end{array}\right) = \left(\begin{array}{c|ccc} \lambda_1 & \star & \cdots & \star \\ \hline 0 & & & \\ \vdots & & \tilde{Q}^{-1}C\tilde{Q} & \\ 0 & & & \end{array}\right) =$

$$\left(\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & \lambda_2 & & \star \\ \vdots & & \ddots & \\ 0 & 0 & & \lambda_n \end{array}\right) \text{ which is an upper triangular matrix.} \qquad \square$$

2.9 *Triangularization*

LECTURE 03/13

Given $A \in Mat(n \times n, \mathbb{C})$, we want to find an upper triangular matrix $T \sim A$.

Summary of algorithm from last class:

1. Find one eigenvalue λ_1 and corresponding eigenvector v_1 of A and **extend** v_1 to a basis $v_1, \cdots, v_n$ of $\mathbb{C}^n$:

   $P \equiv (v_1|v_2|\cdots|v_n)$, then $P^{-1}AP = \left(\begin{array}{c|c} \lambda_1 & u^t \\ \hline 0 & \\ \vdots & B \\ 0 & \end{array}\right)$ where $u \in \mathbb{C}^{n-1}, B \in Mat(n-1 \times n-1, \mathbb{C})$.

2. Apply step 1. to B.

   $\vdots$

   $n-1^{th}$ step: we have an upper triangular matrix $T \sim A$.

Example 2.14. Let $A = \begin{pmatrix} 3 & -3 & -3 \\ 1 & 0 & -2 \\ -1 & 2 & 4 \end{pmatrix}$. We arbitrarily choose x_0 and apply A until we obtain LD vectors:

$$x_0 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \xrightarrow{A} \begin{pmatrix} 3 \\ 1 \\ -1 \end{pmatrix} \xrightarrow{A} \begin{pmatrix} 9 \\ 5 \\ -5 \end{pmatrix} = 5 \underbrace{\begin{pmatrix} 3 \\ 1 \\ -1 \end{pmatrix}}_{Ax_0} - 6 \underbrace{\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}}_{x_0}$$

The last vector is obviously an LC of the two previous ones.

$\implies A^2 x_0 = 5Ax_0 - 6x_0$

$\implies (A^2 - 5A + 6I)x_0 = 0$

$\implies (\lambda - 2)(\lambda - 3) = 0$

$\lambda 3$: $\quad 3I - A = \begin{pmatrix} 0 & 3 & 3 \\ -1 & 3 & 2 \\ 1 & -2 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}$

Indeed, not invertible since $\det(3I - A) = 0$, so $\lambda = 3$ is an eigenvalue. We then solve the system of equations and obtain $E_3 = span\begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}$

Caution: Using this method, AT LEAST ONE OF THE VALUES FOUND IS AN EIGENVALUE. (i.e. not both as in the χ_λ method.) Moreover, there is no guarantee that these are the **only** eigenvalues. To check which λ is an eigenvalue, row-reduce $\lambda I - A$: Non-invertible means λ is an eigenvalue.

We now extend v_1 to a basis of $\mathbb{C}^3$:

Let $v_1 = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}, v_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, v_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$, and $P_1 \equiv \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}$, where $P_1^{-1}AP_1$ is the matrix representation wrt basis (v_1, v_2, v_3).

$v_1 \xrightarrow{A} 3 \cdot v_1 + 0 \cdot v_2 + 0 \cdot v_3$

$v_2 \xrightarrow{A} -3 \cdot v_1 + 3 \cdot v_2 - 1 \cdot v_3$

$v_3 \xrightarrow{A} -3 \cdot v_1 + 1 \cdot v_2 + 1 \cdot v_3$

$\implies B = P_1^{-1}AP = \begin{pmatrix} 3 & -3 & -3 \\ 0 & 3 & 1 \\ 0 & -1 & 1 \end{pmatrix}$, $C \equiv \begin{pmatrix} 3 & 1 \\ -1 & 1 \end{pmatrix}$. We repeat the same process for C.

We calculate the characteristic polynomial and obtain that $\lambda = 2$ is the only eigenvalue, thus $E_2 = span\{w_1\}, w_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$. We extend the basis, so let $w_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ s.t. w_1, w_2 span $\mathbb{C}^2$.

$\tilde{P_2} \equiv \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}$, $\tilde{C} \equiv \tilde{P_2}^{-1}C\tilde{P_2} = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$, since we have that

$w_1 \xrightarrow{C} 2 \cdot w_1 + 0 \cdot w_2$

$w_2 \xrightarrow{C} 1 \cdot w_1 + 2 \cdot w_2$.

We get $P_2 = \left(\begin{array}{c|c} 1 & \\ \hline & \tilde{P_2} \end{array}\right) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix}$, where $P_2^{-1}BP_2$ will be upper triangular.

Next, we look at two methods for computing $T = P_2^{-1}BP_2$.

1. We interpret $P_2^{-1}BP_2$ as a change of coordinates from standard basis to basis consisting of columns of P_2, namely p_1, p_2, p_3 respectively.

   $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \xrightarrow{B} 3 \cdot p_1 + 0 \cdot p_2 + 0 \cdot p_3$

   $\begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} \xrightarrow{B} 0 \cdot p_1 + 2 \cdot p_2 + 0 \cdot p_3$

   $\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \xrightarrow{B} -3 \cdot p_1 + 1 \cdot p_2 + 2 \cdot p_3$

   $\implies T = \begin{pmatrix} 3 & 0 & -3 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$.

   Also, we can see that $P_2^{-1}BP_2 = P_2^{-1}P_1^{-1}AP_1P_2 = (P_1P_2)^{-1}A(P_1P_2) = P^{-1}AP$, $P = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ -1 & -1 & 1 \end{pmatrix}$.

2. See that $P_2^{-1}BP_2 = \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & \tilde{P_2}^{-1} \end{array}\right) \left(\begin{array}{c|c} \lambda_1 & u^t \cdot \tilde{P_2} \\ \hline 0 & C \end{array}\right) \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & \tilde{P_2} \end{array}\right) =$

   $\left(\begin{array}{c|c} \lambda_1 & u^t \cdot \tilde{P_2} \\ \hline 0 & \tilde{P_2}^{-1}C\tilde{P_2} \end{array}\right) = \left(\begin{array}{c|c} \lambda_1 & u^t \cdot \tilde{P_2} \\ \hline 0 & \tilde{C} \end{array}\right)$

   So $T = \left(\begin{array}{c|cc} 3 & 0 & -3 \\ \hline 0 & 2 & 1 \\ 0 & 0 & 2 \end{array}\right)$.

3. Another way to solve is to find all eigenvalues. In this example, $\lambda = 2, \lambda = 3$ are the only eigenvalues (so we have 1-dimensional eigenspaces):$E_3 = span\begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}, E_2 = span\begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$. We extend the basis and take $v_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$.

   $v_1 \xrightarrow{A} 3 \cdot v_1 + 0 \cdot v_2 + 0 \cdot v_3$

   $v_2 \xrightarrow{A} 0 \cdot v_1 + 2 \cdot v_2 + 0 \cdot v_3$

   $v_3 \xrightarrow{A} -3 \cdot v_1 + 1 \cdot v_2 + 2 \cdot v_3$

   $P \equiv \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ -1 & -1 & 1 \end{pmatrix}$

   $P^{-1}AP = \begin{pmatrix} 3 & 0 & -3 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$

To prove A (defined in examples) is not diagonalizable, we find all eigenvalues and find 2 eigenvalues for a 3×3 matrix, thus not diagonalizable.

2.10 Jordan Canonical Form

Theorem 2.26. *Any linear operator T acting unto vector space V separate V into two distinct subspaces W_1 and W_2 such that T acts as an isomorphism unto W_1 and as a nilpotent operator unto W_2: $V = W_1 \bigoplus W_2$*

The goal of the Jordan Canonical Form is to find a basis such that $[J] = [D] + [N]$ where $[J]$ are the Jordan blocks, $[D]$ is a diagonal matrix of eigenvalues and $[N]$ is the standard upper triangular nilpotent matrix. and finally $[T] = \begin{pmatrix} [J]_1 & 0 & ... \\ 0 & [J]_2 & ... \\ ... & ... & ... \end{pmatrix}$,

$[J]_n = \begin{pmatrix} \lambda_n & 1 & 1 & ... \\ 0 & \lambda_n & 1 & ... \\ 0 & 0 & \lambda_n & ... \\ ... & ... & ... & ... \end{pmatrix}$.

Definition 2.19. The **Minimal Polynomial**($m(t)$)

- is factored into the same irreducible factors as the characteristic polynomial($\Delta(t)$), let

  $\Delta(t) = (t - \lambda_1)^{n_1}(t - \lambda_2)^{n_2}...$ then $m(t) = (t - \lambda_1)^{m_1}(t - \lambda_2)^{m_2}...$ such that $m_i \leq n_i$

- $m(t)$ contains the smallest m_i such that

  $m([T]) = ([T] - \lambda_1[Id])^{m_1}([T] - \lambda_2[Id])^{m_2}... = 0$

How to find JCF:

1. Find Minimal Polynomial and eigenvalues
2. The number of distinct eigenvalues is the number of Jordan Blocks
3. Let $U_n = [T] - \lambda_n[Id]$, to find the number of the each n sized-blocks associated use formula [number of block size n] $= 2Null(U^n) - (Null(U^{n-1}) + Null(U^{n+1}))$

3 Applications

3.1 Systems of 1st Order Linear ODEs with Constant Coefficients

LECTURE 03/15

Intro to Matrix Exponentiation

Reminder: The MacLaurin series of $e^x = 1 + x + \frac{x^2}{2} + \cdots = \sum_{k=0}^{\infty} \frac{x^k}{k!}$
Now let $X \in \mathbf{Mat}(n \times n, K)$

$$\implies X^k = \underbrace{X \cdot \cdots \cdot X}_{k \text{ times}} \in \mathbf{Mat}(n \times n, K), \text{ and } \frac{1}{k!}X^k \in \mathbf{Mat}(n \times n, K)$$

also, $\implies \sum_{k=0}^{N} \frac{1}{k!}X^k \in \mathbf{Mat}(n \times n, K)$.
Fact: $\lim_{N\to\infty} \frac{1}{k!}X^k$ exists, and we define the following:

Definition 3.1. $\exp(X) \equiv \sum_{k=0}^{\infty} \frac{1}{k!}X^k$, the **matrix exponential** of X.

Special case: X is a diagonal matrix $X = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$

$$\implies X^k = \begin{pmatrix} \lambda_1^k & & \\ & \ddots & \\ & & \lambda_n^k \end{pmatrix}$$

$$\implies \exp(X) = \sum_{k=0}^{\infty} \frac{1}{k!} \begin{pmatrix} \lambda_1^k & & \\ & \ddots & \\ & & \lambda_n^k \end{pmatrix} = \sum_{k=0}^{\infty} \begin{pmatrix} \frac{1}{k!}\lambda_1^k & & \\ & \ddots & \\ & & \frac{1}{k!}\lambda_n^k \end{pmatrix}$$

$$= \begin{pmatrix} \sum_{k=0}^{\infty} \frac{1}{k!}\lambda_1^k & & \\ & \ddots & \\ & & \sum_{k=0}^{\infty} \frac{1}{k!}\lambda_n^k \end{pmatrix}$$

$$= \begin{pmatrix} e^{\lambda_1} & & \\ & \ddots & \\ & & e^{\lambda_n} \end{pmatrix}$$

Caution: this formula only holds in this particular case. In general,

$$\exp(X) = \begin{pmatrix} e^{a_{11}} & \dots & e^{a_{1n}} \\ & \ddots & \\ e^{a_{n1}} & \dots & e^{a_{nn}} \end{pmatrix}$$

Matrix Exponentiation and Similarity

Let $X \in \mathbf{Mat}(n \times n, K), P \in \mathbf{GL}(n, K)$:

$$P^{-1}\exp(X)P = P^{-1}\left(\sum_{k=0}^{\infty} \frac{1}{k!}X^k\right)P = \sum_{k=0}^{\infty} P^{-1}\left(\frac{1}{k!}X^k\right)P = \sum_{k=0}^{\infty} \frac{1}{k!}\underbrace{(P^{-1}(X^k)P)}_{Aside}$$

$$= \sum_{k=0}^{\infty} \frac{1}{k!}(P^{-1}XP)^k = \exp(P^{-1}XP)$$

Aside:

$$\begin{aligned} P^{-1}(X^2)P &= P^{-1}(X \cdot X)P \\ &= P^{-1}X(PP^{-1})XP \\ &= (P^{-1}XP)(P^{-1}XP) \\ &= (P^{-1}XP)^2 \end{aligned}$$

Applying this idea recursively, we get $P^{-1}(X^k)P = (P^{-1}XP)^k$.

Systems of Linear ODEs:

One-dimensional case:
Let $x = x(t)$. Find the solutions of $\dot{x} = ax,\ a \in \mathbb{R}$ ($\dot{x} = x' = \frac{d}{dt}x(t)$).

One solution is $x(t) = e^{at}$. (Check: $\dot{x} = e^{at} \cdot a = ax$ ✓)
General solution: Let $x = x(t)$ be any solution of $\dot{x} = ax$.
Consider $\frac{x}{e^{at}} = xe^{-at}$. Then:

$$\frac{d}{dt}(xe^{-at}) = \underbrace{\dot{x}}_{ax} e^{-at} + x \cdot e^{-at} \cdot (-a) = axe^{-at} - axe^{-at} = 0$$

Thus, xe^{-at} is a constant, say $xe^{-at} = C \implies x = C \cdot e^{at},\ C \in \mathbb{R}$.
Especially, $x(0) = Ce^0 = C$, i.e. C is the IC of x at $t = 0$.

$$\implies \text{the general solution of } \begin{cases} \dot{x} = ax \\ x(0) = x_0 \in \mathbb{R} \end{cases} \text{ is } x(t) = x_0 \cdot e^{at}.$$

Multivariable case:

Example 3.1. n=2 $\begin{cases} \dot{x_1} &= x_1 - x_2 \\ \dot{x_2} &= 6x_1 - 4x_2 \end{cases}, \quad \begin{matrix} x_0^1 = 1 \\ x_0^2 = 1 \end{matrix}$ i.e. $X_0 = \binom{1}{1}$

$$\begin{pmatrix} \dot{x_1} \\ \dot{x_2} \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ 6 & -4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix};\ \dot{X} = AX;\ \text{IC } X(0) = X_0 \in K^n$$

PROBLEM: the equations are "entangled" and thus difficult to solve. We'll use linear algebra to "disentangle" them. Consider the differential equation of the form $\dot{X} = AX$, where $A \in \mathbf{Mat}(n \times n, K)$ and $X = X(t) = \begin{pmatrix} x_1(t) \\ \cdots \\ x_n(t) \end{pmatrix}$ differentiable with initial condition $X(0) = \begin{pmatrix} x_1(0) \\ \cdots \\ x_n(0) \end{pmatrix} = \begin{pmatrix} x_0^1 \\ \cdots \\ x_0^n \end{pmatrix} = X_0 \in K^n$.

Assume for now that A is diagonalizable, i.e. $D = P^{-1}AP,\ P \in \mathbf{GL}(n, K),\ D = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$.

$B \in \mathbf{Mat}(n \times n, K)$
$B\dot{x} = (\dot{Bx})$

$$\begin{aligned} &\implies PDP^{-1} = A \implies \dot{X} = AX = PDP^{-1}X \\ &\implies (\dot{X} = AX = PDP^{-1}X) \cdot P^{-1} \\ &\implies P^{-1}\dot{X} = DP^{-1}X \\ &\implies (\dot{P^{-1}X}) = DP^{-1}X \end{aligned}$$

Let $Y \equiv P^{-1}X$, then we have $\dot{Y} = DY$.
Let $Y_0 \equiv P^{-1}X_0$

This is a change of coordinates from standard basis to $(u_1, \ldots, u_n)$ where $P = (v_1 \mid \cdots \mid v_n)$

$$\begin{pmatrix} \dot{y_1} \\ \vdots \\ \dot{y_n} \end{pmatrix} = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix},\ Y(0) = Y_0 = \begin{pmatrix} Y_0^1 \\ \vdots \\ Y_0^n \end{pmatrix}$$

$$\begin{aligned} \Leftrightarrow \dot{Y_1} &= \lambda_1 Y_1,\ Y_1(0) = Y_0^1 \implies Y_1(t) = Y_0^1 \cdot e^{\lambda_1 t} \\ &\vdots \\ \dot{Y_n} &= \lambda_n Y_n,\ Y_n(0) = Y_0^n \implies Y_n(t) = Y_0^n \cdot e^{\lambda_n t} \end{aligned}$$

$$\begin{aligned}
\Rightarrow Y &= \begin{pmatrix} e^{\lambda_1 t} & & \\ & \ddots & \\ & & e^{\lambda_n t} \end{pmatrix} = \exp \begin{pmatrix} \lambda_1 t & & \\ & \ddots & \\ & & \lambda_n t \end{pmatrix} \cdot Y_0 \\
&= \exp(tD) \cdot Y_0 \\
\text{and } X &= X_0 \cdot e^{at} = \exp(at) X_0 \\
\implies Y &= P^{-1}X,\ X = PY;\ Y_0 = P^{-1}X_0 \\
\implies X &= PY = P\exp(tD)P^{-1}X_0 = \exp(P(tD)P^{-1})X_0 \\
&= \exp(t\underbrace{PDP^{-1}}_{=A})X_0 \\
&= \exp(tA)X_0
\end{aligned}$$

So for the n-dimensional case we have $X = \exp(tA)X_0$, and for the 1-dimensional case $x = x_0e^{at} = \exp(ta)x_0$, i.e. the same formula!

Example 3.2. (back to previous example):

$$\begin{pmatrix} \dot{x_1} \\ \dot{x_2} \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ 6 & -4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix};\ x_0 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} x_1(0) \\ x_2(0) \end{pmatrix}$$

$A = \begin{pmatrix} 1 & -1 \\ 6 & -4 \end{pmatrix}$ has eigenvalues -2 and -1: $E_{-2} = span\{\begin{pmatrix} 1 \\ 3 \end{pmatrix}\},\ E_{-1} = span\{\begin{pmatrix} 1 \\ 2 \end{pmatrix}\}$

$$\begin{aligned}
\implies P &= \begin{pmatrix} 1 & 1 \\ 3 & 2 \end{pmatrix},\ P^{-1} = -\begin{pmatrix} 2 & -1 \\ -3 & 1 \end{pmatrix} = \begin{pmatrix} -2 & 1 \\ 3 & -1 \end{pmatrix} \\
D &= \begin{pmatrix} -2 & 0 \\ 0 & -1 \end{pmatrix},\ Y = P^{-1}x,\ Y_0 = P^{-1}x_0 = \begin{pmatrix} -2 & 1 \\ 3 & -1 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 2 \end{pmatrix} \\
\dot{Y} &= DY,\ \begin{cases} \dot{Y}_1 = -2Y_1,\ Y_1(0) = -1 \\ \dot{Y}_2 = -Y_2,\ Y_2(0) = 2 \end{cases}
\end{aligned}$$

$$\implies Y_1(t) = e^{-2t},\ Y_2(t) = 2e^{-t}$$

$$\exp(tD) = \begin{pmatrix} e^{-2t} & \\ & e^{-t} \end{pmatrix},\ X(t) = P\begin{pmatrix} e^{-2t} & \\ & e^{-t} \end{pmatrix}P^{-1}X_0$$

$$\implies X(t) = \begin{pmatrix} 1 & 1 \\ 3 & 2 \end{pmatrix}\begin{pmatrix} e^{-2t} & \\ & e^{-t} \end{pmatrix}\begin{pmatrix} -1 \\ 2 \end{pmatrix} = \begin{pmatrix} -e^{-2t}+2e^{-t} \\ -3e^{-2t}+4e^{-t} \end{pmatrix} = \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix}$$

Check:

$$\begin{aligned}
x_1(0) &= -1+2 = 1 \checkmark \\
x_2(0) &= -3+4 = 1 \checkmark \\
\dot{x_1} &= 2e^{-2t} - 2e^{-t} \\
x_1 - x_2 &= -e^{-2t} + 2e^{-t} + 3e^{-2t} - 4e^{-t} \\
&= 2e^{-2t} - 2e^{-t}
\end{aligned}$$

Now, let us consider cases when A is not diagonalizable. If A is not diagonalizable, we can at least triangularize it since $A \sim T$ for T upper-triangular.

We thus have $A = PTP^{-1} \iff T = P^{-1}AT$
$\dot{X} = PTP^{-1}X,\ X(0) = X_0$
$P^{-1}\dot{X} = TP^{-1}X,\ Y \equiv P^{-1}X,\ \dot{Y} = TY,\ Y(0) = Y_0 = P^{-1}X_0.$

$$\begin{aligned}
\dot{y_1} &= a_{11}y_1 + a_{12}y_2 + \cdots + a_{1n}y_n \\
\dot{y_2} &= \qquad\quad a_{22}y_2 + \cdots + a_{2n}y_n \\
&\vdots \\
\dot{y}_{n-1} &= \qquad a_{(n-1)(n-1)}y_{n-1} + a_{(n-1)n}y_n \\
\dot{y_n} &= \qquad\qquad\qquad a_{nn}y_n
\end{aligned}$$

LECTURE 03/20

Recall that previously, we looked at examples where A is diagonalizable, i.e.
$\dot{x} = Ax,\ x(0) = x_0$
$A = PDP^{-1},\ \dot{x} = PDP^{-1}x$
$y \equiv P^{-1}x,\ y_0 \equiv P^{-1}x_0$
$\implies P^{-1}\dot{x} = DP^{-1}x,\ \dot{y} = Dy, y(0) = y_0$

$$\implies \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} e^{\lambda_1 t} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & e^{\lambda_n t} \end{pmatrix}$$

$$\exp(tD) = \begin{pmatrix} e^{\lambda_1 t} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & e^{\lambda_n t} \end{pmatrix}$$

$$\begin{aligned}
x &= P\begin{pmatrix} e^{\lambda_1 t} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & e^{\lambda_n t} \end{pmatrix}P^{-1}x_0 \\
&= \exp(tA)x_0
\end{aligned}$$

This implies that all solutions are linear combinations of $e^{\lambda_1 t}, \cdots, e^{\lambda_n t}$, where $\lambda_1, \cdots, \lambda_n$ are the eigenvalues of A.

Thus, we can solve for y_n directly and then solve from bottom to top.

Example 3.3. $\dot{Y} = \begin{pmatrix} 2 & 1 \\ & 2 \end{pmatrix} Y,\ Y(0) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$

$\dot{y_1} = 2y_1 + y_2,\ y_1(0) = 1$

$\dot{y_2} = \quad 2y_2,\ y_2(0) = 1 \implies y_2 = e^{2t}$

$\dot{y_1} = 2y_1 + e^{2t}$ We solve this nonhomogeneous linear DE:

Homogeneous solution for $\dot{y_1} - 2y_1 = 0$ is Ce^{2t}.

Particular solution of the above DE is $y_1 = (At+B)e^{2t}$. We next solve for initial conditions:

$y_1(0) = 1 \implies B = 1$

$\dot{y_1} - 2y_1 = e^{2t} \implies A = 1$

$y_1 = (t+1)e^{2t}$

We notice that this solution is not a linear combination of exponentials as in the diagonalizable case, so the structure of solutions is very different.

We note that we can also solve this using matrix exponentials:

$$\begin{aligned}
\dot{Y} &= TY,\ Y(0) = Y_0 = P^{-1}X_0 \\
\implies Y &= \exp(tT)Y_0 \qquad (\star) \\
P^{-1}X &= \exp(tT)P^{-1}X_0 \\
X &= P\exp(tT)P^{-1}X_0 = \exp(P(tT)P^{-1})X_0 = \exp(tPTP^{-1})X_0 \\
X(t) &= \exp(tA)X_0
\end{aligned}$$

Proof. of $(\star)$ in the special case where $T = \begin{pmatrix} a & b \\ 0 & a \end{pmatrix}, a, b \in \mathbb{R}$. We show that $\exp(tY)Y_0$ solves $\dot{Y} = TY, Y(0) = Y_0$, where $\exp(X) = \sum_{n=0}^{\infty} \frac{1}{n!} X^n$.

We begin by deriving a formula for $T^n : T^n = \begin{pmatrix} a^n & na^{n-1}b \\ 0 & a^n \end{pmatrix}$. [18]

With this, we can now show:

$$\begin{aligned}
\exp(tT) &= \sum_{n=0}^{\infty} \frac{1}{n!} t^n T^n = \sum_{n=0}^{\infty} \frac{t^n}{n!} \begin{pmatrix} a^n & na^{n-1}b \\ 0 & a^n \end{pmatrix} = \begin{pmatrix} \sum_{n=0}^{\infty} \frac{t^n}{n!} a^n & \sum_{n=0}^{\infty} \frac{t^n}{n!} na^{n-1}b \\ 0 & \sum_{n=0}^{\infty} \frac{t^n}{n!} a^n \end{pmatrix} \\
&= \begin{pmatrix} \sum_{n=0}^{\infty} \frac{ta^n}{n!} & tb \sum_{n=1}^{\infty} \frac{t^{n-1}}{(n-1)!} a^{n-1} \\ 0 & \sum_{n=0}^{\infty} \frac{ta^n}{n!} \end{pmatrix} = \begin{pmatrix} e^{ta} & tbe^{ta} \\ 0 & e^{ta} \end{pmatrix} \\
Y(t) &= \exp(T) \cdot Y_0 = \begin{pmatrix} e^{ta} & tbe^{ta} \\ 0 & e^{ta} \end{pmatrix} \begin{pmatrix} y_{01} \\ y_{02} \end{pmatrix}
\end{aligned}$$

□

[18] *Proof by induction:*
$n = 1 \implies T^1 = \begin{pmatrix} a & b \\ 0 & a \end{pmatrix}$
Inductive step: assume this is true for all n, now show $n \to n+1$:

$$\begin{aligned}
T^{n+1} &= T^n T = \begin{pmatrix} a^n & na^{n-1}b \\ 0 & a^n \end{pmatrix} \begin{pmatrix} a & b \\ 0 & a \end{pmatrix} \\
&= \begin{pmatrix} a^{n+1} & a^n b + na^n b \\ 0 & a^{n+1} \end{pmatrix} = \begin{pmatrix} a^{n+1} & a^n(n+1)a^n b \\ 0 & a^{n+1} \end{pmatrix}
\end{aligned}$$

Example 3.4. $\dot{Y} = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix} Y,\ Y(0) = Y_0 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, so $a = 2,\ b = 1$:

$Y(t) = \begin{pmatrix} (1+t)e^{2t} \\ e^{2t} \end{pmatrix}$

3.2 *Difference Equations with Constant Coefficient*

Difference equations are a discrete version of ODEs.

First order: Let a_n be a sequence recursively defined for given a_0 by $a_{n+1} = c \cdot a_n,\ \forall n \in \mathbb{N}$. Finding a general formula for a_n is straightforward:

$$a_1 = c \cdot a_0 \qquad a_2 = c \cdot a_1 = c^2 \cdot a_0 \qquad \cdots \qquad a_n = c^n a_0$$

Second order: $a_{n+2} = c \cdot a_{n+1} + d \cdot a_n, \ \forall n \in \mathbb{N}$, for $a_0, \ a_1$ known initial conditions. To solve for a second order difference equation, we create a dummy variable and diagonalize:

Example 3.5. Fibonacci sequence: $a_{n+2} = a_{n+1} + a_n, \ a_0 = 0, \ a_1 = 1.$

$$a_{n+2} = c \cdot a_{n+1} + d \cdot a_n$$
$$a_{n+1} = a_{n+1}$$
$$\begin{pmatrix} a_{n+2} \\ a_{n+1} \end{pmatrix} = \begin{pmatrix} c & d \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_{n+1} \\ a_n \end{pmatrix} \implies X_{n+1} = AX_n, \ X_0 = \begin{pmatrix} a_1 \\ a_0 \end{pmatrix}$$
$$X_1 = Ax_0$$
$$X_2 = AX_1 = A^2X_0$$
$$\dots$$
$$X_n = A^nX_0$$

We assume A is diagonalizable: $D = P^{-1}AP, \ A = PDP^{-1}$
$A^n = PD^nP^{-1}$ (from matrix exponentials)
$\implies X_n = PD^nP^{-1}X_0$

LECTURE 03/22

Example 3.6. $a_{n+2} = 4a_{n+1} - 3a_n, \ a_0 = 0, \ a_1 = 1$

$$X_n = A^nX_0, \ A = \begin{pmatrix} 4 & -3 \\ 1 & 0 \end{pmatrix}$$
$$\chi_A = det \begin{pmatrix} \lambda-4 & 3 \\ -1 & \lambda \end{pmatrix} = \lambda^2 - 4\lambda + 3 = (\lambda-3)(\lambda-1)$$
$$E_3 : \begin{pmatrix} -1 & 3 \\ -1 & 3 \end{pmatrix} \sim \begin{pmatrix} -1 & 3 \\ 0 & 0 \end{pmatrix} \implies E_3 = span\{\begin{pmatrix} 3 \\ 1 \end{pmatrix}\}$$
$$E_1 : \begin{pmatrix} -3 & 3 \\ -1 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix} \implies E_1 = span\{\begin{pmatrix} 1 \\ 1 \end{pmatrix}\}$$
$$\implies P = \begin{pmatrix} 3 & 1 \\ 1 & 1 \end{pmatrix} \implies P^{-1} = \frac{1}{2}\begin{pmatrix} 1 & -1 \\ -1 & 3 \end{pmatrix}$$
$$\implies X_n = \frac{1}{2}\begin{pmatrix} 3 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 3^n & \\ & 1 \end{pmatrix}\begin{pmatrix} 1 & -1 \\ -1 & 3 \end{pmatrix} X_0$$
$$\begin{pmatrix} a_{n+1} \\ a_n \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 3 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 3^n & -3^n \\ -1 & 3 \end{pmatrix} X_0$$
$$= \frac{1}{2}\begin{pmatrix} 3^{n+1}-1 & -3^{n+1}+3 \\ 3^n-1 & -3^n+3 \end{pmatrix}\begin{pmatrix} a_1 \\ a_0 \end{pmatrix}$$
$$\implies a_n = \tfrac{1}{2}[(3^n-1)a_1 + (-3^n+3)a_0] = 3^n\tfrac{a_1-a_0}{2} + \tfrac{-a_1+3a_0}{2}$$

1. if $a_0 = 0, \ a_1 = 1 : \ a_n = \frac{3^n-1}{2}$
2. if $a_0 = 1, \ a_1 = 3 : \ a_n = 3^n + 0 = 3^n$
3. if $a_0 = 1, \ a_1 = 1 : \ a_n = 1 \ \forall n \in \mathbb{N}$ constant equation.
   So the solutions "escape to infinity" iff $a_1 \neq a_0$. If $a_1 = a_0$, then the sequence is constant.

In general, if A is diagonalizable with eigenvalues $\lambda_1, \ \lambda_2$, the solutions are LCs of λ_1^n and λ_2^n, i.e. LCs of exponential functions. NOW WHAT IF A IS NOT DIAGONALIZABLE?

Example 3.7. $a_{n+2} = 2a_{n+1} - a_n, \ a_0 = 0, \ a_1 = 1$

$X_{n+1} = \begin{pmatrix} 2 & -1 \\ 1 & 0 \end{pmatrix} = AX_n$ where $A = \begin{pmatrix} 2 & -1 \\ 1 & 0 \end{pmatrix}$

$$\chi_A = det\begin{pmatrix} \lambda-2 & 1 \\ -1 & \lambda \end{pmatrix} = \lambda^2 - 2\lambda + 1 = (\lambda-1)^2 \implies \lambda = 1 \text{ only eigenvalue.}$$
$$E_1 : \begin{pmatrix} -1 & 1 \\ -1 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix} \implies E_1 = span\{\begin{pmatrix} 1 \\ 1 \end{pmatrix}\}$$

$\implies$ A is not diagonalizable, so we triangulate A. Extend $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ by e.g. $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ to a basis for $\mathbb{C}^2$.

Recall $\begin{pmatrix} a & b \\ 0 & a \end{pmatrix}^n = \begin{pmatrix} a^n & na^{n-1}b \\ 0 & a^n \end{pmatrix}$

$$P = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix},\ P^{-1} = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix},\ T = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} \implies T^n = \begin{pmatrix} 1 & -n \\ 0 & 1 \end{pmatrix}$$
$$\implies X_n = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & -n \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} X_0 = \begin{pmatrix} 1+n & -n \\ n & 1-n \end{pmatrix} \begin{pmatrix} a_1 \\ a_0 \end{pmatrix}$$
$$\implies a_n = n \cdot a_1 + (1-n)a_0,\ a_0 = 0,\ a_1 = 1 \implies a_n = n$$

4 Orthogonality and Inner Product Spaces

4.1 Definitions

Definition 4.1. Standard Inner Product on $\mathbb{R}^n$ (also called dot product or standard scalar product). Given $u = \begin{pmatrix} u_1 \\ \cdots \\ u_n \end{pmatrix}$ and $v = \begin{pmatrix} v_1 \\ \cdots \\ v_n \end{pmatrix}$, $\langle u, v\rangle \equiv u_1v_1 + \cdots + u_nv_n \equiv (u_1 \cdots u_n)\begin{pmatrix} v_1 \\ \cdots \\ v_n \end{pmatrix} = u^t \cdot v$

Theorem 4.1. Properties:
*$\langle , \rangle$ is **bilinear**, i.e. linear in each component.*

- $\langle u + w, v\rangle = \langle u, v\rangle + \langle w, v\rangle, \quad \langle k, uv\rangle = k\langle u, v\rangle$
- $\langle u, v + w\rangle = \langle u, v\rangle + \langle u, w\rangle, \quad \langle u, kv\rangle = k\langle u, v\rangle$

$\langle u, u\rangle = u_1^2 + \cdots + u_n^2 = \|u\|^2 = (u_1 \cdots u_n)\begin{pmatrix} u_1 \\ \cdots \\ u_n \end{pmatrix} = u^t \cdot u$
$\implies \|u\| = \sqrt{u_1^2 + \cdots + u_n^2}$
The angle α between vectors u and v is $cos\alpha = \frac{\langle u,v\rangle}{\|u\|\|v\|}$
$\alpha = \frac{\pi}{2} \Leftrightarrow cos\alpha = 0 \Leftrightarrow \langle u, v\rangle = 0$

Definition 4.2. $u, v \in \mathbb{R}^n$ are **orthogonal** iff $\langle u, v\rangle = 0$.

Note that 0 is orthogonal to any $v \in \mathbb{R}^n$.

NOW TO DEFINE THE STANDARD INNER PRODUCT IN $\mathbb{C}^n$:

$\mathbb{C} = \mathbb{C}^1$ $z = x + iy$, $|z| = \sqrt{x^2 + y^2}$ where $x = \mathbf{Re}(z)$, $y = \mathbb{I}m(z)$.
We define $\overline{z} = \overline{x + iy} = x - iy$
$\implies z\overline{z} = (x + iy)(x - iy) = x^2 + y^2 = |z|^2$

$\mathbb{C}^n$ Let $z = \begin{pmatrix} z_1 \\ \cdots \\ z_n \end{pmatrix} = \begin{pmatrix} x_1+iy_1 \\ \cdots \\ x_n+iy_n \end{pmatrix} \in \mathbb{C}^n$
We obtain $\|z\|$ by identifying $\mathbb{C}^2$ with $\mathbb{R}^{2n}$, i.e. identify $z \in \mathbb{C}^n$ with $\begin{pmatrix} x_1 \\ y_1 \\ \cdots \\ x_n \\ y_n \end{pmatrix} \in \mathbb{R}^{2n}$
$\implies \|z\| = \sqrt{x_1^2 + y_1^2 + \cdots + x_n^2 + y_n^2} = \sqrt{\overline{z_1}z_1 + \cdots \overline{z_n}z_n}$
$\implies \|z\|^2 = \overline{z_1}z_1 + \cdots \overline{z_n}z_n = (\overline{z_1} \cdots \overline{z_n})\begin{pmatrix} z_1 \\ \cdots \\ z_n \end{pmatrix} = \overline{z}^t z = z^\star z$

Definition 4.3. Standard Inner Product on $\mathbb{C}^n$. Given $z = \begin{pmatrix} z_1 \\ \cdots \\ z_n \end{pmatrix}$ and $w = \begin{pmatrix} w_1 \\ \cdots \\ w_n \end{pmatrix}$, we define
$\langle z, w\rangle = z^\star w = (\overline{z_1} \cdots \overline{z_n})\begin{pmatrix} w_1 \\ \cdots \\ w_n \end{pmatrix} = \overline{z_1}w_1 + \cdots + \overline{z_n}w_n$

Properties: $\|z\|^2 = \langle z, z\rangle$
Linearity:

2nd argument:
$$\begin{aligned}\langle z, w+v\rangle &= z^\star(w+v) = (\overline{z_1}\cdots\overline{z_n})\begin{pmatrix} w_1+v_1 \\ \cdots \\ w_n+v_n\end{pmatrix}\\ &= \overline{z_1}(w_1+v_1)+\ldots+\overline{z_n}(w_n+v_n)\\ &= (\overline{z_1}w_1+\ldots+\overline{z_n}w_n)+(\overline{z_1}v_1+\ldots+\overline{z_n}z_n)\\ &= \langle z,w\rangle+\langle z,v\rangle\\ \langle z, kw\rangle &= z^\star(kw) = \overline{z_1}(k_1w_1)+\ldots+\overline{z_n}(kw_n)\\ &= k(\overline{z_1}w_1+\ldots+\overline{z_n}w_n) = k\langle z,w\rangle\end{aligned}$$
$\implies \langle , \rangle$ linear in 2nd argument.

1st argument:
$$\begin{aligned}\langle z+v,w\rangle &= (\overline{z_1+v_1})w_1+\ldots+(\overline{z_n+v_n})w_n\\ &= (\overline{z_1}+\overline{v_1})w_1+\ldots+(\overline{z_n}+\overline{v_n})w_n\\ &= (\overline{z_1}w_1+\ldots+\overline{z_n}w_n)+(\overline{v_1}w_1+\ldots+\overline{v_n}w_n)\\ &= \langle z,w\rangle+\langle v,w\rangle\\ \langle kz,w\rangle &= (\overline{kz_1})w_1+\ldots+(\overline{kz_n})w_n\\ &= \overline{k}\overline{z_1}w_1+\ldots+\overline{k}\overline{z_n}w_n = \overline{k}\langle z_n,w_n\rangle\end{aligned}$$

So summary:
$$\langle z,v+w\rangle = \langle z,w\rangle+\langle v,w\rangle, \quad \langle z,kw\rangle = k\langle z,w\rangle$$
$$\langle z+v,w\rangle = \langle z,w\rangle+\langle v,w\rangle, \quad \langle kz,w\rangle = \overline{k}\langle z,w\rangle$$

Definition 4.4. Two vectors $z, w \in \mathbb{C}^n$ are **orthogonal** iff $\langle z,w\rangle = 0$.

Why is orthogonality important in Linear Algebra? For example, V VS over $\mathbb{R}$, $\{v_1,\cdots,v_n\}$ basis, $v \in V$. How do we find $[v]_B$? We need to find $a_1,\cdots,a_n \in K$ s.t. $v = a_1v_1+\cdots+a_nv_n$. So we must find $P = (v_1 \mid \cdots \mid v_n) \implies P\begin{pmatrix} a_1\\ \cdots\\ a_n\end{pmatrix} = a_1v_1+\cdots+a_nv_n = v \implies \begin{pmatrix} a_1\\ \cdots\\ a_n\end{pmatrix} = P^{-1}v$ which is computationally costly.

LECTURE 03/27

Let $v_1,\cdots,v_n \in K^n$ be pairwise orthogonal and LI, $v \in span\{v_1,\cdots,v_n\}$. Then $\exists\, a_1,\cdots,a_n \in K$ s.t. $v = a_1v_1+\cdots+a_nv_n$.
In the general case, finding the a_j is computationally non-trivial.

Example 4.1. Let $P = (v_1 \mid \cdots \mid v_n)$
$P\begin{pmatrix} a_1\\ \cdots\\ a_n\end{pmatrix} = a_1v_1+\cdots+a_nv_n = v \implies \begin{pmatrix} a_1\\ \cdots\\ a_n\end{pmatrix} = P^{-1}v$
i.e. we either have to solve the inhomogeneous linear system $P\begin{pmatrix} a_1\\ \cdots\\ a_n\end{pmatrix} = v$, or compute P^{-1}.

HOWEVER, if $v_1,\cdots,v_n$ are pairwise orthogonal, there is a much faster way:
$$v = a_1v_1+\cdots+a_nv_n$$

RECAP FROM LAST LECTURE:
Standard Inner Product:
On $\mathbb{R}^n$
$\langle u,v\rangle \equiv u^tv = u_1v_1+\cdots+u_nv_n$
On $\mathbb{C}^n$
$\langle u,v\rangle \equiv u^\star v = \overline{u_1}v_1+\cdots+\overline{u_n}v_n$
Norm:
$\|v\| = \sqrt{\langle v,v\rangle}$
Caution: the standard inner product on $\mathbb{C}^n$ is NOT bilinear, it is only sesqui-linear: $\langle ku,v\rangle = \overline{k}\langle u,v\rangle$ while $\langle u,kv\rangle = k\langle u,v\rangle$.

Now taking $\langle v_j, \cdot\rangle$,

$$\begin{aligned}\langle v_j, v\rangle &= \langle v_j, a_1v_1 + \cdots + a_jv_j + \cdots + a_nv_n\rangle \\ &= \langle v_j, a_1v_1\rangle + \cdots + \langle v_j, a_jv_j\rangle + \cdots + \langle v_j, a_nv_n\rangle \\ &= a_1 \underbrace{\langle v_j, v_1\rangle}_{=0} + \cdots + a_j\langle v_j, v_j\rangle + \cdots + a_n \underbrace{\langle v_j, v_n\rangle}_{=0} \\ &= a_j\langle v_j, v_j\rangle \\ \implies a_j &= \frac{\langle v_j, v\rangle}{\langle v_j, v_j\rangle}\end{aligned}$$

i.e. we have the following:

Theorem 4.2. *If we have a vector $v = a_1v_1 + \cdots + a_nv_n$ where $v_1, \cdots, v_n$ are pairwise orthogonal $\implies a_j = \frac{\langle v_j, v\rangle}{\langle v_j, v_j\rangle}$ $\forall i \le j \le n$.*

Definition 4.5. Unit vector. A vector $v \in K^n$ is called a **unit vector** if $\|v\| = \sqrt{\langle v, v\rangle} = 1$

Theorem 4.3. *We can turn a given non-zero vector v into a unit vector by defining the following: $\frac{v}{\|v\|} \equiv \frac{1}{\|v\|}v$*

Proof. In general, $\langle kv, kv\rangle = k\langle kv, v\rangle = k\bar{k}\langle v, v\rangle$ which works in both $\mathbb{R}$ and $\mathbb{C}$ since $\bar{k} = k$ $\forall k \in \mathbb{R}$.

$$\langle kv, kv\rangle = |k|^2\langle v, v\rangle$$

$$\text{Now consider} \frac{1}{\|v\|}v :$$

$$\langle \frac{1}{\|v\|}v, \frac{1}{\|v\|}v\rangle = \frac{1}{\|v\|^2}\langle v, v\rangle = \frac{1}{\cancel{\|v\|^2}}\cancel{\|v\|^2} = 1$$

$$\implies \|\frac{1}{\|v\|}v\| = 1 \quad \text{i.e. } \frac{1}{\|v\|}v \text{ is a unit vector.}$$ □

Definition 4.6. $\{v_1, \cdots, v_k\} \in K^n$ are called

- **orthogonal** if $v_1, \cdots, v_k$ are pairwise orthogonal
- **orthonormal** if $v_1, \cdots, v_k$ are pairwise orthogonal and all v_j are unit vectors.

Example 4.2. $S \equiv \{e_1, \cdots, e_n\}$ is an orthonormal basis for K^n.

NOW LET $\{v_1, \cdots, v_k\}$ be orthonormal and $v \in span\{v_1, \cdots, v_k\}$, $v = a_1v_1 + \cdots + a_nv_n$. Then $a_j = \frac{\langle v_j, v\rangle}{\langle v_j, v_j\rangle} = \frac{\langle v_j, v}{\|v_j\|^2} = \langle v_j, v\rangle$.

i.e. we have the following:

Theorem 4.4. *If $\{v_1, \cdots, v_k\}$ orthonormal, $v = a_1v_1 + \cdots + a_nv_n$, then $a_j = \langle v_j, v\rangle$.*

Definition 4.7. A matrix $A \in \mathbf{Mat}(n \times n, \mathbb{R})$ is called **orthogonal** if the columns of A are **orthonormal**.

Definition 4.8. A matrix $A \in \mathbf{Mat}(n \times n, \mathbb{C})$ is called **unitary** if the columns of A are **orthonormal** with respect to the standard inner product on $\mathbb{C}^n$.

4.2 *Inverses of orthogonal and unitary matrices*

Let $A \in \mathbf{Mat}(n \times n, \mathbb{R})$ be orthogonal.

$$A = (v_1 \mid \cdots \mid v_n)$$

$$A^t = \begin{pmatrix} v_1^t \\ \vdots \\ v_n^t \end{pmatrix} \implies A^t \cdot A = \begin{pmatrix} v_1^t \\ \vdots \\ v_n^t \end{pmatrix} (v_1 \mid \cdots \mid v_n) = (b_{ij})$$

$$\implies b_{ij} = v_i^t \cdot v_j = \langle v_i, v_j \rangle = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j \end{cases}$$

$$\implies A^t \cdot A = I \implies A^{-1} = A^t$$

Now let $A \in \mathbf{Mat}(n \times n, \mathbb{C})$ be unitary.

If $A = (a_{ij}$, then $\overline{A} \equiv (\overline{a_{ij}})$

$$A = (v_1 \mid \cdots \mid v_n), \; A^\star \equiv \overline{A}^t$$

$$A^\star = \begin{pmatrix} v_1^\star \\ \vdots \\ v_n^\star \end{pmatrix} \implies A^\star \cdot A = \begin{pmatrix} v_1^\star \\ \vdots \\ v_n^\star \end{pmatrix} (v_1 \mid \cdots \mid v_n) = I \implies A^{-1} = A^\star$$

Thus we have the following:

Theorem 4.5. *Inverse of orthogonal and unitary matrices:*

- *If $A \in \mathbf{Mat}(n \times n, \mathbb{R})$, A orthogonal, then $A^{-1} = A^t$.*
- *If $A \in \mathbf{Mat}(n \times n, \mathbb{C})$, A unitary, then $A^{-1} = A^\star$.*

Let $U \leq K^n$, $B = \{v_1, \cdots, v_k\}$ a basis for U. We want to find an orthogonal or orthonormal basis B' for U: Gram-Schmidt-Process.

This algorithm requires orthogonal projections: (Imagine a diagram of v projected onto u, $u \neq 0$, with $v_\parallel = proj_u(v)$ the projection onto u and $v_\perp$ the perpendicular part.)

$$v_\parallel = t \cdot u \quad \text{where } t \text{ is to be determined}$$

$$v = v_\parallel + v_\perp = tu + v_\perp$$

$$\begin{aligned} \implies \langle u, v \rangle &= \langle u, tu + v_\perp \rangle \\ &= \langle u, tu \rangle + \underbrace{\langle u, v_\perp \rangle}_{=0} \\ &= t \cdot \langle u, u \rangle \end{aligned}$$

$$\implies proj_u(v) = v_\parallel = \frac{\langle u, v \rangle}{\langle u, u \rangle} u$$

$$\implies v_\perp = v - v_\parallel = v - \frac{\langle u, v \rangle}{\langle u, u \rangle} u$$

Now we can introduce the GRAM-SCHMIDT process:

Let $U \leq K^n$, $B = \{v_1, \cdots, v_k\}$ a basis for U.

Define: $u_1 = v_1$

$$u_2 = v_2 - proj_{u_1}(v_2) = v_2 - \frac{\langle u_1, v_2\rangle}{\langle u_1, u_1\rangle} u_1$$

$$u_3 = v_3 - proj_{u_1}(v_3) - proj_{u_2}(v_3) = v_3 - \frac{\langle u_1, v_3\rangle}{\langle u_1, u_1\rangle} u_1 - \frac{\langle u_2, v_3\rangle}{\langle u_2, u_2\rangle} u_2$$

$\ldots$

$$u_k = v_k - \frac{\langle u_1, v_k\rangle}{\langle u_1, u_1\rangle} u_1 - \ldots - \frac{\langle u_{k-1}, v_k\rangle}{\langle u_{k-1}, u_{k-1}\rangle} u_{k-1}$$

Then: $span\{v_1, \cdots, v_j\} = span\{u_1, \cdots, u_j\}\ \forall 1 \leq j \leq k$

$u_j \neq 0\ \forall 1 \leq j \leq k$

$\{u_1, \cdots, u_j\}$ is orthogonal $\forall 1 \leq j \leq k$

Especially, we get that $span\{u_1, \cdots, u_j\} = span\{v_1, \cdots, v_j\} = U$ and $\{u_1, \cdots, u_j\}$ is orthogonal.

Remark 12. We can find an orthonormal basis by "normalizing" $u_1, \cdots, u_j$: $B'' = \{\frac{1}{\|u_1\|} u_1, \cdots, \frac{1}{\|u_k\|} u_k\}$

Proof. by "finite" induction.

Base: $j = 1 \quad u_1 = v_1 \neq 0$

Inductive: Now assume the result holds for a j with $1 \leq j < k$.
Under that assumption, we will prove the result for $j + 1$.

$$u_{j+1} = v_{j+1} - \underbrace{\frac{\langle u_1, v_{j+1}\rangle}{\langle u_1, u_1\rangle} u_1 - \ldots - \frac{\langle u_j, v_{j+1}\rangle}{\langle u_j, u_j\rangle} u_j}_{\substack{\in\, span\{u_1, \cdots, u_j\} = span\{v_1, \cdots, v_j\} \\ v_{j+1} \notin span\{v_1, \cdots, v_j\} \text{ since } \{v_1, \cdots, v_j\} \text{ is LI}}}$$

$$\implies u_{j+1} \neq 0$$

To show orthogonality, we need to show that u_{j+1} is orthogonal to $u_1, \cdots, u_j$: $u_{j+1} \perp u_i \quad 1 \leq i \leq j$.

$$\langle u_i, u_{j+1}\rangle = \langle u_i, v_j + 1\rangle - \frac{\langle u_1, v_{j+1}\rangle}{\langle u_1, u_1\rangle} \underbrace{\langle u_i, u_1\rangle}_{=0} - \ldots$$

$$- \frac{\langle u_i, v_{j+1}\rangle}{\langle u_i, u_i\rangle} \langle u_i, u_i\rangle - \ldots - \frac{\langle u_j, v_{j+1}\rangle}{\langle u_j, u_j\rangle} \underbrace{\langle u_i, u_j\rangle}_{=0}$$

$$= \langle u_i, v_{j+1} - \frac{\langle u_i, v_{j+1}\rangle}{\cancel{\langle u_i, u_i\rangle}} \cancel{\langle u_i, u_i\rangle} = 0$$

$\implies \{u_1, \cdots, u_{j+1}\}$ is orthogonal and all $u_i \neq 0$
$\implies$ (assignment question) $\{u_1, \cdots, u_{j+1}\}$ is LI.

We still need to show that $span\{u_1, \cdots, u_{j+1}\} = \{v_1, \cdots, v_{j+1}\}$.

Knowing that $span\{u_1, \cdots, u_j\} = span\{v_1, \cdots, v_j\}$

$$\begin{aligned} \implies span\{u_1, \cdots, u_{j+1}\} &= span\{v_1, \cdots, v_j, u_{j+1}\} \\ &= span\{v_1, \cdots, v_j, \underbrace{u_{j+1}}_{\text{LC of } v_1, \cdots, v_{j+1}}, v_{j+1}\} \\ &= span\{v_1, \cdots, v_j, v_{j+1}\} \end{aligned}$$

□

LECTURE 03/29

Example 4.3. Let $B = \left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 3 \end{pmatrix} \right\}$ be a basis for $\mathbb{R}^3$. We want to find a basis B' that is (i) Orthogonal (ii) Orthonormal. We use the Gram-Schmidt method:

$$u_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$$

$$u_2 = \begin{pmatrix} 3 \\ 1 \\ -1 \end{pmatrix} - \frac{\langle \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 \\ 1 \\ -1 \end{pmatrix} \rangle}{\langle \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \rangle} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix} - \frac{4}{2} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix}$$

$$\begin{aligned} u_3 &= \begin{pmatrix} 1 \\ 1 \\ 3 \end{pmatrix} - \frac{\langle \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 3 \end{pmatrix} \rangle}{\langle \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \rangle} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} - \frac{\langle \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 3 \end{pmatrix} \rangle}{\langle \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix} \rangle} \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix} \\ &= \begin{pmatrix} 1 \\ 1 \\ 3 \end{pmatrix} - \frac{2}{2} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} - \frac{-3}{3} \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} \end{aligned}$$

(i) $B' = \{u_1, u_2, u_3\}$ is an orthogonal basis. (ii) We divide each u_i by its norm: $\{\frac{1}{\sqrt{2}}u_1, \frac{1}{\sqrt{3}}u_2, \frac{1}{\sqrt{6}}u_3\}$ i an orthonormal basis.

Moreover, we notice that $P \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{6}} \\ 0 & -\frac{1}{\sqrt{3}} & \frac{2}{\sqrt{6}} \end{pmatrix}$ is orthogonal since $P^{-1} = P^t$, where the columns are pairwise orthogonal unit vectors. We also notice that the same applies for rows:

$$\langle R_1, R_2 \rangle = 0 \qquad \langle R_1, R_3 \rangle = 0 \qquad \langle R_2, R_3 \rangle = 0$$
$$\text{and } \langle R_1, R_2 \rangle = \langle R_1, R_3 \rangle = \langle R_2, R_3 \rangle = 1$$

Theorem 4.6. *Let P be orthogonal, then the row vectors of P are pairwise orthogonal unit vectors.*

Proof. P orthogonal $\implies P^{-1} = P^t$

$$\implies I = PP^{-1} = PP^t = \begin{pmatrix} u_1^t \\ \cdots \\ u_n^t \end{pmatrix} \left(u_1 \mid \cdots \mid u_n \right)$$
$$\implies u_i^t \cdot u_i = 1 = \langle u_i, u_i \rangle = ||u_i||^2 \implies ||u_i|| = 1 \quad \forall 1 \leq i \leq n$$
$$\implies u_i^t \cdot u_j = 0, \quad i \neq j \quad \forall 1 \leq i, j \leq n$$
$$\implies \langle u_i, u_j \rangle = 0 \implies u_i, u_j \text{ orthogonal.}$$

□

Theorem 4.7. *(a) Let P, Q be orthogonal, then P^{-1} and $P \cdot Q$ are orthogonal, thus $\mathcal{O}(n) \equiv \{P \in Mat(n \times n, \mathbb{R}) : P \text{ orthogonal}\}$ is a group.*

(b) Let P, Q be unitary, then P^{-1} and $P \cdot Q$ are unitary, thus $\mathcal{U}(n) \equiv \{P \in Mat(n \times n, \mathbb{R}) : P \text{ unitary}\}$ is a group.

$\mathcal{O}(n), \mathcal{U}(n)$ are subgroups of $\mathbf{GL}(n, \mathbb{R})$.

4.3 *Orthogonal/unitary matrices and diagonalizability (upper-triangularizability)*

We wonder what matrices $\in \mathbf{Mat}(n \times n, \mathbb{R})$ can be diagonalized or upper-triangularized via orthogonal, i.e. for matrices A $\exists P \in \mathcal{O}(n)$: $P^{-1}AP = P^tAP = D$ or $P^tAP = T$.

Same for unitary matrices $A \in \mathbf{Mat}(n \times n, \mathbb{C})$, we want $P^{-1}AP = P^{\star}AP = D$ or $P^{\star}AP = T$.

We begin with upper-triangularizability of matrices $\in \mathbf{Mat}(n \times n, \mathbb{C})$ via unitary matrices.

Theorem 4.8. ***Schur.*** *Every matrix $A \in \mathbf{Mat}(n \times n, \mathbb{C})$ can be upper-triangularized via a unitary matrix.*

Proof. Same as proof of Theorem 2.25, by induction. We first need to find one eigenvalue and its corresponding eigenvector. Next we extend to an arbitrary basis and obtain an orthogonal basis by Gram-Schmidt.

$n = 1$: Any 1×1 matrix is already upper-triangular.

$n - 1 \to n$: Outline of proof: let $A \in \mathbf{Mat}(n \times n, \mathbb{C})$, let λ_1 be eigenvalue of A and let v_1 be a unit eigenvector to λ_1. We then extend $\{v_1\}$ to a basis $\{v_1, u_2, \cdots, u_n\}$ for $\mathbb{C}^n$ and apply Gram-Schmidt to obtain an orthonormal basis $\{v_1, \cdots, v_n\}$ for $\mathbb{C}^n$.
$P = (v_1 | \cdots | v_n)$ is unitary. Then $\exists$ $C \in \mathbf{Mat}((n-1) \times (n-1), \mathbb{C})$:

$$B = P^{-1}AP = P^{\star}AP = \left(\begin{array}{c|c} \lambda & \star \\ \hline 0 & C \end{array}\right)$$

By inductive step, C can also be upper-triangularized via unitary matrix $\tilde{Q} \in \mathcal{U}(n-1)$, i.e. $\tilde{Q}^{-1}C\tilde{Q} = \tilde{Q}^{\star}C\tilde{Q} = \tilde{T}$ upper triangular. (Could INSERT more details).

In summary: $P^{\star}AP = B$ and $Q^{\star}BQ = T$, so $Q^{\star}P^{\star}APQ = T = (PQ)^{\star}A(PQ)$. The product of two unitary matrices is unitary, thus PQ is unitary and A is upper triangular via unitary matrix. □

$(AB)^{\star} = (\overline{AB})^t = (\overline{A}\overline{B})^t = \overline{B}^t\overline{A}^t$
$= B^{\star}A^{\star}$

Corollary 4.9. *If $A \in Mat(n \times n, \mathbb{R})$ has only real eigenvalues, then A can be upper-triangularized via an orthogonal matrix.*

Proof. If all eigenvalues are real, v_1 is real, and $v_2, \cdots, v_n$ can be chosen real $\implies$ all coefficients in previous proof are real $\implies$ PQ is orthogonal $\implies$ A is upper-triangular via orthogonal matrix. □

Diagonalizability

Real case: Under what conditions can $A \in \mathbf{Mat}(n \times n, \mathbb{R})$ be diagonalizable via an orthogonal matrix? Assume A can be diagonalized via $P \in \mathcal{O}(n)$ s.t $P^{-1}AP = P^tAP = D$

$\implies A = PDP^{-1} = PDP^t$

$\implies A^t = (PDP^{-1})^t(P^t)^tD^tP^t = PDP^t$

$\implies A = A^t$, so A must be symmetric.

Thus, for A to be diagonalizable via orthogonal matrix, A needs to be symmetric (necessary but not sufficient condition) and have real eigenvalues.

Theorem 4.10. $\forall A \in \mathbf{Symm}(n \times n, \mathbb{R})$:

(a) All eigenvalues λ of $A \in \mathbb{R}$.

(b) Eigenvectors to distinct eigenvalues of A are orthogonal.

Proof. (a) Let λ be an eigenvalue of A and X its corresponding eigenvector. Consider $X^\star AX$, then, since A is symmetric and real:

$$
\begin{aligned}
&X^\star(AX) = X^\star(\lambda X) = \lambda X^\star X = \lambda||X||^2 \\
&= (X^\star A)X = (X^\star A^t)X = (X^\star A^\star)X = (AX)^\star X = (\lambda X)^\star X \\
&= \overline{\lambda}X^\star X = \overline{\lambda}||X||^2 \\
&\implies \lambda||X||^2 = \overline{\lambda}||X||^2 \\
&\implies (\lambda - \overline{\lambda})||X||^2 = 0 \\
&\quad X \leq 0 \text{ since it is an eigenvector} \\
&\implies \lambda - \overline{\lambda} = 0 \\
&\implies \lambda = \overline{\lambda} \implies \lambda \text{ real}
\end{aligned}
$$

(b) Let $\lambda_1 \leq \lambda_2$ be eigenvalues with x_1, x_2 corresponding eigenvectors, then

$$
\begin{aligned}
&x_1^\star Ax_2 = x_1^\star(Ax_2) = x_1^\star(\lambda_2 x_2) = \lambda_2 x_1^\star x_2 = \lambda_2\langle x_1, x_2\rangle \\
&(x_1^\star A)x_2 = (Ax_1)^\star x_2 = (\lambda_1 x_1)^\star x_2 = \overline{\lambda_1}x_1^\star x_2 = \text{by (a)} = \lambda_1\langle x_1, x_2\rangle \\
&\implies \lambda_2\langle x_1, x_2\rangle = \lambda_1\langle x_1, x_2\rangle \\
&\implies (\lambda_2 - \lambda_1)\langle x_1, x_2\rangle = 0 \\
&\implies \langle x_1, x_2\rangle = 0 \implies x_1, x_2 \text{ orthogonal}
\end{aligned}
$$

□

Remark 13. Let $A \in \mathbf{Symm}(n \times n, \mathbb{R})$, $P \in \mathbf{GL}(n, \mathbb{R})$. In general, $P^{-1}AP$ is not symmetric, i.e. the property of symmetry is, in general, not preserved under linear coordinate transformations.

LECTURE 04/03

Recap from last class:

1. If $A \in \mathbf{Mat}(n \times n, \mathbb{R})$ is orthogonally diagonalizable, then A is symmetric.
2. If A is real symmetric, then all eigenvalues of A are real.

Example 4.4. $A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$, $P = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$

Then $P^{-1}AP = \begin{pmatrix} 1 & 4 \\ 0 & -1 \end{pmatrix} = B$, A is symmetric while B is not.

Theorem 4.11. *Let $A \in \mathbf{Symm}(n \times n, \mathbb{R})$, P orthogonal i.e. $P \in \mathcal{O}(n)$. Then $B = P^{-1}AP = P^tAP$ is symmetric.*

Proof. $B = P^tAP \implies B^t = (P^tAP)^t = P^tA^tP = P^tAP = B$

$\implies B$ is symmetric. □

Theorem 4.12. ***Spectral Theorem (real version)**.*

A matrix $A \in \mathbf{Mat}(n \times n, \mathbb{R})$ is diagonalizable via an orthogonal matrix iff A is symmetric.

Proof. ($\Rightarrow$) already done.

($\Leftarrow$) Let A be real symmetric: then all eigenvalues of A are real. $\implies$ by Corollary 4.9 (Schur, real version) $\exists P \in \mathcal{O}(n)$ s.t. $P^{-1}AP = P^tAP = T$ where T is upper triangular. Since $P \in \mathcal{O}(n)$, T is symmetric. But a matrix that is both symmetric and upper triangular is diagonalizable $\implies$ T is diagonalizable $\implies$ A is diagonalizable via an orthogonal matrix. □

Definition 4.9. A matrix $A \in \mathbf{Mat}(n \times n, \mathbb{C})$ is **hermetian** if $A^\star = A$.

A hermetian matrix is the complex equivalent of a real symmetric matrix.

We might like to conjecture that A is diagonalizable via a unitary matrix iff A is hermetian. However, while it is true that all hermetian matrices ARE unitarily diagonalizable, the converse does not hold.

Definition 4.10. $A \in \mathbf{Mat}(n \times n, \mathbb{C})$ is **normal** if $AA^\star = A^\star A$.

Not all matrices are normal, e.g. $\begin{pmatrix} 1 & 4 \\ 0 & -1 \end{pmatrix}$ is not normal.

Theorem 4.13. *The following classes of matrices are **normal**:*

1. *Real symmetric matrices*
2. *Real skew symmetric matrices*
3. *Real orthogonal matrices*
4. *Hermetian matrices*
5. *Skew hermetian matrices (i.e. $A^\star = -A$)*
6. *Unitary matrices*

Proof. 1. $AA^\star = AA^t = A \cdot A = A^tA = A^\star A$

2. $AA^\star = AA^t = A(-A) = -A^2 = (-A)A = A^tA = A^\star A$

3. $AA^\star = AA^t = AA^{-1} = I = A^{-1}A = A^tA = A^\star A$
4. $AA^\star = A \cdot A = A^\star A$
5. Left to the reader.
6. $AA^\star = AA^{-1} = I = A^{-1}A = A^\star A$ ☐

Remark 14. The class of normal matrices is strictly larger than the union of these 6 classes.

Remark 15. If A is normal, $P \in \mathbf{GL}(n\mathbb{C})$, then $P^{-1}AP$ is in general NOT normal.

Example 4.5. (same as Example 4.4) $A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$, $P = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$

Theorem 4.14. *Let $A \in \mathbf{Mat}(n \times n, \mathbb{C})$ normal, $P \in \mathcal{U}(n)$ (unitary group). Then $P^{-1}AP = P^\star AP$ is normal.*

Proof. Let $B \equiv P^\star AP$. Then:

$$BB^\star = P^\star AP(P^\star AP)^\star = P^\star A \underbrace{P \cdot P^\star}_{=I} A^\star P = P^\star AA^\star P$$

$$B^\star B = (P^\star AP)^\star P^\star AP = P^\star A^\star \underbrace{P \cdot P^\star}_{=I} AP = P^\star \underbrace{A^\star A}_{=AA^\star} P$$

$$\implies BB^\star = B^\star B \implies B \text{ is normal.}$$ ☐

Theorem 4.15. *Let $A \in \mathbf{Mat}(n \times n, \mathbb{C})$ be both normal and upper triangular. Then A is diagonal.*

Proof. by induction on n.

$n = 1$ Nothing to show since ALL 1×1 matrices are normal, upper triangular and diagonalizable.

$n-1 \to n$ $A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ & a_{22} & \cdots & a_{2n} \\ & & \ddots & \\ & & & a_{nn} \end{pmatrix}$, $A^\star = \begin{pmatrix} \overline{a_{11}} & & & \\ \overline{a_{12}} & \overline{a_{22}} & & \\ \vdots & \vdots & \ddots & \\ \overline{a_{1n}} & \overline{a_{2n}} & & \overline{a_{nn}} \end{pmatrix}$

Let $B \equiv AA^\star = (b_{ij}) \implies b_{11} = a_{11}\overline{a_{11}} + a_{12}\overline{a_{12}} + \cdots + a_{1n}\overline{a_{1n}}$

$C \equiv A^\star A = (c_{ij}) \implies c_{11} = a_{11}\overline{a_{11}}$

Since A is normal, $B = C \implies b_{11} = c_{11}$

$\implies a_{11}\overline{a_{11}} = a_{11}\overline{a_{11}} + a_{12}\overline{a_{12}} + \cdots + a_{1n}\overline{a_{1n}}$

$\implies |a_{11}|^2 = |a_{11}|^2 + \underbrace{|a_{12}|^2 + \cdots + |a_{1n}|^2}_{\geq 0 \text{ real}}$

$\implies a_{12} = \cdots = a_{1n} = 0$

$$\implies A = \left(\begin{array}{c|ccc} a_{11} & & 0 & \\ \hline & a_{22} & & \star \\ 0 & & \ddots & \\ & & & a_{nn} \end{array}\right), \quad A^\star = \left(\begin{array}{c|c} \overline{a_{11}} & 0 \\ \hline 0 & \tilde{A}^\star \end{array}\right)$$

$$\implies AA^\star = \left(\begin{array}{c|c} \|a_{11}\|^2 & 0 \\ \hline 0 & \tilde{A}\tilde{A}^\star = \tilde{A}^\star\tilde{A} \end{array}\right) = A^\star A$$

$\implies$ $\tilde{A}$ is normal $(n-1) \times (n-1)$ matrix, $\implies$ by inductive hypothesis $\tilde{A}$ is diagonalizable, $\implies$ A is diagonalizable. ☐

Theorem 4.16. ***Spectral Theorem (complex version).*** *A matrix $A \in$* **Mat**$(n \times n, \mathbb{C})$ *is diagonalizable via a unitary matrix iff A is normal.*

Proof.

($\Rightarrow$) Let $P \in \mathcal{U}(n)$ s.t. $P^{-1}AP = P^{\star}AP = D \implies A = PDP^{\star}$.

$$AA^{\star} = PDP^{\star}(PDP^{\star})^{\star} = PD\underbrace{P^{\star}P}_{=I}D^{\star}P^{\star} = PD\overline{D}P^{\star}$$
$$A^{\star}A = (PDP^{\star})^{\star}PDP^{\star} = PD^{\star}\underbrace{P^{\star}P}_{=I}DP^{\star} = P\overline{D}DP^{\star}$$

Note that ANY two diagonal matrices commute
$\implies D\overline{D} = \overline{D}D \implies AA^{\star} = A^{\star}A$

($\Leftarrow$) Let A be normal. By Theorem 4.8 (Schur), $\exists P \in \mathcal{U}(n)$ s.t. $P^{-1}AP = P^{\star}AP = T$ upper triangular. Since $P \in \mathcal{U}(n)$, T is normal and upper triangular, $\implies$ T is diagonalizable $\implies$ A is diagonalizable via unitary matrix. ☐

Remark 16. *on the real spectral theorem.* We know that all real symmetric matrices are diagonalizable and have only real eigenvalues.

Definition 4.11. A real symmetric matrix is called

1. **positive definite** if ALL eigenvalues of A are positive.
2. **positive semidefinite** if ALL eigenvalues of A are ≥ 0.
3. **negative definite** if ALL eigenvalues of A are < 0.
4. **negative semidefinite** if ALL eigenvalues of A are ≤ 0.

LECTURE 04/05

Example 4.6. *on the Real Spectral Theorem*

$$A = \begin{pmatrix} -3 & 1 & -2 \\ 1 & -3 & -2 \\ -2 & -2 & 0 \end{pmatrix}, \text{ real symmetric.}$$

We want to find $P \in \mathcal{O}(n)$ such that $P^t AP = D$.

1. Find eigenvalues and eigenvectors: $\lambda = -4, \lambda = 2$
   $E_{-4} = span\{\begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}\}$, $\quad E_2 = span\{\begin{pmatrix} -1 \\ -1 \\ 2 \end{pmatrix}\}$ Note that eigenvectors to distinct eigenvalues are orthogonal (LI by thm 2.19). For multiple eigenvalues, we use Gram-Schmidt to find orthogonal bases for the corresponding eigenspaces.

2. Apply Gram-Schmidt on E_{-4}:

   Let $u_1 = \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}$, then $u_2 = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} - \frac{\langle \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, u_1 \rangle}{\langle u_1, u_1 \rangle} \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix} = \frac{1}{5}\begin{pmatrix} -1 \\ 5 \\ 2 \end{pmatrix}$ Note that we can take any scalar multiple (orthogonality is preserved), thus we take $u_2 = \begin{pmatrix} -1 \\ 5 \\ 2 \end{pmatrix}$. We can check that $\langle u_2, \begin{pmatrix} -1 \\ -1 \\ 2 \end{pmatrix} \rangle = 0$.

3. Obtain orthogonal matrix: recall that an orthogonal matrix has orthonormal columns, therefore we need to divide each vector by its norm:
$E_{-4} = span\left\{\frac{1}{\sqrt{5}}\begin{pmatrix}2\\0\\1\end{pmatrix}, \frac{1}{\sqrt{30}}\begin{pmatrix}-1\\5\\2\end{pmatrix}\right\}$, $E_2 = span\left\{\frac{1}{\sqrt{6}}\begin{pmatrix}-1\\-1\\2\end{pmatrix}\right\}$

$$P = \begin{pmatrix} \frac{2}{\sqrt{5}} & -\frac{1}{\sqrt{30}} & \frac{4}{\sqrt{6}} \\ 0 & \frac{5}{\sqrt{30}} & -\frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{30}} & \frac{2}{\sqrt{6}} \end{pmatrix} \in \mathcal{O}(n)$$
$$P^t AP = \begin{pmatrix} -4 & & \\ & -4 & \\ & & 2 \end{pmatrix}$$

4.4 *Real Quadratic Forms*

Application of the *Real Spectral Theorem*.

Definition 4.12. A (real) **quadratic form** is a purely quadratic polynomial in n variables with real coefficients.

For example, $x_1^2 + 4x_1x_2 + x_2^2$ and $x_1^2 + 2x_1x_2 + 2x_2^2$ are quadratic forms.

There are two interesting questions that we would like to answer concerning these polynomials:

1. What is the range of the quadratic form?
2. Can we express it as a sum or difference of pure squares?

Definition 4.13. A quadratic form $Q(x) = Q(x_1, x_2, \cdots, x_n)$ is **positive definite** if $Q(x) \geq 0\ \forall x \in \mathbb{R}^n$ and $Q(x) = 0 \Leftrightarrow x = 0$.
Similar for negative definite.

Example 4.7. $x_1^2 + 4x_1x_2 + x^2$ is not positive definite: Let $x_2 = -x_1$
$Q(x_1, -x_1) = x_1^2 - 4x_1^2 + x_1^2 = -2x_1^2 < 0\ \forall x_1 \neq 0$.

We want to use a more systematic approach:

$$\begin{aligned} Q(x) &= a_{11}x_1^2 + \cdots + a_{nn}x_n^2 + b_{12}x_1x_2 + \cdots + b_{(n-1),n}x_{n-1}x_n \\ &= \sum_{i=1}^{n} a_{ii}x_i^2 + \sum_{i,j=1,i<j}^{n} b_{ij}x_ix_j \end{aligned}$$

Let A be an $n \times n$ matrix:

$$\begin{aligned} x^t Ax &= (x_1 \cdots x_n)\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn}\end{pmatrix}\begin{pmatrix}x_1\\ \vdots \\ x_n\end{pmatrix} \\ &= (x_1 \cdots x_n)\begin{pmatrix} a_{11}x_1 + \cdots + a_{1n}x_n \\ \vdots \\ a_{n1}x_1 + \cdots + a_{nn}x_n \end{pmatrix} \\ &= a_{11}x_1^2 + a_{12}x_1x_2 + \cdots + a_{1n}x_1x_n + \cdots \\ &\quad + a_{n1}x_nx_1 + a_{n2}x_nx_2 + \cdots + a_{nn}x_n^2 \text{ since } i < j: \\ &= a_{11}x_1^2 + \cdots + a_{nn}x_n^2 + (a_{12} + a_{21})x_1x_2 + (a_{13} + a_{31})x_1x_3 + \cdots \\ &\quad + (a_{n-1,n} + a_{n,n-1})x_{n-1}x_n \end{aligned}$$

This is a quadratic form! Conversely, every quadratic form $Q(x)$ can be written in matrix from, i.e. $Q(x) = x^t Ax, b_{ij} = a_{ij} + a_{ji}$.

If we split evenly, i.e. $a_{ij} = a_{ji} = \frac{1}{2}b_{ij}$, the resulting matrix is real symmetric, and therefore diagonalizable by an orthogonal matrix.

We now answer our two questions:

2. Let $Q(x) = x^t Ax, A \in \textbf{Symm}(n \times n, \mathbb{R})$. Let $P \in \mathcal{O}(n)$ such that $P^t AP = D \Leftrightarrow A = PDP^t$.
   $D = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$ and $Q(x) = x^t PDP^t x$, let $P^t x = y \implies x^t P = y^t$:
   $\implies Q(x) = y^t Dy = \lambda_1 y_1^2 + \cdots + \lambda_n y_n^2$
   We expressed the quadratic form as a sum of pure squares. We can thus easily answer our first question:

1. Q is positive definite $\Leftrightarrow$ all $\lambda_i > 0$ i.e. $\Leftrightarrow$ A is positive definite.

Example 4.8.
$x_1^2 + 4x_1x_2 + x_2^2 = Q(x_1, x_2)$
$\implies A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$, and thus:
$Q(x_1, x_2) = \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$

Example 4.9. $Q(x_1, x_2) = x_1^2 + 4x_1x_2 + x_2^2$, $A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$ orthogonally diagonalizable.
$\lambda = -1, 3$.
$E_{-1} = span\left\{\begin{pmatrix} -1 \\ 1 \end{pmatrix}\right\}, E_3 = span\left\{\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right\}$.
Orthonormal basis is thus
$\left\{\frac{1}{\sqrt{2}}\begin{pmatrix} -1 \\ 1 \end{pmatrix}, \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right\}$, $P \equiv \frac{1}{\sqrt{3}}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$
$P^{-1} = P^t = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$, $A = PDP^t$

$$Q(x_1, x_2) = \begin{pmatrix} x_1 & x_2 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 3 & 0 \\ 0 & -1 \end{pmatrix} \cdot \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} x_1 + x_2 \\ x_2 - x_1 \end{pmatrix}$$

This is the matrix of y_1, y_2.

$$Q(y) = \lambda_1 y_1^2 \lambda_2 y_2^2 = \frac{3}{2}(x_1 + x_2)^2 - \frac{1}{2}(x_2 - x_1)^2.$$

As a check, we can expand this result.
Note that this is not positive definite because of the negative eigenvalue $\lambda = -1$.

Remark 17. We actually saw quadratic forms in Calculus 3 as well when classifying critical points with the second derivative test:

$n = 1$ Single variable case : f twice continuously differentiable, $f'(x_0) = 0$. if $f''(x_0) > 0, x_0$ is a local minimum. If $f''(x_0) < 0, x_0$ is a local maximum.

$n - dim$ Multivariable case : f twice continuously differentiable, x_0 is a critical point i.e. $\nabla f(x_0) = 0$. We take the Taylor expansion about x_0:

$$f(x) = f(x_0) + \nabla f(x_0) \cdot (x - x_0) + \frac{1}{2}(x - x_0)H(x_0)(x - x_0)^2 + \cdots$$

where $H = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}\right)$ is the Hessian matrix, symmetric in this case since f is twice continuously differentiable.
This is a quadratic form, with a $\frac{1}{2}$ coefficient due to even split.
Let $P \in \mathcal{O}(n)$ such that:

$$P^t H(x_0)P = D = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$$

$\implies H(x_0) = PDP^t$
$\implies (x - x_0)^t H(x_0)(x - x_0) = y^t Dy = \lambda_1 y_1^2 + \cdots + \lambda_n y_n^2$
So x_0 is a local minimum if all $\lambda_i > 0$, so we have:

- $H(x_0)$ positive definite $\implies$ x_0 is a local minimum,
- $H(x_0)$ negative definite $\implies$ x_0 is local maximum.

4.5 Real Inner Product Spaces

Goal: We want to introduce length and angle to abstract vector spaces. These are important to define convergence (Analysis). We must distinguish the cases $K = \mathbb{R}$ and $K = \mathbb{C}$.
$K = \mathbb{R}$: important properties of standard inner product on $\mathbb{R}^n$: $\langle x, y\rangle = x^t y$

- Bilinear
- Positive definite: $\langle x, x\rangle \geq 0$ ($\langle x, x\rangle = 0 \Leftrightarrow x = 0$)
- symmetric: $\langle x, y\rangle = \langle y, x\rangle$

We use these axioms to define inner product in abstract vector space: LECTURE 04/10

Definition 4.14. Let V vector space over $\mathbb{R}$. $\langle , \rangle : V \times V \to \mathbb{R}$ is a **real inner product space** on V iff: $\langle , \rangle$ is bilinear (i.e. linear in each argument), symmetric ($\forall\ x, y \in V : \langle x, y\rangle = \langle y, x\rangle$) and positive definite ($\forall\ x \in V : \langle x, x\rangle \geq 0 \wedge \langle x, x\rangle = 0$ iff $x = 0$).

Example 4.10.

1. The standard inner product $\langle x, y\rangle \equiv x^t y$ on $\mathbb{R}^n$ is an inner product.
2. Let $V = \mathbb{R}^n$, A positive definite (and symmetric) matrix. Then: $\langle X, Y\rangle = X^t AY$ is an inner product.

   Proof. Bilinearity: $$\begin{aligned}\langle X + U, Y\rangle &= (X + U)^t AY = X^t AY + U^t AY \\ &= \langle X, Y\rangle + \langle U, Y\rangle \\ \text{and: } \langle kX, Y\rangle &= (kX)^t AY = kX^t AY \\ &= k\langle x, y\rangle\end{aligned}$$

   Similarly we can show linearity in second argument.

   Symmetry: $$\begin{aligned} &\langle Y, X\rangle = Y^t AX \\ \Longrightarrow\ &(\underbrace{Y^t AX}_{\in\mathbb{R}})^t = X^t \underbrace{A^t}_{=A} Y = X^t AY \\ \Longrightarrow\ &Y^t AX = X^t AY\end{aligned}$$

   Positive definite: $$\langle X, X\rangle = X^t \underbrace{A}_{pos.\,def} X = \begin{cases} \geq 0 & \forall\ x \in \mathbb{R}^b \\ = 0 & \text{iff } x = 0\end{cases}$$

   $\Longrightarrow \langle , \rangle$is positive definite. $\square$

3. Let $V = C([a, b])$ the VS of all continuous functions on $[a, b]$. $\langle f, g\rangle \equiv \int_a^b f(x)g(x)dx$ is an inner product.

Proof. Bilinearity: $\langle f_1+f_2,g\rangle = \int_a^b [f_1(x)+f_2(x)]g(x)dx$

$$= \int_a^b f_1(x)g(x)+f_2(x)g(x)dx$$

$$= \langle f_1,g\rangle + \langle f_2,g\rangle$$

and: $\langle kf,g\rangle = \int_a^b kf(x)g(x)dx$

$$= k\int_a^b f(x)g(x) = k\langle f,g\rangle$$

Similarly we can show linearity in second argument.

Symmetry: $\langle g,f\rangle = \int_a^b g(x)f(x)dx = \int_a^b f(x)g(x)dx = \langle f,g\rangle$

Positive definite: $\langle f,f\rangle = \int_a^b \underbrace{f^2(x)}_{\geq 0}dx \quad \geq 0 \quad \forall f \in C([a,b])$

And let $\langle f,f\rangle = 0 \implies \int_a^b f^2(x)dx = 0$. This holds if $f \equiv 0$ on $[a,b]$. Then $\exists c \in [a,b]$ s.t. $f(c) \neq 0$

INSERT picture 10:34

$\implies f^2(c)\rangle 0$ ($\langle d$). Then $\epsilon > 0$ s.t. $f^2(x)\rangle \frac{d}{2}$ on $(c-\epsilon, c+\epsilon)$

$\implies \int_a^b f^2(x)dx \geq \frac{d}{2}2\epsilon = d\epsilon > 0$

$\implies \langle f,f\rangle = 0$ iff $f \equiv 0 \implies \langle,\rangle$ is positive definite. □

Remark 18.

- $\langle,\rangle$ on $C([a,b])$ induces an inner product e.g. on P_n and P (VS of all polynomials).
- $\langle,\rangle$ is not an inner product of the set $F([a,b])$ of all real valued functions on $[a,b]$: while $\langle,\rangle$ is still bilinear and symmetric, it fails to be positive definite.

  For example, $f(x) \equiv \begin{cases} 1 & if\ x=0 \\ 0 & if\ x \neq 0 \end{cases}$

  Then $f \neq 0$ but $\langle f,f\rangle = \int_a^b f^2(x)dx = 0 \implies \langle,\rangle$ is not positive def.

4.6 *Complex Inner Product Spaces*

Reminder: Properties of standard inner product on $\mathbb{C}^n$:

$\langle z,w\rangle \equiv z^\star w$

Linear in 2. argument.

Sesquilinear in 1. argument: $\langle z_1+z_2,w\rangle = \langle z_1,w\rangle + \langle z_2,w\rangle$

but $\langle kz,w\rangle = \overline{k}\langle z,w\rangle$

$\langle w,z\rangle = w^\star z = (\overline{w_1}\ldots\overline{w_n})\begin{pmatrix} z_1 \\ \vdots \\ z_n\end{pmatrix} = z_1\overline{w_1}+\cdots+z_n\overline{w_n}$

$\langle z,w\rangle = z^\star w = (\overline{z_1}\ldots\overline{z_n})\begin{pmatrix} w_1 \\ \vdots \\ w_n\end{pmatrix} = \overline{z_1}w_1+\cdots+\overline{z_n}w_n$

Note that $z_1\overline{w_1}+\cdots+z_n\overline{w_n} = \overline{\overline{z_1}w_1+\cdots+\overline{z_n}w_n}$

$\implies \langle w,z\rangle = \overline{\langle z,w\rangle}$ "Conjugate symmetry".

Positive definite: $\langle z,z\rangle = z^\star z = (\overline{z_1}\dots\overline{z_n})\begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix}$

$$= z_1\overline{z_1} + \cdots + z_n\overline{z_n} = \underbrace{\|z_1\|^2 + \cdots + \|z_n\|^2}_{\geq 0,\ \in \mathbb{R}}$$

and $\langle z,z\rangle = 0$ iff $\|z_1\|^2 + \cdots + \|z_n\|^2 = 0$ iff $z_1 = \cdots = z_n = 0$
$\implies \langle,\rangle$ is positive definite.

Definition 4.15. Let V be a VS over $\mathbb{C}$. $\langle,\rangle : V \times V \to \mathbb{C}$. $(V, \langle,\rangle)$ is a **complex inner product space** if it satisfies the following:

1. $\langle,\rangle$ is sesquilinear.
2. Conjugate symmetry: $\langle z,w\rangle = \overline{\langle w,z\rangle}\ \forall\, z,w \in V$
3. Positive definite: $\langle z,z\rangle \geq 0\ \forall\ z \in V$ ("the left side is real and non negative") and $\langle z,z\rangle = 0 \Leftrightarrow z = 0$

Caution: some books define sesquilinearity with linearity in the 1. argument (inc. Wikipedia).

Example 4.11. $V = \mathbb{C}^n$, A hermetian, positive definite matrix (eigenvalues of the hermetian matrices are real; positive definite if all eigenvalues are positive). Then $\langle z,w\rangle \equiv z^\star Aw$ is an inner product on $\mathbb{C}^n$.

4.7 *Length in Inner Product Spaces over K*

$x \in V$. Define $\|x\| \equiv \sqrt{\langle x,x\rangle} \geq 0$.
MOTIVATION: $V = \mathbb{R}^n$ with standard inner product:
By Pythagoras, $\left\|\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}\right\| = \sqrt{x_1^2 + \cdots + x_n^2} = \sqrt{\langle x,x\rangle}$

Definition 4.16. Let V be a VS over K. The **norm** $\|\cdot\| : V \to \mathbb{R}_0^+ = [0,\infty)$ is s.t.

1. $\|x\| \geq 0\ \forall x \in V$ and $\|x\| = 0 \Leftrightarrow x = 0$.
2. $\|kx\| = |k| \cdot \|x\|\ \forall k \in K\ \forall x \in V$
3. Triangle inequality: $\|x+y\| \leq \|x\| + \|y\|\ \forall x,y \in V$

Theorem 4.17. *Let V inner product space over K, $\|x\| \equiv \sqrt{\langle x,x\rangle}\ \forall x \in V$. Then $\|\cdot\|$ is a norm on V.*

Proof.

1. $\|x\| = \sqrt{\langle x,x\rangle} \geq 0$ and $\|x\| = 0 = \sqrt{\langle x,x\rangle} \Leftrightarrow \langle x,x\rangle = 0 \Leftrightarrow x = 0$
2. $\|kx\| = \sqrt{\langle kx,kx\rangle} = \sqrt{k\langle kx,x\rangle} = \sqrt{k\bar{k}\langle x,x\rangle}$
   $= |k|\sqrt{\langle x,x\rangle} = |k| \cdot \|x\|$
3. Difficult. We will prove the Cauchy-Schwarz Inequality first. $\square$

Theorem 4.18. ***Cauchy-Schwarz Inequality***. *Let V be an inner product space over K. Then $|\langle x,y\rangle| \leq \|x\| \cdot \|y\| \ \forall x,y \in V$.*

Proof.
Lemma:

$$\begin{aligned}
\langle a+\lambda b, a+\lambda b\rangle &= \langle a+\lambda b, a\rangle + \langle a+\lambda b, \lambda b\rangle \\
&= \langle a+\lambda b, a\rangle + \lambda\langle a+\lambda b, b\rangle \\
&= \langle a,a\rangle + \langle \lambda b, a\rangle + \lambda\left(\langle a,b\rangle + \langle \lambda b, b\rangle\right) \\
&= \langle a,a\rangle + \overline{\lambda}\underbrace{\langle b,a\rangle}_{=\overline{\langle a,b\rangle}} + \lambda\left(\langle a,b\rangle + \overline{\lambda}\langle b,b\rangle\right) \\
&= \langle a,a\rangle + \lambda\langle a,b\rangle + \bar{\lambda}\overline{\langle a,b\rangle} + \lambda\bar{\lambda}\langle b,b\rangle
\end{aligned}$$

Taking $Y_\perp$ as the perpendicular part of the projection of Y onto X, $Y_\perp = Y - \frac{\langle X,Y\rangle}{\langle X,X\rangle}X$. To find $\|Y_\perp\|^2$:

$$\begin{aligned}
0 \leq \|Y_\perp\|^2 &= \langle Y - \frac{\langle X,Y\rangle}{\langle X,X\rangle}X, Y - \frac{\langle X,Y\rangle}{\langle X,X\rangle}X\rangle \\
&\text{using the Lemma: } \lambda = -\frac{\langle X,Y\rangle}{\langle X,X\rangle} \\
&= \langle Y,Y\rangle - \frac{\langle X,Y\rangle}{\langle X,X\rangle}\langle Y,X\rangle - \frac{\overline{\langle X,Y\rangle}}{\langle X,X\rangle}\overline{\langle Y,X\rangle} + \frac{\langle X,Y\rangle}{\langle X,X\rangle}\cdot\frac{\overline{X,Y}}{\cancel{\langle X,X\rangle}}\cdot\cancel{\langle X,X\rangle} \\
&= \langle Y,Y\rangle - \frac{\langle X,Y\rangle\langle X,Y\rangle}{\langle X,X\rangle} - \cancel{\frac{\langle X,Y\rangle\langle X,Y\rangle}{\langle X,X\rangle}} + \cancel{\frac{\langle X,Y\rangle\langle X,Y\rangle}{\langle X,X\rangle}} \\
&= \|Y\|^2 - \frac{|\langle X,Y\rangle|^2}{\|X\|^2}
\end{aligned}$$

$$\implies 0 \leq \|Y\|^2 - \frac{|\langle X,Y\rangle|^2}{\|X\|^2}$$

$$\implies 0 \leq \|X\|^2\|Y\|^2 - |\langle X,Y\rangle|^2 \implies |\langle X,Y\rangle|^2 \leq \|X\|^2\cdot\|Y\|^2 \qquad \square$$

LECTURE 04/12

We are interested in knowing under what conditions equality occurs in Cauchy-Schwarz.

Lemma 4.19. *Equality holds in Cauchy-Schwarz whenever x,y are linearly dependent (i.e. $x \parallel y$).*

Proof. For $x = 0$: $0 = 0$.
For $x \neq 0$: Similar to Cauchy-Schwarz proof:
$0 \leq ||y_\perp||^2 = \|y\|^2 - \frac{|\langle x,y\rangle|^2}{\|x\|^2} \implies \|y\|^2 - \frac{|\langle x,y\rangle|^2}{\|x\|^2} = 0 \Leftrightarrow ||y_\perp|| = 0$.
We can conclude x,y must be linearly dependent since
$y_\perp = \lambda x - \frac{\langle x,\lambda x\rangle}{\langle x,x\rangle}x = \lambda x - \lambda\frac{\langle x,x\rangle}{\langle x,x\rangle}x = \lambda x - \lambda x = 0$
$\implies$ equality in Cauchy-Schwarz, since if x,y are linearly independent, $y_\perp \neq 0$ and we get a strict inequality in Cauchy-Schwarz.
So equality in Cauchy-Schwarz $\Leftrightarrow$ x,y linearly dependent. $\square$

We can use Cauchy-Schwarz to prove the triangle inequality, but first, we present a lemma:

Lemma 4.20. *Let $z \in \mathbb{C}$, then:*

(a) $z + \bar{z} = 2\mathbf{Re}(z)$.

(b) $|\mathbf{Re}(z)| \leq |z|$.

Proof. Let $z = x + iy$ and $\bar{z} = x - iy$, then

(a) $z + \bar{z} = x + iy + x - iy = 2x = 2\mathbf{Re}(z)$.

(b) $|\mathbf{Re}(z)| = |x| = \sqrt{x^2} \leq \sqrt{x^2 + y^2} = |z|$. □

Theorem 4.21. ***Triangle Inequality****. Let V be an inner product space over K. $\|x + y\| \leq \|x\| + \|y\|$, $\forall x, y \in V$.*

Proof.

$$\begin{aligned}
\|x+y\|^2 &= \langle x+y, x+y\rangle = \langle x,x\rangle + \langle x,y\rangle + \langle y,x\rangle + \langle y,y\rangle \\
&= \|x\|^2 + \langle x,y\rangle + \overline{\langle x,y\rangle} + \|y\|^2 \\
&= \|x\|^2 + 2\mathbf{Re}(\langle x,y\rangle) + \|y\|^2 \quad \text{by Lemma 4.20 (a)} \\
&\leq \|x\|^2 + 2|\mathbf{Re}(\langle x,y\rangle)| + \|y\|^2 \\
&\leq \|x\|^2 + 2|\langle x,y\rangle| + \|y\|^2 \quad \text{by Lemma 4.20 (b)} \\
&\leq \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 \quad \text{by Cauchy-Schwarz} \\
\|x+y\|^2 &\leq (\|x\| + \|y\|)^2 \\
\implies \|x+y\| &\leq \|x\| + \|y\|,
\end{aligned}$$

where $\|\cdot\|$ is a norm on any inner product space. □

4.8 *Angle in Inner Product Spaces over K*

To define angle for any inner product space, we use the definition commonly used in $\mathbb{R}^n$ as motivation:
$x, y \in \mathbb{R}^n, \langle x, y\rangle = x^t y$ is standard inner product on $\mathbb{R}^n$, then $\cos\theta = \frac{\langle x,y\rangle}{\|x\|\|y\|}$, where $0 \leq \theta \leq \pi$ is the angle between x and y.

Definition 4.17. Let V be a real inner product space with inner product $\langle , \rangle$, let $x, y \neq 0$ in V. Let θ be uniquely determined **angle** $\in [0, \pi]$, such that:

$$\cos\theta = \frac{\langle x, y\rangle}{\|x\|\|y\|}$$

We note that in $\cos\theta = \frac{\langle x,y\rangle}{\|x\|\|y\|}$, the fraction must $\in [-1, 1]$ by Cauchy-Schwarz. Moreover, we note that the angle between x, y is equal to the angle between y, x since $\langle , \rangle$ is symmetric.

Definition 4.18. Let V be a real inner product space. $\langle x, y\rangle$ are **orthogonal** $\Leftrightarrow \langle x, y\rangle = 0$. For $x, y \neq 0 \implies \theta(x, y) = \frac{\pi}{2}$.

This definition of orthogonality can carry over to complex inner product spaces.

Theorem 4.22. ***Pythagoras****. Let V be inner product space, $x, y \in V$ orthogonal, then $\|x + y\|^2 = \|x\|^2 + \|y\|^2$.*

Proof.
$$\begin{aligned}\|x+y\|^2 &= \langle x+y, x+y\rangle = \langle x+y, x+y\rangle \\ &= \langle x,x\rangle + \langle x,y\rangle + \langle y,x\rangle + \langle y,y\rangle \\ &= \|x\|^2 + \|y\|^2 \quad \text{since } (x \perp y)\end{aligned}$$
□

We note that Pythagoras applies to real as well as complex inner product spaces since we only used the definition of orthogonality.

4.9 *Fourier Series (speed topic)*

Consider V vector space of all continuous functions on $\mathbb{R}$ with inner product $\int_{-\pi}^{\pi} f(x)g(x)dx$.

Theorem 4.23. *Consider* $S \equiv \{1, \sin x, \cos x, \sin(2x), \cos(2x), \cdots\}$. *(Notice that* $1 = \cos(0 \cdot x)$.*) Functions* $\in S$ *are pairwise orthogonal.*

See Assignment 3, Exercise 3.

Proof. Using trigonometric identities and integration by parts, we can show that:

$$\int_{-\pi}^{\pi} \sin(nx)\cos(mx)dx = 0 \ \ \forall n \in \mathbb{N},\ m \in \mathbb{N}_0$$
$$\int_{-\pi}^{\pi} \sin(nx)\sin(mx)dx = 0 \ \ \text{if } m \neq n,\quad n,m \in \mathbb{N}$$
$$\int_{-\pi}^{\pi} \cos(nx)\cos(mx)dx = 0 \ \ \text{if } m \neq n,\quad n,m \in \mathbb{N}_0$$

We note that elements of S are only orthogonal, not orthonormal:

$$\int_{-\pi}^{\pi} \sin^2(nx)dx = \pi \ \ \forall n \in \mathbb{N}$$
$$\int_{-\pi}^{\pi} \cos^2(nx) = \pi \ \ \forall n \in \mathbb{N} \quad (-2\pi \text{for} n = 0)$$
$$\int_{-\pi}^{\pi} 1dx = 2\pi$$
□

Recall that in $\mathbb{R}^n$ with standard inner product with $v_1, \cdots, v_n$ pairwise orthogonal and non-zero, and $v = a_1v_1, \cdots, a_nv_n$, then we have that $a_j = \frac{\langle v_j, v\rangle}{\langle v_j, v_j\rangle}$. In the proof of this fact, we only use linearity in the second argument and positive definiteness, thus, this carries over to any inner product space.

Let $f(x)$ be a linear combination of the members of S, i.e.

$$f(x) = c_0 + [a_1\cos x + \cdots + a_n\cos(nx)] + [b_1\sin x + \cdots + b_n\sin(nx)]$$

With $f(x)$ known, we want to recover the coefficients, then:

$$a_n = \frac{\langle\cos(nx), f(x)\rangle}{\langle\cos(nx), \cos(nx)\rangle} = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos(nx)dx$$
$$b_n = \frac{\langle\sin(nx), f(x)\rangle}{\langle\sin(nx), \sin(nx)\rangle} = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin(nx)dx$$
$$c_0 = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)dx$$

We use the formulas above as motivation for the definition of the (infinite) Fourier series.

Definition 4.19. Let f be periodic of period 2π and continuous. For a_n, b_n, c_0 defined as above, we consider the infinite series, the **Fourier series of** f:

$$f(x) = c_0 + \sum_{n=1}^{\infty} \Big(a_n \cos(nx) + b_n \sin(nx) \Big)$$

Theorem 4.24. ***Dirichlet***. *Let f be periodic of period 2π, piecewise continuous, and piecewise continuously differentiable on $[-\pi, \pi]$, then the Fourier series of f converges at all points where f is continuous, and it represents f.*

Example 4.12. Let $f(x) = x^2, [-\pi, \pi]$ extended periodically. It satisfies the Dirichlet conditions and thus its Fourier series represents it:

$$x^2 = \frac{1}{3}\pi^2 + 4\sum_{n=1}^{\infty} \frac{(-1)^n}{n^2} \cos(nx) \ \forall x \in \mathbb{R}$$

The b_n's are all 0 since f is even. Especially, at discontinuity $x = \pi$:

$$x^2 = \frac{1}{3}\pi^2 + 4\sum_{n=1}^{\infty} \frac{(-1)^n}{n^2}(-1)^n = \frac{1}{3}\pi^2 + 4\sum_{n=1}^{\infty} \frac{1}{n^2}$$

$$\implies \sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}$$