Theorem 1 (Primary Decomposition Thm.)**.** *Let $T : V \to V$ be a linear operator on a finite-dimensional vector space over $\mathbb{F}$. Let*

$$m_T(z) = p_1(z)^{e_1} p_2(z)^{e_2} \cdots p_m(z)^{e_m}$$

be the unique factorization of the minimal polynomial of T into a product of distinct monic prime powers. Then

1. *$V = U_1 \oplus U_2 \oplus \cdots \oplus U_m$, where $U_j = \operatorname{Ker} p_j(T)^{e_j}$ for $1 \le j \le m$.*
2. *The projection of E_j of V on U_j along the sum of the other U_i's is of the form $q_j(T)$ for some polynomial $q_j(z)$.*
3. *Each of the U_j are T-invariant.*
4. *Any linear operator $V \to V$ that commutes with T carries each U_j to itself.*
5. *Any T-invariant subspace W of V can be written as the direct sum*

$$W = (W \cap U_1) \oplus (W \cap U_2) \oplus \cdots \oplus (W \cap U_m).$$

6. *The minimal polynomial of T restricted to U_j is $p_j(z)^{e_j}$.*

Proof.

(1) & (2) Define $s_j(z) = \frac{m_T(z)}{p_j(z)^{e_j}}$, for $1 \le j \le m$. The ideal in $\mathbb{F}[z]$ generated by $s_1(z), s_2(z), \ldots s_m(z)$ has a single generator $d(z)$ because $\mathbb{F}[z]$ is a PID. By construction of the s_j, no p_j divides $d(z)$. In fact, if $d(z)$ is irreducible then it divides some p_j, but the p_j are prime. We conclude that $d(z)$ is a unit. Hence, there exist polynomials $t_1(z), t_2(z), \ldots, t_m(z)$ such that

$$1 = s_1(z)t_1(z) + s_2(z)t_2(z) + \cdots + s_m(z)t_m(z).$$

Define $E_j := s_j(T)t_j(T)$. Then

$$I = E_1 + E_2 + \cdots + E_m$$

and

$$\begin{aligned} E_i E_j &= s_i(z)t_i(z)s_j(z)t_j(z) \\ &= m_T(z)g(z) \end{aligned}$$

for some polynomial $g(z)$. We now have that $E_iE_j(T) = m_T(T)g(T) = 0$. This is enough to insure that the E_i are all projections. If $\operatorname{Im} E_j = U_j$, then $V = U_1 \oplus U_2 \oplus \cdots \oplus U_m$.

We need to show that the U_j are as defined in the theorem. That is, we need to show that

$$\operatorname{Im} E_j = \operatorname{Ker} p_j(T)^{e_j}.$$

By construction of the $s_j(z)$, we can write the minimal polynomial

$$m_T(z) = s_j(z) p_j(z)^{e_j}.$$

Hence,

$$\begin{aligned} p_j(T)^{e_j} E_j &= t_j(T) s_j(T) p_j(T)^{e_j} \\ &= t_j(T) m_T(T) \\ &= 0. \end{aligned}$$

Suppose that $w \in \operatorname{Im} E_j$. Then

$$p_j(T)^{e_j} w = p_j(T)^{e_j} E_j v = 0$$

for some $v \in V$. Hence, $\operatorname{Im} E_j \subseteq \operatorname{Ker} p_j(T)^{e_j}$.

For the reverse inclusion, suppose that $v \in \operatorname{Ker} p_j(T)^{e_j}$. For $i \neq j$,

$$S(z) = s_i(z) t_i(z) = \left(\prod_{r \neq i,j} p_r(z)^{e_r} \right) t_i(z) p_j(z)^{e_j}.$$

The operator $S(T)$ is the projection E_i. We observe that $E_i v = 0$ for all $i \neq j$. We now have that $v = E_j v$ and conclude that $\operatorname{Ker} p_j(T) \subseteq \operatorname{Im} E_j$.

We've proven statements (1) and (2). We now show that these two statements imply the others.

(3) The Ej are all polynomial in T, as we saw in the discussion above. Hence, the E_j commute with T. Then $E_j T(U_j) = T E_j(U_j) = T(U_j)$, but $E_j T(U_j) \subseteq U_j$ so $T(U_j) \subseteq U_j$, and therefore the U_j are T-invariant.

(4) Suppose that $S : V \to V$ is a linear operator which commutes with T. Since each E_j is polynomial in T, each E_j commutes with S. By the same argument given in (3), the U_j are S-invariant.

(5) We first observe that

$$(W \cap U_1) \oplus (W \cap U_2) \oplus \cdots \oplus (W \cap U_m)$$

is certainly contained in W. Suppose $w \in W$. By assuption, W is T-invariant so $Tw \in W$ and $E_j w \in W$ because E_j is polynomial in T. We then write

$$w = E_1 w + E_2 w + \cdots + E_m w$$

and because each $E_j w \in U_j$ we conclude that $E_j w \in U_j \cap W$, for $1 \leq j \leq m$.

(6) Let $m_j(z)$ be the minimal polynomial of T restricted to U_j. Define T_j to be T restricted to U_j. Since $p_j(T)^{e_j} E_j = 0$ on U_j, we have that $p_j(T)^{e_j}$ is the zero operator on U_j. This implies that $m_j(z)$ is some power of $p_j(z)$. Consider the fact that,

$$0 = m_j(T_j)E_j = m_j(T_j)s_j(T_j)r_j(T_j)$$

on U_j. The operator $m_j(T)E_j$ on $U_i = \operatorname{Im} E_i$ when $i \neq j$ is zero on U_i for all $i \neq j$ because $E_j E_i = 0$. We conclude that $m_j(T)E_j$ is identically zero on V. This gives us that $m_T(z)$ divides

$$m_j(z)s_j(z)t_j(z) = m_j(z)\left(1 - \sum_{i \neq j} s_i(z)t_i(z)\right)$$

and therefore $p_j(z)^{e_j}$ divides the right-hand-side. The irreducibility of p_j implies that $p_j^{e_j}$ divides m_j. We conclude that $m_j(z) = p_j(z)^{e_j}$. □

1 Cyclic Operators

Consider the operator on $\mathbb{C}^3$ given by the matrix

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

One can check that the eigenvalues of this matrix are $\lambda = 1$ and $\lambda = 2$. By direct computation, it can be seen that $(A - I)(A - 2I)$ is not the zero matrix. Hence, the characteristic polynomial and the minimal polynomial correspond. We then get the decomposition

$$V_A \cong \frac{\mathbb{F}[z]}{\langle (x-1)^2 \rangle} \oplus \frac{\mathbb{F}[z]}{\langle x - 2 \rangle}$$

which is a cyclic $\mathbb{F}[z]$-module.

2 Examples

Example 1. Consider the matrix

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

We need to find the eigenvalues of M. That is, we need to find λ such that $(M - \lambda I)$ is not injective.

$$\begin{pmatrix} 1-\lambda & 1 & 1 & 1 \\ 0 & 1-\lambda & 0 & -1 \\ 0 & 0 & 1-\lambda & 1 \\ 0 & 0 & 0 & 1-\lambda \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{pmatrix} = \begin{pmatrix} (1-\lambda)v_1 + v_2 + v_3 + v_4 \\ (1-\lambda)v_2 - v_4 \\ (1-\lambda)v_3 + v_4 \\ (1-\lambda)v_4 \end{pmatrix} = 0.$$

The last is satisfied when $\lambda = 1$ or $v_4 = 0$.

$\lambda = \mathbf{1}$:

$$\begin{pmatrix} 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{pmatrix} = \begin{pmatrix} v_2 + v_3 + v_4 \\ -v_4 \\ v_4 \\ 0 \end{pmatrix}$$

which readily gives us that $v_4 = 0$, $v_2 = -v_3$, and v_1 is free. The eigenspace $E_{\lambda=1}$ is spanned by $(1, 1, -1, 0)^T$ and $(0, 1, -1, 0)^T$. We can directly compute that $(M - I)^2$ is the zero operator. We ahve two possibilities for the Jordan form of M:

$$\left(\begin{array}{cc|cc} 1 & 1 & & \\ & 1 & & \\ \hline & & 1 & 1 \\ & & & 1 \end{array}\right) \text{ or } \left(\begin{array}{cc|c|c} 1 & 1 & & \\ \hline & & 1 & \\ \hline & & & 1 \end{array}\right)$$

Computing the kernels of $M - I$ and $(M - I)^2$:

$$\dim \operatorname{Ker}(M - I) = 2$$
$$\dim \operatorname{Ker}(M - I)^2 = 4$$

which gives us that

# of Jordan Blocks that are 1×1 or larger

$$\dim \operatorname{Ker}(M - I) = 2$$

# of Jordan Blocs that are 2×2 or larger

$$\dim \operatorname{Ker}(M - I)^2 - \dim \operatorname{Ker}(M - I) = 4 - 2 = 2.$$

Hence, there are two 2×2 Jordan blocks, and we want the first of our possibilities above.

Example 2. Consider the matrix

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

We need to find the eigenvalues of C

$$0 = (C - \lambda I)v = \begin{pmatrix} 1-\lambda & 0 & 0 \\ -1 & 1-\lambda & 0 \\ 0 & 0 & 1-\lambda \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} (1-\lambda)v_1 \\ -v_1 + (1-\lambda)v_2 \\ (1-\lambda)v_3 \end{pmatrix}$$

which is satisfied when $\lambda = 1$ and $v_1 = 0$.

$\lambda = \mathbf{1}$:

$$\begin{pmatrix} 0 & 0 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} 0 \\ -v_1 \\ 0 \end{pmatrix}$$

which implies that $v_1 = 0$ and the eigenspace $E_{\lambda=1}$ is spanned by $(0,1,0)^T$ and $(0,0,1)^T$.

Note that $(C - I)^2$ is the zero operator. We have the following for the size of the Jordan blocks.

# of Jordan blocks that are 1×1 or larger:

$$\dim \operatorname{Ker}(C - I) = 2$$

# of Jordan blocks which are 2×2 or larger:

$$\dim \operatorname{Ker}(C - I)^2 - \dim \operatorname{Ker}(C - I) = 3 - 2 = 1$$

Hence, C is similar to the matrix

$$\left(\begin{array}{cc|c} 1 & 1 & \\ & 1 & \\ \hline & & 1 \end{array}\right)$$

3 V as an $\mathbb{F}[z]$-module

We've been studying linear operators of a finite-dimensional vector space V into itself. We constructed the companion matrix while viewing V as an $\mathbb{F}[z]$-module. Our setup was as follows:

- $\mathbb{F}$ is an algebraically closed field.
- $\mathbb{F}[z]$ is the ring of polynomials in a single indeterminate, with coefficients in $\mathbb{F}$.
- $T : V \to V$ is an linear operator of V that we want to understand.

Theorem 2 (Elementary Divisor Decomposition)**.** *Viewing V as an $\mathbb{F}[z]$-module with respect to the linear operator T, V_T is the direct sum of a finite number of cyclic modules*

$$V_T \cong \frac{\mathbb{F}[z]}{p_1^{e_1}} \oplus \frac{\mathbb{F}[z]}{p_2^{e_2}} \oplus \frac{\mathbb{F}[z]}{p_m^{e_m}}$$

where the $p_j^{e_j}$ are positive powers of primes in $\mathbb{F}[z]$, which are not necessarily distinct.

Since $\mathbb{F}$ is algebraically closed, we can assume that each of the p_j take the form $(z - \lambda_j)$, for some $\lambda_j \in \mathbb{F}$. The product of the p_j is the characteristic polynomial. One could then relax the condition that the field $\mathbb{F}$ is algebraically closed to $\mathbb{F}$ containing all of the eigenvalues of T.

Not every matrix is diagonalizable. The motivation for obtaining the Jordan form is to obtain a matrix, similar to the one we're studying, that is as close to diagonal as possible. The linear operator T acting on V is equivalent to z action on the $\mathbb{F}[z]$-module V_T.

From the elementary divisor form of the decomposition theorem, we need a particularly nice basis for each of the cyclic powers $\mathbb{F}[z]/p_j^{e_j}$. For our situation, we are trying to understand $\mathbb{F}[z]/(z - \lambda)^k$. We know that the standard basis for this vector space is

$$\{1, \bar{z}, \bar{z}^2, \ldots, \bar{z}^{k-1}\}.$$

Note that we can write $z = \lambda + (z - \lambda)$. The action of z gives us the following mapping:

$$\begin{aligned}
(\bar{z}-\lambda)^{k-1} &\longmapsto \lambda(\bar{z}-\lambda)^{k-1} \\
(\bar{z}-\lambda)^{k-2} &\longmapsto \lambda(\bar{z}-\lambda)^{k-2} + (\bar{z}-\lambda)^{k-1} \\
&\vdots \\
(\bar{z}-\lambda)^{2} &\longmapsto \lambda(\bar{z}-\lambda)^{2} + (\bar{z}-\lambda)^{3} \\
(\bar{z}-\lambda) &\longmapsto \lambda(\bar{z}-\lambda) + (\bar{z}-\lambda)^{2} \\
1 &\longmapsto \lambda + (\bar{z}-\lambda)
\end{aligned}$$

We now have the matrix

$$J_{\lambda,k} = \begin{pmatrix} \lambda & 1 & & & & \\ & \lambda & 1 & & & \\ & & \lambda & 1 & & \\ & & & \ddots & \ddots & \\ & & & & \lambda & 1 \\ & & & & & \lambda \end{pmatrix}$$

which is an elementary Jordan block of size k with eigenvalue λ.

Going back to an arbitrary operator T, with characteristic polynomial $c_T(z)$, we can decompose V as an $\mathbb{F}[z]$-module

$$V_T \cong$$

where $c_T(z) = (z-\lambda_1)^{e_1}(z-\lambda_2)^{e_2} \cdots (z-\lambda_m)^{e_m}$. We repeat, for emphasis, that the λ_j are not necessarily distinct. We can represent T as a block matrix of elementary Jordan blocks

$$\begin{pmatrix} J_{k_1,\lambda_1} & & & \\ & J_{k_2,\lambda_2} & & \\ & & \ddots & \\ & & & J_{k_m,\lambda_m} \end{pmatrix}$$

This matrix is uniquely determined up to permutation of the Jordan blocks.

References

[1] Sheldon Axler, *Linear Algebra Done Right*. Springer, 2nd Edition, 1997

[2] David S. Dummit, Richard M. Foote, *Abstract Algebra*. John Wiley & Sons, Inc., 3rd Edition, 2004

[3] Anthony W. Knapp, *Basic Algebra*. Birkhäuser, 1st edition, 2006