M304T: Linear Algebra

Canonical Form

Triangular form

Dr. R. Bhuvaneswari

BMS College for Women

October 2020

Lemma

If $W \subset V$ is invariant under T, then T induces a linear transformation $\overline{T}$ on V/W defined by

$$(v + W)\overline{T} = vT + W \quad \text{for } v \in V.$$

If T satisfies a polynomial $q(x)$ in $F[x]$ then so does $\overline{T}$.
((i.e) $\overline{T}$ also satisfies $q(x)$).
If $p_1(x)$ is a minimal polynomial for $\overline{T}$ over F and $p(x)$ is a minimal polynomial for T over F, then

$$p_1(x) | p(x).$$

Proof:

Let $\overline{V} = V/W$.

The elements of $\overline{V}$ are the cosets $v + W$ of W in V.

Given $\overline{v} \in \overline{V}$, $\overline{v} = v + W$ for some $v \in V$.

Define $\overline{T} : \overline{V} \to \overline{V}$ by

$$\overline{v}\,\overline{T} = vT + W$$

for $\overline{v} = v + W \in \overline{V}$.

Now we prove that $\overline{T}$ is a linear transformation on $\overline{V}$.

$\overline{T}$ is well-defined:

Let $\overline{v}_1 = v_1 + W$ and $\overline{v}_2 = v_2 + W$ be two elements in $\overline{V}$ such that $\overline{v}_1 = \overline{v}_2$.

$$\implies v_1 + W = v_2 + W$$

$$\implies v_1 - v_2 \in W$$

$$\implies (v_1 - v_2)T \in WT \subset W \quad (\because W \text{ is invariant under } T)$$

$$\implies v_1T - v_2T \in W$$

$$\implies v_1T + W = v_2T + W$$

$$\implies \overline{v}_1\overline{T} = \overline{v}_2\overline{T}.$$

Thus $\overline{T}$ is well-defined.

$\overline{T}$ is a linear transformation:

Let $\overline{v}_1 = v_1 + W$ and $\overline{v}_2 = v_2 + W$ be two elements in $\overline{V}$. Then

$$\begin{aligned}
(\overline{v}_1 + \overline{v}_2)\overline{T} &= (v_1 + W + v_2 + W)\overline{T} \\
&= (v_1 + v_2 + W)\overline{T} \\
&= (v_1 + v_2)T + W \\
&= v_1T + v_2T + W (\because T \text{ is linear}) \\
&= v_1T + W + v_2T + W \\
&= \overline{v}_1\overline{T} + \overline{v}_2\overline{T}.
\end{aligned}$$

Thus $\overline{T}$ preserves addition.

For $\alpha \in F$ and Let $\overline{v} = v + W \in \overline{V}$,

$$\begin{aligned}
(\alpha\overline{v})\overline{T} &= (\alpha v + W)\overline{T} \\
&= (\alpha v)T + W \\
&= vT\alpha + W (\because T \text{ is linear}) \\
&= (vT + W)\alpha \\
&= \overline{v}\,\overline{T}\alpha.
\end{aligned}$$

Thus $\overline{T}$ preserves scalar multiplication.

Hence $\overline{T}$ is a linear transformation on $\overline{V}$.

Now we prove that for any polynomial $q(x) \in F[x]$, $\overline{q(T)} = q(\bar{T})$. Let $\overline{v} = v + W \in \overline{V}$. Then

$$\begin{aligned}
\bar{v}\,\overline{T^2} &= vT^2 + W \\
&= (vT)T + W \\
&= (vT + W)\overline{T} \\
&= (\overline{v}\,\overline{T})\overline{T} \\
&= \overline{v}\,\overline{T}^2.
\end{aligned}$$

Therefore,

$$\overline{T^2} = (\overline{T})^2.$$

Similarly, we can get

$$\overline{T^k} = (\overline{T})^k$$

for any integer $k \geqslant 0$.

Let $q(x) = a_0 + a_1 x + a_1 x^2 + \ldots + a_k x^k \in F[x]$. Then

$$q(T) = a_0 + a_1 T + a_1 x T^2 + \ldots + a_k T^k$$

$$\implies q(\bar{T}) = a_0 + a_1 \bar{T} + a_1 (\bar{T})^2 + \ldots + a_k (\bar{T})^k$$

$$\implies q(\bar{T}) = a_0 + a_1 \bar{T} + a_1 \overline{T^2} + \ldots + a_k \overline{T^k}$$

$$\implies q(\bar{T}) = \overline{q(T)}.$$

If T satisfies a polynomial $q(x) \in F[x]$ then $q(T) = 0$ implies that

$$q(\bar{T}) = \overline{q(T)} = \overline{0} = 0.$$

Thus $\bar{T}$ also satisfies $q(x)$.

Given : $p_1(x)$ is the minimal polynomial for $\overline{T}$ over F and $p(x)$ is the minimal polynomial for T over F.

Then $p_1(\overline{T}) = 0$ and $p(T) = 0$.

Since $p(T) = 0$, we have $p(\bar{T}) = 0$.

Note that $p_1(x)$ is the minimal polynomial for $\bar{T}$ and $p(\bar{T}) = 0$.

Thus $p_1(x) | p(x)$.

Note

All the characteristic roots of T which lie in F are roots of the minimal polynomial of T over F.

Therefore, We say that all the characteristic roots of T are in F if all the roots of the minimal polynomial of T over F lie in F.

Theorem

If $T \in A(V)$ has all its characteristic roots in F, then there is a basis of V in which the matrix of T is triangular.

Proof: We prove this theorem by induction on the dimension of V over F.

If $dim(V) = 1$, then every element in $A(V)$ is a scalar, and so the theorem is true.

Assume that the theorem is true for all vector spaces over F of dimension $n - 1$.

Let V be a vector space of dimension n over F.

Let the linear transformation T on V has all its characteristic roots in F.

We will prove that there is a basis of V in which matrix of T is triangluar.

Let $\lambda_1 \in F$ be a characteristic root of T.

Then there exists a nonzero vector v_1 in V such that

$$v_1 T = \lambda_1 v_1.$$

Let $W = \{\alpha v_1 : \alpha \in F\}$.

Then W is a one-dimensional subspace of V generated by v_1.

Note that $WT \subset W$.

Therefore, W is invariant under T.

Let $\overline{V} = V/W$. Then

$$dim(\overline{V}) = dim(V) - dim(W) = n - 1.$$

By the above lemma, the linear transformation T on V induces a linear transformation $\overline{T}$ on $\bar{V}$ defined by

$$(v + W)\bar{T} = vT + W$$

such that the minimal polynomial of $\bar{T}$ divides the minimal polynomial of T.

Given all the characteristic roots of T are in F.

Then all the roots of minimal polynomial for T lie in F.

Thus all the roots of minimal polynomial for $\bar{T}$ (which are roots of minimal polynomial of T) lie in F and hence all the characteristic roots of $\bar{T}$ also lie in F.

Since $dim(\bar{V}) = n - 1$ and the linear transformation $\bar{T}$ on $\bar{V}$ has all its roots in F, by induction hypothesis, there exists a basis $\overline{v}_2, \overline{v}_3, \cdots, \overline{v}_n$ of $\bar{V}$ over F such that the matrix of $\bar{T}$ in the basis is triangular.

Further

$$
\begin{aligned}
\overline{v}_2\overline{T} &= \alpha_{22}\overline{v}_2 \\
\overline{v}_2\overline{T} &= \alpha_{32}\overline{v}_2 + \alpha_{33}\overline{v}_3 \\
&\vdots \\
\overline{v}_i\overline{T} &= \alpha_{i2}\overline{v}_2 + \alpha_{i3}\overline{v}_3 + \cdots + \alpha_{ii}\overline{v}_i \\
&\vdots \\
\overline{v}_n\overline{T} &= \alpha_{n2}\overline{v}_2 + \alpha_{n3}\overline{v}_3 + \cdots + \alpha_{nn}\overline{v}_n
\end{aligned}
$$

Since $\overline{v}_i \in \bar{V} = V/W$, $\exists \ v_i \in V$ such that $\overline{v}_i = v_i + W$.

Let $v_2, v_3, \ldots, v_n$ be the elements of V mapping into $\overline{v}_2, \overline{v}_3, \cdots, \overline{v}_n$ respectively.

claim: $v_1, v_2, \ldots, v_n$ is a basis for V.

Let $\alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n = 0$ where $\alpha_i \in F$.

Then

$$W = 0 + W = (\alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n) + W$$

$$\implies \alpha_1(v_1 + W) + \alpha_2(v_2 + W) + \ldots + \alpha_n(v_n + W) = W$$

(ie) $\alpha_2\overline{v}_2 + \ldots + \alpha_n\overline{v}_n = \overline{0} = W$, the zero element of V/W.

(Note that $\alpha_1 v_1 \in W$.)

Since $\overline{v}_2, \overline{v}_3, \cdots, \overline{v}_n$ form a basis for $\overline{V}$ over F, we have

$$\alpha_2 = 0, \alpha_3 = 0, \ldots, \alpha_n = 0.$$

Now $\alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n = 0$ implies that

$$\alpha_1 v_1 = 0$$

Since $v_1 \neq 0$, we get $\alpha_1 = 0$.

Thus $\alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n = 0$ implies

$$\alpha_1 = 0, \alpha_2 = 0, \ldots, \alpha_n = 0.$$

Hence $v_1, v_2, \ldots, v_n$ is linearly independent vectors in V.

Since $dim(V) = n$, $v_1, v_2, \ldots, v_n$ forms a basis for V over F.

Now we prove that in this basis the matrix of T is triangular.

Since $\overline{v}_2 \overline{T} = \alpha_{22} \overline{v}_2$, we have

$$v_2 T + W - \alpha_{22}(v_2 + W) = \overline{0}$$

$$\implies v_2 T - \alpha_{22} v_2 \in W$$

Since the elements of W are scalar multiples of v_1, we have

$$v_2 T - \alpha_{22} v_2 = \alpha_{21} v_1$$

for some $\alpha_{21} \in F$.

$\therefore$, $v_2 T = \alpha_{21} v_1 + \alpha_{22} v_2$

Similarly,in general, for any $i \geqslant 2$,

$$\overline{v}_i \overline{T} = \alpha_{i2}\overline{v}_2 + \alpha_{i3}\overline{v}_3 + \cdots + \alpha_{ii}\overline{v}_i$$

implies that

$$v_i T - \alpha_{i2} v_2 - \alpha_{i3} v_3 - \ldots - \alpha_{ii} v_i \in W$$

$$\implies v_i T - \alpha_{i2} v_2 - \alpha_{i3} v_3 - \ldots - \alpha_{ii} v_i = \alpha_{i1} v_1$$

for $\alpha_{i1} \in F$.

$$\implies v_i T = \alpha_{i1} v_1 + \alpha_{i2} v_2 + \alpha_{i3} v_3 + \ldots + \alpha_{ii} v_i$$

Since W is invariant under T, $v_1 \in W \implies v_1 T \in W$

$$\implies v_1 T = \alpha_{11} v_1$$

Thus $v_1, v_2, \ldots, v_n$ is a basis of V such that

$$\begin{aligned} v_1 T &= \alpha_{11} v_1 \\ v_2 T &= \alpha_{21} v_1 + \alpha_{22} v_2 \\ &\vdots \\ v_i T &= \alpha_{i1} v_1 + \alpha_{i2} v_2 + \cdots + \alpha_{ii} v_i \\ &\vdots \\ v_n T &= \alpha_{n1} v_1 + \alpha_{n2} v_2 + \cdots + \alpha_{nn} v_n \end{aligned}$$

That is, $v_i T$ is a linear combination only of v_i and its predecessors in the basis.

Thus the matrix of T in this basis $v_1, v_2, \ldots, v_n$ is triangular.

Therefore, the theorem is true for any vector space V of dimension n.

Hence the theorem is true for any finite dimensional vector space.

Above theorem in terms of matrices.

Theorem

Suppose the matrix $A \in F_n$ has all its characteristic roots in F. Then there is a matrix $C \in F_n$ such that CAC^{-1} is a triangular matrix.

Proof: Assume the matrix $A \in F_n$ has all its characteristic roots in F. Then A defines a linear transformation T in $F^{(n)}$ whose matrix in the basis

$$v_1 = (1, 0, 0, \cdots, 0), v_2 = (0, 1, 0, \cdots, 0), \cdots, v_n = (0, 0, 0, \cdots, 1)$$

is precisely A.

Note that the characteristic roots of T are the characteristic root of A.

Therefore, all the characteristic roots of T are in F.

By the above theorem, there is a basis of $F^{(n)}$ in which the matrix of T, $m(T)$ is triangular.

Thus A is a matrix of T in one basis and $m(T)$ is a matrix of T in another basis.

By change of basis theorem, there exists an invertible matrix $C \in F_n$ such that

$$m(T) = CAC^{-1}.$$

Thus there exists $C \in F_n$ such that CAC^{-1} is triangular.

Note

If a linear transformation T (or a matrix A) is brought to triangular form over F, then the characteristic roots of T (or A) are the main diagonal entries on the triangular form.

Caylety Hamilton Theorem

Theorem

If V is an $n-$ dimensional vector space over F and if $T \in A(V)$ has all its characteristic roots in F, then T satisfies a polynomial of degree n over F.

Proof: Since T has all its characteristic roots in F, we can find a basis $v_1, v_2, \ldots, v_n$ is a basis of V over F in which matrix of T is triangular.

That is, we can find a basis $v_1, v_2, \ldots, v_n$ of V over F such that

$$\begin{aligned} v_1 T &= \alpha_{11} v_1 \\ v_2 T &= \alpha_{21} v_1 + \alpha_{22} v_2 \\ &\vdots \\ v_i T &= \alpha_{i1} v_1 + \alpha_{i2} v_2 + \cdots + \alpha_{ii} v_i \\ &\vdots \\ v_n T &= \alpha_{n1} v_1 + \alpha_{n2} v_2 + \cdots + \alpha_{nn} v_n \end{aligned}$$

We know that in triangular form, the characteristic roots are the entries in the main diagonal.

Let $\lambda_1, \lambda_2, \cdots, \lambda_n$ be the characteristic roots of T. Then $\lambda_i = a_{ii}$ for each $i = 1$ to n.

Therefore, we get

$$
\begin{aligned}
v_1 T &= \lambda_1 v_1 \\
v_2 T &= \alpha_{21} v_1 + \lambda_2 v_2 \\
&\vdots \\
v_i T &= \alpha_{i1} v_1 + \alpha_{i2} v_2 + \cdots + \alpha_{i(i-1)} v_{i-1} + \lambda_i v_i \\
&\vdots \\
v_n T &= \alpha_{n1} v_1 + \alpha_{n2} v_2 + \cdots + \alpha_{n1}(n-1) v_{n-1} + \lambda_n v_n
\end{aligned}
$$

Which implies that

$$
\begin{aligned}
v_1 T - \lambda_1 v_1 &= 0 \\
(i.e) v_1 (T - \lambda_1) &= 0 \\
v_2 (T - \lambda_2) &= \alpha_{21} v_1 \\
&\vdots \\
v_i (T - \lambda_i) &= \alpha_{i1} v_1 + \alpha_{i2} v_2 + \cdots + \alpha_{i(i-1)} v_{i-1} \\
&\vdots \\
v_n (T - \lambda_n) &= \alpha_{n1} v_1 + \alpha_{n2} v_2 + \cdots + \alpha_{n1} (n-1) v_{n-1}
\end{aligned}
$$

Since $v_1(T - \lambda_1) = 0$, we have

$$v_2(T - \lambda_2)(T - \lambda_1) = \alpha_{21} v_1(T - \lambda_1) = 0.$$

We know that

$$(T - \lambda_1)(T - \lambda_2) = (T - \lambda_2)(T - \lambda_1)$$

Therefore,

$$v_1(T - \lambda_2)(T - \lambda_1) = v_1(T - \lambda_1)(T - \lambda_2) = 0$$

Continuing this procedure, we get for $i = 1$ to n,

$$
\begin{aligned}
v_1(T - \lambda_i)(T - \lambda_{i-1}) \cdots (T - \lambda_1) &= 0 \\
v_2(T - \lambda_i)(T - \lambda_{i-1}) \cdots (T - \lambda_1) &= 0 \\
&\vdots \\
v_i(T - \lambda_i)(T - \lambda_{i-1}) \cdots (T - \lambda_1) &= 0
\end{aligned}
$$

For $i = n$, the matrix $S = (T - \lambda_n)(T - \lambda_{n-1}) \cdots (T - \lambda_1)$ satisfies $v_1 S = v_2 S = \ldots = v_n S = 0$.

Therefore, S annihilates all the vectors in the basis of V and so S must annihilate all elements of V.

That is $vS = 0$ for all $v \in V$.

$\implies S = 0$.

Let $p(x) = (x - \lambda_1)(x - \lambda_2) \cdots (x - \lambda_n) \in F[x]$ be a polynomial of degree n.

Then $p(T) = S = 0$.

Thus T satisfies a polynomial of degree n over F.

Remark

Not every linear transformation on V over F has all its characteristic roots in F.

It depends on the field F.

For example, if $F = \mathbb{R}$, then then minimal polynomial of

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

over F is $x^2 + 1$ has no roots in F.

So in general we can not assume all characteristic roots of $T \in A(V)$ lie in F.