Canonical Forms
Linear Algebra Notes

Satya Mandal

October 25, 2005

1 Introduction

Here $\mathbb{F}$ will denote a field and V will denote a vector space of dimension $\dim(V) = n$. (*In this note, unless otherwise stated,* $n = \dim(V)$)

We will study operatores T on V. The goal is to investigate if we can find a basis $e_1, \ldots, e_n$ such that

$$the\ matrix\ of\ T = diagonal(\lambda_1, \lambda_2, \ldots, \lambda_n)$$

is a diagonal matrix. This will mean that

$$T(e_i) = \lambda_i e_i.$$

2 Characteristic Values

2.1 Basic Definitions and Facts

Here we will discuss basic facts.

2.1 (Definition.) Let V be a vector space over a field $\mathbb{F}$ and $T \in L(V, V)$ be linear operator.

1. A scalar $\lambda \in \mathbb{F}$ is said to be a **characterisitic value** of T, if

   $$T(e) = \lambda e \quad for\ some \quad e \in V \quad with \ \ e \neq 0.$$

   A characterisitic value is also known an **eigen value.**

2. This non-zero element $e \in V$ above is called a **characterisitic vector** of T associated to λ. A characterisitic vector is also known an **eigen vector.**

3. Write
$$N(\lambda) = \{v \in V : T(v) = \lambda v\}.$$
Then $N(\lambda)$ is a subspace of V and is said to be the **characterisitic space** or **eigen space** of T associated to λ.

2.2 (Lemma.) Let V be a vector space over a field $\mathbb{F}$ and $T \in L(V,V)$ Then T is singular if and only if $\det(T) = 0$.

Proof. ($\Rightarrow$): We have $T(e_1) = 0$ for some $e_1 \in V$ with $e_1 \neq 0$. We can extend e_1 to a basis $e_1, e_2, \ldots, e_n$ of V. Let A be the matrix of T with respect to this basis. Since, $T(e_1) = 0$, the first row of A is zero. Therefore,
$$\det(T) = \det(A) = 0.$$
So, this implication is established.

($\Leftarrow$): Suppose $det(T) = 0$. Let $e_1, e_2, \ldots, e_n$ be a basis of V and A be the matrix of T with respect to this basis. So
$$\det(T) = \det(A) = 0.$$
Therefore, A is not invertible. Hence $AX = 0$ for some non-zero column
$$X = \begin{pmatrix} c_1 \\ c_2 \\ \cdots \\ c_n \end{pmatrix}.$$

Write $v = c_1 e_1 + c_2 e_2 + \cdots + c_n e_n$. Since not all c_i is zero, $v \neq 0$. Also,
$$T(v) = \sum c_i T(e_i)$$
$$= (T(e_1), T(e_2), \ldots, T(e_n)) \begin{pmatrix} c_1 \\ c_2 \\ \cdots \\ c_n \end{pmatrix} = (e_1, e_2, \ldots, e_n) A \begin{pmatrix} c_1 \\ c_2 \\ \cdots \\ c_n \end{pmatrix} = 0..$$

So, T is singular.

The following are some equivalent conditions.

2.3 (Theorem.) Let V be a vector space over a field $\mathbb{F}$ and $T \in L(V, V)$ be linear operator. Let $\lambda \in \mathbb{F}$ be a scalar. Then the following are equivalent:

1. λ is a characteristic value of T.

2. The operator $T - \lambda I$ is singular (or is not invertible).

3. $\det(T - \lambda I) = 0$.

Proof. ((1) $\Rightarrow$ (2)): We have $(T - \lambda I)(e) = 0$ for some $e \in V$ with $e \neq 0$. So, $T - \lambda I$ is singular and (2) is established.

((2) $\Rightarrow$ (1)): Since $T - \lambda I$ is singular, we have $(T - \lambda I)(e) = 0$ for some $e \in V$ with $e \neq 0$. Therefore, (1) is established.

((2) $\Leftrightarrow$ (3)): Immediate from the above lemma.

2.4 (Definition.) Let $A \in \mathbb{M}_n(\mathbb{F})$ be an $n \times n$ matrix with entries in a field $\mathbb{F}$.

1. A scalar $\lambda \in \mathbb{F}$ is said to be a **characteristic value** of A if $\det(A - \lambda I_n) = 0$. Equivalently, $\lambda \in \mathbb{F}$ is said to be a characteristic value of A if the matrix $(A - \lambda I_n)$ is not invertible.

2. The monic polynomial $\det(XI - A)$ is said to be the **characteristic polynomial** of A. Therefore, characteristic values of A are the roots of the characteristic polynomial of A.

2.5 (Lemma.) Let $A, B \in \mathbb{M}_n(\mathbb{F})$ be two $n \times n$ matrices with entries in a field $\mathbb{F}$. If A and B are similar, then they have some characteristic polynomials.

Proof. Suppose A, B are similar matrices. Then $A = PBP^{-1}$. The characteristic polynomial of $A = \det(XI - A) = \det(XI - PBP^{-1}) = \det(P(XI - B)P^{-1}) = \det(XI - B) =$ the characteristic polynomial of B.

2.6 (Definitions and Facts) Let V be a vector space over a field $\mathbb{F}$ and $T \in L(V, V)$ be linear operator.

1. Let A be the matrix of T with respect to some basis E of V. We define the **characteristic polynomial of** T to be the characteristic polynomial of A. Note that this polynomial is well defined by the above lemma.

2. We say that T is **diagonalizable** if the there is a basis $e_1, \ldots, e_n$ such that each e_i is a characteristic value of T. In this case, $T(e_i) = \lambda_i e_i$ for some $\lambda_i \in \mathbb{F}$. Hence, with respect to this basis, the matrix of $T = Diagonal(\lambda_1, \lambda_2, \ldots, \lambda_n)$

   Depending on how many of these eigen values λ_i are distinct, we can rewrite the matrix of T.

3. Also note, if T is diagonilazable as above, the characteric polynomial of $T = (X - \lambda_1)(X - \lambda_2) \cdots (X - \lambda_n)$, which is completely factorizable.

4. Suppose T is diagonalizable, as above. Depending on how many of these eigen values λ_i are distinct, we can rewrite the matrix of T.

   Now sSuppose T is diagonalizable $c_1, c_2, \ldots, c_r$ are the distinct eigen values of T. Then the matrix of T with respect to some basis of V looks like:

$$\begin{pmatrix} c_1 I_{d_1} & 0 & \cdots & 0 \\ 0 & c_2 I_{d_2} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & c_r I_{d_r} \end{pmatrix}$$

   where I_k is the identity matrix of order k. So, $d_1 + d_2 + \cdots + d_r = n = \dim(V)$.

   In this case, the characteristic polynomial of

$$T = (X - c_1)^{d_1} (X - c_2)^{d_2} \cdots (X - c_r)^{d_r}.$$

   Further,

$$d_i = \dim(N(c_i)).$$

   (see 3 of 2.1.)

2.7 (Read Examples) Read Example 1 and 2 in page 184

2.2 Decomposition of V

2.8 (Definition) *Suppose V is vector space of a field $\mathbb{F}$ with $\dim(V) = n$. Let $f(X) = a_0 + a_1X + a_2X^2 + \cdots + a_rX^r \in \mathbb{F}[X]$ be polynomial and $T \in L(V, V)$ be linear operator. Then, by definition,*

$$f(T) = a_0Id + a_1T + a_2T^2 + \cdots + a_rT^r \in L(V, V)$$

is an operator. So, $L(V, V)$ becomes a module over $\mathbb{F}[X]$.

2.9 (Remark) Suppose V is vector space of a field $\mathbb{F}$ with $\dim(V) = n$. Let $T \in L(V, V)$ be linear operator. Let $f(X)$ be a characteristic polynomial of T. We have understanable interest how $f(T)$ works.

2.10 (Lemma) Suppose V is vector space of a field $\mathbb{F}$ with $\dim(V) = n$. Let $T \in L(V, V)$ be linear operator. Let $f(X) \in \mathbb{F}[X]$ be any polynomial. Suppose

$$T(v) = \lambda v$$

for some $v \in V$ and $\lambda \in \mathbb{F}$. Then

$$f(T)(v) = f(\lambda)v.$$

The proof is obvious. This means if λ is an eigen value of T then $f(\lambda)$ is an eigen value of $f(T)$

2.11 (Lemma) Suppose V is vector space of a field $\mathbb{F}$ with $\dim(V) = n$. Let $T \in L(V, V)$ be linear operator. Suppose $c_1, \ldots, c_k$ are the distinct eigen values of T. Let

$$W_i = N(c_i)$$

be the eigen space of T associated to c_i. Write

$$W = W_1 + W_2 + \cdots + W_k.$$

Then

$$\dim(W) = \dim(W_1) + \dim(W_2) + \cdots + \dim(W_k).$$

Indeed, if

$$E_i = \{e_{ij} \in W_i : j = 1, \ldots, d_i\}$$

is a basis of W_i, then

$$E = \{e_{ij} \in W_i : j = 1, \ldots, d_i; j = 1, \ldots, k\}$$

is basis of W.

Proof. We only need to prove the last part. So, let

$$\sum \lambda_{ij} e_{ij} = 0$$

for some scalar $\lambda_{ij} \in \mathbb{F}$.

Write

$$\omega_i = \sum_{j=1}^{d_i} \lambda_{ij} e_{ij}.$$

Then $\omega_i \in W_i$ and

$$\omega_1 + \omega_2 + \cdots + \omega_k = 0 \qquad (I).$$

We will first prove that $\omega_i = 0$.

Since

$$T(e_{ij}) = c_i e_{ij}$$

for any polynomial $f(X) \in \mathbb{F}[X]$, we have

$$f(T)(e_{ij}) = f(c_i) e_{ij}.$$

Therefore,

$$f(T)(\omega_i) = \sum_{j=1}^{d_i} \lambda_{ij} f(T)(e_{ij}) = \sum_{j=i}^{d_i} \lambda_{ij} f(c_i) e_{ij} = f(c_i)\omega_i \qquad (II).$$

Now let

$$g(X) = \frac{\prod_{i=2}^{k} (X - c_i)}{\prod_{i=2}^{k} (c_1 - c_i)}$$

Note $g(X)$ is a polynomial. Also note this definition/ expression makes sense because $c_1, \ldots, c_k$ are distinct. And also $g(c_1) = 1$ and $g(c_2) = f_1(c_3) = \cdots = g(c_k) = 0$.

Use (II) and apply to (I). We get

$$0 = g(T)(\sum_{i=1}^{k} \omega_i) = \sum_{i=1}^{k} g(T)(\omega_i) = \sum_{i=1}^{k} g(c_i)\omega_i. = \omega_1$$

Similarly, $\omega_i = 0$ for $i = 1. \ldots, k$. This means

$$0 = \omega_i = \sum_{j=i}^{d_i} \lambda_{ij} e_{ij}$$

Since E_i is a basis, $\lambda_{ij} = 0$ for all i, j and the proof is complete.

Following is the final theorem in this section.

2.12 (Theorem) Suppose V is vector space of a field $\mathbb{F}$ with $\dim(V) = n$. Let $T \in L(V, V)$ be linear operator. Suppose $c_1, \ldots, c_k$ are the distinct eigen values of T. Let

$$W_i = N(c_i)$$

be the eigen space of T associated to c_i. Then the following are equivalent:

1. T is diagonalizable.
2. The characteristic polynomial for T is
$$f = (X - c_1)^{d_1}(X - c_2)^{d_2} \cdots (X - c_k)^{d_k}$$
and $\dim(W_i) = d_i$ for $i = 1, \ldots, k$.
3. $\dim(W_1) + \dim(W_2) + \cdots + \dim(W_k) = \dim(V)$.

Proof. ((1) $\Rightarrow$ (2)): This is infact obvious. If $c_1, \ldots, c_k$ are the distinct eigen values and since T is diagonalizable, the matrix of T is as in (4) of (2.6). Therefore, we can compute the characteristic polynomial using this matrix and (2) is established.

((2) $\Rightarrow$ (3)): We have $\dim(V) = degree(f)$. Therefore,

$$\dim(V) = d_1 + d_2 + \cdots + d_k = \dim(W_1) + \dim(W_2) + \cdots + \dim(W_k).$$

Hence (3) is established.

((3) $\Rightarrow$ (1)): Write $W = W_1 + \cdots + W_k$. Then, by lemma 2.11

$$\dim(W) = \dim(W_1) + \dim(W_2) + \cdots + \dim(W_k).$$

Therefore, by (3), $\dim(V) = \dim(W)$. Hence (3) is established and the proof is complete.

In fact, I would like to restate the "final theorem" 2.12 in terms of direct sum of linear subspaces. So, I need to define direct sum of vector spaces.

2.13 (Definition) Let V be a vector space over $\mathbb{F}$ and $V_1, V_2 \ldots, V_k$ be subspaces of V. We say that V is **direct sum of** $V_1, V_2, \ldots, V_k$, if each element $x \in V$ can be written uniquely as

$$x = \omega_1 + \omega_2 + \cdots + \omega_k$$

with $\omega_i \in V_i$.

Equivalently, if

1. $V = V_1 + V_2 + \cdots + V_k$, and
2. $\omega_1 + \omega_2 + \cdots + \omega_k = 0$ with $\omega_i \in V_i$ implies that $\omega_i = 0$ for $i = 1, \ldots, k$.

If V is **direct sum of** $V_1, V_2 \ldots, V_k$ then we write

$$V = V_1 \oplus V_2 \oplus \cdots \oplus V_k.$$

Following is a proposition on direct sum decomposition.

2.14 (Proposition) Let V be a vector space over $\mathbb{F}$ with $\dim(V) = n < \infty$. Let $V_1, V_2 \ldots, V_k$ be subspaces of V Then

$$V = V_1 \oplus V_2 \oplus \cdots \oplus V_k$$

if and only if $V = V_1 + V_2 + \cdots + V_k$ and

$$\dim(V) = \dim(V_1) + \dim(V_2) + \cdots + \dim(V_k).$$

Proof. ($\Rightarrow$): Obvious.
($\Leftarrow$): Let $E_i = \{e_{ij} : j = 1, \ldots, d_i\}$ be basis of V_i. Let $E = \{e_{ij} : j = 1, \ldots, d_i; i = 1, \ldots, k\}$. Since $V = V_1 + V_2 + \cdots + V_k$, we have $V = Span E$. Since $\dim(V) = cardinlity(E)$, we have E forms a basis of V. Now it follows that if $\omega_1 + \cdots + \omega_k = 0$ with $\omega_i \in W_i$ then $\omega_i = 0 \quad \forall i$. This completes the proof.

Now we restate the final theorem 2.12 in terms of direct sum.

2.15 (Theorem) Suppose V is vector space of a field $\mathbb{F}$ with $\dim(V) = n$. Let $T \in L(V, V)$ be linear operator. Suppose $c_1, \ldots, c_k$ are the distinct eigen values of T. Let

$$W_i = N(c_i)$$

be the eigen space of T associated to c_i. Then the following are equivalent:

1. T is diagonalizable.
2. The characteristic polynomial for T is

   $$f = (X - c_1)^{d_1}(X - c_2)^{d_2} \cdots (X - c_k)^{d_k}$$

   and $\dim(W_i) = d_i$ for $i = 1, \ldots, k$.
3. $\dim(W_1) + \dim(W_2) + \cdots + \dim(W_k) = \dim(V)$.
4. $V = W_1 \oplus W_2 \oplus \cdots \oplus W_k$.

Proof. Clearly, we proved

$$(1) \Longleftrightarrow (2) \Longleftrightarrow (3).$$

We will prove $(3) \Longleftrightarrow (4)$.
$((4) \Rightarrow (3))$: This part is obvious because we can combine bases of W_i to get a basis of V.
$((3) \Rightarrow (4))$:Write $W = W_1 + W_2 + \cdots + W_k$. Because of (4) and by lemma 2.11, $\dim(W) = \sum \dim(W_i) = \dim(V)$. Therefore, $V = W = W_1 + W_2 + \cdots + W_k$.

Since $\dim(V) = \sum \dim(W_i)$, by proposition 2.14 $V = W_1 \oplus W_2 \oplus \cdots \oplus W_k$ and the proof is complete.

3 Annihilating Polynomials

Suppose K is a commutative ring and M be a $K-$module. For $x \in M$, we define annihiltor of x as

$$ann(x) = \{\lambda \in K : \lambda x = 0\}.$$

Note that $ann(x)$ is an ideal of K. (*That means*

$$ann(x) + ann(x) \subseteq ann(x) \quad and \quad K * ann(x) \subseteq ann(x)).$$

We shall consider annihilator of a linear operator, as follows.

3.1 Minimal (monic) polynomials

3.1 (Facts) Let V be a vector space over a field $\mathbb{F}$ with $\dim(V) = n$.

Recall, we have seen that $M = L(V, V)$ is a $\mathbb{F}[X]-$module. For $f(X) \in \mathbb{F}[X]$ and $T \in L(V, V)$, scalar multiplication is defined by $f * T = f(T) \in L(V, V)$

1. So, for a linear operator $T \in L(V, V)$, the annihilator of T is:

   $$ann(T) = \{f(X) \in \mathbb{F}[X] : f(T) = 0\}$$

   is an ideal of the polynomial ring $\mathbb{F}[X]$.

2. Note that $ann(T)$ is a non-zero proper ideal. It is non-zero because $\dim(L(V, V)) = n^2$ and hance

   $$1, T, T^2, \ldots, T^{n^2}$$

   is a linearly dependent set.

3. Also recall that any ideal I of $\mathbb{F}[X]$ is a principal ideal, which means that $I = \mathbb{F}[X]p$ where p is the non-zero monic in I polynomial of least degree.

4. Therefore,
$$ann(T) = \mathbb{F}[X]p(X)$$
where $p(X)$ is the monic polynomial of least degree such that $p(T) = 0$.

   This polynomial $p(X)$ is defined to be the **minimal monic polynomial (MMP)** for T.

5. Let us consider similar concepts for square matrices.
   (a) For an $n \times n$ matrix A, we define annihilator $ann(A)$ of A and **minimal monic polynomial** of A is a similar way.
   (b) Suppose two $n \times n$ matrices A, B similar an d $A = PBP^{-1}$. Then for a polynomial $f(X) \in \mathbb{F}[X]$ we have
   $$f(A) = Pf(B)P^{-1}.$$
   (c) Therefore $ann(A) = ann(B)$.
   (d) Hence A and B have SAME minimal monic polynomial.

3.2 Comparison of minimal monic and characteristic polynomials:

Given a linear operator T we can think of two polynomials - the minimal monic polynomial and the characteristic polynomial of T. We will compare them.

3.2 (Theorem) *Let V be a vector space over a field $\mathbb{F}$ with* $\dim(V) = n$. *Suppose $p(X)$ is the minimal monic polynomial of T and $g(X)$ is the characteristic polynomial of T. Then p, g have the same roots in* $\mathbb{F}$. (Although multiplicity may differ.)

Same statement holds for matrices.

Proof. We will prove, for $c \in \mathbb{F}$,
$$p(c) = 0 \Longleftrightarrow g(c) = 0.$$

Recall $g(X) = \det(XI - A)$, where A is the matrix of T with respect to some basis. Also by theorem 2.3, $g(c) = 0$ is and only if $cI - T$ is singular.

Now suppose $p(c) = 0$. So, $p(X) = (X - c)q(x)$. for some $q(X) \in \mathbb{F}$. Since $degree(q) < degree(p)$, by minimality of p we have $q(T) \neq 0$. Let $v \in V$ be such that $v \neq 0$ and $e = q(T)(v) \neq 0$. Since, $p(T) = 0$, we have $(T - cI)q(T) = 0$. Hence $0 = (T - cI)q(T)(v) = (T - cI)(e)$. So, $(T - cI)$ is singular and hence $g(c) = 0$. This estblishes the proof of this part.

Now assume that $g(c) = 0$. Therefore, $T - cI$ is singular. So, there is vector $e \in V$ with $e \neq 0$ such that $T(e) = ce$. Applying this equation to p we have

$$p(T)(e) = p(c)e$$

(see lemma 2.10). Since $p(T) = 0$ and $e \neq 0$, we have $p(c) = 0$ and the proof is complete.

The above theorem raises the question if these two polynomial are same? Answer is, not in general. But MMP devides the ch-polynomial as follows.

3.3 (Caley-Hamilton Theorem) *Let V be a vector space over a field $\mathbb{F}$ with $\dim(V) = n$. Suppose $Q(X)$ is the characteristic polynomial of T. Then $Q(T) = 0$.*

In particular, if $p(X)$ is the minimal monic polynomial of T, then

$$p \mid Q.$$

Proof. Write

$$K = \mathbb{F}[T] = \{f(T) : f \in \mathbb{F}[X], T \in L(V, V)\}.$$

Observe that

$$\mathbb{F} \subseteq K \subseteq L(V, V).$$

are subrings. Note $Q(T) \in K$ and we will prove $Q(T) = 0$.

Let $e_1, \ldots, e_n$ be a basis of V and $A = (a_{ij})$ be the matrix of T. So, we have

$$(T(e_1), T(e_2), \dots, T(e_n)) = (e_1, e_2, \dots, e_n)A. \qquad (I)$$

Consider the following matrix, with entries in K :

$$B = \begin{pmatrix} T - a_{11}I & -a_{12}I & -a_{13}I & \cdots & -a_{1n}I \\ -a_{21}I & T - a_{22}I & -a_{23}I & \cdots & -a_{2n}I \\ -a_{31}I & -a_{32}I & T - a_{33}I & \cdots & -a_{3n}I \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ -a_{n1}I & -a_{n2}I & -a_{n3}I & \cdots & T - a_{nn}I \end{pmatrix}.$$

Note that the

$$Q(X) = \det(I_n X - A).$$

Therefore (*I think this is the main point to understandin this proof.*),

$$Q(T) = \det(I_n T - A) = \det(B).$$

The above equation (I) says that

$$(e_1, e_2, \dots, e_n)B = (0, 0, \dots, 0).$$

Multiply this equation by $Adj(B)$, and we get

$$(e_1, e_2, \dots, e_n)BAdj(B) = (0, 0, \dots, 0)Adj(B) = (0, 0, \dots, 0).$$

Therefore,

$$(e_1, e_2, \dots, e_n)(\det(B))I_n = (0, 0, \dots, 0).$$

Therefore,

$$(e_1, e_2, \dots, e_n)(Q(T))I_n = (0, 0, \dots, 0).$$

This implies that

$$Q(T)(e_i) = 0 \quad \forall i = 1, \dots, n.$$

Hence $Q(T) = 0$ and the proof is complete.

4 Invatiant Subspaces

4.1 (Definition) Let V be a vector space over the field $\mathbb{F}$ and $T : V \to V$ be a linear operator. A subspace W of V is said to be **invariant under** T if

$$T(W) \subseteq W.$$

4.2 (Examples) Let V be a vector space over the field $\mathbb{F}$ and

$$T : V \to V$$

be a linear operator.

1. (Trivial Examples) Then V and $\{0\}$ are invariant under T.
2. Suppose e be an eigen vector of T and $W = \mathbb{F}e$. Then W is invariant under T.
3. Suppose λ be an eigen value of T and $W = N(\lambda)$ be the eigen space of λ. Then W is invariant under T.

4.3 (Remark) Let V be a vector space over the field $\mathbb{F}$ and

$$T : V \to V$$

be a linear operator. Suppose W is an invariant subspace T. Then the restriction map

$$T_{|W} : W \to W$$

is an well defined linear operator on W. So, the following diagram

$$\begin{array}{ccc} W & \xrightarrow{T_{|W}} & W \\ \downarrow & & \downarrow \\ V & \xrightarrow{T} & V \end{array}$$

commutes.

4.4 (Remark) Let V be a vector space over the field $\mathbb{F}$ with $\dim(V) = n < \infty$. Let

$$T : V \to V$$

be a linear operator. Suppose W is an invariant subspace T and

$$T_{|W} : W \to W$$

is the restriciton of T.

1. Let p be the characteristic polynomial of T and q be the characteristic polynomial of $T_{|W}$. Then $q \mid p$.

2. Also let P be the minimal (monic) polynomial of T and Q be the minimal (monic) polynomial of $T_{|W}$. Then $Q \mid P$.

Proof. Proof of (2) is easier. Since $P(T) = 0$ we also have $P(T_{|W}) = 0$. Therefore

$$P(X) \in ann(T_{|W}) = \mathbb{F}[\mathbb{X}]Q(X).$$

Hence $Q \mid P$ and proof of (2) is complete.

To prove (1), let $E = \{e_1, e_2, \ldots, e_r\}$ be a basis of W. Extend this basis to a basis $\mathcal{E} = \{e_1, e_2, \ldots, e_r, e_{r+1}, \ldots, e_n\}$ of V. Let A be the matrix of T with respect to $\mathcal{E}$ and B be the matrix of $T_{|W}$ with respect to E. So, we have

$$(T(e_1), \ldots, T(e_r)) = (e_1, \ldots, e_r)B$$

and

$$(T(e_1), \ldots, T(e_r), T(e_{r+1}), \ldots, T(e_n)) = (e_1, \ldots e_r, e_{r+1}, \ldots, , e_n)A.$$

Therefore, A can be written as blocks as follows:

$$A = \begin{pmatrix} B & C \\ 0 & D \end{pmatrix}$$

So,

$$P(X) = \det(I_n X - A) = \det(I_r X - B)\det(I_{n-r} X - D)$$

and

$$Q(X) = \det(I_r X - B).$$

Therefore $Q \mid P$. The proof is complete.

4.5 (Definitions and Remarks) 1. Suppose $\mathbb{F}$ is a field. Recall an $n \times n$ matrix $A = (a_{ij})$ is call an **upper triangular** matrix if $a_{ij} = 0$ for all i, j with $1 \leq i < j \leq n$. Similarly, we define **lower triangular** matrices.

2. Now let V be a vector space over $\mathbb{F}$ with $\dim V = n < \infty$. A linear operator $T : V \to V$ is said to be **triangulable,** if there is a basis $E = \{e_1, \ldots, e_n\}$ of V such that the matrix of V is an (upper) triangular matrix. (*Note that it does not make a difference if we say "upper" or "lower" trinagular. To avoid confusion, we will assume upper triangular.*)

3. Now suppose a linear operator T is triangulable. So, for a basis $E = \{e_1, \ldots, e_n\}$ we have $(T(e_1), \ldots, T(e_n)) = (e_1, \ldots, e_n)A$ for some triangular matrix $A = (a_{ij})$. We assume that A is upper triangular. For $1 \leq r \leq n$, write $W_r = span(e_1, \ldots, e_r)$. Then W_r is invariant under T.

4. (**Factorization.**) Suppose $T \in L(V, V)$ is triangulable. So, the matrix of T, with respect to a basis $e_1, \ldots, e_n$, is an upper triangular matrix $A = (a_{ij})$. Note that the characterictic polynomial q of T is given by

$$q(X) = \det(IX - A) = (X - a_{11})(X - a_{22}) \cdots (X - a_{nn}).$$

Therefore, q is completely factorizable. So, we have

$$q(X) = (X - c_1)^{d_1}(X - c_2)^{d_2} \cdots (X - c_k)^{d_k}.$$

where $d_1 + d_2 + \cdots + d_k = \dim V$ and $c_1, \ldots, c_k$ are the distinct eigen values of T.

Also, since the minimal monic polynomial p of T divides q, it follows that p is also completely factorizable. Therefore,

$$p(X) = (X - c_1)^{r_1}(X - c_2)^{r_2} \cdots (X - c_k)^{r_k}.$$

where $r_i \leq d_i$ for $i = 1, \ldots, k$.

4.6 (Theorem) Let V be a vector space over $\mathbb{F}$ with with finite dimension $\dim V = n$ and $T : V \to V$ be a linear operator on V. Then T is triangulable if and only if the minimal polynomial p of T is a product of linear factors.

Proof. ($\Rightarrow$): We have already shown in (4) of Remark 4.5, that if T is triangulable then the MMP p factors into linear factors.

($\Leftarrow$): Now assume that the MMP p factors as

$$p(X) = (X - c_1)^{r_1}(X - c_2)^{r_2} \cdots (X - c_k)^{r_k}.$$

Let q denote the characteristic polynomial of T. Since p and q have the same roots, $q(c_1) = q(c_2) = \cdots = q(c_k) = 0$. Now we will split the proof into several steps.

Step-1: Write $\lambda_1 = c_1$. By (2.3), λ_1 is an eigen value of T. So, there is a non-zero vector $e_1 \in V$ such that $T(e_1) = \lambda_1 e_1$. Write $W_1 = Span(e_1)$.

Step-2: Extend e_1 to a basis $e_1, E_2, \ldots, E_n$ of V. Write $V_1 = Span(E_2, \ldots, E_n)$. Note that

1. e_1 is linearly independent and $\dim W_1 = 1$.
2. W_1 is invariant under T.
3. $\dim V_1 = n - 1$ and $V = W_1 \oplus V_1$.

Let $v \in V_1$ and $T(v) = \lambda_1 e_1 + \lambda_2 E_2 + \cdots + \lambda_n E_n$, for some $\lambda_1, \ldots, \lambda_n \in \mathbb{F}$. Define $T_1(v) = \lambda_2 E_2 + \cdots + \lambda_n E_n \in V_1$. Then

$$T_1 : V_1 \to V_1$$

is a well defined linear operator on V_1. Diagramatically, T_1 is given by

$$\begin{array}{ccc} V_1 & \xrightarrow{T_1} & V_1 \\ \downarrow & & \uparrow{\scriptstyle pr} \\ V = W_1 \oplus V_1 & \xrightarrow{T} & V = W_1 \oplus V_1 \end{array}$$

where $pr : V = W_1 \oplus V_1 \to V_1$ is the projection map. Let p_1 be the MMP of T_1. Now, we proceed to prove that $p_1 \mid p$.

$$Claim : \quad ann(T) \subseteq ann(T_1).$$

To prove this claim, let A be the matrix of T with respect to $e_1, E_2, \ldots, E_n$ and B be the matrix of T_1 with respect to $E_2, \ldots, E_n$. Since W_1 is invariant under T, we have

$$A = \begin{pmatrix} \lambda_1 & C \\ 0 & B \end{pmatrix}.$$

Therefore, the matrix of T^m is given by

$$A^m = \begin{pmatrix} \lambda_1^m & C_m \\ 0 & B^m \end{pmatrix}.$$

For some matrix C_m. Therefore, for a polynomial $f(X) \in \mathbb{F}[X]$ that matrix $f(A)$ of $f(T)$ is given by

$$f(A) = \begin{pmatrix} f(\lambda_1) & C_* \\ 0 & f(B) \end{pmatrix}.$$

some matrix C_*. So, if $f(X) \in ann(T)$ then $f(T) = 0$. Hence $f(A) = 0$. This implies $f(B) = 0$ and hence $f(T_1) = 0$. So, $ann(T) \subseteq ann(T_1)$ and the claim is established.

Therefore, $p_1 \mid p$. So, p_1 satisfies the hypothesis of the theorem. So, there is a an element $e_2 \in V_1$ such that $T_1(e_2) = \lambda_2 e_2$ where $(X - \lambda_2) \mid p_1 \mid p$.

Also follows that $T(e_2) = ae_1 + \lambda_2 e_2$.

Step-3 Write $W_2 = Span(e_1, e_2)$.

Note that

1. e_1, e_2 are linearly independent and $\dim W_2 = 2$.
2. W_2 is invariant under T.
3. Also
$$(T(e_1), T(e_2)) = (e_1, e_2) \begin{pmatrix} \lambda_1 & a_{12} \\ 0 & \lambda_2 \end{pmatrix}.$$

Step-4 If $W_2 \neq V$ (that is if $2 < n$), the process will continue. We extend e_1, e_2 to a basis $e_1, e_2, E_3, \ldots, E_n$ of V (Well, they are different E_i, not the same as in previous steps.) Write $V_2 = Span(E_3, \ldots, E_n)$. Note

1. $\dim(V_2) = n - 2$
2. $V = W_2 \oplus V_2$.

As in the previous steps, define $T_2 : V_2 \to V_2$as in the diagram (you should define explicitly):

$$\begin{array}{ccc} V_2 & \xrightarrow{T_2} & V_2 \\ \downarrow & & \uparrow{\scriptstyle pr} \\ V = W_2 \oplus V_2 & \xrightarrow{T} & V = W_2 \oplus V_2 \end{array}$$

where $pr : V = W_2 \oplus V_2 \to V_2$ is the projection map.

Let p_2 be the MMP of T_2. Using same argument, we will prove $p_2 \mid p$. Then we can find $\lambda_3 \in \mathbb{F}$ and $e_3 \in V_2$ such that $T_3(e_3) = \lambda_3 e_3$ where $(X - \lambda_3) \mid p_2 \mid p$. Therefore $T(e_3) = a_{13}e_2 + a_{23}e_2 + \lambda_3 e_3$.

So, we have

$$(T(e_1), T(e_2), T(e_3)) = (e_1, e_2, e_3) \begin{pmatrix} \lambda_1 & a_{12} & a_{13} \\ 0 & \lambda_2 & a_{23} \\ 0 & 0 & \lambda_3 \end{pmatrix}.$$

Final Step: The process continues for n steps and we get linearly independent set (basis) $e_1, e_2, \dots, e_n$ such that

$$(T(e_1), T(e_2), T(e_3), \dots, T(e_n)) =$$

$$(e_1, e_2, e_3, \dots, e_n) \begin{pmatrix} \lambda_1 & a_{12} & a_{13} & \dots & a_{1n} \\ 0 & \lambda_2 & a_{23} & \dots & a_{2n} \\ 0 & 0 & \lambda_3 & \dots & a_{3n} \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \lambda_n \end{pmatrix}.$$

This completes the proof.

Recall a field $\mathbb{F}$ is said to be a an **algebraically closed field** if every non-constant polynomial $f \in \mathbb{F}[X]$ has a root in $\mathbb{F}$. It follows that k is an algebraically closed field if and only if every non-constant polynomial $f \in \mathbb{F}[X]$ product linear polynomials.

4.7 (Theorem) Suppose $\mathbb{F}$ is an algebraically closed field. Then every $n \times n$ matrix over $\mathbb{F}$ is similar to a triangular matrix.

Proof. Consider the operation

$$T : \mathbb{F}^n \to \mathbb{F}^n$$

such that $T(X) = AX$. Now use the above theorem.

4.8 (Theorem) Let V be a vector space over $\mathbb{F}$ with with finite dimension $\dim V = n$ and $T : V \to V$ be a linear operator on V. Then T is diagonalizable if and only if the minimal polynomial p of T is of the form

$$p = (X - c_1)(X - c_2)\cdots(X - c_k)$$

where $c_1, c_2, \ldots, c_k$ are the distinct eigen values of T.

Proof. ($\Rightarrow$): Suppose T is diagonalizable. Then, there is a basis $e_1, \ldots, e_n$ of V such that

$$(T(e_1), T(e_2), \ldots, T(e_n) =$$

$$(e_1, \ldots, e_n)\begin{pmatrix} c_1 I_{d_1} & 0 & 0 & \ldots & 0 \\ 0 & c_2 I_{d_2} & 0 & \ldots & 0 \\ 0 & 0 & c_3 I_{d_3} & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & c_n I_{d_k} \end{pmatrix}.$$

Write $g(X) = (X - c_1)(X - c_2)\cdots(X - c_k)$ we will prove $g(T) = 0$. For, $i = 1, \ldots, d_1$ we have $(T - c_1)(e_i) = 0$. Therefore,

$$g(T)(e_i) = (T - c_1)(T - c_2)\cdots(T - c_k)(e_1) = 0.$$

Similarly, $g(T)(e_i) = 0$ for all $i = 1, \ldots, n$. So, $g(T) = 0$. Hence $p \mid g$. Since $c_1, \ldots, c_k$ are roots of both, we have $p = g$. Hence this part of the proof is complete.

($\Leftarrow$): We we asssume that $p(X) = (X - c_1)(X - c_2)\cdots(X - c_k)$ and prove that T is digonalizable. Let $W_i = N(c_i)$ be a eigen space of c_i. Let $W = \sum_{i=1}^{k} W_i$ be the sum of eigen spaces. Assume that $W \neq V$.

Now we will repeat some protions of the proof of theorem 4.6 and get a contradiction. Let $e_1, \ldots, e_m$ be a basis of W and $e_1, \ldots, e_m, E_{m+1}, \ldots, E_n$ be a basis of V. Write $V' = Span(E_{m+1}, \ldots, E_n)$. Note

1. W is invariant under T.

2. $V = W \oplus V'$.

Define $T' : V' \to V'$ according to the diagram:

$$\begin{array}{ccc} V' & \xrightarrow{\quad T' \quad} & V' \\ \downarrow & & \uparrow {\scriptstyle pr} \\ V = W \oplus V' & \xrightarrow{\ T\ } & V = W \oplus V' \end{array}$$

where $pr : V = W \oplus V' \to V'$ is the projection map.

As in the prrof of theorem 4.6, the MMP p' of T' divides p. Therefore, there is an element $e \in V'$ such that $T'(e) = \lambda e$ for some $\lambda \in \{c_1, c_2, \ldots, c_k\}$. We assume $\lambda = c_1$. Hence

$$T(e) = a_1 e_1 + \cdots + a_n e_m + c_1 e$$

where $a_i \in \mathbb{F}$. We can rewrite this equation as

$$T(e) = \beta + c_1 e$$

where $\beta = \omega_1 + \omega_2 + \cdots + \omega_k \in W$ and $\omega_i \in W_i$. So,

$$(T - c_1)(e) = \beta.$$

Since $T(W) \subseteq W$, for $h(X) \in \mathbb{F}[X]$ we have $h(T)(\beta) \in W$. Write

$$p = (X - c_1)q \qquad and \qquad q(X) - q(c_1) = h(X)(X - c_1).$$

So,

$$(q(T) - q(c_1))(e) = h(T)(T - c_1)(e) = h(T)(\beta)$$

is in W. Also

$$0 = p(T)(e) = (T - c_1)q((T)(e)$$

Therefore $q((T)(e) \in W_1 \subseteq W$. So, $q(c_1)e = q((T)(e) - (q(T) - q(c_1))(e)$ is in W. Since $q(c_1) \neq 0$ we get $e \in W$. This is a contradiction and the proof is complete.

5 Simultaneous Triangulation and Diagonilization

Suppose $\mathcal{F} \subseteq L(V, V)$ is a family of linear operators on a vector space V a field $\mathbb{F}$. We say that $\mathcal{F}$ is a **commuting family** if $TU = UT$ for all $U, T \in \mathcal{F}$.

In this section we try to find a basis E of V so that, for all T in a family $\mathcal{F}$ the matrix of T is diagonal (or triangular) with respect to E. Following are the main theorems.

5.1 (Theorem) Let V be a finite dimensional vector space with $\dim V = n$ over a field $\mathbb{F}$. Let $\mathcal{F} \subseteq L(V, V)$ be a commuting and triangulable family of operators on V. Then there is a basis $E = \{e_1, \ldots, e_n\}$ such that, for every $T \in \mathcal{F}$, the matrix of T with respect to E is a triangular matrix.

Proof. The proof is some fairly similar to that of theorem 4.6. We will omit the proof. You can work it out when you need.

Following is the matrix version of the above theorem.

5.2 (Theorem) Let $\mathcal{F} \subseteq M_{nn}(\mathbb{F})$ be a commuting and triangulable family of $n \times n$ matrices. Then there is an invertible matrix P such that, for every $A \in \mathcal{F}$, we have PAP^{-1} is an upper triangular matrix.

5.3 (Theorem) Let V be a finite dimensional vector space with $\dim V = n$ over a field $\mathbb{F}$. Let $\mathcal{F} \subseteq L(V, V)$ be a commuting and diagonalizable family of operators on V. Then there is a basis $E = \{e_1, \ldots, e_n\}$ such that, for every $T \in \mathcal{F}$, the matrix of T with respect to E is a diagonal matrix.

Proof. We will omit the proof. You can work it out when you need.

6 Direct Sum

Part of this section we already touched. We gave the definition 2.13 of **direct sum** of subspaces. Following is an exercise. Note that we can make the same definition for any subspace W.

6.1 (Exercise) Let V be a finite dimensional vector space over a field $\mathbb{F}$. Let $W_1, \ldots, W_k$ be subspaces of V. Then $V = W_1 \oplus W_2 \oplus \cdots \oplus W_k$ if and only if $V = W_1 + W_2 + \cdots + W_k$ and for each $j = 2, \ldots, k$, we have

$$(W_1 + \cdots + W_{j-1}) \cap W_j = \{0\}.$$

6.2 (Examples) (1) $\mathbb{R}^2 = \mathbb{R}e_1 \oplus \mathbb{R}e_2$ where $e_1 = (1, 0), e_2 = (0, 1)$.
(2) Let $V = \mathbb{M}_{nn}(\mathbb{F})$. Let U be the subspace of all upper triangular matrices. Let L be subspace of all strictly lower triangular matrices (that means diagonal entries are zero). Then $V = U \oplus L$.
(3) Recall theorem 2.15 that V is direct sum of eigen spaces of diagonizable operators T.

We used the word 'projection' before in the context of direct sum. Here we define projections.

6.3 (Definition) Let V be a finite dimensional vector space over a field $\mathbb{F}$. An linear operator $E : V \to V$ is said to be a **projection** if $E^2 = E$.

6.4 (observations) Let V be a finite dimensional vector space over a field $\mathbb{F}$. Let $E : V \to V$ be a projection. Let $R = range(E)$ and $N = N_E$ be the null space of E. Then

1. For $v \in V$, we have $x \in R \Leftrightarrow E(x) = x$.
2. $V = N \oplus R$.
3. For $v \in V$, we have $v = (v - E(v)) + E(v) \in N \oplus R$.
4. Let $V = W_1 \oplus W_2 \cdots \oplus W_k$ be direct sum of subspaces W_i. Define operators $E_i : V \to V$ by

$$E_i(v) = v_i \quad where \quad v = v_1 + \cdots + v_k, \quad v_i \in W_i.$$

Note E_i are well defined projections with

$$range(E_i) = W_i \quad and \quad N_{E_i} = \overline{W_i}$$

where $\overline{W_i} = W_1 \oplus \cdots \oplus W_{i-1} \oplus W_{i+1} \oplus \cdots \oplus W_k$.

Following is a theorem on projections.

6.5 (Theorem) Let V be a finite dimensional vector space over a field $\mathbb{F}$. Suppose $V = W_1 \oplus W_2 \oplus \cdots \oplus W_k$ be direct sum of subspaces W_i. Then there are k linear operators $E_1, \ldots, E_k$ on V such that

1. each E_i is a projection (i. e. $E_i^2 = E_i$).
2. $E_i E_j = 0 \quad \forall \quad i \neq j$.
3. $E_1 + E_2 + \cdots + E_k = I$.
4. $range(E_i) = W_i$.

Conversely, if $E_1, \ldots, E_k$ are k linear operators on V satisfying all the conditions (2)-(3) above then E_i is a projection (i.e. (1) holds) and with $W_i = E_i(V)$ we have $V = W_1 \oplus W_2 \oplus \cdots \oplus W_k$.

Proof. The proof is easy and left as an exercise. First, try with $k = 2$ operators, if you like.

Homework: page 213, Exercise 1, 3, 4-7, 9.

7 Invariant Direct Sums

This section deals with some of the very natural concepts. Suppose V is a vector space over a filed $\mathbb{F}$ and $V = W_1 \oplus W_2 \oplus \cdots \oplus W_k$. where W_i are subspaces. Suppose for each $i = 1, \ldots, k$ we are given linear operators $T_i \in L(W_i, W_i)$ on W_i, Then we can define a linear operator $T : V \to$ such that

$$T(\sum_{i=1}^{k} v_i) = \sum_{i=1}^{k} T_i(v_i) \quad for \quad v_i \in W_i.$$

So the restriction $T_{|W_i} = T_i$. This means that the diagram

$$\begin{array}{ccc} W_i & \xrightarrow{T_i} & W_i \\ \downarrow & & \downarrow \\ V & \xrightarrow{T} & V \end{array}$$

commute.

Conversely, Suppose V is a vector space over a filed $\mathbb{F}$ and $V = W_1 \oplus W_2 \oplus \cdots \oplus W_k$. where W_i are subspaces. Let $T \in L(V, V)$ be a linear operator. Assume that W_i are invariant under T. Then we can define linear operatop $T_i : W_i \to W_i$ by $T_i(v) = T(v)$ for $v \in W_i$. Therefore, the above diagram commutes and T can be reconstructed from $T_1, \ldots, T_k$, in the same way as above.

8 Primary Decomposition

We studied linear operators T on V under the assumption that the characteristic polynomial q or the MMP p splits completely in to linear factors. In this section we will not have this assumtion. Here we will exploit the fact q, p have unique factorization.

8.1 (Primary Decomposition Theorem) Let V be a vector space over $\mathbb{F}$ with finite dimension $\dim V = n$ and $T : V \to V$ be a linear operator on V. Let p be the minimal monic polynomial (MMP) of T and

$$p = p_1^{r_1} p_2^{r_2} \cdots p_k^{r_k}$$

where $r_i > 0$ and p_i are distinct irreducible monic polynomials in $\mathbb{F}[X]$. Let

$$W_i = \{v \in V : p_i(T)^{r_i}(v) = 0\}$$

be the null space of $p_i(T)^{r_i}$. Then

1. $V = W_1 \oplus \cdots \oplus W_k$;
2. each W_i is invariant under T;
3. Let $T_i = T_{|W_i} : W_i \to W_i$ be the operator on W_i induced by T. Then the minimal monic polynomial of T_i is $p_i^{r_i}$.

Proof. Write

$$f_i = \frac{p}{p_i^{r_i}} = \prod_{j \neq i} p_j^{r_j}.$$

Note that $f_1, f_2, \ldots, f_k$ have no common factor. So

$$CGD(f_1, f_2, \ldots, f_k) = 1.$$

Therefore

$$f_1 g_1 + f_2 g_2 + \cdots + f_k g_k = 1$$

for some $g_i \in \mathbb{F}[X]$.

For $i = 1, \ldots, k$, let $h_i = f_i g_i$ and $E_i = h_i(T)) \in L(V, V)$. Then

$$E_1 + E_2 + \cdots + E_k = \sum h_i(T) = Id. \qquad (I)$$

Also, for $i \neq j$ note that $p \mid h_i h_j$. Since $p(T) = 0$ we have

$$E_i E_j = h_i(T) h_j(T) = 0. \qquad (II)$$

Write $W_i' = E_i(V)$ the range of E_i. By converse part of theorem 6.5, it follows that $V = W_1' \oplus \cdots \oplus W_k'$.

By (I), we have $T = TE_1 + TE_2 + \cdots + TE_k$. So

$$T(W_i') = T(E_i(V)) = \sum_{j=1}^{k} TE_j E_i(V) = TE_i^2(V) = TE_i(V) =$$

$$E_i T(V) \subseteq E_i(V) = W_i'.$$

Therefore, W_i' is invariant under T. We will show that $W_i' = W_i$ is the null space of $p_i(T)^{r_i}$.

We have

$$p_i(T)^{r_i}(W_i') = p_i(T)^{r_i} f_i(T) g_i(T)(V) = p(T) g_i(T)(V) = 0.$$

So, $(W_i') \subseteq W_i$ the null space of $p_i(T)^{r_i}$.

Now suppose $w \in W_i$. So, $p_i(T)^{r_i}(w) = 0$. For $j \neq i$, we have $p_i^{r_i} \mid f_j g_j = h_j$ and hence, $E_j(v) = h_j(T)(w) = 0$. Therefore $w = \sum_{j=1}^{k} E_j(w) = E_i(w)$ is in W_i'. So, $W_i \subseteq W_i'$. Therfore $W_i = W_i'$ and (1) and (2) are established.

Now $T_i : W_i \to W_i$ is the restriction of T to W_i. It remains to show that MMP of T_i is $p_i^{r_i}$. It is enough to show this for $i = 1$ or that is MMP of T_1 is $p_1^{r_1}$.

We have $p_1(T_1)^{r_1} = 0$, because W_1 is the null space $p_1(T)^{r_1}$. Therefore $p_1^{r_1} \in ann(T_1)$.

Now suppose $g \in ann(T_1)$. So, $g(T_1) = 0$. Then

$$g(T) f_1(T) = g(T) \prod_{j=2}^{k} p_j^{r_j}.$$

Since $g(T)_{|W_1} = g(T_1) = 0$, we have $g(T)$ vanishes on W_1 and also for $j = 2, \ldots, k$ we have $p_j(T)^{r_j}$ vanished on W_j. Therefore, $g(T) f_1(T) = 0$. Hence $p \mid g f_1$. Hence $p^{r_1} = \frac{p}{f_1} \mid g$. Therefore p^{r_1} is the MMP of T_1 and the proof is complete.

Remarks. (1) Note that the projections $E_i = h_i(T)$ in the above theorem are polynomials in T.

(2) Also think what it means if some (or all) of the irreducible factors $p_i = (X - \lambda_i)$ are linear.