11 The Jordan Decomposition

11.1 Review

Recall this theorem from last time.

> **Theorem 11.1**
> Considering a transformation $T : V \to V$, there must exist a basis $\vec{v}_1, \cdots, \vec{v}_n$ such that the matrix of T (in this basis) is
> $$A = \begin{pmatrix} J_{a_1}(\lambda_1) & 0 & \cdots & 0 \\ 0 & J_{a_2}(\lambda_2) & \ddots & 0 \\ \vdots & \ddots & J_{a_i}(\lambda_i) & 0 \\ 0 & 0 & 0 & J_{a_n}(\lambda_n) \end{pmatrix},$$
> where $J_{a_i}(\lambda_i)$ are the Jordan blocks.

A special case is when all the $a_i = 1$. Then,

$$A = \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_r \end{pmatrix}$$

is a diagonal matrix.[35]

11.2 The Jordan Decomposition, Continued

The **characteristic polynomial** of the matrix A will be

$$p_A(t) = (t - \lambda_1)^{a_1} \cdots (t - \lambda_r)^{a_r},$$

where it is possible to have repeated λ_i. As a result, it is not possible to determine the Jordan decomposition simply from the characteristic polynomial, since there are different ways to take a repeated root and split it up into Jordan blocks. (If all the roots of the characteristic polynomial are distinct, the Jordan form is uniquely determined.)

However, the characteristic polynomial does provide some information. For a fixed eigenvalue λ,

$$\sum_{J_{a_i}(\lambda)} a_i = \text{exponent of } (t - \lambda) \text{ in } p_A(t).$$

> **Example 11.2** ($n = 4$)
> For example, when $n = 4$, consider a matrix where $p_A(t) = t^4$. There are multiple possible Jordan forms; in particular, it can be split up as $4, 3 + 1, 2 + 2, 2 + 1 + 1,$ or $1 + 1 + 1 + 1 :$
> $$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 & 0 & \\ 0 & 0 & 1 & \\ 0 & 0 & 0 & \\ & & & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 & & \\ 0 & 0 & & \\ & & 0 & 1 \\ & & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 & & \\ 0 & 0 & & \\ & & 0 & \\ & & & 0 \end{pmatrix}, \begin{pmatrix} 0 & & & \\ & 0 & & \\ & & 0 & \\ & & & 0 \end{pmatrix}.$$

For a given Jordan block, there is one eigenvector. Fixing λ again, this tells us that

$$\dim(\ker(\lambda I - A))$$

is equal to the number of blocks with λ along the diagonal.

Up to reordering of the basis vectors, the Jordan decomposition is unique.

---

[35]In the textbook, Artin puts the 1s below the diagonal in a Jordan block. Conventionally, the 1s are above the diagonal, but it doesn't make a difference, because reversing the order of the vectors $\vec{e}_1, \cdots \vec{e}_a$ to $\vec{e}_a, \cdots \vec{e}_1$ moves the 1s from above the diagonal to below the diagonal. The difference is notational.

> **Example 11.3**
>
> Take $J_4(0) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$. Under $J_4(0)$, each basis vector maps to the next basis vector, and there is one chain of length 4:
>
> $$\vec{e}_4 \mapsto \vec{e}_3 \mapsto \vec{e}_2 \mapsto \vec{e}_1 \mapsto \vec{0}.$$
>
> As a result, applying $J_4(0)$ multiple times will eventually send all vectors to zero; that is, in this case, $J_4(0)^4 = 0$.
>
> On the other hand, consider $J_{2,2}(0) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} J_2(0) & 0 \\ 0 & J_2(0) \end{pmatrix}$.
>
> Applying the operator to the basis vectors yields two chains of length 2:
>
> $$\vec{e}_2 \mapsto \vec{e}_1 \mapsto \vec{0}$$
>
> $$\vec{e}_4 \mapsto \vec{e}_3 \mapsto \vec{0}$$
>
> In this case as well, the operator will map every vector to zero upon repeated application.

In general, for $\lambda \neq 0$, $(\lambda I - T)\vec{e}_i$ is not necessarily zero (it is zero only if $\vec{e}_i$ is an eigenvector), but for some large enough n,

$$(\lambda I - T)^n \vec{e}_i = 0.^{36}$$

In Example 11.3, there was a *chain* of length 4 for the first matrix, while in the second matrix, we had two chains of length 2.

> **Note 11.4**
>
> The Jordan decomposition theorem is powerful because *any* square matrix has a Jordan decomposition. On the other hand, most matrices are diagonalizable, and any matrix will be ε away from a diagonalizable matrix, and the Jordan decomposition is unnecessary. Only in the zero percent of the time[a] when the characteristic polynomial has repeated roots is it necessary.
>
> ---
> [a]This concept is fleshed out in **measure theory**.

11.3 Proof of Jordan Decomposition Theorem

The proof of the Jordan decomposition theorem is quite involved and relatively tricky, so the important part for the rest of class is understanding the style of proof, rather than the exact details. This proof will break down the theorem inductively into smaller and smaller pieces.

Let's start with a couple of definitions that will help us with the proof.

> **Definition 11.5**
>
> Given a vector space V and a linear transformation $T : V \to V$, a *subspace* $W \subseteq V$ is called T-**invariant** if $T(\vec{w}) \in W$ for all $\vec{w} \in W$.

For example, if the vector space V is the space of polynomials of degree at most 3, and the subspace W is the space of polynomials of degree at most 2, W will be $T-$invariant under the linear operator T that is taking the derivative.

---

[36]A vector that is killed not necessarily immediately but eventually by $\lambda I - T$ is known as a **generalized eigenvector**; there is a question about them on the problem set.

> **Definition 11.6**
> Given a vector space V and two subspaces $W, W' \subseteq V$, we say that V is the **direct sum** of W and W', notated $V = W \oplus W'$, if every $\vec{v} \in V$ can be written *uniquely* as
> $$\vec{v} = \vec{w} + \vec{w}',$$
> where $\vec{w} \in W$ and $\vec{w}' \in W'$.

For example, if $V = \mathbb{R}^3$, every vector can be written as the sum of some vector in the $z-$direction and some vector lying in the $xy-$plane.

Equivalently, there must exist a basis
$$\{\vec{w}_1, \cdots, \vec{w}_r, \vec{w}'_1, \cdots, \vec{w}'_r\}$$
of V such that $\{\vec{w}_1, \cdots, \vec{w}_r\}$ is a basis of W and $\{\vec{w}'_1, \cdots, \vec{w}'_r\}$ is a basis of W'. This is also sometimes called a **splitting** of V, since V has been split up into two subspaces.

> **Theorem 11.7**
> If $\dim W + \dim W' = \dim V$, and $W \cap W' = \{\vec{0}\}$, then it must be the case that $V = W \oplus W'$.[a]
>
> ---
> [a] This can be proved using the characterization in terms of bases, and is related to a homework problem.

> **Definition 11.8**
> Given a splitting $V = W \oplus W'$ and a linear operator $T : V \to V$, we say that this splitting is T-*invariant* if W and W' are $T-$invariant.

In a basis for W and W', the matrix for T must be block-diagonal; that is, of the form
$$\begin{pmatrix} \star & 0 \\ 0 & \star \end{pmatrix}$$
where each $\star$ is some matrix; this is because vectors in W or W' will be mapped back to other vectors in W or W'.

Conversely, if T is block diagonal in some basis, it automatically provides a T-invariant splitting of V. The span of the collection of basis vectors in the first block becomes a T-invariant subspace W and the second one becomes W'. Essentially, these definitions provide a characterization of linear transformations being **block-diagonal**, *without* having to pick a basis.

Now, we can finally start proving the Jordan Decomposition Theorem. Roughly, the proof follows an induction argument on the dimension of V, where a vector space is split up into two smaller dimensional T-invariant subspaces for the operator T, both of which will then have Jordan decompositions by the inductive hypothesis, which will provide the Jordan decomposition of the original vector space. Essentially, we want to break it down to the case of a singular eigenvalue, considering matrices that look like those in Example 11.3, relying on the fact that repeatedly applying these operators will eventually take any vector to zero.

> **Definition 11.9**
> A linear transformation T is **nilpotent** if there exists some $m \geq 0$ such that $T^m = 0$.[a]
>
> ---
> [a] The Jordan block $J_m(0)$ is nilpotent with exponent m.

Proof. This proof has several steps.

- **Step 0.** Over complex vector spaces, there will always exist an eigenvalue, so let λ be some eigenvalue of T. Because λI is already diagonal, we can replace T with $T - \lambda I$, so that 0 can be assumed to be one of the eigenvalues. Essentially, if the Jordan decomposition theorem is true for $T - \lambda I$, by adding the diagonal matrix λI, the Jordan decomposition theorem will become true for T.

- **Step 1.**

Rough Sketch. After this simplification, we will *zero in*[37] on the 0 eigenvalue. We show that there exists a $T-$invariant splitting $V = W \oplus U$ such that

$$T_w : W \to W$$

is nilpotent and

$$T_u : U \to U$$

is invertible. For a nilpotent operator, the only possible eigenvalues are 0[38], while for an invertible operator, there are only nonzero eigenvalues[39], so this splitting separates out the eigenvectors will eigenvalue $\lambda = 0$.

By assumption, there exists a zero eigenvalue, and so $\dim W \geq 1$, and then $\dim U \lneq n$. Since $\dim U < n$, by the inductive hypothesis, there is a Jordan decomposition for U. However, since $\dim W$ could be equal to n ($\dim U$ could be 0), the inductive hypothesis does not apply and so we must still show that there is a Jordan decomposition for nilpotent operators.

Full Proof. Now, we still need to show that this splitting exists. Consider the vector space V; $TV = \operatorname{im} T$ lies inside of V (it cannot possibly take V to a higher-dimensional space), and so we obtain the chain

$$V \supset TV \supset T(TV) \supset T(T(T(V)) \supset \cdots .$$

The dimension can only drop finitely many times[40], since $T^i(V)$ cannot have negative dimension, so there exists some stable dimension m between $\dim T$ and 0 such that

$$T^m V = T^{m+1} V = T^{m+2} V = \cdots .$$

Let

$$U := T^m V = \operatorname{im}(T^m) \text{ and } W = \ker(T^m).$$

First, T is nilpotent on W because $W = \ker(T^m)$, so $(T|_W)^m = 0$, which is the definition of being nilpotent. Also, $T|_U$ is invertible because $U = \operatorname{im}(T|_U)$, so $T|_U$ is surjective from U to itself, which implies that it is invertible. Lastly, $W \cap U = \{v \in U : T^m = 0\}$, by definition, which is precisely the zero vector, because T is invertible on U so it maps only the zero vector to the zero vector. Using the rank-nullity theorem, $\dim \ker T^m + \dim \operatorname{im} T^m = \dim V$, so by Theorem 11.7, $W \oplus U$ is in fact a splitting.

- **Step 2.** Now, we prove that if T is nilpotent, it has a Jordan decomposition. We have a vector space V, a linear operator $T : V \longrightarrow V$, and some m such that $T^m = 0$.

To do so, we will find by induction on the dimension a basis of V for which T acts in "chains" as in Example 11.3. Let $W = \operatorname{im} T \subsetneq V$. By induction, there exists such a basis $\{\vec{e}_i\}$ for W where T acts in chains.

---
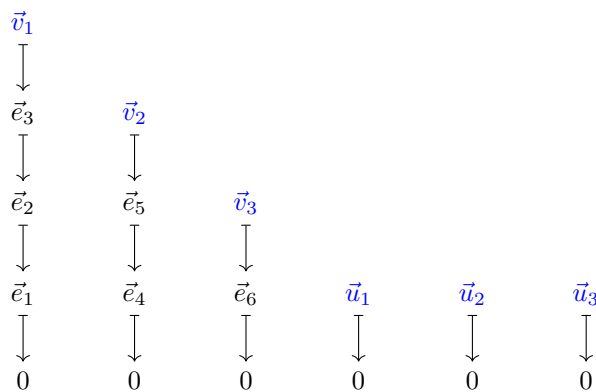
[37]Ha ha

[38]Consider a nilpotent operator A. Then there is some n such that $A^n = 0$. If v is an eigenvector for A with eigenvalue λ, $A^n v = \lambda^n v = 0$, so $\lambda^n = 0$ and thus λ must also be zero.

[39]Assume 0 is an eigenvalue of an invertible operator A, corresponding to an eigenvector v. Then $Av = 0$ for some nonzero v; then both the vector 0 and the vector v map to 0 and thus A is not one-to-one or invertible.

[40]In fact, this argument relies on the fact that V is finite-dimensional!

For each chain, we insert a preimage $\vec{v}_i$ of the top vector in each chain, where the $\vec{v}_i$ are not vectors in W but rather vectors that *map* to vectors in W. These exist since W is the image of T, and so every vector in W is the image of some other vector under T. Additionally, we add vectors in $\ker T$, $\vec{u}_i$, which all map to zero since they are in the kernel. This produces a bunch of chains for V, starting for a bunch of chains for W.

We claim that $\mathcal{B} = \{\vec{e}_i\} \cup \{\vec{v}_j\} \cup \{\vec{u}_k\}$ is a basis for V. It is linearly independent because applying T to any linear dependence would give a dependence between basis vectors of W (since $T(\vec{v}_i) = \vec{e}_j$ and $T(\vec{u}_k) = 0$.) Also, where c is the number of chains, the number of vectors in $\mathcal{B}$ is $\dim(W) + \dim(\ker(T)) - c + c$, which is precisely the dimension of V, and thus $\mathcal{B}$ is in fact a basis.

For this particular example illustrated in the figure, the Jordan blocks for W have size 3, 2, and 1, and for V these are extended to size 4, 3, and 2, along with three more blocks of size 1.

The schematic of the argument is more important than the exact argument itself, but we still have to do the whole thing. :)

$\square$

Resource: Algebra I Student Notes
Fall 2021
Instructor: Davesh Maulik
Notes taken by Jakin Ng, Sanjana Das, and Ethan Yang