8 Jordan canonical form

Throughout this section, we will work with an operator $T \in \mathcal{L}(V)$, and assume that its characteristic polynomial $\chi_T(X)$ factors as a product of linear factors, so that

$$\chi_T(X) = \prod_{i=1}^{r} (X - \lambda_i)^{e_i}.$$

This always happens if V is a complex vector space, and sometimes happens if V is a real vector space.

Recall that we have the following concepts. Let T be an operator on a finite-dimensional vector space. Choose a basis for T, ultimately irrelevant. We have defined the

- Characteristic polynomial: $\det(X \,\mathrm{id} - [T])$.
- Minimal polynomial; the smallest polynomial such that $\mathrm{minpoly}(T) = 0$.

The characteristic polynomial encodes information about the eigenvalues. In fact, λ is an eigenvalue if and only if $\chi_T(\lambda) = 0$.

It turns out that the zeros of the minimal polynomial are the same; we have $\chi_T(\lambda) = 0$ if and only if $\mu_T(\lambda) = 0$.

Let $\lambda_1, \cdots, \lambda_r$ be the distinct eigenvalues of T. Let $U_i = \mathrm{null}(T - \lambda_i)$, and write $\chi_T(X) = \prod (X - \lambda_i)^{e_i}$. We know that T is diagonlizable if and only if $e_i = \dim U_i$ for each i.

Our goal is to assess what happens when this fails. In doing so, we'll get a canonical form for an operator, whether or not it's diagonaliable. As a side effect, we'll be able to compute the minimal polynomial of an operator, too.

Before going too far, it's nice to have the following results:

Lemma If $U \subseteq V$ and $T \in \mathcal{L}(V)$ and U invariant under T, then $\mathrm{charpoly}_{T|_U} \mid \mathrm{minpoly}_{T|_V}$.

Proof Choose a basis, and then use facts about determinants. □

Lemma If $U \subseteq V$ and $T \in \mathcal{L}(V)$ and U invariant under T, then $\mathrm{minpoly}_{T|_U} \mid \mathrm{minpoly}_{T|_V}$.

Proof Sketch, anyway. Let f and g be the respective minimal polynomials. Divide g by f, so that $g(X) = a(X)f(X) + r(X)$, $\deg r < \deg f$. Then $r(T)$ annihilates U, and thus must be zero; otherwise, this would contradict minimality of f. □

8.1 Generalized eigenvectors

There are two sources of difficulty in diagonalizing matrices. One is that, depending on which numbers we work with, a given polynomial may not factor, e.g., $X^2 + 1$ working over the reals. By fiat, we have eliminated this possibility.

The other is somehow more intrinsic, and not so easily fixed. Consider $T \in \mathcal{L}(\mathbb{R}^2)$ with matrix $[T]_{\text{Std}} = \begin{pmatrix} 3 & 1 \\ 0 & 3 \end{pmatrix}$. Then 3 is the only eigenvalue, but $\dim \text{null}(T - 3\,\text{id}) = 1$. What's going on here?

In this example, the *algebraic multiplicity* of the eigenvalue is 2, reflecting the fact that the characteristic polynomial is $(X-2)^2$. Still, if you take the operator $(T-3)^2$, then *everything* is annihilated.

In general, if $T \in \mathcal{L}(V)$ and λ is an eigenvalue of T, a vector $v \in V$ is called a *generalized eigenvector of T* if

$$(T - \lambda\,\text{id})^j v = 0$$

for some j.

So, we need to get good at understanding iterates of operators.

Lemma

a. For all k, null $S^k \subseteq$ null S^{k+1}.

b. If null $S^k =$ null S^{k+1}, then for all j null $S^k =$ null S^{k+j}.

Proof First part is easy. For the last, suppose we know that null $S^k =$ null S^{k+j}; we will try to show that null $S^k =$ null S^{k+j+1}. One inclusion is free; for the other, we need to show that if $S^{k+j+1}(v) = 0$, then $S^k = 0$. But

$$\begin{aligned} 0 &= S^{k+j+1}(v) \\ &= S^{k+1}(S^j v) \\ S^j v &\in \text{null } S^{k+1} \\ S^j v &\in \text{null } S^k \\ v &\in \text{null } S^{j+k} \\ &= \text{null } S^k \end{aligned}$$

□

Proposition If $n = \dim V$, then null $S^n =$ null $S^{n+1} = \cdots$.

Proof Let k be the first value at which null $S^k =$ null S^{k+1}. Then we have dim null $S <$ dim null $S^2 \cdots <$ dim null $S^k \leq n$, so that $k \leq n$. □

We can do something similar with images of operators.

Lemma $S \in \mathcal{L}(V)$, $\dim V = n$.

a. For all k, $\text{im}(S^k) \supseteq \text{im}(S^{k+1})$.

b. im $S^n =$ im $S^{n+1} = \cdots$.

Proof For the first part, if $v \in \operatorname{im} S^{k+1}$, then $v = S^{k+1}(w) = S^k(Sw) \in \operatorname{im} S^k$.

For the second, we know that if $j \geq n$, then $\operatorname{im} S^j = \operatorname{im} S^n$. But

$$\begin{aligned} \dim V &= \dim \operatorname{im} S^j + \dim \operatorname{null} S^j \\ &= \dim \operatorname{im} S^j + \dim \operatorname{null} S^n \\ &= \dim \operatorname{im} S^n + \dim \operatorname{null} S^n \\ \dim \operatorname{im} S^n &= \dim \operatorname{im} S^j \end{aligned}$$

□

Corollary S acts invertibly on $\operatorname{im} S^n$.

Lemma If $S \in \mathcal{L}(V)$, $\dim V = n$, then there is a direct sum decomposition

$$V = \operatorname{im} S^n \oplus \operatorname{null} S^n.$$

Proof Considering the operator S^n, we know that $\dim \operatorname{im} S^n + \dim \operatorname{null} S^n = n$. To show that we get a direct sum, we need to show that the intersection is trivial, so that the dimension of the space spanned by them is n, and then the sum is direct.

So, suppose $v \in \operatorname{im} S^n \cap \operatorname{null} S^n$. Then $v = S^n(w)$ for some w, and $S^n(v) = 0$.

$$\begin{aligned} S^n(v) &= 0 \\ S^{2n}(w) &= 0 \end{aligned}$$

But $\operatorname{null} S^n = \operatorname{null} S^{2n}$, so

$$\begin{aligned} S^n(w) &= 0 \\ v &= 0 \end{aligned}$$

□

Theorem Let λ be an eigenvalue of an operator $T \in \mathcal{L}(V)$. Then there is a decomposition

$$V = \operatorname{null}(T - \lambda \operatorname{id})^n \oplus U$$

where each summand is invariant under T, and λ is not an eigenvalue for the action of T on U.

Proof Consider $S = T - \lambda \operatorname{id}$, and apply the previous result. We get

$$V = \operatorname{null}(S^n) \oplus \operatorname{im} S^n$$

and we call the last part U. Now, if $(T - \lambda \operatorname{id})^n(v) = 0$, then $T(T - \lambda \operatorname{id})^n(v) = 0$. This shows that one summand is invariant under T, and thus the other is, too. Moreover, since S is invertible on $\operatorname{im} S^n$, λ is not an eigenvalue for the action of T on $\operatorname{im} S^n$. □

Note that $\text{null}(T - \lambda\,\text{id})^n$ is exactly the space of *generalized* eigenvectors associated to λ.

Corollary In the above decomposition, we have

$$\text{charpoly}_T(X) = (X - \lambda)^{\dim \text{null}(T-\lambda\,\text{id})^n}\,\text{charpoly}_{T|_U}(X)$$
$$\text{minpoly}_T(X) = \text{minpoly}(T|_{\text{null}(T-\lambda\,\text{id})^n})\,\text{minpoly}(T|_U)$$

Proof For the first one, use the fact that the decomposition of V forces a block-diagonal structure on $[T]$; now use properties of determinants.

For the second, the minimal polynomial clearly divides the product. Then use the fact that they have no factor in common... □

Theorem If $T \in \mathcal{L}(V)$ a complex vector space, let $\lambda_1, \cdots, \lambda_r$ be the distinct eigenvalues of T. Then

$$V = \oplus\,\text{null}(T - \lambda_i\,\text{id})^n.$$

Proof Let λ_1 be the first such eigenvalue; write $V = \text{null}(T - \lambda_1)^n \oplus U_1$. Note that λ_1 is not an eigenvalue of T acting on U_1. Now recurse; do the same thing for U_1. □

8.2 Bases for nilpotent actions

We work with an operator S, which we think of as being $T - \lambda$. Our goal is to understand the generalized eigenspace associated to λ, that is, the nullspace of S^n.

We can build up the space as follows.

Suppose $v \in V$, $S^j(v) = 0$, $S^{j-1}(v) \neq 0$.

Lemma The set $\{S^{j-1}(v), S^{j-2}(v), \cdots, S(v), v\}$ is linearly independent.

Proof We will prove by induction that $\{S^{j-1}(v), \cdots, S^{j-i}(v)\}$ is linearly independent. This is obvious for $i = 1$, since we're guaranteed that $S^{j-1}(v) \neq 0$.

Otherwise, assume we've secured the statement for $i - 1$, and that

$$\begin{aligned}
a_{j-1}S^{j-1}(v) + a_{j-1}S^{j-2}(v) + \cdots + a_{j-i+1}S^{j-i+1}(v) + a_{j-i}S^{j-i}(v) &= 0\\
S^{i-1}(a_{j-1}S^{j-1}(v) + a_{j-1}S^{j-2}(v) + \cdots + a_{j-i+1}S^{j-i+1}(v) + a_{j-i}S^{j-i}(v)) &= 0\\
a_{j-i}S^{j-1}(v) &= 0\\
a_{j-i} &= 0\\
a_{j-1}S^{j-1}(v) + a_{j-1}S^{j-2}(v) + \cdots + a_{j-i+1}S^{j-i+1}(v) &= 0
\end{aligned}$$

so that each $a_i = 0$, by induction. □

For such a v, note that the space $U_v = \text{span}\{S^{j-1}(v), \cdots, S^2(v), S(v), v\}$ is invariant under S. The matrix of $S|_{U_v}$ is a *nilpotent Jordan block*, which we should draw now but I'm too damned lazy.

Lemma The minimal polynomial for $S|_{U_v}$ is X^j.

Proof □

8.3 Computing nilpotent Jordan forms

Let $K_j = \text{null}(S^j)$. There is always an inclusion $K_{j-1} \subseteq K_j$. Let L_j be a complement so that $K_j = K_{j-1} \oplus L_j$. Then L_j is the set of elements which are annihilated by S^j, but not by S^{j-1}.

Lemma Suppose $j \geq 2$. Then $S(L_j) \subseteq L_{j-1}$. If $v \in L_j$ and $S(v) = 0$, then $v = 0$.

Corollary If $v_1, \cdots, v_r \in L_j$ are linearly independent, then so are $S(v_1), \cdots, S(v_r)$.

Let $k_j = \dim K_j$, $\ell_j = \dim L_j$. Then $\ell_j = k_j - k_{j-1}$.

We can read off the Jordan blocks, in the following sense.

Suppose that the chain stabilizes at m, so that $K_m = K_{m+1} = \cdots$.

Then there are:

- ℓ_m Jordan blocks of size m.
- $\ell_{m-1} - \ell_m$ Jordan blocks of size $m-1$
- $\ell_{m-2} - \ell_{m-1}$ Jordan blocks of size $m-2$.
- $\cdots$
- $\ell_k - \ell_{k+1}$ Jordan blocks of size k.

The point is that, at each stage, we can fill up part of L_k using the images of elements from L_{k+1}; the rest come from Jordan blocks of exact size k.

Moreover, the number of Jordan blocks of size *at least* k is just ℓ_k.

Example Consider the operator $S \in \mathcal{L}(\mathbb{R}^5)$ with matrix

$$[S] = \begin{bmatrix} 58 & -224 & 511 & -214 & 4 \\ 16 & -62 & 139 & -58 & 1 \\ 6 & -24 & 50 & -20 & 0 \\ 13 & -52 & 110 & -44 & 0 \\ -4 & 12 & -38 & 20 & -2 \end{bmatrix}$$

Then you can compute directly that $S^5 = 0$. Moreover, the dimensions of the nullspaces of various iterates of S are as fllows:

j	$k_j = \dim K_j = \dim \operatorname{null}(S^j)$
4	5
3	5
2	4
1	2
0	0

Let's see what we can say about the Jordan block structure of this thing. Note that that already $A^3 = 0$; this means that the minimal polynomial divides X^3. Moreover, $A^2 \neq 0$, so that the minimal polynomial is exactly X^2.

Now, $\dim L_3 = 1$. Let $v^{(3)}$ be a basis for it. Then the images of $v^{(3)}$ under S are in the various spaces L_j, and we have

$$\begin{array}{cc} L_3 & v^{(3)} \\ L_2 & S(v^{(3)}) \\ L_1 & S^2(v^{(3)}) \end{array}$$

(In fact, you can check this directly, using $v^{(3)} =$??)

Now, $\dim L_2 = 4 - 2 = 2$, but we've only found one vector in it so far. Let $v^{(2)}$ be an element of L_2 linearly independent from $v^{(3)}$. Then we have

$$\begin{array}{cc} L_3 & v^{(3)} \\ L_2 & S(v^{(3)}), v^{(2)} \\ L_1 & S^2(v^{(3)}), S(v^{(2)}) \end{array}$$

In fact, this is everything; for now we have found 2 independent elements in L_1, but $\dim L_1 = 2$.

So, consider the basis $\mathcal{B} = \{S^2(v^{(3)}), S(v^{(3)}), v^{(3)}, S(v^{(2)}), v^{(2)}\}$. With respect to this basis, we have

$$[S]_{\mathcal{B}} = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

The characteristic polynomial of this operator is X^5, but its minimal polynomial is X^3.

8.4 Return to generalized eigenvalues

So, let $\lambda_1, \cdots, \lambda_r$ be the distinct eigenvalues of T.

We have seen that there is a direct sum decomposition

$$V = \oplus(\mathrm{null}((T - \lambda_j)^n)).$$

Let $U_j = \mathrm{null}(T - \lambda_j)^n$ be the generalized eigenspace associated to λ_j. Then using the previous, we may find a basis for U_j such that

$$[T - \lambda_j|_{U_j}]$$

is a Jordan block with zero's on the diagonal, so that $[S|_{U_j}] = \cdots$. Moreover, $[T|_{U_j}]$ is in Jordan form with λ_j on the diagonal.

Example Consider

$$T = \begin{bmatrix} 60 & -224 & 511 & -214 & 4 \\ 16 & -60 & 139 & -58 & 1 \\ 6 & -24 & 52 & -20 & 0 \\ 13 & -52 & 110 & -42 & 0 \\ -4 & 12 & -38 & 20 & 0 \end{bmatrix}.$$

Its characteristic polynomial is $\chi_T(X) = (X - 2)^5$. The work done before shows that there is a basis such that

$$[S]_{\mathcal{B}} = \begin{pmatrix} 2 & 1 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 2 \end{pmatrix}$$

Assembling these bases for the various U_j, we have shown that in some basis $[T]$ is in Jordan normal form.

And this is sort of the point of the whole thing. Supppose you promise to always order eigenvalues in a certain way. (It's obvious how to do this if the eigenvalues are real; it requires choices if the eigenvalues are complex.) Suppose further that you promise to order the Jordan blocks from, say, largest to smallest.

Then two matrices are similar – that is, they represent the same operator in different bases – if and only if they have the same Jordan form.

Here are some more results about Jordan form:

Theorem Suppose $T \in \mathcal{L}(V)$, $\dim V = n$. Let $\lambda_1, \cdots, \lambda_r$ be the distinct eigenvalues, and suppose that

$$\chi_T(X) = \prod (X - \lambda_i)^{e_i}$$

where $\sum e_i = n$.

Let $V_i = \ker(T - \lambda_i\,\mathrm{id})$, and let $U_i = \ker((T - \lambda_i\,\mathrm{id})^n)$. Let the sizes of the Jordan blocks of U_i be given by $s_{i,1} \geq s_{i,2} \geq \cdots \geq s_{i,b_i}$.

- V_i is the eigenspace associated to λ_i.
- U_i is the generalized eigenspace associated to λ_i.
- The minimal polynomial of T is
$$\prod(X - \lambda_i)^{s_{i,1}}.$$
- T is diagonalizable if any of the following holds:
  i. For each i, $\dim V_i = e_i$
  ii. For each i, $U_i = V_i$
  iii. The size of each Jordan block is $s_{i,j} = 1$
  iv. The minimal polynomial of T is
$$\prod(X - \lambda_i).$$
- Let $T' \in \mathcal{L}(V)$ be another linear operator. Then T and T' are similar if and only if $\chi_T(X) = \chi_{T'}(X)$ and if they have the same Jordan block structure.

Example Suppose you know the following information about an operator $T \in \mathcal{L}(\mathbb{R}^{20})$:

- $\chi_T(X) = (X-2)^{16}(X+3)^4$
- Let $S_2 = (T-2)$, $S_{-3} = (T+3)$. Suppose we have the following information:

i	$\dim \operatorname{null}((S_2)^i)$
5	16
4	16
3	14
2	11
1	6
0	0

i	$\dim \operatorname{null}((S_{-3})^i)$
5	4
4	4
3	4
2	4
1	3
0	0

What is the Jordan canonical form of this matrix?