General Vector Spaces

8.1 Fields

Def 8.1.2 A **field** consists of

1. a nonempty set $\mathbb{F}$
2. an operation of *addition* $a+b$ for all $a, b \in \mathbb{F}$
3. an operation of *multiplication* ab for all $a, b \in \mathbb{F}$

Axioms:

1. Closure under addition
2. Commutative Law for addition
3. Associative Law for addition
4. Existence of the Additive Identity 0
5. Existence of Additive Inverse
6. Closure under multiplication
7. Commutative Law for multiplication
8. Associative Law for multiplication
9. Existence of Multiplicative Identity 1
10. Existence of Multiplicative Inverse a^{-1}
11. Distributive Law

Prop 8.1.5 Additional properties of field

1. Uniqueness of Additive Identity
2. Uniqueness of Additive Inverse
3. Uniqueness of Multiplicative Identity
4. Uniqueness of Multiplicative Inverse
5. For any $a \in \mathbb{F}$, $a0 = 0$ and $(-1)a = -a$
6. For any $a, b \in \mathbb{F}$, if $ab = 0$, then $a = 0$ or $b = 0$.

Prop 8.1.11 Define the trace of A, denoted $tr(A)$, as the sum of the entries on the diagonal of A.

1. If A and B are $n \times n$ matrix over $\mathbb{F}$, then $tr(A+B) = tr(A) + tr(B)$
2. If $c \in F$ and A is an $n \times n$ matrix over $\mathbb{F}$, then $tr(cA) = c \cdot tr(A)$
3. If C and D are $m \times n$ and $n \times m$ matrices, respectively over $\mathbb{F}$, then $tr(CD) = tr(DC)$

8.2 Vector Spaces

Def 8.2.2. A vector space consists of

1. a field $\mathbb{F}$, where the elements are called *scalars*
2. a nonempty set V, where the elements are called *vectors*
3. an operation of *vector addition* $u+v$ for all $u, v \in V$
4. an operation of *scalar multiplication* cu for all $c \in \mathbb{F}, u \in V$

Axioms:

1. Closure under VA
2. Commutative Law for VA
3. Associative Law for VA
4. Existence of the Zero Vector
5. Existence of Additive Inverse
6. Closure under SM
7. For all $b, c \in \mathbb{F}$ and $u \in V$, $b(cu) = (bc)u$
8. For all $u \in V$, $1u = u$
9. (Distributive Law I) For all $c \in \mathbb{F}$ and $u, v \in V$, $c(u + v) = cu + cv$
10. (Distributive Law II) For all $b, c \in \mathbb{F}$ and $u \in V$, $(b + c)u = bu + cu$

Prop 8.2.4 Additional properties of a vector space

1. Uniqueness of the Zero Vector
2. Uniqueness of the Additive Inverse
3. For all $u \in V$, $0\mathbf{u} = \mathbf{0}$ and $(-1)u = -u$
4. For all $c \in \mathbb{F}$, $c\mathbf{0} = \mathbf{0}$
5. If $cu = 0$ where $c \in \mathbb{F}$ and $u \in V$, then $c = 0$ or $u = 0$.

8.3 Subspaces

Def 8.3.2 A subset W of a vector space V is called a **subspace** of V if W is itself a vector space using the same VA and SM as in V.

Note: $\{\mathbf{0}\}$ and V are called *trivial subspaces* of V. Other subspaces are called *proper subspaces* of V.

Thm 8.3.4 A subset W of V is a subspace of V only if it satisfies property 4, 1, 6 of Def 8.2.2

Remark 8.3.5. A nonempty subset W of V is a subspace of V iff for all $a, b \in \mathbb{F}$ and $u, v \in W$, then $au + bv \in W$

Thm 8.3.8 The intersection of 2 subspaces is a subspace.

Note: The union of two subspaces might not be a vector space, e.g. $W_1 = \{(x, 0) \mid x \in \mathbb{F}\}$ and $W_2 = \{(0, y) \mid y \in \mathbb{F}\}$

Def 8.3.11 Let W_1 and W_2 be subspaces of a vector space V. The **sum** of W_1 and W_2 is defined to be the set $W_1 + W_2 = \{w_1 + w_2 \mid w_1 \in W_1, w_2 \in W_2\}$

Thm 8.3.12 The sum of subspaces is a subspace.

Remark 8.3.14 Let W_1 and W_2 be subspaces of a vector space V. Then $W_1 + W_2$ is the smallest subspace of V that contains both W_1 and W_2.

Tut 1. Let W_1 and W_2 be subspaces of a vector space V, then $W_1 \cup W_2$ is a subspace of V iff $W_1 \subseteq W_2$ or $W_2 \subseteq W_1$.

8.4 Linear Spans and Linear Independence

Let V be a vector space over a field $\mathbb{F}$ and B be a nonempty subset of V

Def 8.4.2 Let $v_1, v_2, \ldots, v_m \in V$. For any scalars $c_1, c_2, \ldots c_m$, the vector $c_1v_1 + c_2v_2 + \ldots c_mv_m$ is called a **linear combination** of $v_1, v_2, \ldots, v_m$. (Note: m is finite)

Thm 8.4.3 The set W of all linear combinations of (finite) vectors taken from B is a subspace of V

Def 8.4.4 The subspace W in Thm 8.4.3 is called the *subspace of V spanned by B* and we write $W = \text{span}_{\mathbb{F}}(B)$. We also say W is a *linear span* of B and B *spans* W. Note that $B \subseteq W$

Def 8.4.8

1. Let $v_1, v_2, \ldots, v_k \in V$. The vectors are **linearly independent** if the vector equation $c_1v_1 + c_2v_2 + \ldots c_kv_k = \mathbf{0}$ has only the trivial solution $c_1 = c_2 = \cdots = c_k = 0$
2. B is **linearly independent** if for every finite subset $\{v_1, v_2, \ldots, v_k\}$ of B, $v_1, v_2, \ldots, v_k$ are linearly independent.

Remark 8.4.9 Linear independence is used to determine whether there are redundant vectors in a set.

8.5 Bases and Dimensions

Def 8.5.1 A subset B of a vector space V is called a **basis** for V if B is lin. ind. and B spans V

A vector space V is called *finite dimensional* if it has a basis consisting of finitely many vectors; otherwise, V is called *infinite dimensional*

Remark 8.5.2

1. For convenience, the empty set $\emptyset$ is defined as the basis for a zero space
2. Every vector space has a basis. Proof by Zorn's Lemma

Lemma 8.5.5 Let V be a finite dimensional vector space and $B = \{v_1, v_2, \ldots, v_n\}$ a basis for V. Any vector $u \in V$ can then be expressed *uniquely* as a linear combination of vectors in B.

Def 8.5.6 Let V be a finite dimensional vector space over a field $\mathbb{F}$ where V is not a zero space

1. A basis $B = \{v_1, v_2, \ldots, v_n\}$ for V is called an *ordered basis* if the vectors in B have a fixed order s.t. v_1 is the first vector, v_2 is the second vector etc.
2. Let $B = \{v_1, v_2, \ldots, v_n\}$ be an ordered basis for V and let $u \in V$. If $u = c_1v_1 + c_2v_2 + \ldots c_nv_n$ for $c_1, c_2, \ldots, c_n \in \mathbb{F}$ then the coefficients $c_1, c_2, \ldots, c_n$ are called the **coordinates** of u relative to the basis B. In particular the vector $(u)_B = (c_1, c_2, \ldots, c_n)$ or $[u]_B = (c_1, c_2, \ldots, c_n)^T$ in $\mathbb{F}^n$ is called the **coordinate vector** of u relative to the basis B.

Lemma 8.5.7 Let V be a finite dimensional vector space over a field $\mathbb{F}$ where V is not a zero space and let B be its ordered basis.

1. For any $u, v \in V$, $u = v$ iff $(u)_B = (v)_B$
2. For any $v_1, v_2, \ldots v_r \in V$ and $c_1, c_2, \ldots c_r \in \mathbb{F}$, $(c_1v_1 + c_2v_2 + \cdots c_rv_r)_B = c_1(v_1)_B + c_2(v_2)_B + \cdots c_r(v_r)_B$

Thm 8.5.10 Let V be a vector space with a basis of n vectors. Then

1. any subset of V with more than n vectors is always linearly dependent
2. any subset of V with less than n vectors cannot span V.

Def 8.5.11 The **dimension** of a finite dimensional vector space V over a field $\mathbb{F}$ denoted by $dim_{\mathbb{F}}(V)$ is defined to be the number of vectors in a basis for V. In addition, we define the dimension of a zero space to be 0.

Thm 8.5.13 Let V be a finite dimensional vector space and B a subset of V. The following are equivalent:

1. B is a basis for V
2. B is linearly independent and $|B| = dim(V)$
3. B spans V and $|B| = dim(V)$

Thm 8.5.15 Let W be a subspace of a finite dimensional vector space V. Then

1. $dim(W) \leq dim(V)$
2. if $dim(W) = dim(V)$, then $W = V$.

Remark:

1. Row equivalent matrices have the same rowspace but may not have the same column space
2. Let A be a matrix and R a row-echelon form of A. A basis for the column space of A can be obtained by taking the columns of A that correspond to the pivot columns in R. A basis for the rowspace of A can be obtained by taking the set of nonzero rows in R (or the corresponding rows in A).

Thm 8.5.17 Let V be a finite dimensional vector space. Suppose C is a linearly independent subset of V. Then there exists a basis B for V s.t. $C \subseteq B$.

To extend a basis, put the current basis in rows, do Gaussian elimination, and add (e_i) if column i is not a pivot column.

Tut 2. Let W_1 and W_2 be finite dimensional subspaces of a vector space.

1. Let B_1, B_2 be basis for W_1, W_2 respectively. Then $\operatorname{span}(B_1 \cup B_2) = W_1 + W_2$
2. $\dim(W_1 + W_2) = \dim(W_1) + \dim(W_2) - \dim(W_1 \cap W_2)$

8.6 Direct Sums of Subspaces

Let W_1 and W_2 be subspaces of a vector space V.

Def 8.6.3 We say that the subspace $W_1 \oplus W_2$ is a **direct sum** of W_1 and W_2 if every vector $u \in W_1 + W_2$ can be expressed **uniquely** as $u = w_1 + w_2$ where $w_1 \in W_1$ and $w_2 \in W_2$.

Theorem 8.6.5 $W_1 + W_2$ is a direct sum iff $W_1 \cap W_2 = \{\mathbf{0}\}$

Tut 1 Q4. $W_1 + W_2$ is the smallest subspace of V that contains both W_1 and W_2.

Tut 1 Q5. $W_1 \cup W_2$ is a subspace of V iff $W_1 \subseteq W_2$ or $W_2 \subseteq W_1$

Theorem 8.6.7 Suppose $W_1 + W_2$ is a direct sum.

1. If B_1 and B_2 are bases for W_1 and W_2 respectively, then $B_1 \cup B_2$ is a basis for $W_1 \oplus W_2$
2. If both W_1 and W_2 are finite dimensional, then dim($W_1 \oplus W_2$) = dim(W_1) + dim(W_2)

Remark 8.6.8 Even if $W_1 + W_2$ is not a direct sum, it is still true that span($B_1 \cup B_2$) = $W_1 + W_2$.

Def 8.6.9 We can generalize the definition of **sum** and **direct sum** to k subspaces of V.
$W_1 + \cdots + W_k$ is a direct sum iff $W_1 \cap W_2 = \{\mathbf{0}\}, (W_1 + W_2) \cap W_3 = \{\mathbf{0}\}, \ldots$, etc

8.7 Cosets and Quotient Spaces

Let W be a subspace of a vector space V.

Def 8.7.1 For $\mathbf{u} \in V$, the set $W + \mathbf{u} = \{w + \mathbf{u} \mid w \in W\}$ is called the **coset of W containing $\mathbf{u}$**

Thm 8.7.3

1. For any $v, w \in V$, the following are equivalent:
   1. $v \in W + w$
   2. $w \in W + v$
   3. $v - w \in W$
   4. $W + v = W + w$
2. For any $v, w \in V$, either $W + v = W + w$ or $(W + v) \cap (W + w) = \emptyset$

Lemma 8.7.5

1. Suppose $u_1, u_2, v_1, v_2 \in V$ s.t. $W + u_1 = W + u_2$ and $W + v_1 = W + v_2$. Then $W + (u_1 + v_1) = W + (u_2 + v_2)$
2. Suppose $u_1, u_2 \in V$ s.t. $W + u_1 = W + u_2$. Then $W + cu_1 = W + cu_2$ for all $c \in \mathbb{F}$

Def 8.7.6

1. We define the *addition* of two cosets by $(W + u) + (W + v) = W + (u + v)$ for $u, v \in V$.
2. We define the *scalar multiplication* of a coset by $c(W + u) = W + cu$ for $c \in \mathbb{F}$ and $u \in V$.

Thm 8.7.8 Denote the set of all cosets of W in V by $V/W = \{W + u \mid u \in V\}$. Then V/W is a vector space over $\mathbb{F}$ using the addition and scalar multiplication defined in Def 8.7.6

Its zero vector is $W (= W + \mathbf{0})$.

Def 8.7.9 The vector space V/W is called the **quotient space of V modulo W**

Thm 8.7.11 Let $\{w_1, w_2, \ldots, w_m\}$ be a basis for W.

1. For $v_1, v_2, \ldots, v_k \in V$, $\{v_1, v_2, \ldots, v_k, w_1, w_2, \ldots, w_m\}$ is a basis for V iff $\{W + v_1, W + v_2, \ldots, W + v_k\}$ is a basis for V/W
2. dim(V/W) = dim(V) - dim(W)

9 General Linear Transformations

9.1 Linear Transformations

Let V and W be vector spaces over the same field $\mathbb{F}$.

Def 9.1.2 A *linear transformation* $T : V \to W$ is a mapping from V to W s.t.

1. For all $u, v \in V$, $T(u+v) = T(u) + T(v)$
2. For all $c \in \mathbf{F}$ and $u \in V$, $T(cu) = cT(u)$.

If $W = V$, the linear transformation $T : V \to V$ is called a *linear operator* on V
If $W = \mathbb{F}$, the linear transformation $T : V \to \mathbb{F}$ is called a *linear functional* on V

Remark 9.1.3 A mapping $T : V \to W$ is a linear transformation iff $T(au + bv) = aT(u) + bT(v)$ for all $a, b \in \mathbb{F}$ and $u, v \in V$.

Prop 9.1.5 If $T : V \to W$ is a linear transformation, then $T(\mathbf{0}) = \mathbf{0}$.

Remark 9.1.6 Suppose V has a basis B and let T be a linear transformation.

- T is completely determined by the images of vectors from B, i.e. $T(u)$ is completely determined by $T(v_1), \ldots T(v_m)$ for $v_1, \ldots, v_m \in V$
- We can define a linear transformation $S : V \to W$ by using the basis B, i.e. setting the value of $S(v)$ for all $v \in B$

9.2 Matrices for Linear Transformations

- $[u]_B$ denotes *column* coordinate vector relative to B. $(u)_B$ denotes *row* coordinate vector relative to B.
- Let $T : V \to W$ be a linear transformation where V and W are finite dimensional vector spaces over a field $\mathbb{F}$ s.t. $n = dim(V) \geq 1$ and $m = dim(W) \geq 1$.
- B and C are ordered bases for V and W respectively

Thm 9.2.1 Let $T : V \to W$ be a linear transformation. For any ordered bases B and C for V and W respectively, there exists an $\dim(W) \times \dim(V)$ matrix $\mathbf{A}$ s.t. $[T(u)]_C = \mathbf{A}[u]_B$ for all $u \in V$

Def 9.2.2. Let $B = \{v_1, v_2, \ldots, v_n\}$. The matrix $\mathbf{A} = ([T(v_1)]_C \quad [T(v_2)]_C \quad \cdots \quad [T(v_n)]_C)$ is called the *matrix for T relative to the ordered bases B and C*. This matrix $\mathbf{A}$ is usually denoted by $[T]_{C,B}$.

Note that $[T(u)]_C = [T]_{C,B}[u]_B$ for all $u \in V$

If $W = V$ and $C = B$, we simply denote $[T]_{B,B}$ by $[T]_B$ and called it the matrix for T relative to the ordered basis B.

Lemma 9.2.3 Let T_1 and T_2 be a linear transformation. For any B and C, we have $T_1 = T_2$ iff $[T_1]_{C,B} = [T_2]_{C,B}$

Thm 9.2.6 Suppose $I_V : V \to V$. Suppose B and C are two ordered bases for V. For any $u \in V$, the matrix $[I_V]_{C,B}$

- converts $[u]_B$ to $[u]_C$, i.e. $[u]_C = [I_V(u)]_C = [I_V]_{C,B}[u]_B$.

- is called the **transition matrix** from B to C.
- is invertible, and its inverse is the transition matrix from C to B, i.e. $[I_V]_{B,C}$

9.3 Compositions of Linear Transformations

Let $S : U \to V$ and $T : V \to W$ be linear transformations. Suppose U, V, W are finite dimensional where $dim(U), dim(V), dim(W) \geq 1$.

Thm 9.3.1 Then the composition mapping $T \circ S : U \to W$, defined by $(T \circ S)(u) = T(S(u))$ for $u \in U$ is also a linear transformation.

Thm 9.3.3 Let A, B, C be ordered bases for U, V, W respectively. Then $[T \circ S]_{C,A} = [T]_{C,B}[S]_{B,A}$

Def 9.3.5 Let T be a linear operator. For any nonnegative integer m, define

$$T^m = \begin{cases} I_v & \text{if} \quad m = 0 \\ \underbrace{T \circ T \circ \cdots \circ T}_{m \text{ times}} & \text{if} \quad m \geq 1 \end{cases}$$

Corollary 9.3.6 Let T be a linear operator. Let B be an ordered basis. Then $[T^m]_B = ([T]_B)^m$

Lemma 9.3.8 Let T be a linear operator. Let B, C are two ordered bases for V and P the transition matrix from B to C, i.e. $P = [I_V]_{C,B}$. Then $[T]_B = P^{-1}[T]_C P$

Def 9.3.9 Let $\mathbb{F}$ be a field and $A, B \in M_{n \times n}(\mathbb{F})$. B is said to be **similar** to A if there exists an invertible matrix $P \in M_{n \times n}(\mathbb{F})$ s.t. $B = P^{-1}AP$

Thm 9.3.10 Let T be a linear operator on V and let C be an ordered basis for V. Then an $n \times n$ matrix D over $\mathbb{F}$ is similar to $[T]_C$ iff there exists an ordered basis B for V s.t. $D = [T]_B$.

9.4 The Vector Space $L(V, W)$

Let V and W be vector spaces over the same field $\mathbb{F}$. Suppose V, W are finite dimensional where $dim(V), dim(W) \geq 1$ with B and C be ordered basis for V and W respectively.

Def 9.4.1

1. Let $T_1, T_2 : V \to W$ be linear transformations. We define a mapping $T_1 + T_2 : V \to W$ by $(T_1 + T_2)(u) = T_1(u) + T_2(u)$ for $u \in V$.
2. Let $T : V \to W$ be a linear transformation and $c \in \mathbb{F}$ We define a mapping $cT : V \to W$ by $(cT)(u) = cT(u)$ for $u \in V$.

Both mappings $T_1 + T_2$ and cT are linear transformations.

Prop 9.4.3

1. If $T_1, T_2 : V \to W$ are linear transformations, then $[T_1 + T_2]_{C,B} = [T_1]_{C,B} + [T_2]_{C,B}$
2. If $T : V \to W$ be a linear transformation and $c \in \mathbb{F}$, then $[cT]_{C,B} = c[T]_{C,B}$

Remark 9.4.4: Matrices and linear transformations have a lot of similarities. The observations above show their relations in addition and scalar multiplication.

- Thm 9.3.3 shows that the composition of linear transformations is equivalent to matrix multiplication.
- Thm 9.6.6 shows a corresponding analog of matrix inverse in linear transformations.

Thm 9.4.5 Let $L(V, W)$ be the set of all linear transformations from V to W. Then $L(V, W)$ is a vector space over $\mathbb{F}$ with addition and scalar multiplication defined in Def 9.4.1.

Furthermore, if V and W are finite dimensional, then dim($L(V, W)$) = dim(V) dim(W)

Def 9.4.6 The vector space $L(V, \mathbb{F})$ is called the **dual space** of V and is denoted by V^*

By Thm 9.4.5. dim(V^*) = dim(V)

9.5 Kernels and Ranges

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

Def 9.5.1

1. The subset Ker(T) = $\{u \in V \mid T(u) = 0\}$ of V is called the **kernel** of T. Ker(T) is also known as the **nullspace** of T and denoted by $N(T)$
2. The subset $R(T) = \{T(u) \mid u \in V\}$ of W is called the **range** of T

Thm 9.5.2 Ker(T) is a subspace of V , and $R(T)$ is a subspace of W.

Def 9.5.4

1. If Ker(T) is finite dimensional, then dim(Ker(T)) is called the **nullity** of T and is denoted by nullity(T)
2. If R(T) is finite dimensional, then dim(R(T)) is called the **rank** of T and is denoted by rank(T)

Lemma 9.5.6 Suppose V and W are finite dimensional with dim(V) ≥ 1 and dim(W) ≥ 1. For any ordered bases B and C for V and W respectively

1. $\{[u]_B \mid u \in \mathrm{Ker}(T)\}$ is the nullspace of $[T]_{C,B}$ and nullity(T) = nullity($[T]_{C,B}$)
2. $\{[u]_C \mid u \in \mathrm{R}(T)\}$ is the column space of $[T]_{C,B}$ and rank(T) = rank($[T]_{C,B}$)

Thm 9.5.7 (Dimension Theorem for Linear Transformations). Let $T : V \to W$ where V and W are finite dimensional. Then, rank(T) + nullity(T) = dim(V)

Thm 9.5.9 Suppose B and C are subsets of V s.t. B is a basis for Ker(T) and $\{T(v) \mid v \in C\}$ is a basis for R(T) and for any $v, v' \in C$ if $v \neq v'$, then $T(v) \neq T(v')$. Then $B \cup C$ is a basis for V.

Def 9.5.11 Let $f : A \to B$ be a mapping.

1. f is **injective** or *one-to-one* if $\forall z \in B$, there exists at most one $x \in A$ s.t. $f(x) = z$.
2. f is **surjective** or *onto* if $\forall z \in B$, there exists at least one $x \in A$ s.t. $f(x) = z$.
3. f is **bijective** if it is both injective and surjective.

Prop 9.5.12

1. T is injective iff Ker(T) = $\{\mathbf{0}\}$ iff nullity(T) = 0

2. T is surjective iff R(T) = W.

Note:

1. Let $S : U \to V$ and $T : V \to W$ be linear transformations. $Ker(S) \subseteq Ker(T \circ S)$ and $R(T \circ S) \subseteq R(T)$
2. Let $S, T : V \to W$ be linear transformations. Then $R(S+T) \subseteq R(S) + R(T)$ and $Ker(S) \cap Ker(T) \subseteq Ker(S+T)$

9.6 Isomorphism

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

Def 9.6.1 The linear transformation $T : V \to W$ is called an **isomorphism** from V onto W if T is bijective.

Def 9.6.3 A mapping $T : V \to W$ is bijective iff there exists a mapping $S : W \to V$ s.t. $S \circ T = I_V$ and $T \circ S = I_W$ where I_V and I_W are identity operators on V and W respectively. The mapping S is known as the *inverse* of T and is denoted by T^{-1}. Thus a bijective mapping is also called an *invertible mapping*

Thm 9.6.4 If T is an isomorphism, then T^{-1} is a linear transformation and hence is also an isomorphism.

Thm 9.6.6 Suppose V and W are finite dimensional with dim(V) = dim(W) ≥ 1. Let B and C be ordered bases for V and W respectively.

1. T is an isomorphism iff $[T]_{C,B}$ is an invertible matrix
2. If T is an isomorphism, $[T^{-1}]_{B,C} = ([T]_{C,B})^{-1}$

Thm 9.6.8 Let $S : W \to V$ and $T : V \to W$ be linear transformations s.t. $T \circ S = I_W$.

1. S is injective and T is surjective.
2. If V and W are finite dimensional and dim(V) = dim(W), then S and T are isomorphisms, $S^{-1} = T$ and $T^{-1} = S$.

Def 9.6.10 Let V and W be vector spaces over a field $\mathbb{F}$. If there exists an isomorphism from V onto W, then V is said to be **isomorphic** to W and we write $V \cong_{\mathbb{F}} W$ or simply $V \cong W$.

Thm 9.6.13 Let V and W be finite dimensional vector spaces over the same field. Then V is isomorphic to W iff dim(V) = dim(W)

Example: $\mathcal{M}_{n\times n}(\mathbf{F}) \cong_{\mathbb{F}} \mathbb{F}^{mn}$, $\mathcal{P}_n(\mathbb{F}) \cong_{\mathbb{F}} \mathbb{F}^{n+1}$, $\mathbb{C}^n \cong_{\mathbb{R}} \mathbb{R}^{2n}$

Thm 9.6.15 (The First Isomorphism Theorem). Let $T : V \to W$ be a linear transformation. Then $V/\mathrm{Ker}(T) \cong R(T)$

10 Multilinear Forms and Determinants

10.1 Permutations

Def 10.1.2 A **permutation** σ of $\{1, 2, \ldots, n\}$ is a bijective mapping from $\{1, 2, \ldots, n\}$ to $\{1, 2, \ldots, n\}$. We usually represent σ by $\begin{pmatrix} 1 & 2 & \ldots & n \\ \sigma(1) & \sigma(2) & \ldots & \sigma(n) \end{pmatrix}$. The set of all permutations of $\{1, 2, \ldots, n\}$ is denoted by S_n. Note that $|S_n| = n!$

Notation 10.1.4

1. For $\sigma, \tau \in S_n$, $\sigma\tau = \sigma \circ \tau$ is also a permutation.
2. For $\alpha, \beta \in \{1, 2, \ldots, n\}$, let $\phi_{\alpha,\beta}$ denote the permutation of $\{1, 2, \ldots, n\}$ s.t.

$$\phi_{\alpha,\beta}(k) = \begin{cases} k & \text{if } k \neq \alpha, \beta \\ \alpha & \text{if } k = \alpha \\ \beta & \text{if } k = \beta \end{cases}$$

   This permutation is called the **transposition** of α and β. Note that $\phi_{\alpha,\beta} = \phi_{\beta,\alpha}$ and $\phi^{-1}_{\alpha,\beta} = \phi_{\alpha,\beta}$

Lemma 10.1.6

1. $\{\sigma^{-1} \mid \sigma \in S_n\} = S_n$
2. For any $\tau \in S_n$, $\{\tau\sigma \mid \sigma \in S_n\} = \{\sigma\tau \mid \sigma \in S_n\} = S_n$

Lemma 10.1.7 For every $\sigma \in S_n$, there exists $\alpha_1, \alpha_2, \ldots \alpha_k \in \{1, 2, \ldots, n\}$ s.t. $\sigma = \sigma_{\alpha_1,\alpha_1+1}\sigma_{\alpha_2,\alpha_2+1}\cdots\sigma_{\alpha_k,\alpha_k+1}$

Def 10.1.9 Let $\sigma \in S_n$. An **inversion** occurs in σ if $\sigma(i) > \sigma(j)$ for $i < j$. If the total number of inversions in σ is even, σ is called an **even** permutation; otherwise σ is an **odd** permutation.

The **sign** (or **parity**) of σ, denoted as $\text{sgn}(\sigma)$ is defined to be 1 if σ is even and -1 if σ is odd.

Thm 10.1.11 For any $\sigma, \tau \in S_n$, $\text{sgn}(\sigma\tau) = \text{sgn}(\sigma)\text{sgn}(\tau)$. Moreover, $\text{sgn}(\phi_{\alpha,\beta}) = -1$

Corollary 10.1.12

1. If $\sigma \in S_n$ is a product of k transpositions, then $\text{sgn}(\sigma) = (-1)^k$
2. A permutation is even (respectively, odd) if it is a product of even (respectively, odd) number of transpositions.
3. For any $\sigma \in S_n$, $\text{sgn}(\sigma^{-1}) = \text{sgn}(\sigma)$

10.2 Multilinear Forms

Def 10.2.1 Let V be a vector space over a field $\mathbb{F}$. A mapping $T : V^n \to \mathbb{F}$ is called a **multilinear form** on V if for each $i, 1 \leq i \leq n$,
$T(u_1, \ldots u_{i-1}, av + bw, u_{i+1}, \ldots, u_n) = aT(u_1, \ldots u_{i-1}, v, u_{i+1}, \ldots, u_n) + bT(u_1, \ldots u_{i-1}, w, u_{i+1}, \ldots, u_n)$ for all $a, b \in \mathbb{F}$ and $u_1, \ldots u_{i-1}, u_{i+1}, \ldots, u_n, v, w \in V$

A multilinear form T on V is called **alternative** if $T(u_1, u_2, \ldots u_n) = 0$ whenever $u_\alpha = u_\beta$ for some $\alpha \neq \beta$

Define $P : \mathcal{M}_{n\times n}(\mathbb{F}) \to \mathbb{F}$ by

$$P(A) = \sum_{\sigma \in S_n} a_{\sigma(1),1}a_{\sigma(2),2}\cdots a_{\sigma(n),n}$$

for $A = (a_{ij}) \in \mathcal{M}_{n\times n}(\mathbb{F})$. The value $P(A)$ is known as the **permanent** of A.

Thm 10.2.3 Let $T : V^n \to \mathbb{F}$ be an alternative multilinear form on a vector space V. Then for all $\sigma \in S_n$ and $u_1, u_2, \ldots, u_n \in V$, we have $T(u_1, u_2, \ldots, u_n) = \operatorname{sgn}(\sigma) \cdot T(u_{\sigma(1)}, u_{\sigma(2)}, \ldots, u_{\sigma(n)})$

Remark 10.2.4 Let $T : V^n \to \mathbb{F}$ be a multilinear form on a finite dimensional vector space V over a field $\mathbb{F}$. Fix a basis $\{v_1, v_2, \ldots, v_m\}$ for V. Take any $u_1, u_2, \ldots u_n \in V$, let

$$
\begin{aligned}
u_1 &= a_{11}v_1 + a_{21}v_2 + \cdots + a_{m1}v_m \\
u_2 &= a_{12}v_1 + a_{22}v_2 + \cdots + a_{m2}v_m \\
&\vdots \\
u_n &= a_{1n}v_n + a_{2n}v_2 + \cdots + a_{mn}v_m
\end{aligned}
$$

where $a_{11}, a_{12}, \ldots, a_{mn} \in \mathbb{F}$

1. Let $\mathcal{F}$ be the set of all mapping from $\{1, 2, \ldots n\}$ to $\{1, 2, \ldots, m\}$. We have

$$
T(u_1, u_2, \ldots, u_n) = \sum_{f\in\mathcal{F}} a_{f(1),1}a_{f(2),2}\cdots a_{f(n),n}\, T(v_{f(1)}, v_{f(2)}, \ldots, v_{f(n)})
$$

2. Suppose T is an alternative form.
   1. If $m < n$, then T is a zero mapping.
   2. If $m \geq n$, then (10.3) still holds if we change the set $\mathcal{F}$ to the set of all injective mapping. In particular, when $m = n$ we have $T(u_1, u_2, \ldots, u_n) = \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma)a_{\sigma(1),1}a_{\sigma(2),2}\cdots a_{\sigma(n),n}\, T(v_1, v_2, \ldots, v_n)$

10.3 Determinants

Def 10.3.1 A mapping $D : M_{n\times n}(\mathbb{F}) \to \mathbb{F}$ is called a *determinant function* on $M_{n\times n}(\mathbb{F})$ if it satisfies the following axioms:

1. By regarding the columns of matrices in $M_{n\times n}(\mathbb{F})$ as vectors in $\mathbb{F}^n$, D is a multilinear form on $\mathbb{F}^n$.
2. $D(A) = 0$ if $A \in M_{n\times n}(\mathbb{F})$ has two identical columns, i.e. as a multilinear form on $\mathbb{F}^n$, D is alternative.
3. $D(I_n) = 1$

Theorem 10.3.2 There exists one and only one determinant function on $\mathcal{M}_{n\times n}(\mathbb{F})$ and it is the function det : $\mathcal{M}_{n\times n}(\mathbb{F}) \to \mathbb{F}$ defined by

$$
\det(A) = \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma)a_{\sigma(1),1}a_{\sigma(2),2}\cdots a_{\sigma(n),n}
$$

for $A = (a_{ij}) \in \mathcal{M}_{n\times n}(\mathbb{F})$. This formula is known as the classical definition of determinants.

Lemma 10.3.4 Let $A \in \mathcal{M}_{n\times n}(\mathbb{F})$. Then $\det(A) = \det(A^T)$

Thm 10.3.5 (Cofactor expansions). Let $A = (a_{ij}) \in \mathcal{M}_{n\times n}(\mathbb{F})$. Define $\tilde{A_{ij}}$ to be the $(n-1)\times(n-1)$ matrix obtained from A by deleting the ith row and the jth column. Then for any $\alpha = 1, 2, \ldots, n$ and $\beta = 1, 2, \ldots, n$,

$$
\begin{aligned}
\det(A) &= a_{\alpha 1}A_{\alpha 1} + a_{\alpha 2}A_{\alpha 2} + \cdots + a_{\alpha n}A_{\alpha n} \\
&= a_{1\beta}A_{1\beta} + a_{2\beta}A_{2\beta} + \cdots + a_{n\beta}A_{n\beta}
\end{aligned}
$$

where $A_{ij} = (-1)^{i+j} \det(\tilde{A_{ij}})$.

11 Diagonalization and Jordan Canonical Forms

11.1 Eigenvalues and Diagonalization

Let T be a linear operator on a finite dimensional vector space V with $\dim(V) \geq 1$.

Def 11.1.2 Let V be a vector space. A nonzero vector $u \in V$ is called an *eigenvector* of T if $T(u) = \lambda u$ for some scalar *eigenvalue* λ.

Def 11.1.4 $det(T)$ is the determinant of the matrix $[T]_B$ where B is any ordered basis for V.

Remark 11.1.5 The determinant of T is independent of the choice of basis B.

Thm 11.1.6 For a scalar λ, let $\lambda I_V - T$ be the linear operator defined by $(\lambda I_V - T)(u) = \lambda u - T(u)$ for some $u \in V$

1. λ is an eigenvalue of T iff $det(\lambda I_V - T) = 0$. (λ is a solution to the charateristic polynomial of T)
2. $u \in V$ is an eigenvector of T associated with λ iff u is a nonzero vector in the **eigenspace** $\mathrm{Ker}(T - \lambda I_V)$

Notation 11.1.7

1. Denote the characteristic polynomial of T, $c_T(x) = det(xI_V - T)$
2. Denote the eigenspace of A associated with λ as $E_\lambda(A)$

Remark 11.1.8 For a basis B of V, we have $c_T(x) = c_{[T]_B}(x)$, a monic polynomial of degree $\dim(V)$

Def 11.1.10 T is **diagonalizable** if there exists an ordered basis B for V s.t. $[T]_B$ is a diagonal matrix

Thm 11.1.11 T is diagonalizable iff V has a basis B s.t. every vector in B is an eigenvector of T .

Algorithm 11.1.12 Determining whether the linear operator T is diagonalizable.

1. Find a basis C for V and compute $A = [T]_C$
2. Write $c_A(x) = \prod_{i=1}^{k} (x - \lambda_i)^{r_i}$ where λ_i are distinct and $\sum_{i=1}^{k} = dim(V)$
3. For each eigenvalue λ_i, find a basis B_{λ_i} for the eigenspace $E_{\lambda_i}(T)$. If $|B_{\lambda_i}| < r_i$ for some i, then T is not diagonalizable
4. $B = \bigcup_{i=1}^{k} B_{\lambda_i}$ is a basis for V and $D = [T]_B$ is a diagonal matrix. Note that $D = P^{-1}AP$ where $P = [I_V]_{C,B}$ is the transition matrix from B to C.

If we let C be the standard bases, then columns of P are eigenvectors of T.

11.2 Triangular Canonical Forms

Lemma 11.2.2 Suppose A is an $r \times m$ matrix, B is an $r \times n$ matrix, C is an $s \times m$ matrix, D is an $s \times n$ matrix, E is an $m \times t$ matrix, F is an $m \times u$ matrix, G is an $n \times t$ matrix, H is an $n \times u$ matrix, then

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} E & F \\ G & H \end{pmatrix} = \begin{pmatrix} AE + BG & AF + BH \\ CE + DG & CF + DH \end{pmatrix}$$

Thm 11.2.3 (Triangular Canonical Forms). Let $\mathbb{F}$ be a field.

1. Let $A \in \mathcal{M}_{n\times n}(\mathbb{F})$. If the characteristic polynomial $c_A(x)$ can be factorized over linear factors over $\mathbb{F}$, then there exists an invertible matrix $P \in \mathcal{M}_{n\times n}(\mathbb{F})$ s.t. $P^{-1}AP$ is an upper triangular matrix.
2. Let T be a linear operator on a finite dimensional vector space V with $\dim(V) \geq 1$. If the characteristic polynomial $c_T(x)$ can be factorized over linear factors over $\mathbb{F}$, then there exists an ordered basis B for V s.t. $[T]_B$ is an upper triangular matrix.

Tut 7

1. A linear operator T on a finite dimensional vector space V is **triangularizable** if there exists an ordered basis B for V s.t. $[T]_B$ is a triangular matrix. Then T is triangularizable iff its characteristic polynomial can be factorized into linear factors.

11.3 Invariant Subspaces

Def 11.3.1 Let V be a vector space and $T : V \to V$ a linear operator. A subspace W of V is said to be **T-invariant** if $T(u)$ is contained in W for all $u \in W$, i.e. $T[W] = \{T(u) \mid u \in W\} \subseteq W$.

If W is a T-invariant subspace of V, the linear operator $T|_W : W \to W$ defined by $T|_W(u) = T(u)$ for $u \in W$ is called the **restriction** of T on W.

Prop 11.3.3 Let S and T be linear operators on V. Suppose W is a subspace of V which is both S-invariant and T-invariant. Then

1. W is $(S \circ T)$-invariant and $(S \circ T)|_W = S|_W \circ T|_W$
2. W is $(S + T)$-invariant and $(S + T)|_W = S|_W + T|_W$
3. for any scalar c, W is cT-invariant and $(cT)|_W = c(T|_W)$

Discussion 11.3.4 Suppose W is a T-invariant subspace of V with $\dim(W) \geq 1$. Let $\dim(W) = m$ and $\dim(V) = n \geq m$. Let C be an ordered basis of W and B a basis for V extended from C. Then, $[T]_B = \begin{pmatrix} A_1 & A_2 \\ 0 & A_3 \end{pmatrix}$ where $A_1 = [T|_W]_C$, and $(A_2\ A_3)^T$ is the coordinate vector w.r.t. B of the image of the basis extension under T.

Lemma 11.3.6 Let D be a square matrix s.t. $D = \begin{pmatrix} A & B \\ 0 & C \end{pmatrix}$ where both A and C are square matrices. Then $det(D) = det(A)det(C)$

Thm 11.3.7 Let T be a linear operator on a finite dimensional vector space V. Suppose W is a T-invariant subspace of V with $\dim(W) \geq 1$, then $c_{T|_W}(x) \mid c_T(x)$

Thm 11.3.10 Let T be a linear operator on a finite dimensional vector space V. Take a nonzero vector $u \in V$. Suppose **the T-cyclic subspace** W = span$\{u, T(u), T^2(u), \ldots\}$

generated by u is finite dimensional.

1. $\dim(W) = k$ where k is the smallest positive integer s.t. $T^k(u)$ is a linear combination of $u, T(u), \ldots, T^{k-1}(u)$
2. Suppose $\dim(W) = k$
   1. $\{u, T(u), \ldots, T^{k-1}(u)\}$ is a basis for W.
   2. If $T^k(u) = a_0 u + a_1 T(u) + \cdots a_{k-1} T^{k-1}(u)$ where $a_0, a_1, \cdots, a_{k-1} \in \mathbb{F}$, then $c_{T|_W}(x) = -a_0 - a_1 x - \cdots - a_{k-1} x^{k-1} + x^k$

Comment: The T-cyclic subspace, which is T-invariant, is a very useful invariant subspace as it helps to find a basis B s.t. $[T]_B$ is in a simpler form. See Discussion 11.3.4

Discussion 11.3.12 Suppose $V = W_1 \oplus W_2 \oplus \cdots \oplus W_k$ where W_t are T-invariant subspaces of V with $\dim(W_t) = n_t \geq 1$ for $t = 1, 2, \ldots, k$. For each t, let $C_t = \{v_1^{(t)}, v_2^{(t)}, \ldots, v_{n_t}^{(t)}\}$ be an ordered basis for W_t.

Let $[T|_{W_t}]_{C_t} = A_t$. Using $B = C_1 \cup C_2 \cup \cdots \cup C_k$ as an ordered basis for V, we obtain

$$[T]_B = \begin{pmatrix} A_1 & 0 & & 0 \\ 0 & A_2 & & 0 \\ & & \ddots & \\ 0 & 0 & & A_k \end{pmatrix}$$

Furthermore, $c_T(x) = \prod_{i=1}^{k} c_{A_i}(x) = \prod_{i=1}^{k} c_{T|_{W_i}}(x)$.

11.4 Cayley-Hamilton Theorem

Notation 11.4.1 Let $\mathbb{F}$ be a field and let $p(x) = a_0 + a_1 x + \cdots a_m x^m$ where $a_0, a_1, \ldots a_m \in \mathbb{F}$

1. For a linear operator T on a vector space V over $\mathbb{F}$, we use $p(T)$ to denote the linear operator $a_0 I_V + a_1 T + \cdots + a_m T^m$ on V.
2. For an $n \times n$ matrix A over $\mathbb{F}$, we use $p(A)$ to denote the $n \times n$ matrix $a_o I_n + a_1 A + \cdots + a_m A^m$

Lemma 11.4.2 Let T be a linear operator on a vector space V over $\mathbb{F}$ and A be an $n \times n$ matrix over $\mathbb{F}$. In the following $p(x), q(x) \in \mathcal{P}(\mathbb{F})$

1. Suppose V is finite dimensional where $\dim(V) = n \geq 1$. For any ordered basis B for V, $[p(T)]_B = p([T]_B)$
2. If W is a T-invariant subspace of V, then W is also a $p(T)$-invariant subspace of V and $p(T)|_W = p(T|_W)$
3. Polynomial addition, scalar multiplication and polynomial multiplication also works if we substitute x for T (slightly different form: $u(T) = p(T) \circ q(T) = q(T) \circ p(T)$ for $u(x) = p(x)q(x)$) and A.

Thm 11.4.4 (Cayley-Hamilton Theorem)

1. Let T be a linear operator on a finite dimensional vector space V where $\dim(V) \geq 1$. Then $c_T(T) = O_V$, where O_V is the zero operator on V.
2. Let A be a square matrix. Then $c_A(A) = \mathbf{0}_n$.

11.5 Minimal Polynomials

Let T be a linear operator on a finite dimensional vector space V over $\mathbb{F}$ where $\dim(V) \geq 1$.

Def 11.5.2 The **minimal polynomial** $m_T(x)$ of T is the monic polynomial $p(x)$ of smallest degree s.t. $p(T) = O_V$, i.e. if $q(x)$ is a nonzero polynomial over $\mathbb{F}$ s.t. $q(T) = O_V$, then $\deg(q(x)) \geq \deg(p(x))$

The existence of a minimal polynomial is guaranteed by Cayley-Hamilton Thm.

The minimal polynomial for a zero mapping is $m_{O_V}(x) = x$

Lemma 11.5.5

1. Let $p(x)$ be a polynomial over $\mathbb{F}$. Then $p(T) = O_V$ iff $p(x)$ is divisible by the minimal polynomial of T.
2. If W is a T-invariant subspace of V with $\dim(W) \geq 1$, then the minimal polynomial of T is divisible by the minimal polynomial of $T|_W$
3. Suppose λ is an eigenvalue of T s.t. $c_T(x) = (x-\lambda)^r q(x)$ where $x - \lambda \nmid q(x)$. Then $m_T(x) = (x-\lambda)^s q_1(x)$ where $1 \leq s \leq r$ and $q_1(x) \mid q(x)$

Thm 11.5.7 Let T be a linear operator on a vector space V. Suppose W_1 and W_2 are T-invariant subspace of V.

1. $W_1 + W_2$ is T-invariant.
2. If W_1 and W_2 are finite dimensional with $\dim(W_1) \geq 1$ and $\dim(W_2) \geq 1$, $m_{T|_{W_1+W_2}}(x) = \operatorname{lcm}(m_{T|_{W_1}}(x), m_{T|_{W_2}}(x))$

Thm 11.5.8 Suppose $c_T(x) = \prod_{i=1}^{k}(x-\lambda_i)^{r_i}$ where $\lambda_1, \lambda_2, \ldots, \lambda_k$ are distinct eigenvalues of T. Then $m_T(x) = \prod_{i=1}^{k}(x-\lambda_i)^{s_i}$ where $1 \leq s_i \leq r_i$ for all i. Define $K_{\lambda_i}(T) = Ker((T - \lambda_i I_v)^{s_i})$ for $i = 1, 2, \ldots, k$. Then, $V = K_{\lambda_1}(T) \oplus K_{\lambda_2}(T) \cdots \oplus K_{\lambda_k}(T)$

1. $E_{\lambda_i}(T) \subseteq K_{\lambda_i}(T)$
2. $K_{\lambda_i}(T)$ is a T-invariant subspace of V.
3. $m_{T|_{K_{\lambda_i}(T)}}(x) = (x-\lambda_i)^{s_i}$
4. $c_{T|_{K_{\lambda_i}(T)}}(x) = (x-\lambda_i)^{r_i}$
5. $\dim(K_{\lambda_i}(T)) = r_i$

Thm 11.5.10 Let $c_T(x) = \prod_{i=1}^{k}(x-\lambda_i)^{r_i}$ where $\lambda_1, \lambda_2, \ldots, \lambda_k$ are distinct eigenvalues of T. The following are equivalent:

1. T is diagonalizable
2. $m_T(x) = \prod_{i=1}^{k}(x-\lambda_i)$
3. $\dim(E_{\lambda_i}(T)) = r_i$ for $i = 1, 2, \ldots, k$
4. $V = E_{\lambda_1}(T) \oplus E_{\lambda_2}(T) \cdots \oplus E_{\lambda_k}(T)$

Corollary 11.5.11 Let W be a T-invariant subspace of V with $dim(W) \geq 1$. If T is diagonalizable, then $T|_W$ is also diagonalizable.

HW4. Let W be a T-cyclic subspace of V. Then $m_{T|_W}(x) = c_{T|_W}(x)$

PYP .

- (2013/2014S1) Let A be an invertible $n \times n$ matrix. $c_{A^{-1}}(x) = x^n[c_A(0)]^{-1}c_A(1/x)$ and $m_{A^{-1}}(x) = x^k[m_A(0)]^{-1}m_A(1/x)$ where $k = deg(m_A(x))$
- (2018/2019S1) Let $p(x)$ and $q(x)$ be polynomials over $\mathbb{F}$ s.t. $gcd(p(x), q(x)) = 1$, i.e. exist polynomials $a(x), b(x)$ s.t. $a(x)p(x) + b(x)q(x) = 1$. For any nonzero $v \in Ker(p(T))$, then $q(T)(v) \neq 0$

11.6 Jordan Canonical Forms

Let T be a linear operator on a finite dimensional vector space V over $\mathbb{F}$ where $\dim(V) \geq 1$.

Def 11.6.2 Let λ be a scalar. The **Jordan block** of order t associated with λ is a $t \times t$ matrix

$$J_t(\lambda) = \begin{pmatrix} \lambda & 1 & & 0 & \\ & \lambda & 1 & & \\ & & \ddots & \ddots & \\ & 0 & & \ddots & 1 \\ & & & & \lambda \end{pmatrix}$$

Lemma 11.6.3 Given a Jordan Block $J = J_t(\lambda)$, $c_J(x) = m_J(x) = (x - \lambda)^t$

Thm 11.6.4 Suppose $c_T(x)$ can be factorized into linear factors over $\mathbb{F}$, then there exists an ordered basis B for V s.t. $[T]_B = J$ with

$$J = \begin{pmatrix} J_{t_1}(\lambda_1) & & 0 & \\ & J_{t_2}(\lambda_2) & & \\ & & \ddots & \\ & 0 & & J_{t_m}(\lambda_m) \end{pmatrix}$$

where $\lambda_1, \lambda_2, \ldots \lambda_m$ are (not necessarily distinct) eigenvalues of T.

Remark 11.6.5. Let $A \in \mathcal{M}_{n\times n}(\mathbb{F})$. Suppose $c_A(x)$ can be factorized into linear factors over $\mathbb{F}$, applying Thm 11.6.4 to $T = L_A$ implies that we can find an invertible matrix $A \in \mathcal{M}_{n\times n}(\mathbb{F})$ s.t. $P^{-1}AP = J$

Def 11.6.6 For a linear operator T of finite dimensional vector space V, if $\exists$ an ordered basis B s.t. $[T]_B = J$ (see 11.6.4), we say that T has a Jordan canonical form J.

Similarly, for a square matrix A, if there exists an invertible matrix P s.t. $P^{-1}AP = J$, we say that A has a Jordan canonical form J.

Remark 11.6.8. Jordan canonical forms is unique up to the ordering of the Jordan blocks.

Thm 11.6.9 Suppose a linear operator T of finite dimensional space V has a Jordan canonical form J (as seen in 11.6.4)

1. $c_T(x) = \prod_{i=1}^{m} (x - \lambda_i)^{t_i}$
2. $m_T(x)$ is the least common multiple of $\{(x - \lambda_i)^{t_i} \mid i = 1, 2, \ldots, m\}$
3. For every eigenvalue λ of T, $\dim(E_\lambda(T))$ is the total number of Jordan blocks associated with λ in the matrix J.

12 Inner Product Spaces

In this chapter, we only focus on real and complex vector spaces.

Notation 12.1.2 Let A be a complex matrix. We use $\overline{A}$ to denote the **conjugate** of A. Define the $A^* = \overline{A^T}$ as the **conjugate transpose** of A. Then

1. $(A + B)^* = A^* + B^*$
2. $(AC)^* = C^* A^*$
3. $(cA)^* = \overline{c} A^*$

Def 12.1.3 Let V be a vector space over $\mathbb{F}$. An **inner product** on V is a mapping which assigns to each ordered pair of vectors $u, v \in V$ a scalar $\langle u, v \rangle \in \mathbb{F}$ s.t. the following axioms are satisfied:

1. For all $u, v \in V$, $\langle u, v \rangle = \overline{\langle v, u \rangle}$
2. For all $u, v, w \in V$, $\langle u + v, w \rangle = \langle u, w \rangle + \langle v, w \rangle$
3. For all $c \in \mathbb{F}$ and $u, v \in V$, $\langle cu, v \rangle = c\langle u, v \rangle$ (we can derive that $\langle u, cv \rangle = \overline{c}\langle u, v \rangle$)
4. $\langle 0, 0 \rangle = 0$ and for all nonzero $u \in V$, $\langle u, u \rangle > 0$. In particular $\langle 0, u \rangle = 0$

Def 12.1.5 A vector space V equipped with an inner product is called an *inner product space*

- The usual inner product on $\mathbb{C}^n$ is defined as uv^*
- Consider the vector space $C([a, b])$ the set of continuous function on the closed interval $[a, b]$, then an inner product on $C([a, b])$ is $\langle f, g \rangle = \frac{1}{b-a} \int_a^b f(t)g(t)dt$
- Let V be the set of all real infinite sequences (a_n) s.t. $\sum_{n=1}^{\infty} a_n^2$ converges, then an inner product is $\langle (a_n), (b_n) \rangle = \sum_{n=1}^{\infty} a_n b_n$. This space is known as the l_2-space

12.2 Norms and Distances

Def 12.2.2 Let V be an inner product space

1. For $u \in V$, the **norm** (or *length*) of u is defined to be $\|u\| = \sqrt{\langle u, u \rangle}$
2. For $u, v \in V$, the **distance** between u and v is $d(u, v) = \|u - v\|$

Thm 12.2.4 Let V be an inner product space over $\mathbb{F}$

1. $\|0\| = 0$, and for any nonzero $u \in V$, $\|u\| > 0$
2. For any $c \in \mathbb{F}$ and $u \in V$, $\|cu\| = |c|\|u\|$
3. **(Cauchy-Schwarz Ineq)** For any $u, v \in V$, $|\langle u, v \rangle| \leq \|u\|\|v\|$
4. **(Triangle Ineq)** For any $u, v \in V$, $\|u + v\| \leq \|u\| + \|v\|$

12.3 Orthogonal and Orthonormal Bases

Discussion 12.3.1. u and V are perpendicular to each other iff $\langle u, v\rangle = 0$.

Def 12.3.2 Let V be an inner product space

1. 2 vectors $u, v \in V$ are **orthogonal** to each other if $\langle u, v\rangle = 0$
2. Let W be a subspace of V. A vector u is **orthogonal** (or perpendicular) to W if u is orthogonal to all vectors in W.
3. A subset B of V is **orthogonal** if the vectors in B are pairwise orthogonal.
4. A subset B of V is **orthonormal** if B is orthogonal and all vectors in B are unit vectors

Lemma 12.3.3 Let V be an inner product space over $\mathbb{F}$

1. Let $W = \text{span}(B)$ where $B \subseteq V$. For $u \in V$, u is orthogonal to W iff u is orthogonal to every vectors in B.
2. If B is an orthogonal set of nonzero vectors from V, then B is always linearly independent
3. Suppose V is finite dimensional where $\dim(V) \geq 1$. Let B be an ordered orthonormal basis for V. Then $\langle u, v\rangle = (u)_B((v)_B)^* = ([u]_B)^T\overline{[v]}_B$

Note: if $\mathbb{F} = \mathbb{R}$, then $\langle u, v\rangle = (u)_B \cdot (v)_B$

Remark 12.3.4

1. Suppose V is a finite dimensional inner product space. To determine whether a set B of nonzero vectors from V is an orthogonal (orthonormal) basis for V, we only need to check that (1) B is orthogonal (orthonormal) and (2) $|B| = \dim(V)$
2. By Lemma 12.3.3.3, a finite dimensional real inner product space is essentially the same as the Euclidean space

Thm 12.3.6 Let V be a finite dimensional inner product space. If $B = \{w_1, w_2, \ldots, w_n\}$ is an orthonormal basis for V, then for any vector $u \in V$, $u = \langle u, w_1\rangle w_1 + \langle u, w_2\rangle w_2 + \cdots + \langle u, w_n\rangle w_n$.

Thm 12.3.7 (Gram-Schmidt Process). Suppose $\{u_1, u_2, \ldots, u_n\}$ is a basis for a finite dimensional inner product space V. Let

$$\begin{aligned} v_1 &= u_1 \\ v_2 &= u_2 - \tfrac{\langle u_2, v_1\rangle}{\langle v_1, v_1\rangle} v_1 \\ &\vdots \\ v_n &= u_n - \tfrac{\langle u_n, v_1\rangle}{\langle v_1, v_1\rangle} v_1 - \tfrac{\langle u_n, v_2\rangle}{\langle v_2, v_2\rangle} v_2 - \cdots - \tfrac{\langle u_n, v_{n-1}\rangle}{\langle v_{n-1}, v_{n-1}\rangle} v_{n-1} \end{aligned}$$

Then $\{v_1, v_2, \ldots, v_n\}$ is an orthogonal basis for V.

12.4 Orthogonal Complements & Projections

Let V be an inner product space and W a subspace of V.

Def 12.4.1 The **orthogonal complement** of W is defined to be the set $W^\perp = \{v \in V \mid \langle v, u\rangle = 0 \; \forall u \in W\} \subseteq V$

Thm 12.4.3

1. $W^\perp$ is a subspace of V
2. $W \cap W^\perp = \{0\}$, i.e. $W + W^\perp$ is a direct sum
3. If W is finite dimensional, then $V = W \oplus W^\perp$
4. If V is finite dimensional, then dim(V) = dim(W) + dim($W^\perp$)

Thm 12.4.6

1. $W \subseteq (W^\perp)^\perp$.
2. If W is finite dimensional, then $W = (W^\perp)^\perp$

Def 12.4.8 Suppose $V = W \oplus W^\perp$, i.e. every $u \in V$ can be uniquely expressed as $u = w + w'$ where $w \in W$ and $w' \in W^\perp$. The vector w is called the **orthogonal projection** of u onto W and is denoted by $\text{Proj}_W(u)$

Prop 12.4.9. The mapping $\text{Proj}_W : V \to V$ is a linear operator and is called the orthogonal projection of V onto W.

Thm 12.4.11 Let W be finite dimensional. If $B = \{w_1, w_2, \dots, w_k\}$ is an orthonormal basis for W, then for any vector $u \in V$, $\text{Proj}_W(u) = \langle u, w_1\rangle w_1 + \langle u, w_2\rangle w_2 + \cdots + \langle u, w_k\rangle w_k$ and $\text{Proj}_{W^\perp}(u) = u - \text{Proj}_W(u)$

Thm 12.4.13 (Best Approximation) Suppose $V = W \oplus W^\perp$. Then for any $u \in V$, $d(u, \text{Proj}_W(u)) \leq d(u, w)$ for all $w \in W$, i.e. $\text{Proj}_W(u)$ is the best approximation of u in W.

12.5 Adjoints of Linear Operators

Let V be an inner product space over $\mathbb{F}$ and let T be a linear operator on V

Def 12.5.1 A linear operator T^* is called the **adjoint** of T if $\langle T(u), v\rangle = \langle u, T^*(v)\rangle$ for all $u, v \in V$

Note:

1. the classical adjoint of a matrix is a completely different concept.
2. I_V, 0_V and L_A are its own adjoint.
3. We can derive that $\langle u, T(v)\rangle = \langle T^*(u), v\rangle$

Thm 12.5.4

1. The adjoint of T is unique if it exists
2. Suppose V is finite dimensional where $\dim(V) \geq 1$
   1. T^* always exists
   2. If B is an ordered orthonormal basis for V, then $[T^*]_B = ([T]_B)^*$
   3. rank(T) = rank(T^*) and nullity(T) = nullity(T^*)

Prop 12.5.7 Let $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. Suppose S and T are linear operators on V s.t. S^* and T^* exists. Then

1. $(S + T)^* = S^* + T^*$
2. for any $c \in \mathbb{F}$, $(cT)^* = \bar{c}T^*$
3. $(S \circ T)^* = T^* \circ S^*$
4. $(T^*)^* = T$

5. if W is a subspace of V that is both T- and T^*- invariant, then $(T|_W)^* = T^*|_W$.

Def 12.5.8 Let $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$.

1. Suppose T^* exists, T is invertible and $T^{-1} = T^*$, i.e. $T \circ T^* = T^* \circ T = I_V$
   1. If $\mathbb{F} = \mathbb{C}$, then T is called a **unitary operator**
   2. If $\mathbb{F} = \mathbb{R}$, then T is called a **orthogonal operator**
2. Let A be an invertible matrix over $\mathbb{F}$ s.t. $A^{-1} = A^*$, i.e. $AA^* = A^*A = I$
   1. If $\mathbb{F} = \mathbb{C}$, then A is called a **unitary matrix**
   2. If $\mathbb{F} = \mathbb{R}$, then A is called a **orthogonal matrix** (only real square matrices)

An orthogonal matrix is also a unitary matrix.

Prop 12.5.9 Let V be finite dimensional where $\dim(V) \geq 1$. Take any ordered orthonormal basis B for V. If $\mathbb{F} = \mathbb{C}$ (or $\mathbb{R}$), then T is unitary (orthogonal) iff $[T]_B$ is a unitary (orthogonal) matrix

Thm 12.5.11 Let V be finite dimensional where $\dim(V) \geq 1$. The following are equivalent

1. T is unitary (when $\mathbb{F} = \mathbb{C}$) or orthogonal (when $\mathbb{F} = \mathbb{R}$)
2. For all $u, v \in V$, $\langle T(u), T(v) \rangle = \langle u, v \rangle$
3. For all $u \in V$, $\|T(u)\| = \|u\|$
4. There exists an orthonormal basis $\{w_1, w_2, \ldots, w_n\}$ for V, where $n = \dim(V)$ s.t. $\{T(w_1), T(w_2), \ldots, T(w_n)\}$ is also orthonormal.

Thm 12.5.14 Let A be an $n \times n$ complex matrix. Suppose $\mathbb{C}^n$ is equipped with the usual inner product. The following statements are equivalent

1. A is unitary
2. The rows of A form an orthonormal basis for $\mathbb{C}^n$
3. The columns of A form an orthonormal basis for $\mathbb{C}^n$

Thm 12.5.15 Let V be a complex finite dimensional inner product space where $\dim(V) \geq 1$. If B and C are ordered orthonormal bases for V, then the transition matrix from B to C is a unitary matrix, i.e. $[I_V]_{B,C} = ([I_V]_{C,B})^{-1} = ([I_V]_{C,B})^*$.

PYP

- (2012/2013 S2).
  - $\text{Ker}(T^* \circ T) = \text{Ker}(T)$
  - Given $b \in V$, then $x = u$ is a solution to $(T^* \circ T) = T^*(b)$ iff $T(u)$ is the orthogonal projection of b onto $R(T)$.

12.6 Unitary and Orthogonal Diagonalization

Let V an inner product space over $\mathbb{F}$ and let T be a linear operator on V

Def 12.6.1 Suppose T^* exists.

1. T is called a **self-adjoint operator** if $T = T^*$
2. T is called a **normal operator** if $T \circ T^* = T^* \circ T$

Let A be a complex square matrix.

1. A is called a **Hermitian matrix** if $A = A^*$
2. A is called a **normal matrix** if $AA^* = A^*A$

Remark:

- Normal operator is similar to unitary/orthogonal operator except that it doesnt need to be identity matrix)
- Self-adjoint operator / Hermitian matrix is equal to symmetric matrices under $\mathbb{R}$
- All self-adjoint operators, orthogonal operators and unitary operators are normal.
- All Hermitian matrices, real symmetric matrices, unitary matrices and orthogonal matrices are normal.

Prop 12.6.2 Let V be finite dimensional with $\dim(V) \geq 1$. Take an ordered orthonormal basis B for V and let $A = [T]_B$.

1. If $\mathbb{F} = \mathbb{C}$ (or $\mathbb{R}$), then T is self-adjoint iff A is a Hermitian (symmetric) matrix.
2. T is normal iff A is a normal matrix

Lemma 12.6.4 Suppose $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$ and T a normal operator on V.

1. For all $u, v \in V$, $\langle T(u), T(v) \rangle = \langle T^*(u), T^*(v) \rangle$
2. For any $c \in \mathbb{F}$, the linear operator $T - cI_V$ is normal
3. If u is an eigenvector of T associated with λ, then u is an eigenvector of T^* associated with $\bar{\lambda}$
4. If u and v are eigenvector of T associated with λ and μ, respectively, where $\lambda \neq \mu$, then u and v are orthogonal.

Remark 12.6.5. Lemma 12.6.4 holds if we replace V with $\mathbb{C}^n$ (equipped with the usual inner product) and T by an $n \times n$ normal matrix A.

Def 12.6.7 Suppose $\mathbb{F} = \mathbb{C}$ (or $\mathbb{R}$).

1. Suppose there exists an ordered orthonormal basis B for V s.t. $[T]_B$ is a diagonal matrix, then T is called **unitarily (orthogonally) diagonalizable**
2. A complex (real) square matrix A is called **unitarily (orthogonally) diagonalizable** if there exists a unitary (orthogonal) matrix P s.t. P^*AP is a diagonal matrix.

Thm 12.6.9 && Thm 12.6.12

1. Let V be a complex (real) finite dimensional inner product space where $\dim(V) \geq 1$. A linear operator T on V is unitarily (orthogonally) diagonalizable if and only if T is normal (self-adjoint).
2. A complex (real) square matrix A is unitarily (orthogonally) diagonalizable if and only if A is normal (symmetric).

To find an ordered orthonormal basis B so that the matrix $[T]_B$ is a diagonal matrix, we just union the bases of all eigenspaces of $[T]_C$ where C is any orthonormal basis for V and normalize each bases.

Tut 10 Q5.

- T is self-adjoint iff all its eigenvalues are real.
- A linear operator P is **positive definite** if P is self-adjoint and $\langle P(u), u\rangle > 0$. P is positive definite iff all its eigenvalues are (nonzero) positive real numbers.

PYP 2018/2019S1.

- If T is an invertible linear operator, then $T^* \circ T$ is unitarily diagonalizable and all its eigenvalues are nonzero real positive numbers.