MATH 361 - Week 1 Notes

Jasraj Sandhu

January 2024

Review

Definition (Vector Space)

A vector space V is a set, together with two operations (addition and scalar multiplication) satisfying for all vectors $u, v, w \in V$ and scalars α, β,

(1) $v + w = w + v$

(2) $v + (u + w) = (v + u) + w$

(3) $\exists 0 \in V$ such that $v + 0 = v$ for all v

(4) For all v, there exists $-v$ such that $v + (-v) = 0$

(5) $1 \cdot v = v$

(6) $(\alpha\beta) \cdot v = \alpha \cdot (\beta \cdot v)$

(7) $\alpha \cdot (v + w) = \alpha \cdot v + \alpha \cdot w$

(8) $(\alpha + \beta) \cdot v = \alpha \cdot v + \beta \cdot v$

- Properties (1) - (4) suggest that V is an abelian group under $+$. (An abelian group is a group in which the law of composition is commutative).

- **Note:** We will usually write αv instead of $\alpha \cdot v$.

- Here, the **scalars** belong to a **field** $\mathbb{F}$ (almost always $\mathbb{C}$, sometimes $\mathbb{R}$, also $\mathbb{Z}$ (mod p) where p is a prime, $\mathbb{Q}$, p-adics).

- We prefer $\mathbb{C}$ because of the **fundamental theorem of algebra**:

  - Every complex polynomial $p(z) = a_0 + a_1 z + \ldots + a_n z^n$ of degree n has n roots in $\mathbb{C}$. $\mathbb{R}$ is not algebraically complete since e.g. $x^2 + 1$ has no real roots.

- One important way this comes up is **eigenvalues**. The real matrix $\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ has characteristic polynomial $z^2 + 1$, which has no real roots. Recall that the characteristic polynomial of an $n \times n$ matrix A is defined as $C_A(\lambda) = \det(\lambda I_n - A)$.

- Over $\mathbb{C}$, any $n \times n$ matrix has n complex eigenvalues (possibly repeated).

- If V and W are both vector spaces, a map $T : V \to W$ is said to be **linear** if

  (i) $T(v_1 + v_2) = Tv_1 + Tv_2$

  (ii) $T(\alpha v) = \alpha T v$

  We typically write Tv instead of $T(v)$.

- **Aside:** Vector spaces form a **category** and the linear maps are the **morphisms** in the category.

- We associate two important spaces to each T:

  (1) nullspace/kernel

  $$\text{null } T = \ker T = \{v \in V : Tv = 0\} \ ,$$

  which is a subspace of V (the domain).

  (2) image/range

  $$\operatorname{Im} T = \text{ran } T = \{Tv : v \in V\}$$

  which is a subspace of W (the codomain).

- We use **linear map**, **linear transformation**, and **operator** interchangeably.

1. Let $V = \{f : [0,1] \to \mathbb{C} : f \text{ is differentiable and } f' \text{ is continuous on } [0,1]\}$ and $W = \{g : [0,1] \to \mathbb{C} : g \text{ is continuous}\}$. Define $T : V \to W$ by $Tf := f' \in W$ (here f is in V). Then

$$\begin{aligned} \ker T &= \{v \in V : Tv = 0\} \\ &= \{f : f' \equiv 0\} \\ &= \{f : f(x) = c \text{ for all } x, \text{ where } c \text{ is some constant}\} \end{aligned}$$

and

$$\begin{aligned} \text{ran } T &= \{Tv : v \in V\} \\ &= W \ . \end{aligned}$$

- All notions from MATH 311 concerning $\mathbb{R}$-vector spaces transfer to $\mathbb{C}$-vector spaces:
  - independence/span
  - subspaces and the subspace test
  - basis and dimension
  - rank (= dimension of range/image): Note that the image space of a matrix is the same as the column space of a matrix. Furthermore, for any matrix A, $\text{rank}(A) = \dim(\text{Col}(A)) = \dim(\text{Row}(A))$.
  - nullity (= dimension of kernel/nullspace)
  - rank-nullity

Rank-Nullity Theorem

Suppose V and W are $\mathbb{C}$-vector spaces with $\dim V = n$, $\dim W = m$, and a linear map $T : V \to W$. Then

$$n = \text{rank}\, T + \text{nullity}\ T\ .$$

- **Note:** Not all our vector spaces are finite dimensional. For example, continuous functions on $[0, 1]$ are infinite dimensional. That is, there is no finite list $\{f_1, \ldots, f_n\}$ of continuous functions so that **any** continuous function f can be expressed as a linear combination of the f_i.

- If V and W are both finite dimensional, then any linear map $T : V \to W$ can be expressed with a matrix with respect to bases selected for V and W.
  - Let $\{b_1, \ldots, b_n\}$ be a basis for V and $\{c_1, \ldots, c_m\}$ be a basis for W. Note that $\dim V = n$ and $\dim W = m$.
  - Find scalars $a_{11}, a_{21}, \ldots, a_{m1}$ so that

    $$Tb_1 = a_{11}c_1 + a_{21}c_2 + \ldots + a_{m1}c_m \ ,$$

    where $Tb_1 \in W$. Continue for $b_2, \ldots, b_n$ by finding a_{ij} so that

    $$Tb_j = a_{1j}c_1 + a_{2j}c_2 + \ldots + a_{mj}c_m$$

    for $j = 2, \ldots, n$.
  - Then $\big[T\big] = \big[a_{ij}\big]_{\substack{i=1,\ldots,m \\ j=1,\ldots,n}}$ is the $m \times n$ matrix representation of T with respect to these bases.
- For us, we "confuse" the notions of a linear map and any of its matrix representations by assuming that a linear map between finite dimensional spaces is a matrix acting on $\mathbb{C}$ where $n = \dim V$.

- **Fact:** If $\dim V = n$, then V is **isomorphic** (as a $\mathbb{C}$-vector space) to $\mathbb{C}^n$. That is, there is an injective ($\ker T = \{\vec{0}_V\}$) and surjective ($\text{ran } T = \mathbb{C}^n$) linear map $T : V \to \mathbb{C}^n$.

- To find T, let $\{b_1, \ldots, b_n\}$ be a basis for V and define

$$T(c_1 b_1 + \ldots + c_n b_n) = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} .$$

- We can show that T is an isomorphism.

  Proof. First, we show that T is injective; that is, $\ker T = \{\vec{0}_V\}$. Suppose $\vec{x} \in \ker T$. We know that

$$\ker T = \{\vec{v} \in V \mid T\vec{v} = \vec{0}_{\mathbb{C}^n}\} .$$

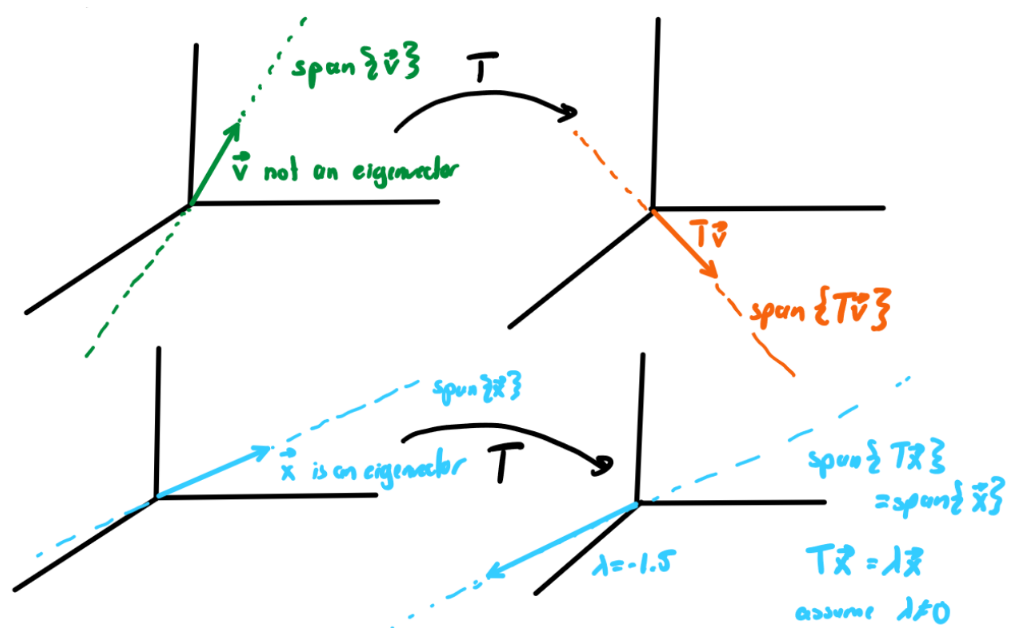
  Since $\vec{x} \in \ker T$, this means that $T\vec{x} = \vec{0}_{\mathbb{C}^n}$. $\square$

Spectral Theory

- Suppose $\dim V = n$ and $T : V \to V$ is linear (i.e. T is represented as an $n \times n$ matrix). $\lambda \in \mathbb{C}$ is an **eigenvalue** for T if there is a non-zero $\vec{x} \in V$ such that $T\vec{x} = \lambda\vec{x}$. Here, $\vec{x}$ is an **eigenvector** associated to λ.

- We call $\ker(T - \lambda I)$ the **eigenspace** for an eigenvalue λ. Every non-zero vector $\vec{x} \in \ker(T - \lambda I)$ is an eigenvector since $(T - \lambda I)\vec{x} = 0$ iff $T\vec{x} = \lambda I\vec{x} = \lambda\vec{x}$.

- Eigenvectors span the invariant lines (through the origin) for T.



- We find eigenvalues by computing the roots λ of the **characteristic equation** $C_A(z) = \det(A - Iz)$ (or $\det(Iz - A)$) and then compute the eigenspace by solving the system

$$(A - \lambda I)\vec{x} = \vec{0}\ .$$

  Note here that A is **any** matrix representation for T.

- We saw last time that the matrix $\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ has no **real** eigenvalues, but it does have **complex** eigenvalues $\lambda = \pm i$.

- Since $C_A(z)$ is a degree n polynomial, the fundamental theorem of algebra guarantees that it has n roots in $\mathbb{C}$. Therefore, we have the following theorem.

> **Theorem**
>
> Each linear map $T : V \to V$ has exactly $\dim V$ complex eigenvalues (including repetition).

> **Definition**
>
> A linear map $T : V \to V$ with $\dim V = n$ is called **diagonal** if there is a basis $\{\vec{b}_1, \ldots, \vec{b}_n\}$ for V and $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ with
>
> $$T\vec{b}_j = \lambda_j \vec{b}_j$$
>
> for all $j = 1, \ldots, n$. That is, if we take the matrix representation for A with respect to the basis $\{\vec{b}_1, \ldots, \vec{b}_n\}$, we have
>
> $$\mathrm{diag}(\lambda_1, \ldots, \lambda_n) := \begin{bmatrix} \lambda_1 & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & \lambda_n \end{bmatrix} = A\ .$$
>
> Note that necessarily $\{\vec{b}_1, \ldots, \vec{b}_n\}$ is a basis of **eigenvectors** for V.

- For an actual $n \times n$ complex matrix, we reserve the use of the word "diagonal" if the matrix is already expressed as a diagonal $A = \mathrm{diag}\{\lambda_1, \ldots, \lambda_n\}$.

- We say that an $n \times n$ matrix A is **diagonalizable** if there is a basis of $\mathbb{C}^n$ consisting of eigenvalues for A.

> **Theorem**
>
> An $n \times n$ matrix A is diagonalizable if and only if there is an invertible S with
>
> $$S^{-1}AS = D$$
>
> is a diagonal matrix

Proof. First suppose $\exists S$ so that $S^{-1}AS = D$ is diagonal. Write

$$D = \begin{bmatrix} \lambda_1 & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & \lambda_n \end{bmatrix}$$

and

$$S = \begin{bmatrix} \vec{b}_1 & \vec{b}_2 & \dots & \vec{b}_n \end{bmatrix} .$$

Since S is invertible, $\{\vec{b}_1, \dots, \vec{b}_n\}$ is a basis for $\mathbb{C}^n$. Since $S^{-1}AS = D$, we have that $AS = SD$. That is,

$$\begin{aligned} AS &= SD \\ A\begin{bmatrix} \vec{b}_1 & \dots & \vec{b}_n \end{bmatrix} &= \begin{bmatrix} \vec{b}_1 & \dots & \vec{b}_n \end{bmatrix} \begin{bmatrix} \lambda_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda_n \end{bmatrix} \\ \begin{bmatrix} A\vec{b}_1 & \dots & A\vec{b}_n \end{bmatrix} &= \begin{bmatrix} \lambda_1\vec{b}_1 & \dots & \lambda_n\vec{b}_n \end{bmatrix} \qquad (*) \end{aligned}$$

Comparing columns gives $A\vec{b}_j = \lambda_j\vec{b}_j$ for all $j = 1, \dots, n$. So, A is diagonal with respect to this basis $\{\vec{b}_1, \dots, \vec{b}_n\}$.

Conversely, suppose there exists a basis of eigenvectors $\{\vec{b}_1, \dots, \vec{b}_n\}$ corresponding to eigenvalues $\{\lambda_1, \dots, \lambda_n\}$... $\square$

1. Let $N = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ (this is an example of a **nilpotent** matrix). N is not diagonalizable! N is upper triangular, and so its diagonals entries $\{0, 0\}$ must be it's eigenvalues. If N were diagonalizable, then there would exist an invertible matrix S so that

$$S^{-1}NS = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} ,$$

which implies that

$$N = S \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} S^{-1} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} .$$

This is a contradiction since

$$N = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} .$$

A similar calculation shows that any matrix of the form

$$\begin{bmatrix} \lambda_1 & 1 & 0 & 0 & \dots & 0 \\ 0 & \lambda_2 & 1 & 0 & \dots & 0 \\ 0 & 0 & \lambda_3 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & 0 & 0 & \lambda_{n-1} & 1 \\ 0 & 0 & 0 & 0 & 0 & \lambda_n \end{bmatrix}$$

is never diagonalizable.

2. (Exponential of a Matrix) Recall the function $e^z = \sum_{n=0}^{\infty} \frac{z^n}{n!}$. This converges absolutely for all $z \in \mathbb{C}$. For an $n \times n$ matrix A, define

$$e^A := \sum_{n=0}^{\infty} \frac{A^n}{n!} \; .$$

(To define this, each partial sum $I + A + \frac{A^2}{2!} + \ldots + \frac{A^n}{n!}$ is well-defined. We can appeal to a future fact about norms to deduce convergence.) Now, suppose A is diagonalizable and write

$$S^{-1}AS = D = \begin{bmatrix} \lambda_1 & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & \lambda_n \end{bmatrix} \; .$$

Then

$$\begin{aligned} (S^{-1}AS)^m &= S^{-1}A^m S \\ &= D^m \\ &= \begin{bmatrix} \lambda_1^m & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & \lambda_n^m \end{bmatrix} \; , \end{aligned}$$

which implies that

$$\begin{aligned} \frac{A^m}{m!} &= \frac{SD^mS^{-1}}{m!} \\ &= S\frac{D^m}{m!}S^{-1} \\ &= S \begin{bmatrix} \lambda_1^m/m! & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & \lambda_n^m/m! \end{bmatrix} S^{-1} \; . \end{aligned}$$

Finally,

$$\begin{aligned} e^A &= \sum_{m=0}^{\infty} \frac{A^m}{m!} \\ &= \sum_{m=0}^{\infty} S \begin{bmatrix} \lambda_1^m/m! & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & \lambda_n^m/m! \end{bmatrix} S^{-1} \\ &= S \left(\sum_{m=0}^{\infty} \begin{bmatrix} \lambda_1^m/m! & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & \lambda_n^m/m! \end{bmatrix} \right) S^{-1} \end{aligned}$$

$$= S \begin{bmatrix} e^{\lambda_1} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & e^{\lambda_n} \end{bmatrix} S^{-1} \ .$$

Theorem

Suppose A is an $n \times n$ consisting of n **distinct** eigenvalues. Then A is diagonalizable.

Remark: The converse is false. For example, consider the 2×2 matrix $\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$. Here, the eigenvalues are $\lambda_1 = \lambda_2 = 2$, which are not distinct.

Proof. Let the eigenvalues of an $n \times n$ matrix A be denoted $\lambda_1, \ldots, \lambda_n$ and $\{\vec{b}_1, \ldots, \vec{b}_n\}$ be the associated eigenvectors. We must show that $\{\vec{b}_1, \ldots, \vec{b}_n\}$ is a basis. It is enough to show linear independence. We can prove by contradiction. Suppose to the contrary that the set $\{\vec{b}_1, \ldots, \vec{b}_n\}$ is linearly dependent. There is a minimal set for $j \in \{2, \ldots, n\}$ so that $\{\vec{b}_1, \ldots, \vec{b}_{j-1}\}$ is linearly independent, but $\{\vec{b}_1, \ldots \vec{b}_j\}$ is linearly dependent. That is, the j^{th} vector $\vec{b}_j$ is the vector that causes linear dependence. We can find $\alpha_1, \ldots, \alpha_j$ with $\alpha_j \neq 0$ and

$$\alpha_1 \vec{b}_1 + \ldots + \alpha_j \vec{b}_j = \vec{0} \ . \tag{*}$$

Multiplying (*) by A gives

$$\begin{aligned} A(\alpha_1 \vec{b}_1 + \ldots + \alpha_j \vec{b}_j) &= A\vec{0} \\ \begin{bmatrix} \lambda_1 & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & \lambda_n \end{bmatrix} (\alpha_1 \vec{b}_1 + \ldots + \alpha_j \vec{b}_j) &= \vec{0} \\ \alpha_1 \lambda_1 \vec{b}_1 + \ldots + \alpha_j \lambda_j \vec{b}_j &= \vec{0} \ . \end{aligned} \tag{1}$$

Also, multiplying (*) by λ_j gives

$$\alpha_1 \lambda_j \vec{b}_1 + \ldots + \alpha_j \lambda_j \vec{b}_j = 0 \ . \tag{2}$$

Then subtracting (2) from (1) gives

$$\alpha_1(\lambda_1 - \lambda_j)\vec{b}_1 + \ldots + \alpha_{j-1}(\lambda_{j-1} - \lambda_j)\vec{b}_{j-1} = \vec{0} \ ,$$

which implies that $\alpha_i(\lambda_i - \lambda_j) = 0$ for $i = 1, \ldots, j-1$, since we assumed the set $\{\vec{b}_1, \ldots, \vec{b}_{j-1}\}$ is linearly independent. Since the λ_i are all distinct, this implies that $\alpha_1 = \alpha_2 = \ldots = \alpha_{j-1} = 0$. Then (*) becomes $\alpha_j \vec{b}_j = \vec{0}$. This is a contradiction, since we assumed that α_j and $\vec{b}_j$ were both non-zero. $\square$