The Jordan Canonical Form: Part 1

Let V be a finite-dimensional vector space over a field K and let T be a linear operator on V which satisfies a polynomial identity of the form

$$(T-a_1)^{k_1}(T-a_2)^{k_2}\cdots(T-a_m)^{k_m}=0$$

with $a_1, a_2, \ldots, a_m$ distinct scalars. Such an identity always exists if $K=\mathbb{C}$. In this and the following lecture we shall prove the following result

Theorem 1. *There exists a basis of V such that the matrix of T is in block-diagonal form with Jordan blocks. If a is an eigenvalue of T, t the sequence $(t_0, t_1, ..., t_n, ...)$ with $t_i = \dim\ker(T-a)^i$ and*

$$(s_0, s_1, ..., s_n, ...) = -R(L-1)^2(t),$$

where R, L are the left and right-shift operators on $\mathbb{R}^\infty$, then s_i is the number of Jordan blocks of size i with eigenvalue a.

By way of an example, Let T be the linear operator on $\mathbb{F}_2^6$ whose matrix with respect the standard basis of $\mathbb{F}_2^6$ is

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

We have

$$A-1=\begin{bmatrix}0&1&1&1&0&0\\0&0&1&1&0&1\\0&0&0&0&1&1\\0&0&0&0&1&1\\0&0&0&0&0&0\\0&0&0&0&0&0\end{bmatrix},\quad (A-1)^2=\begin{bmatrix}0&0&1&1&0&1\\0&0&0&0&0&0\\0&0&0&0&0&0\\0&0&0&0&0&0\\0&0&0&0&0&0\\0&0&0&0&0&0\end{bmatrix},\quad (A-1)^3=0.$$

It follows that $(T-1)^3=0$ and 1 is the only eigenvalue of T. We also have $\mathrm{rank(T-1)}=3$, $\mathrm{rank(T-1)}^2=1$, $\mathrm{rank(T-1)}^3=0$ so that

$$t_1=\dim\ker(T-1)=3,\ t_2=\dim\ker(T-1)^2=5,\ t_3=\dim\ker(T-1)^3=6.$$

Hence t is the sequence $(0,3,5,6,6,....,6,...)$. Now

$$(L-1)(t)=(3,2,1,0,0,...,0,...),\quad (L-1)^2(t)=(-1,-1,-1,0,0,...,0,...)$$

and so $-R(L-1)^2(t)=(0,1,1,1,0,0,...,0,...)$ which, according to Theorem 1, implies that there is one Jordan block of size 1, one of size 2 and one of size 3. Hence there is a basis of $\mathbb{F}_2^6$ such that the matrix of T with respect to this basis is

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

If W is a T-invariant subspace of V and $f = (f_1, f_2, ..., f_n)$ is a basis of W, the matrix (with respect to this basis) of the restriction of T to W is the Jordan matrix $J_n(a)$ iff

$$T(f_1) = af_1,\ T(f_2) = af_2 + f_1, ..., T(f_i) = af_i + f_{i-1}, ..., T(f_n) = af_n + f_{n-1}$$

or, equivalently,

$$(T-a)(f_1) = 0,\ (T-a)(f_2) = f_1, ..., (T-a)(f_i) = f_{i-1}, ..., (T-a)(f_n) = f_{n-1}.$$

For such a basis we have $f_i = (T-a)^{n-i}(f_n)$ with $f_n \in \mathrm{Ker}((T-a)^n) - \mathrm{Ker}((T-a)^{n-1})$. Conversely, if $g \in \mathrm{Ker}((T-a)^n) - \mathrm{Ker}((T-a)^{n-1})$ the sequence

$$g,\ (T-a)(g),\ (T-a)^2(g), ..., (T-a)^{n-1}(g)$$

is a basis for a T-invariant subspace of V such that the matrix of this mapping with respect to the basis

$$f_1 = (T-a)^{n-1}(g),\ f_2 = (T-a)^{n-2}(g), ..., (T-a)(g), g$$

is the Jordan matrix $J_n(a)$. The vector g is called a **cyclic vector of cycle length n** for the eigenvalue a. Each Jordan block corresponds a cyclic vector. The subspace generated by a cyclic vector g and its images under the powers of T is called the **cyclic subspace generated by** g.

We now illustrate how to find cyclic vectors that give a decomposition into a direct sum of cyclic subspaces in the case of Example 1. We first find bases for $\ker(T-1)$, $\ker(T-1)^2$, $\ker(T-1)^3$:

$$\begin{aligned}\mathrm{Ker}(T-1)\ &= \mathrm{Span}(e_1, e_3+e_4, e_2+e_3+e_5+e_6),\\ \mathrm{Ker}((T-1)^2) &= \mathrm{Span}(e_1, e_2, e_5, e_3+e_4, e_3+e_6),\\ \mathrm{Ker}((T-1)^3) &= \mathrm{Span}(e_1, e_2, e_3, e_4, e_5, e_6).\end{aligned}$$

The next step is to complete the basis of $\mathrm{Ker}(T-1)^2$ to a basis of $(T-1)^3$. We find that $g_1 = e_6$ completes the given basis of $\mathrm{Ker}((T-1)^2)$ to a basis of $\mathrm{Ker}((T-1)^3)$. Now $(T-1)(e_6) = e_2+e_3+e_4$ is in the kernel of $(T-1)^2$ but not in the kernel of $T-1$. Thus

$$e_1,\ e_3+e_4,\ e_2+e_3+e_5+e_6,\ e_2+e_3+e_4$$

is linearly independent and we can complete this sequence to a basis of $\mathrm{Ker}((T-1)^2)$ with the vector $g_2 = e_5$. Now $(T-1)^2(g_1) = e_1$,$(T-1)(g_2) = e_3+e_4$ are in the kernel of $T-1$ and are linearly independent. We complete these two vectors to a basis of $\ker(T-1)$ by means of the vector $g_3 = e_2+e_3+e_5+e_6$. Now, the sequence of vectors

$$g_1 = e_6,\ (T-1)(g_1) = e_2+e_3+e_4,\ (T-1)^2(g_1) = e_1,$$

$$g_2 = e_5,\ (T-1)(g_2) = e_3+e_4,\ g_3 = e_2+e_3+e_5+e_6$$

is linearly independent and the basis

$$f_1 = g_3,\ f_2 = (T-1)(g_2),\ f_3 = g_2,\ f_4 = (T-1)^2(g_1),\ f_5 = (T-1)(g_1),\ f_6 = g_1$$

yields the above Jordan canonical form for T. If $v_1, v_2, ..., v_n \in V$ and W is a subspace of V, we say

that the sequence $v_1, v_2, ..., v_n$ is linearly independent mod W if

$$a_1v_1 + a_2v_2 + \cdots + a_nv_n \in W \implies a_1 = a_2 = ... = a_n = 0.$$

This is equivalent to saying the the images of the vectors v_i in the quotient space V/W form a linearly independent sequence. Similarly, we say that $v_1, v_2, ..., v_n$ generate V mod W if every $v \in V$ can be written in the form $v = a_1v_1 + a_2v_n + \cdots + a_nv_n$ with $w \in W$. This is equivalent to saying that the images of the vectors v_i in V/W span V/W.

Lemma 2. *If* $\mathrm{Ker}((\mathrm{T}-\mathrm{a})^{\mathrm{i}}) = \mathrm{Ker}((\mathrm{T}-\mathrm{a})^{\mathrm{i}+1})$ *then* $\mathrm{Ker}((\mathrm{T}-\mathrm{a})^{\mathrm{i}+1}) = \mathrm{Ker}((\mathrm{T}-\mathrm{a})^{\mathrm{i}+2})$.

Proof. Let $v \in \mathrm{Ker}((\mathrm{T}-\mathrm{a})^{\mathrm{i}+2})$. Then $(T-a)(v) \in \mathrm{Ker}((\mathrm{T}-\mathrm{a})^{\mathrm{i}+1}) = \mathrm{Ker}((\mathrm{T}-\mathrm{a})^{\mathrm{i}})$ which implies that $(T-a)^{i+1}(v) = (T-a)^i(T-a)(v) = 0$ and hence that $v \in \mathrm{Ker}((\mathrm{T}-\mathrm{a})^{\mathrm{i}+1})$. □

This lemma shows that, for an eigenvalue a of T, there is an integer $p > 0$ such that

$$0 = t_0 < t_1 < \cdots < t_p = t_{p+1} = t_{p+2} = \cdots ,$$

where $t_i = \dim(T-a)^i$.

Lemma 3. *If* $i \geq 2$ *and* $v \in \mathrm{Ker}((\mathrm{T}-\mathrm{a})^{\mathrm{i}}) - \mathrm{Ker}((\mathrm{T}-\mathrm{a})^{\mathrm{i}-1})$ *then*

$$(T-a)(v) \in \mathrm{Ker}((\mathrm{T}-\mathrm{a})^{\mathrm{i}-1}) - \mathrm{ker}((\mathrm{T}-\mathrm{a})^{\mathrm{i}-2}).$$

Proof. If $v \in \mathrm{Ker}((\mathrm{T}-\mathrm{a})^{\mathrm{i}})$ and $(T-a)(v) \in \mathrm{Ker}((\mathrm{T}-\mathrm{a})^{\mathrm{i}-2})$ then

$$(T-a)^{i-1}(v) = (T-a)^{i-2}(T-a)(v) = 0$$

which implies that $v \in \mathrm{Ker}((\mathrm{T}-\mathrm{a})^{\mathrm{i}-1})$. □

This Lemma is simply the assertion that the linear mapping

$$S_{i-1} : \mathrm{Ker}((\mathrm{T}-\mathrm{a})^{\mathrm{i}})/\mathrm{Ker}((\mathrm{T}-\mathrm{a})^{\mathrm{i}-1}) \to \mathrm{Ker}((\mathrm{T}-\mathrm{a})^{\mathrm{i}-1})/\mathrm{Ker}((\mathrm{T}-\mathrm{a})^{\mathrm{i}-2})$$

defined by $S_{i-1}(v + \mathrm{Ker}((\mathrm{T}-\mathrm{a})^{\mathrm{i}-1}) = (\mathrm{T}-\mathrm{a})(\mathrm{v}) + \mathrm{Ker}((\mathrm{T}-\mathrm{a})^{\mathrm{i}-1})$ is injective. This yields the following result:

Lemma 4. *If* $i \geq 2$ *and* $v_1, v_2, ..., v_n \in \mathrm{Ker}(\mathrm{T}-\mathrm{a})^{\mathrm{i}}$ *is linearly independent mod* $\mathrm{Ker}((\mathrm{T}-\mathrm{a})^{\mathrm{i}-1})$ *then*

$$(T-a)(v_1),\ (T-a)(v_2), ..., (T-a)(v_n) \in \mathrm{Ker}((\mathrm{T}-\mathrm{a})^{\mathrm{i}-1})$$

is a linearly independent sequence mod $\mathrm{Ker}((\mathrm{T}-\mathrm{a})^{\mathrm{i}-2})$.

If $r = (r_0, r_1, ..., r_i, ...) = (L-1)(t)$ then

$$r_i = \dim(\mathrm{Ker}(\mathrm{T}-\mathrm{a})^{\mathrm{i}+1}) - \dim(\mathrm{Ker}(\mathrm{T}-\mathrm{a})^{\mathrm{i}}) = \dim(\mathrm{Ker}((\mathrm{T}-\mathrm{a})^{\mathrm{i}+1})/\mathrm{Ker}((\mathrm{T}-\mathrm{a})^{\mathrm{i}}).$$

Lemma 2 shows that r is a decreasing sequence of natural numbers which are zero for $i > p$, i.e,

$$r_0 \geq r_1 \geq r_2 \geq \cdots \geq r_p = r_{p+1} = r_{p+2} = ... = 0.$$

Theorem 1 states that the number of Jordan blocks of size $i \geq 1$ is

$$-(r_i - r_{i-1}) = r_{i-1} - r_i = \dim(\mathrm{Ker}((\mathrm{T}-\mathrm{a})^{\mathrm{i}})/\mathrm{Ker}((\mathrm{T}-\mathrm{a})^{\mathrm{i}-1}) - \dim(\mathrm{Ker}((\mathrm{T}-\mathrm{a})^{\mathrm{i}+1})/\mathrm{Ker}((\mathrm{T}-\mathrm{a})^{\mathrm{i}}))$$