# Linear Algebra 1

Labix

January 15, 2024

**Abstract**

While the concepts of linear algebra lies in the heart of abstract algebra, it is so important that it has been investigate separately from abstract algebra.

# Contents

# 1   Vector Spaces

## 1.1   Introduction to Vector Spaces

Although the complete development of fields is given in an abstract algebra course, we give the definition of a field here for completeness.

---

**Definition 1.1.1: Fields**

A field $(\mathbb{F}, +, \cdot)$ is a triple where $+ : \mathbb{F} \times \mathbb{F} \to \mathbb{F}$ and $\cdot : \mathbb{F} \times \mathbb{F} \to \mathbb{F}$ such that they satisfy the following rules if $a, b, c \in \mathbb{F}$:
$(\mathbb{F}, +)$ is an abelian group.

- $a + (b + c) = (a + b) + c$

- There exists $0 \in \mathbb{F}$ such that $a + 0 = 0 + a = a$

- There exists $-a \in \mathbb{F}$ such that $a + (-a) = (-a) + a = 0$

- $a + b = b + a$

$(\mathbb{F} \setminus \{0\}, \cdot)$ is an abelian group.

- $a \cdot (b \cdot c) = (a \cdot b) \cdot c$

- There exists $1 \in \mathbb{F}$ such that $a \cdot 1 = 1 \cdot a = a$

- There exists $a^{-1} \in \mathbb{F}$ such that $a \cdot a^{-1} = a^{-1} \cdot a = 1$ if $a \neq 0$

- $a \cdot b = b \cdot a$

Distributive law.

- $a \cdot (b + c) = (a \cdot b) + (a \cdot c)$

---

**Definition 1.1.2: Vector Space**

A vector space $V$ over a field $\mathbb{F}$ is a set of elements $V$ together with $0 \in V$ and two binary opeartions $+ : V \times V \to V$ and $\cdot : F \times V \to V$, vector addition and scalar multiplcation respectively, satisfying the following with $a, b \in \mathbb{F}$ and $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{V}$.
$(V, +)$ is an abelian group.

- $\mathbf{u} + \mathbf{v} \in V$

- $\mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w}$

- There exists a vector $\mathbf{0}_V$ such that $\mathbf{0}_V + \mathbf{u} = \mathbf{u}$

- There exists an additive inverse $-\mathbf{u}$ such that $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$

- $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$

$\mathbb{F}$ acts on $V$ as a group action, with an identity in $V$.

- $a \cdot \mathbf{u} \in V$

- $a(b\mathbf{u}) = (ab)\mathbf{u}$

- There exists a vector $1_V$ such that $1_v \cdot \mathbf{u} = \mathbf{u}$

Distributive laws.

- $a(\mathbf{u} + \mathbf{v}) = a\mathbf{u} + a\mathbf{v}$

- $(a + b)\mathbf{u} = a\mathbf{u} + b\mathbf{u}$

---

> ### Proposition 1.1.3
>
> Let $a \in \mathbb{F}$ and $u \in V$ be a vector space over $\mathbb{F}$.
>
> - $a \cdot \mathbf{0}_V = \mathbf{0}_V$
>
> - $0 \cdot \mathbf{u} = \mathbf{0}_V$
>
> - $(-a)\mathbf{v} = -(a\mathbf{v}) = a(-\mathbf{v})$
>
> - $a\mathbf{v} = \mathbf{0}_V \implies a = 0$ or $\mathbf{v} = \mathbf{0}_V$
>
> - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
>
> *Proof.*
>
> - $a(\mathbf{0}_V) = a(\mathbf{0}_V + \mathbf{0}_V) = a\mathbf{0}_V + a\mathbf{0}_V$. Adding the additive inverse of $a\mathbf{0}_V$ on both sides gives our result.
>
> - $0\mathbf{u} = (0+0)\mathbf{u} = 0\mathbf{u} + 0\mathbf{u}$. Adding the additive inverse of $0\mathbf{u}$ on both sides gives our result.
>
> - Naturally $-(a\mathbf{v})$ is the inverse of $a\mathbf{v}$. Consider $a\mathbf{v} + a(-\mathbf{v})$.
>   $a\mathbf{v} + a(-\mathbf{v}) = a(\mathbf{v} - \mathbf{v}) = a\mathbf{0}_V = \mathbf{0}_V$. Thus $a(-\mathbf{v})$ is also the inverse of $a\mathbf{v}$ and $-(a\mathbf{v}) = a(-\mathbf{v})$. The same could be done to the third item with the other distributive law.
>
> - Suppose that $a \neq 0$. Then $\mathbf{v} = (a^{-1}a)\mathbf{v} = a^{-1}(a\mathbf{v}) = 0$.
>
> $\square$

> ### Proposition 1.1.4
>
> The additive identity, multiplicative identity, additive inverse of a vector space is unique.
>
> - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
>
> *Proof.* Suppose that $\mathbf{e}$ and $\mathbf{f}$ are additive identities. Then $\mathbf{e} + \mathbf{f} = \mathbf{e}$ and $\mathbf{e} + \mathbf{f} = \mathbf{f}$. Thus $\mathbf{e} = \mathbf{f}$. Suppose that $\mathbf{e}$ and $\mathbf{f}$ are multiplicative identities. Then $\mathbf{ef} = \mathbf{e}$ and $\mathbf{ef} = \mathbf{f}$ and $\mathbf{e} = \mathbf{f}$. Let $a \in V$. Suppose that $b, c \in V$ are additive inverses of $a$. Then
>
> $$a + b = a + c \implies b + a + b = b + a + c$$
> $$\implies (b+a) + b = (b+a) + c$$
> $$\implies b = c$$
>
> $\square$

## 1.2   Basis and Dimension

> ### Definition 1.2.1: Linearly Independent
>
> We say that a set of vectors $\{v_1, \ldots, v_n\}$ of a vector space $V$ over $\mathbb{F}$ are linearly independent if
>
> $$\sum_{k=1}^{n} a_k v_k = 0$$
>
> for $a_1, \ldots, a_n \in \mathbb{F}$ implies $a_1 = \cdots = a_n = 0$.

---

**Definition 1.2.2: Span**

We say that a set of vectors $\{v_1, \ldots, v_n\}$ of a vector space $V$ over $\mathbb{F}$ spans $V$ if for all $v \in V$, there exists $a_1, \ldots, a_n \in \mathbb{F}$ such that

$$v = \sum_{k=1}^{n} a_k v_k$$

---

**Definition 1.2.3: Basis**

We say that a set of vectors $\{v_1, \ldots, v_n\}$ of a vector space forms a basis for $V$ if they are linearly independent and spans $V$.

---

**Definition 1.2.4: Dimension**

We say that the dimension of a vector space $V$ is the number of elements in a basis of $V$. If a basis has $n$ elements, then we say that $\dim(V) = n$.

If $n$ is a finite number, then we say that $V$ is finite dimensional.

---

We have yet to shown that the dimension of a vector space is well defined since we do not know whether the cardinality of any two bases are the same. Therefore we have the following important theorem for finite dimensional vector space.

---

**Theorem 1.2.5: Steinitz Exchange Lemma**

Let $U, W$ be finite subsets of a finite dimensional vector space $V$. If $U$ is a set of linearly independent vectors and $W$ spans $V$, then

- $|U| \leq |W|$

- There exists a set $W' \subset W$ with $|W'| = |W| - |U|$ such that $U \cup W'$ spans $V$.

- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

*Proof.* Take $U = \{u_1, \ldots, u_m\}$ and $W = \{w_1, \ldots, w_n\}$. We will show that after reordering elements of $W$, we will have a set $\{u_1, \ldots, u_m, w_m + 1, \ldots, w_n\}$ that it spans $V$. We proceed by induction on $m$. Suppose that $m = 0$. In this case, $|U| \leq |W|$ necesssarily holds and by construction, $W$ already spans $V$.

Now suppose that the proposition is true for $m - 1$. By the induction hypothesis, we may reorder elements of $W$ so that $\{u_1, \ldots, u_{m-1}, w_m, \ldots, w_n\}$ spans $V$. Since $u_m \in V$, there exists $a_1, \ldots, a_n$ such that

$$u_m = \sum_{k=1}^{m-1} a_k u_k + \sum_{k=m}^{n} a_k w_k$$

At least one of $a_m, \ldots, a_n$ must be nonzero else the equality will contradict the linear independence of $u_1, \ldots, u_m$. This must mean that $m \leq n$.

Now by reordering $a_m w_m, \ldots, a_n w_n$, we may assume that $a_m \neq 0$. Thus we have that

$$w_m = \frac{1}{a_m} \left( u_m - \sum_{k=1}^{m-1} a_k u_k - \sum_{k=m+1}^{n} a_k w_k \right)$$

This means that $w_m$ lies in the span of $\{u_1, \ldots, u_m, w_{m+1}, \ldots, w_n\}$. Since this span contains each of the vectors $u_1, \ldots, u_{m-1}, w_m, \ldots, w_n$, by the inductive hypothesis it spans $V$.   $\square$

---

Clearly this implies that linearly independent sets of vectors must have cardinality less than sets of vectors that span $V$. By taking the highest cardinality of such linearly indendent set of vectors, and the

lowest cardinality of such sets of vectors that span $V$, we necessarily have that they are equal and thus is exactly the dimension of $V$.

We will discuss about dimensions and infinite dimensional vector spaces more in functional analysis. For the rest of the notes we will mostly go with finite dimensional vector spaces.

We now give a criterion with matrices to find whether a set of vectors span $V$ or whether they are linearly independent.

---

**Theorem 1.2.6**

Let $V$ be a vector space of dimension $n$ and $S = \{v_1, \ldots, v_n\} \subset V$. Then

- Elements of $S$ are linearly independent if and only if the row echelon form of $\begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix}$ has a leading one in every column

- Elements of $S$ span $V$ if and only if the row echelon form of $\begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix}$ has no zero rows

- $S$ is a basis of $V$ if and only if the row echelon form of $\begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix}$ is equal to the identity

---

## 1.3   Vector Subspaces

---

**Definition 1.3.1: Vector Subspaces**

A subset $U$ of a vector space $V$ is called a subspace of $V$ if $U$ is also a vector space.

---

**Proposition 1.3.2: Subspace Criterion**

$U$ is a subspace of $V$ if and only if $U$ is closed under vector addition and scalar multiplication and contains the zero vector.

- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

*Proof.* Suppose that $U$ is a subspace of $V$. Then necessarily $U$ is closed under vector addition and scalar multiplication and contains the zero vector.

Now suppose that the latter conditions are fulfilled by a subset $U$ of $V$. Then it is easy to see that $U$ satisfies all the criteria for being a vector space. □

---

**Proposition 1.3.3**

If $U_1$ and $U_2$ are subspaces of $V$ then $U_1 \cap U_2$ is also a subspace.

- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

*Proof.* Suppose that $\mathbf{v}, \mathbf{w} \in U_1 \cap U_2$. Then $\mathbf{v}, \mathbf{w} \in U_1$ and $U_2$. Since $U_1, U_2$ are subspaces, $\mathbf{v} + \mathbf{w} \in U_1$ and $U_2$ thus $\mathbf{v} + \mathbf{w} \in U_1 \cap U_2$. The proof is similar for scalar multiplication. □

---

**Definition 1.3.4: Sum of Subspaces**

Let $U, W$ be subspaces of the vector space $V$. Then define

$$U + W = \{\mathbf{u} + \mathbf{w} : \mathbf{u} \in U \text{ and } \mathbf{w} \in W\}$$

---

**Proposition 1.3.5**

Let $U, W$ be subspaces of a vector space $V$. Then $U + W$ is the smallest subspace of $V$ containing $U$ and $W$.

- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

*Proof.* We first show that $U + W$ is indeed a subspace of $V$. Suppose that $v \in U + W$. Then there exists $u \in U$ and $w \in W$ such that $v = u + w$. Then since $U$ and $W$ are closed individually under vector addition and scalar multiplication, any product and addition in $U + W$ can be decomposed into a sum of vectors in $U$ and $W$ and thus the new vector will also be able to be decomposed into $U$ and $W$ and thus lie in $U + W$.

Now suppose that $S$ is a subspace of $V$ containing $U$ and $W$. This means that any linear combination of elements of $U$ and $W$ are contained in $S$ thus $U + W \subseteq S$. This means that if any subspace containing $U$ and $W$ must also contain $U + W$, which means that $U + W$ is the smallest subspace containing $U$ and $W$. □

**Definition 1.3.6: Independent Subspaces**

Let $W_1, \ldots, W_n$ be subspaces of a vector space $V$. We say that $W_1, \ldots, W_n$ are independent if no vector of $W_i$ is a linear combination of the remaining subspaces for every $i \in \{1, \ldots, n\}$

**Definition 1.3.7: Direct Sum**

A vector space is the direct sum of its subspaces

$$V = W_1 \oplus \cdots \oplus W_n$$

if $W_1, \ldots, W_n$ are independent and $V = W_1 + \cdots + W_n$.

**Corollary 1.3.8**

If $V = W_1 \oplus \cdots \oplus W_n$ then

$$\dim(V) = \sum_{k=1}^{n} \dim(W_k)$$

- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

*Proof.* Each basis of $W_k$ are not contained in any other linear combination of all the basis of $W_1, \ldots, W_{k-1}, W_{k+1}, \ldots, W_n$. This means that the set of all the basis of $W_1, \ldots, W_n$ are linearly independent. Since they each span $W_k$ independently, the set of all the basis of $W_1, \ldots, W_n$ will span $W_1 \oplus \cdots \oplus W_n$ and thus is a basis of $V$. Thus we are done. □

## 1.4　Row and Column Ranks

The final section is devoted to matrices as we will soon see that matrices are particularly useful in a lot of things.

**Definition 1.4.1: Row Space**

Let $A_{m \times n}$ be a matrix. The row space of $A$ is the subspace of $\mathbb{F}^m$,

$$\text{span}\{r_1, \ldots, r_m\}$$

where $r_i$ are the rows of $A$. The row rank of $A$ is defined to be the dimension of the row space of $A$.

**Definition 1.4.2: Column Space**

Let $A_{m \times n}$ be a matrix. The column space of $A$ is the subspace of $\mathbb{F}^n$,

$$\text{span}\{c_1, \ldots, c_m\}$$

where $c_i$ are the columns of $A$. The column rank of $A$ is defined to be the dimension of the column space of $A$.

**Lemma 1.4.3**

Applying row operations does not change the row space, row rank of a matrix and column rank of a matrix.

**Theorem 1.4.4**

The row rank of a matrix is equal to the column rank.

We can now define the rank of a matrix without problem.

**Definition 1.4.5: Rank of a Matrix**

Define the rank of a matrix to be its row rank or column rank.

**Proposition 1.4.6**

Let $A$ be a $n \times n$ matrix. Then the following are equivalent.

- The rank of $A$ is $n$

- $A$ is invertible

- The rows of $A$ form a linearly independent set

- The columns of $A$ form a linearly independent set

# 2 Linear Maps

## 2.1 Properties of Linear Maps

### Definition 2.1.1: Linear Transformation

Let $V, W$ be vector spaces over $\mathbb{F}$. A linear transformation or linear map $T$ from $V$ to $W$ is a function $T : V \to W$ such that

- $T(v_1 + v_2) = T(v_1) + T(v_2)$ for all $v_1, v_2 \in V$

- $T(kv) = kT(v)$ for all $k \in \mathbb{F}$, $v \in V$

### Lemma 2.1.2

Let $T : V \to W$ be a linear map.

- $T(0_v) = 0_w$

- $T(-v) = -T(v)$ for all $v \in V$

*Proof.* Suppose that $v \in V$. Then $T(0 \cdot v) = 0 \cdot T(v) = 0$. Also we have that $T(0 - v) = T(0) - T(v) = -T(v)$. □

### Proposition 2.1.3

If $T : U \to V$ and $S : V \to W$ are linear then $S \circ T : U \to W$ is also linear.

*Proof.* Let $au + bv \in U$.

$$S \circ T(au + bv) = S(aT(u) + bT(v))$$
$$= a(S \circ T(u)) + b(S \circ T(v))$$

□

## 2.2 Isomorphisms

### Definition 2.2.1: Isomorphic Linear Maps

A linear map $T : V \to W$ is said to be an isomorphism if $T$ is bijective. In this case we also say that $V$ and $W$ are isomorphic.

### Theorem 2.2.2

Let $T : V \to W$ be an isomorphism of vector spaces $V, W$ over $F$. Then its inverse map $T^{-1} : W \to V$ is a linear map.

### Theorem 2.2.3

Let $T : V \to W$ be a linear map. Then the following are equivalent.

- $T$ is isomorphic

- If $v_1, \ldots, v_n \in V$ is a basis of $V$ then $T(v_1), \ldots, T(v_n) \in W$ is a basis of $W$

---

### Corollary 2.2.4

Every finite dimensional vector space is isomorphic to $\mathbb{R}^n$ for some $n \in \mathbb{N} \setminus \{0\}$.

- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

*Proof.* Direct consequence from the above. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

---

This corollary is especially important since it tells use that we only really need to study all of $\mathbb{R}^n$ to study all of finite dimensional spaces. Once we have our results on $\mathbb{R}^n$, we can translate it via an isomorphism.

---

### Proposition 2.2.5

Let $V, W$ be vetor spaces. The set of all linear maps from $V$ to $W$ forms a vector space. Denote it as $\mathcal{L}(V, W)$

---

## 2.3  Kernels and Images

---

### Definition 2.3.1: Images and Kernels

Let $T : U \to V$ be a linear map. The image of $T$ is defined as

$$\text{im}(T) = \{\mathbf{v} \in V | T(\mathbf{u}) = \mathbf{v}, \forall \mathbf{u} \in U\}$$

The kernel of $T$ is defined as
$$\ker(T) = \{\mathbf{u} \in U | T(\mathbf{u}) = \mathbf{0}_V\}$$

---

### Theorem 2.3.2

Let $T : U \to V$ be a linear map. Then

- $\text{im}(T)$ is a subspace of $V$
- $\ker(T)$ is a subspace of $U$

- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

*Proof.* Let $T : U \to V$ be a linear map.

- We prove that $\text{im}(T)$ is a subspace of $V$. Let $\mathbf{u}, \mathbf{v} \in \text{im}(T)$ and $a \in \mathbb{F}$. Since $\mathbf{u}, \mathbf{v} \in \text{im}(T)$, there exists $\mathbf{u}_0, \mathbf{v}_0 \in U$ such that $T(\mathbf{u}_0) = \mathbf{u}$ and $T(\mathbf{v}_0) = \mathbf{v}$. Note that $a\mathbf{u}_0 \in U$ and $\mathbf{u}_0 + \mathbf{v}_0 \in U$. Consider $T(a\mathbf{u}_0)$. We have $T(a\mathbf{u}_0) = aT(\mathbf{u}_0) = a\mathbf{u}$. Thus $a\mathbf{u} \in \text{im}(T)$. Similarly, $T(\mathbf{u}_0 + \mathbf{v}_0) = T(\mathbf{u}_0) + T(\mathbf{v}_0) = \mathbf{u} + \mathbf{v}$. Thus $\mathbf{u} + \mathbf{v} \in \text{im}(T)$. By the subspace criterion $\text{im}(T)$ is a subspace of $V$.

- We now prove that $\ker(T)$ is a subspace of $U$. Suppose that $u, v \in \ker(T)$ and $a, b \in \mathbb{F}$. Then $T(au + bv) = aT(u) + bT(v) = 0$. Thus $au + bv \in \ker(T)$.

$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

---

### Theorem 2.3.3

Let $A$ be the matrix representing a linear transformation. Let $B$ be the row reduced form of $A$. Then a basis of the image of the linear transformation is given by the columns in $A$ that has leading one in columns in $B$.

---

**Definition 2.3.4: Rank and Nullity**

Let $T : U \to V$ be a linear map.

- $\dim(\text{im}(T))$ is said to be the rank of $T$.

- $\dim(\ker(T))$ is said to be the nullity of $T$.

**Theorem 2.3.5: Rank Nullity Theorem**

Let $T : U \to V$ be a linear map. Then

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

**Theorem 2.3.6**

Let $T : U \to V$ be a linear map, where $\dim(U) = n$, $\dim(V) = m$. Let $e_1, \ldots, e_n$ be a basis of $U$. Then the rank of $T$ is equal to the largest size of a linearly independent subset of $T(e_1), \ldots, T(e_n)$.

## 2.4 Role of Matrices

**Definition 2.4.1: Matrix of a Linear Map**

Let $T : U \to V$ be a linear map where $\dim(U) = n$ and $\dim(V) = m$. Let $\mathbf{e}_1, \ldots, \mathbf{e}_n$ be the standard basis of $U$ and $\{\mathbf{f}_1, \ldots, \mathbf{f}_m\}$ the standard basis of $V$. Let

$$T(\mathbf{e}_i) = \sum_{k=1}^{m} \alpha_{ki} \mathbf{f}_k$$

for $i \in \{1, \ldots, n\}$ Define the matrix of this linear map to be

$$\begin{pmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{m1} & \alpha_{m2} & \cdots & \alpha_{mn} \end{pmatrix}$$

**Theorem 2.4.2**

Let $T : U \to V$ be a linear map. Let $A$ be the matrix of a linear map. Let $\mathbf{v} \in U$. Then the coordinates of $T(\mathbf{v})$ are given by

$$T(\mathbf{v}) = A\mathbf{v}$$

**Theorem 2.4.3**

The rank of a matrix equals the rank of any map that it represents.

**Theorem 2.4.4**

The composition of linear maps is represented by the matrix product of its representatives.

**Theorem 2.4.5**

$T$ is isomorphic if and only if its matrix is nonsingular.

# 3   Eigenspaces

Eigenspaces are invariants of a linear map. Every vector in the eigenspace will only be scaled while maintaining its direction.

## 3.1   Eigenvalues and Eigenvectors

**Definition 3.1.1: Eigenvalues and Eigenvectors**

Let $T : V \to V$ be a linear map, where $V$ is a vector space over $\mathbb{F}$. Suppose that for some non-zero vector $v \in V$, and some scalar $\lambda \in \mathbb{F}$, we have $T(v) = \lambda v$. Then $v$ is called an eigenvector of $T$, and $\lambda$ is called the eigenvalue of $T$.

Notice that $\lambda$ can in fact be 0. If this is the case, then the eigenvectors are just the vectors in the kernel.

**Definition 3.1.2: Eigenspace**

Let $\lambda$ be an eigenvalue of a linear map $T$. The set of all eigenvectors belonging to $\lambda$ is called an eigenspace of $T$ with respect to $\lambda$, denoted $E_\lambda$.

**Lemma 3.1.3**

Let $\lambda$ be an eigenvalue of $A$. Then
$$E_\lambda = \ker(A - \lambda I)$$

- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

*Proof.* Clearly since $Av = \lambda v$ for any eigenvector $v$ of $\lambda$, we also have that $(A - \lambda I)v = 0$ which means that $v \in \ker(A - \lambda I)$. $\qquad\square$

**Proposition 3.1.4**

Let $\lambda_1, \ldots, \lambda_r$ be distinct eigenvalues of $T : V \to V$, and let $v_1, \ldots, v_r$ be the corresponding eigenvectors. Then $v_1, \ldots, v_r$ are linearly independent.

As we can see, distinct eigenvalues are linearly independent. Considering the span of each eigenvectors, we can clearly see that each of their spans are independent.

**Proposition 3.1.5**

If $\lambda_1, \ldots, \lambda_n$ are distinct eigenvalues of a matrix $A$, then $E_{\lambda_1}, \ldots, E_{\lambda_n}$ are independent.

- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

*Proof.* Clear from the fact that the basis of eigenspaces of different eigenvalues are linearly independent. $\qquad\square$

**Definition 3.1.6: Characteristic Polynomial**

Let $A$ be an $n \times n$ matrix.
$$c_A(x) = \det(A - xI_n)$$
is called the characteristic polynomial of $A$.

**Proposition 3.1.7**

Let $A$ be an $n \times n$ matrix. Then $\lambda$ is an eigenvalue of $A$ if and only if

$$c_A(\lambda) = 0$$

> **Definition 3.1.8: Invariant Subspaces**
>
> Let $T : V \to V$ be a linear transformation. Let $U$ be a subspace of $V$. We say that $U$ is $T$-invariant if
> $$v \in U \implies T(v) \in U$$
> for all $v \in U$ or equivalently, $T(U) \subseteq U$.

> **Theorem 3.1.9**
>
> Eigenspaces is an invariant subspace under its linear transformation.

The main result of this subsection, stated that eigenspaces remain invariant under the linear transformation. Clearly this depends on the linear transformation. We will also show that this fact is also unchanged when considering different basis for the linear transformation.

## 3.2 Change of Basis

> **Definition 3.2.1: Change of Basis Matrix**
>
> Let $V$ be a vector space and $B, B'$ are two basis of $V$. A change of basis linear map is a linear map $T : V \to V$ such that $T : V_B \to V_{B'}$, meaning the old basis is mapped to the new basis.

> **Proposition 3.2.2**
>
> Let $V$ be a vector space and $v_1, \ldots, v_n$, $v'_1, \ldots, v'_n$ two distinct basis of $V$. Then
> $$v_k = p_{k1} v'_1 + \cdots + p_{kn} v'_n$$
> for all $k \in \{1, \ldots, n\}$ and for any vector $x$ in the basis $v_1, \ldots, v_n$, the vector in the other basis $x'$ is given by $x' = Px$ with the invertible matrix $P$
> $$\begin{pmatrix} p_{11} & \cdots & p_{1n} \\ \vdots & \ddots & \vdots \\ p_{n1} & \cdots & p_{nn} \end{pmatrix}$$

> **Theorem 3.2.3**
>
> Let $V, W$ be vector spaces. Let $V$ consists of two different basis $B$ and $B'$ with a map $P : V_B \to V_{B'}$. Similarly for $W$ we have $C$ and $C'$ and $Q : W_C \to W_{C'}$. Suppose $A : V_B \to W_C$ is a linear map. Then $A' : V_{B'} \to W_{C'}$ is given by
> $$A' = QAP^{-1}$$

> **Definition 3.2.4: Similar Matrices**
>
> We say that two matrices $A, B \in M_{n \times n}(\mathbb{R})$ are similar if there exists an invertible matrix $P \in M_{n \times n}(\mathbb{R})$ such that $B = PAP^{-1}$.

> **Lemma 3.2.5**
>
> The relation of similarity in matrices is an equivalent relation in $M_{n \times n}(\mathbb{R})$.

Similar matrices will play an important role. We will soon see that every matrix will be similar to relatively nice matrix so that their properties can be investigated, as well as making computations significantly easier.

## 3.3    Diagonalization

We now show a very nice kind of matrices, diagonal matrices that will come into play with linear maps. Our goal is to attempt to classify, by similarity, of every matrix into a diagonal one. We will soon see that this is not possible, and thus giving the last section of these notes meaning.

---

**Definition 3.3.1: Diagonalizable Linear Maps**

An linear map $T$ is diagonalizable if there exists a basis the matrix representation of $T$ is linear.

---

**Proposition 3.3.2: Diagonalizable Matrices**

A linear map $T$ represented by $A$ is diagonalizable if there exists an invertible $P$ and a diagonal matrix $D$ such that $P^{-1}AP = D$. In that case, $P$ consists of eigenvectors of $T$ and the diagonals of $D$ are the eigenvalues of $A$.

---

**Theorem 3.3.3**

If the linear map $T : V \to V$ has $n$ distinct eigenvalues where $\dim(V) = n$, then $T$ is diagonalizable.

---

Although not stated in the theorem, this does not mean that linear maps without $n$ distinct eigenvalues are not diagonalizable. However by taking the contrapositive, we see that not every linear map is diagonalizable because clearly, not every linear map has $n$ distinct eigenvalues.

---

**Theorem 3.3.4**

Let $T \in \mathcal{L}(V)$. Let $\lambda_1, \ldots, \lambda_m$ be the distinct eigenvalues of $T$. Then following are equivalent.

- $T$ is diagonalizable

- $V$ has a basis consisting of eigenvalues of $T$

- $V = E_{\lambda_1} + \cdots + E_{\lambda_m}$

- $\dim(V) = \dim(E_{\lambda_1}) + \cdots + \dim(E_{\lambda_m})$

---

# 4 The Jordan Canonical Form

In the last section, we looked into what kinds of matrices can have "nice" looking matrix under some basis. We now provide a less "nice" looking form of a similar matrix. However, every matrix can be reduced to this relatively "nice" looking form, as long as the field is algebraically closed. This form is called the Jordan Normal Form.

## 4.1 The Minimal Polynomial

---

**Theorem 4.1.1**

Let $\mathbb{F}$ be a field. Let $A$ be a $n \times n$ matrix over $\mathbb{F}$. Then there is some non-zero polynomial $p \in \mathbb{F}[x]$ of degree at most $n^2$ such that $p(A) = \mathbf{0}_n$.

---

*Proof.* Note that $\{I, A, \ldots, A^{n^2}\}$ is linearly dependent in the vector space of $n \times n$ matrices. Thus there exists constant $c_0, \ldots, c_{n^2}$ that are not all zero such that

$$c_0 I + \cdots + c_{n^2} A^{n^2} = \mathbf{0}_n$$

Thus $p(x) = c_0 + c_1 x + \cdots + c_{n^2} x^{n^2}$ is our desired polynomial. □

---

**Theorem 4.1.2**

Let $A_{n \times n}$ be a matrix over $\mathbb{F}$ representing the linear map $T : V \to V$. Then

- There is a unique monic non-zero polynomial $p(x)$ with minimal degree and coefficients in $\mathbb{F}$ such that $p(A) = \mathbf{0}_n$

- If $q(x)$ is any polynomial with $q(A) = \mathbf{0}_n$, then $p | q$

---

*Proof.* By the previous theorem, there exists a polynomial such that $p(A) = 0$. Divide the polynomial by $c_{n^2}$ gives us the desired monic polynomial. Suppose that $p_1, p_2$ are distinct monic polynomials that are minimal such that $p_1(A) = 0$ and $p_2(A) = 0$, then $p = p_1 - p_2$ is a non zero polynomial with a smaller degree and $p(A) = 0$, contradicting the minimality of degree. Thus $p$ is unique.

Let $p(x)$ be the minimal polynomial in the above proof. Let $q(A) = 0$. By division algorithm there exists some $r$ with smaller degree than $p$ such that $q = sp + r$. If $r$ is non-zero, then $r(A) = q(A) - s(A)p(A) = 0$, contradiction of minimality, thus $r = 0$ and $p | q$. □

---

**Definition 4.1.3: The Minimal Polynomial**

The unique monic non-zero polynomial $\mu_A(x)$ of minimal degree with $\mu_A(A) = \mathbf{0}_n$ is called the minimal polynomial of $A$.

---

**Proposition 4.1.4**

Similar matrices have the same minimal polynomial.

---

*Proof.* Similar matrices represent the same linear map, thus both have their minimal polynomial same as $T$, the linear map. □

---

**Proposition 4.1.5**

Let $D$ be a diagonal matrix with $\{d_1, \ldots, d_r\}$ its unqiue diagonal entries, then

$$\mu_D(x) = (x - d_1) \cdots (x - d_r)$$

- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

*Proof.* For any diagonal matrix,

$$p(D) = \begin{pmatrix} p(d_{11}) & 0 & \cdots & 0 \\ 0 & p(d_{22}) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & p(d_{nn}) \end{pmatrix}$$

Thus $p(D) = 0$ if and only if $p(d_{kk}) = 0$ for $k \in \{1, \ldots, n\}$. Thus the smallest-degree monic polynomial vanishing at these points is clearly the polynomial above. $\square$

---

**Corollary 4.1.6**

Every diagonalizable matrix has its minimal polynomial a product of distinct linear factors.

- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

*Proof.* Since diagonalizable matrix is similar to some diagonal matrix and they both have the same minimal polynomial, by the above proposition it is a product of distinct linear factors. $\square$

---

We will later see that in fact, the above criterion is a neccessary and sufficient condition: $A$ is diagonalizable if and only if the minimal polynomial is a product of distinct linear factors.

## 4.2  Cayley-Hamilton Theorem

---

**Theorem 4.2.1: Cayley-Hamilton**

Let $c_A(x) = \det(A - xI)$ be the characteristic polynomial of the $n \times n$ matrix $A$ over a field $\mathbb{F}$, then $c_A(A) = 0$.

- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

*Proof.* Firstly note that if $P(x) = \sum_{i=1}^{n} P_i x^i$ and $Q(x) = \sum_{j=1}^{m} Q_j x^j$ are polynomials with matrix coefficients where the matrix is $n \times n$, and $R(x) = \sum_{k=1}^{n+m} R_k x^k$ is the product of the two polynomials with $R_k = \sum_{i+j=k} P_i Q_j$, then if $M$ is a $n \times n$ matrix that commmutes with all of $Q_j$, then we have

$$R(M) = P(M)Q(M)$$

This can be seen by expanding the sums out.

Now take $Q(x) = A - xI$ and $P(x) = \text{adj}(Q)$. Then we have
$P(x)Q(x) = \det(A - xI)I = c_A(x)I$ by property of the adjoint. And since $A$ commutes with all coefficients of the polynomial of $Q$, we have

$$c_A(A)I = P(A)Q(A) = P(A) \cdot 0 = 0$$

Thus $c_A(A) = 0$. $\square$

---

**Corollary 4.2.2**

For any $A_{n \times n}$ over $\mathbb{F}$, we have $\mu_A | c_A$, and $\deg(\mu_A) \leq n$.

- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

*Proof.* This is clear since $c_A(A) = 0$ and $\mu_A$ is the minimal polynomial such that $\mu_A(A) = 0$ by division with remainder. Since $\deg(c_A) = n$, $\deg(\mu_A) \leq n$. $\square$

This lemma may help with finding out the minimal polynomial.

---

**Lemma 4.2.3**

Let $\lambda$ be an eigenvalue of $A$. Then $\mu_A(\lambda) = 0$.

- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

*Proof.* Let $v$ be an eigenvector of the eigenvalue $\lambda$ of $A$. Trivially $\mu_A(A)v = 0$. But also since

$$A^n v = \lambda^n v$$

we have $0 = \mu_A(A)v = \mu_A(\lambda)v$. Since $v$ is nonzero we must have $\mu_A(\lambda) = 0$. $\square$

---

In general, to deduce the formula for the minimal polynomial, we follow three steps.

Step 1: Find out the eigenvalues of the matrix.
Step 2: List out the possibilities of the minimal degree. This is done using the fact that $\mu_A(\lambda) = 0$ and $\deg(\mu_A) \leq n$.
Step 3: Plug in the matrix to find out which polynomial has its root at $A$.

There is another method to find out the formula using the following lemma.

---

**Lemma 4.2.4**

Let $T : V \to V$ be a linear map. Let

$$V = W_1 \oplus \cdots \oplus W_k$$

be the direct sum of invariant subspaces, meaning $W_1, \ldots, W_k$ are invariant subspaces of $T$. Let $\mu_i(x)$ be the minimal polynomial of $T|_{W_i}$. Then

$$\mu_T(x) = \mathrm{lcm}(\mu_1, \ldots, \mu_k)$$

---

Using this, we derive a better algorithm to find the minimal polynomial:

Step 1: Take $v \neq 0$ an eigenvector and set $W = \mathrm{span}\{v, T(v), T^2(v), \ldots\}$. Then $W$ is invariant under $T$. Let $d$ be the minimal positive integer such that $v, T(v), \ldots, T^d(v)$ are linearly dependent. Then $v, T(v), \ldots, T^{d-1}(v)$ are linearly independent. Then we know that $\mu_T(x)$ has degree larger than $d$ since else $\mu_T(x)v$ will never be 0. Then there is a nontrivial linear dependency relation of the form

$$T^d(v) + c_{d-1}T^{d-1}(v) + \cdots + c_1 T(v) + c_0 v = 0$$

Step 2: Consider the polynomial

$$x^d + c_{d-1}x^{d-1} + \cdots + c_1 x + c_0$$

Then this is precisely the minimal polynomial.

## 4.3 Generalized Eigenspace

---

**Definition 4.3.1: Generalized Eigenvector**

Let $T : V \to V$. Fix $k \in \mathbb{N}$. A non zero vector $v$ such that

$$(T - \lambda I)^k v = 0$$

---

is called a generalized eigenvector of $T$ with respect to the eigenvalue $\lambda$. Also we define

$$N_k(T, \lambda) = \{v \in V | (T - \lambda I)^k v = 0\} = \ker((T - \lambda I)^k)$$

to be the generalized eigenspace of index $k$ of $T$ with respect to $\lambda$. The set of all generalized eigenvector regardless of the index, is defined to be

$$G(T, \lambda) = \{v \in V | (T - \lambda I)^k v = 0 \text{ for some } k \in \mathbb{N}\} = \bigcup_{k=1}^{\infty} N_k(T, \lambda)$$

### Proposition 4.3.2

The dimensions of corresponding generalized eigenspaces of similar matrices are the same.

- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

*Proof.* This is true since generalized eigenspaces are defined without explicitly defining a basis for the linear map. Thus similar matrices that induce the same linear map will have the same dimensions for generalized eigenspaces. $\square$

### Proposition 4.3.3

Let $T : V \to V$ be a linear map with eigenvalue $\lambda$. Then

$$N_1(T, \lambda) \subseteq N_2(T, \lambda) \subseteq N_3(T, \lambda) \subseteq \ldots$$

- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

*Proof.* Trivially if $v \in \ker(A - \lambda I)^i$. This means that $(A - \lambda I)^i v = 0$ and $(A - \lambda I)^{i+1} v = 0$. Thus $N_i(T, \lambda) \subseteq N_{i+1}(T, \lambda)$ for any $i$ and we are done. $\square$

### Proposition 4.3.4

Let $\lambda$ be an eigenvalue of $T : V \to V$. There exists some $n \in N$ such that

$$N_n(T - \lambda I) = N_{n+1}(T - \lambda I) = \ldots$$

Denote $d(\lambda)$ the smallest of such $n$.

- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

*Proof.* $d(\lambda) \leq \dim(V)$. $\square$

### Proposition 4.3.5

Let $\lambda$ be an eigenvalue of $T : V \to V$. Then

$$G(T, \lambda) = N_{\dim(V)}(T, \lambda)$$

### Proposition 4.3.6

Let $T : V \to V$ with eigenvalues $\lambda_1, \ldots, \lambda_m$. Let $v_i \in G(T, \lambda_i)$ for $i \in \{1, \ldots, m\}$. Then $v_1, \ldots, v_m$ are linearly independent.

### Theorem 4.3.7

Let $V$ be a vector space of an algebraically closed field $F$. Let $T : V \to V$. Let $\lambda_1, \ldots, \lambda_m$ be distinct eigenvalues of $T$. Then

- $V = G(T, \lambda_1) \oplus \cdots \oplus G(T, \lambda_m)$

- $G(T, \lambda_j)$ is invariant under $T$.

## 4.4   Jordan Canonical Form

### Definition 4.4.1: Jordan Chain

A Jordan Chain of length $k$ is a sequence of nonzero vectors $v_1, \ldots, v_k$ such that

$$Av_1 = \lambda v_1$$
$$Av_2 = \lambda v_2 + v_1$$
$$\vdots$$
$$Av_k = \lambda v_k + v_{k-1}$$

for some eigenvalue $\lambda$ of $A$.

### Corollary 4.4.2

Let $v_1, \ldots, v_k$ be a Jordan Chain of $\lambda$. Then $v_i \in N_i(A, \lambda)$ for $i \in \{1, \ldots, k\}$ and $(T - \lambda I)(v_i) = v_{i-1}$ except for $i = 1$.

- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

*Proof.* The result is immediate from substitution in the Jordan Chains.  $\square$

### Proposition 4.4.3

The vectors in a Jordan chain are linearly independent.

### Proposition 4.4.4

The subspace spanned by a Jordan Chain is invariant under its linear map.

- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

*Proof.* Note that we just have to find out where $v_1, \ldots, v_k$ are mapped to since they are a basis of our subspace. But $T(v_i) = \lambda v_i + v_{i-1}$ for $i \in \{2, \ldots, k\}$ which is a linear combination of our basis, we must have that the subspace is invariant.  $\square$

### Definition 4.4.5: Jordan Block of Degree $k$

Define the Jordan block of degree $k$ to be the $k \times k$ matrix

$$\gamma_{ij} = \begin{cases} \lambda & \text{if } j = i \\ 1 & \text{if } j = i + 1 \\ 0 & \text{otherwise} \end{cases}$$

This means the diagonal of the matrix is $\lambda$ and the super diagonal is 1. It is denoted as $J_{\lambda, k}$

### Corollary 4.4.6

The matrix of $T$ with respect to the basis $v_1, \ldots, v_n$ is a Jordan Block if and only if $v_1, \ldots, v_n$ is a Jordan Chain.

- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

*Proof.* Let $v_1, \ldots, v_k$ be a Jordan Chain. Our matrix should be in the form

$$\begin{pmatrix} T(v_1) & T(v_2) & \cdots & T(v_k) \end{pmatrix}$$

Calculating each column gives

$$\begin{pmatrix} \lambda & 1 & 0 & 0 & 0 \\ 0 & \lambda & 1 & 0 & 0 \\ 0 & 0 & \ddots & \ddots & 0 \\ 0 & 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & 0 & \lambda \end{pmatrix}$$

For the other side, it is easy to simply compute $v_1, \ldots, v_k$ out with the matrix. $\square$

### Definition 4.4.7: Jordan Basis

A Jordan basis is a basis consisting of one or more Jordan chains strung together.

### Lemma 4.4.8

A Jordan Basis is indeed a basis.

- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

*Proof.* We naturally assume the string of Jordan Chains consists of $n$ vectors in total, corresponding to $\dim(V) = n$. Thus we just have to show linear independence. But this is also trivial. We have shown that vectors in the same Jordan Chain are independent, and vectors in different $G(T, \lambda)$ are proven to be linearly independent. $\square$

### Definition 4.4.9: Direct Sum

Let $A \in F^{n \times n}$ and $B \in F^{m \times m}$. Define the direct sum to be

$$A \oplus B = \begin{pmatrix} A & 0_{n \times m} \\ 0_{m \times n} & B \end{pmatrix}$$

### Lemma 4.4.10

Let $B, C$ be square matrices.

$$(B \oplus C)^n = B^n \oplus C^n$$

### Corollary 4.4.11

The matrix of $T$ with respect to a Jordan Basis is the direct sum

$$J_{\lambda_1, k_1} \oplus \cdots \oplus J_{\lambda_s, k_s}$$

of Jordan Blocks.

### Theorem 4.4.12

Let $A$ be a matrix over an algebraically closed field. Then there exists a Jordan basis for $A$. Moreover, $A$ is similar to some $J$ which is a direct sum of Jordan Blocks.

- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

*Proof.* We will construct the basis with 3 methods. They each contribute to part of the Jordan Basis. Firstly, we will obtain a basis of a subspace. We induct on $n$. The case of $n = 1$ is trivial.

Now suppose $T : V \to V$ is a linear map with $\dim(V) = n$. We want to find a restriction of $T$ that is an automorphism to apply the induction hypothesis. Fix $\lambda$ to be one of the eigenvalues of $T$. This will be used throughout the entire proof. This $\lambda$ is possible because the ground field is algebraically closed. Now I claim that $U = \text{im}(T - \lambda I)$ is invariant under $T$. If this is true, then $T|_U : U \to U$ can be used to apply induction hypothesis. So all we have to show is that $U$ is $T$-invariant and that $\dim(U) < \dim(V)$. The second item must be true by the rank nullity theorem. There must be an eigenvector in $\ker(T - \lambda I)$. Thus $\ker(T - \lambda I) \geq 1$ which implies that $\dim(U) < n$ by the rank nullity theorem. Now we prove that $U$ is invariant. Let $u \in U$, I show that $T(u) \in U$. If $u \in U$, then $u = (T - \lambda I)(v)$ for some $v \in V$, hence

$$T(u) = T(T - \lambda I)(v) = (T - \lambda I)(T(v)) \in \text{im}(T - \lambda I) = U$$

Thus we have proven that $T|_U : U \to U$. Apply induction hypothesis here to obtain a Jordan Basis for $T_U$. Call that Jordan Basis $e_1, \ldots, e_m$.

We now construct our second set of vectors. Recall that a Jordan Basis is a string of $l$ Jordan Chains. Let $v_1, \ldots, v_k$ denote one of the Jordan Chains. We can extend this Jordan Chain one more by setting $(T - \lambda I)^{k+1}(v_{k+1}) = 0$. Do the same thing for everey Jordan Chain, and relabel them to $w_1, \ldots, w_l$. As a side note, the two set of vectors we have now still form a Jordan Basis because we simply extended every Jordan Chain one more.

For the final set of vectors, observe that the first vector of each of the $l$ Jordan Chains are eigenvectors of $T|_U$ with its eigenvalue being $\lambda$. This is because by definition of Jordan Chains, $T|_U(v_1) = \lambda v_1$. Also note that those $l$ vectors are linearly independent. Thus the first vectors of each of the $l$ Jordan Chains span an $l$ dimensional subspace of the eigenspace of $\lambda$. Recall that the eigenspace of $\lambda$ has dimension $\dim(V) - \dim(U) = \dim(\ker(T - \lambda I))$. To minimize notation let $m = \dim(U)$. Thus by extension theorem we can extend the basis of the $l$ dimensional subspace to $\ker(T - \lambda I)$. Call the extension vectors $w_{l+1}, \ldots, w_{n-m}$. As a side note, these $n - m - l$ vectors each Jordan Chains of length 1. Thus we have complete our last set of vectors.

We have $n$ vectors
$$e_1, \ldots, e_m, w_1, \ldots, w_l, w_{l+1}, \ldots, w_{n-m}$$
Thus we only need to prove that they are linearly independent. Let $x = \sum_{k=1}^m \beta_m e_m$. Let

$$\sum_{i=1}^{n-m} \alpha_i w_i + x = 0$$

Applying $T - \lambda I$ on both sides give

$$\sum_{i=1}^{l} \alpha_i (T - \lambda I) w_i + (T - \lambda I)(x) = 0$$

Since $w_{l+1}, \ldots, w_{n-m}$ is in $\ker(T - \lambda I)$, they become 0. Now recall that our construction of $w_1, \ldots, w_l$ is made by extending our Jordan Chains. So applying $(T - \lambda I)$ moves down our Jordan Chain. This means that $(T - \lambda I)x$ no longer contains the last term of each Jordan Chain and are linear combinations of $\{e_1, \ldots, e_m\} \setminus \{\text{Last Term of each Jordan Chain}\}$, while all of the $(T - \lambda I)(w_1), \ldots, (T - \lambda I)(w_l)$ are all last members of ecah Jordan Chain. From the fact that $e_1, \ldots, e_m$ are a basis, we have $\alpha_1 = \cdots = \alpha_l = 0$.

Our sum now becomes $(T - \lambda I)x = 0$. Which means that $x \in \ker(T - \lambda I)$. Now our original sum becomes
$$\sum_{i=l+1}^{n-m} \alpha_i w_i + x = 0$$

By construction, $w_{l+1}, \ldots, w_{n-m}$ extends a basis of the eigenspace of $T|_U$ for $\lambda$, thus $\alpha_{l+1} = \cdots = \alpha_{n-m} = 0$. Also since $e_1, \ldots, e_m$ is a basis of $U$, we have $\beta_1 = \cdots = \beta_m = 0$. Finally by the above corollary, in a Jordan Basis, the matrix of $T$ is a direct sum of Jordan Blocks. $\qquad\square$

**Lemma 4.4.13**

If $A, B$ are similar, then they have the same JCF up to reordering of the Jordan Blocks by direct sum.

**Theorem 4.4.14**

Let $\lambda$ be an eigenvalue of a matrix $A$. Let $J$ be the Jordan Canonical Form of $A$. Then

- The number of Jordan Blocks of $J$ with eigenvalue $\lambda$ is equal to $\dim(\ker(A - \lambda I))$

- Let $k > 0$. Then number of Jordan Blocks of $J$ with eigenvalue $\lambda$ of degree at least $i$ is equal to $\dim(N_i(A, \lambda)) - \dim(N_{i-1}(A, \lambda))$

---

*Proof.* Since similar matrices have the same dimensions for their generalized eigenspaces corresponding to their eigenvalue, WLOG take $A = J = J_{\lambda_1, k_1} \oplus \cdots \oplus J_{\lambda_s, k_s}$. However, note that the dimension of $N_i(A \oplus B, \lambda)$ is equal to $\dim(N_i(A, \lambda)) + \dim(N_i(B, \lambda))$. So we just have to prove the theorem for a single Jordan Block.

Since $(J_{\lambda, k} - \lambda I)^i$ has a single diagonal line of ones $i$ places above the diagonal for $i < k$, and is 0 for $i \geq k$, the dimension of its kernel is $i$ for $0 \leq i \leq k$ and k for $i \geq k$. $\qquad\square$

**Corollary 4.4.15**

The JCF of a matrix is unique up to a reordering of the Jordan Blocks.

---

*Proof.* The above theorem says that the number of Jordan Blocks associated with $\lambda$ is determined by the nullity of $A$, and the size of every Jordan Block is determined by the dimension of the generalized eigenspaces. $\qquad\square$

## 4.5   Results of the Jordan Normal Theorem

**Lemma 4.5.1**

Let $M = A \oplus B$. Then
$$c_M(x) = c_A(x) c_B(x)$$
and
$$\mu_M(x) = \gcd(\mu_A(x), \mu_B(x))$$

**Proposition 4.5.2**

Let $A$ have JCF $J$. Let $\lambda$ be an eigenvalue of $A$. Consider the Jordan Blocks in $J$ related to $\lambda$. The string of Jordan Chains of these Jordan Blocks form a basis for $G(A, \lambda)$.

**Theorem 4.5.3**

Let $T : V \to V$ and $\lambda_1, \ldots, \lambda_m$ be the set of eigenvalues of $T$. Then the characteristic polynomial of $T$ is
$$c_A(x) = (-1)^n \prod_{k=1}^{m} (x - \lambda_k)^{a_k}$$
where $a_k$ is the sum of the degrees of the Jordan Blocks of $T$ of eigenvalue $\lambda_k$

### Theorem 4.5.4

Let $T : V \to V$ and $\lambda_1, \ldots, \lambda_m$ be the set of eigenvalues of $T$. Then the minimal polynomial of $T$ is

$$\mu_A(x) = \prod_{k=1}^{m} (x - \lambda_k)^{b_k}$$

where $b_k$ is the largest among the degrees of the Jordan Blocks of $T$ of eigenvalue of $\lambda_k$. Also, we have $d(\lambda_k) = b_k$

### Theorem 4.5.5

Let $T : V \to V$ and $\lambda_1, \ldots, \lambda_m$ be the set of eigenvalues of $T$. Then $T$ is diagonalizable if and only if $\mu_A(x)$ has no repeated factors.

### Theorem 4.5.6

Let $A, B \in M_{n \times n}(\mathbb{C})$. Then $A$ and $B$ are similar if and only if the following two are true:

- $A$ and $B$ have the same set of eigenvalues

- $\dim(\ker(A - \lambda I)^i) = \dim(\ker(B - \lambda I)^i)$ for all $i$ and eigenvalues $\lambda$

To finish this section, we show the process of determining the Jordan Canonical form of a matrix. The steps are usually as follows:

Step 1: Find out the nullity of $A - \lambda I$ as this gives us the number of Jordan Blocks with eigenvalue $\lambda$.
Step 2: To find out the number of Jordan blocks with eigenvalue $\lambda$ and size at least $i$, we calculate $\dim(\ker(A - \lambda I_n)^i) - \dim(\ker(A - \lambda I)^{i-1})$.

To find out the change of basis matrix, meaning the Jordan Chains, we do the following step:

Step 3: Find out the last one in the chain $v_k$ by solving $(A - \lambda I)^k v_k = 0$ while restricting $v_k$ such that $(A - \lambda I)^{k-1} v_k \neq 0$, and then proceed to find out $v_{k-1}$ by $(A - \lambda I)^{k-1} v_k = v_{k-1}$ and vice versa.

There are also extra information that we can use to determine the JCF:
The degree of $(x - \lambda)$ in $\mu_A$ indicates the maximum size of Jordan Blocks with eigenvalue $\lambda$.
The degree of $(x - \lambda)$ in $c_A$ indicates the total size of used in the JCF of all Jordan Blocks with eigenvalue $\lambda$.

We give an example of finding the JCF of a matrix.

### Example 4.5.7

Find the Jordan Canonical Form of

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

- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

*Proof.* We begin by finding out the eigenvalues of $A$. We have that
$c_A(x) = \det(A - xI) = (1 - x)^2(2 - x)$. This means that the eigenvalues are 1 and 2. Now we begin with step 1.

For eigenvalue 1, we have that row reduced form of $A - I$ is

$$\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

Thus the nullity of $A - I$ is 1. This means that the number of Jordan blocks with eigenvalue 1 is 1. Using information from $c_A$, we know that the total size used for the eigenvalue 1 is 2. This means that there is exactly one Jordan block of size 2 in the JCF of $A$.

This leaves the fact that the remaining Jordan block of size 1 being the eigenvalue 2.

With this, we complete the JCF of $A$ with

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

To compute the basis, or the change of basis matrix $P$, we use step 3. Since $A$ has one Jordan block for eigenvalue 1, we need to find one string of Jordan chain. This chain needs to have length 2 since the size of the Jordan block is 2. (If there are multiple of Jordan blocks of the same eigenvalues, the end vector of the Jordan chains needs to be linearly indendent). We begin by finding the ending of the chain, $v_2$ by using the fact that $(A - I)^2 v_2 = 0$ and $(A - I)v_2 = v_1 \neq 0$. We have that

$$(A - I)^2 = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 1 & 1 \end{pmatrix}$$

We choose that $v_2 = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}$. Now we have

$$v_1 = (A - I)v_2 = \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix}$$

Finally, we choose $v_3 \in \ker(A - 2I)$. But row reducing $A - 2I$ gives

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

We can choose $v_3 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$. This means that

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

$\square$

## 4.6    Functions of Matrices

We start of the last section with a formula for the $n$th power of a Jordan Block.

**Lemma 4.6.1**

Let $J_{\lambda,k}$ be a Jordan Block. Then

$$J_{\lambda,k}^n = \begin{pmatrix} \lambda^n & n\lambda^{n-1} & \cdots & \binom{n}{k-2}\lambda^{n-k+2} & \binom{n}{k-1}\lambda^{n-k+1} \\ 0 & \lambda^n & \cdots & \binom{n}{k-3}\lambda^{n-k+3} & \binom{n}{k-2}\lambda^{n-k+2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \lambda^n & n\lambda^{n-1} \\ 0 & 0 & \cdots & 0 & \lambda^n \end{pmatrix}$$

When used together with the fact that powers of matrices can be distributed through direct sums, we obtain the formula for finding powers of matrices. Namely if $A$ has Jordan canonical form $J$ and $A = PJP^{-1}$, then $A^n = PJ^nP^{-1}$.

We now define functions of matrices in terms of this decomposition.

**Definition 4.6.2: Functions of Matrices**

Let $f$ be a function over $\mathbb{C}$. For every matrix $A \in M_{n \times n}(\mathbb{C})$, define $f(A)$ by

$$f(A) = Pf(J)P^{-1}$$

where $f(J) = f(J_{\lambda_1,k_1}) \oplus \cdots \oplus f(J_{\lambda_t,k_t})$ is the direct sum of Jordan blocks. And finally, for each Jordan block $J_{\lambda,k}$, define $f(J_{\lambda,k})$ by

$$f(J_{\lambda,k}) = \begin{pmatrix} f(\lambda) & f'(\lambda) & \cdots & \frac{1}{(k-2)!}f^{(k-2)}(\lambda) & \frac{1}{(k-1)!}f^{(k-1)}(\lambda) \\ 0 & f(\lambda) & \cdots & \frac{1}{(k-3)!}f^{(k-3)}(\lambda) & \frac{1}{(k-2)!}f^{(k-2)}(\lambda) \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & f(\lambda) & f'(\lambda) \\ 0 & 0 & \cdots & 0 & f(\lambda) \end{pmatrix}$$

Note that taking $f(x) = x^n$, this coincides with the power of a Jordan block, which is presumably what motivated the this definition of matrix functions.