Jordan Normal Form via Exercises

Taking linear algebra as an undergraduate, I got pretty lost when we discussed the Jordan Normal Form. As such, I recently needed to revisit this topic before I could understand certain results about Lie algebras. This note presented a much cleaner explanation than the other sources that I had looked at, which inspired me to jot down this framework (with a few tweaks befitting my own understanding) into a hastily written set of exercises on the topic. Several of the exercises are equipped with a footnote containing a hint.

Generalized Eigenspaces

Let $A : V \to V$ be a linear operator on a finite-dimensional vector space over $\mathbb{C}$.

Definition. For any $\lambda \in \mathbb{C}$, the generalized λ-eigenspace of A is defined to be

$$V_\lambda = \{v \in V : (A - \lambda I)^k v = 0 \text{ for some } k \in \mathbb{N}\}$$

Any non-zero $v \in V_\lambda$ is called a generalized λ-eigenvector. A λ-chain of length k is a sequence of non-zero vectors $v_1, \ldots, v_k \in V$ such that $(A - \lambda I)v_i = v_{i-1}$ for all $i = 2, \ldots, k$ and $(A - \lambda I)v_1 = 0$.

1. If $v_1, \ldots, v_k$ is a λ-chain, then $\{v_1, \ldots, v_k\}$ is a linearly independent subset of V_λ.[1]

Jordan Normal Form Theorem. There is a basis of V that is a union of chains. The number of λ-chains of length k in this basis depends only on our original operator A.

To prove this theorem, we first want to break up V into its generalized eigenspaces.

2. There is some $k \in \mathbb{N}$, such that $(A - \lambda I)^k v = 0$ for all $v \in V_\lambda$ (or equivalently $V_\lambda = \ker(A - \lambda I)^k$).[2]
3. If $V_\lambda = \ker(A - \lambda I)^k$, then $V = \ker(A - \lambda I)^k \oplus \operatorname{im}(A - \lambda I)^k$.[3]
4. For any $k \in \mathbb{N}$, the subspaces $\ker(A - \lambda I)^k$ and $\operatorname{im}(A - \lambda I)^k$ are A-invariant.[4]
5. There exists an A-invariant subspace $W \subseteq V$ such that $V = V_\lambda \oplus W$. Moreover:
   (a) If $\mu \neq \lambda$, then $A - \mu I$ restricts to an invertible operator on V_λ.[5]
   (b) If $\mu \neq \lambda$, then $V_\mu \subseteq W$.[6]
6. If the eigenvalues of A are $\lambda_1, \ldots, \lambda_n$, then we have $V = V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_n}$.

Each generalized eigenspace of A is A-invariant, so to prove existence in the Jordan Normal Form Theorem, it now suffices to show that each V_λ admits a basis that is a union of λ-chains. To prove this, we now restrict to a single V_λ and set $B = A - \lambda I$. By what we have shown above, this operator B has the following property:

Definition. A linear operator $B : V \to V$ is said to be nilpotent if $B^k = 0$ for some $k \in \mathbb{N}$.

[1] If $a_j v_j + \cdots + a_1 v_1 = 0$, what happens when you apply $(A - \lambda I)^{j-1}$ to this linear combination?
[2] If $v_1, \ldots, v_n$ is a basis of V_λ, then there exist $k_1, \ldots, k_n \in \mathbb{N}$ such that $(A - \lambda I)^{k_i} v_i = 0$ for each $i = 1, \ldots, n$.
[3] First show that their intersection is trivial.
[4] Notice that A and $A - \lambda I$ commute.
[5] If $v \in \ker(A - \mu I)$, then v is μ-eigenvector of A. What can we conclude about $(A - \lambda I)v$? What about $(A - \lambda I)^k v$?
[6] Any $u \in V_\mu$ can be written as $u = v + w$, with $v \in V_\lambda$ and $w \in W$. What happens if we apply $(A - \mu I)^k$ to each side?

Nilpotence and Flags

Definition. A weak flag in a vector space U is any ascending collection of subspaces, going from 0 to U:

$$0 = N_0 \subseteq N_1 \subseteq \cdots \subseteq N_\ell = U.$$

Fix a finite set of vectors $\mathcal{S} = \{v_1, \dots, v_m\} \in N_j$. We will say that $\mathcal{S}$ is j-spanning if $N_j = N_{j-1} + \operatorname{span}(\mathcal{S})$. We will say that $\mathcal{S}$ is j-independent if the only $a_1, \dots, a_m \in \mathbb{C}$ satisfying

$$a_1 v_1 + \cdots + a_m v_m \in N_{j-1}$$

are $a_1 = \cdots = a_m = 0$. If the set $\mathcal{S}$ is both j-spanning and j-independent, we will call it a j-basis.

7. Any j-independent set can be enlarged to form a j-basis.

8. For each $j = 1, \dots, \ell$, suppose that $\mathcal{S}_j$ is a j-basis. Then $\mathcal{S}_1 \cup \cdots \cup \mathcal{S}_\ell$ is a basis of U.

Suppose $B : U \to U$ is a nilpotent operator with $B^\ell = 0$. For each $i = 0, 1, \dots, \ell$, we define $N_i = \ker(B^i)$.

9. This yields a weak flag $0 = N_0 \subseteq N_1 \subseteq \cdots \subseteq N_\ell = U$ such that $B(N_i) \subseteq N_{i-1}$ for each $i = 1, \dots, \ell$.

10. If $\mathcal{S}$ is j-independent, then $B(\mathcal{S})$ is $(j-1)$-independent.[7]

11. There is a basis $\mathcal{S}$ of the vector space U, such that if $u \in \mathcal{S}$ and $B^i u \neq 0$, then $B^i u \in \mathcal{S}$.[8]

We now return to the situation of a linear operator $A : V \to V$, a generalized eigenspace V_λ and the nilpotent operator $B : V_\lambda \to V_\lambda$ given by $B = A - \lambda I$. We still consider the weak flag $N_i = \ker(B^i)$ defined above.

12. The generalized eigenspace V_λ admits a basis that is a union of λ-chains of A.

13. Given a basis of V_λ that is a union of λ-chains, the number of chains of length k is $\dim N_k - \dim N_{k-1}$.[9]

We now have all of the pieces necessary to conclude the result of the Jordan Normal Form Theorem!

Commuting Operators

Closely related to the Jordan Normal Form Theorem is another result, describing a decomposition of A:

Jordan-Chevalley Decomposition Theorem. Given a linear operator $A : V \to V$, there is a unique way to write $A = B + D$, where D is diagonalizable, B is nilpotent and $BD = DB$.

Before we prove this, we need to investigate a couple of more general results about commuting operators. Consider two operators $S : U \to U$ and $T : U \to U$ that satisfy $ST = TS$.

14. If S and T are nilpotent, then $S + T$ is nilpotent.[10]

[7] If $a_1 B v_1 + \cdots + a_m B v_m \in N_{j-2}$, then $a_1 v_1 + \cdots + a_m v_m \in N_{j-1}$.
[8] Given a j-basis $\mathcal{S}_j$, we can enlarge $B(\mathcal{S}_j)$ to a $(j-1)$-basis. Start with an ℓ-basis and iterate this process.
[9] If n_k is the number of chains of length k, then we have $\dim N_j = n_j + n_{j+1} + \cdots + n_\ell$ for each $j = 1, \dots, \ell$.
[10] If $S^n = 0$ and $T^n = 0$, then $(S + T)^{2n} = 0$.

15. The eigenspaces of S are T-invariant.

16. Now suppose that S and T are diagonalizable and let $U = U_{\mu_1} \oplus \cdots \oplus U_{\mu_m}$ be the decomposition into eigenspaces of S. For each $i = 1, \ldots, m$, let $P_i : U \to U_{\mu_i}$ be the projection map out of this direct sum.

    (a) For every $i = 1, \ldots, m$, we have $TP_i = P_i T$.[11]

    (b) If $u \in U$ is an eigenvector of T, then $P_i u$ is a common eigenvector of S and T.

    (c) Every eigenvector of T can be written as a sum of common eigenvectors of S and T.

    (d) The vector space U is spanned by common eigenvectors of S and T.

    (e) There is a basis for U in which both S and T are diagonal.

    This proves that if S and T are diagonalizable, then so is $S + T$ (under the assumption $ST = TS$).

Recall the decomposition $V = V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_n}$ into generalized eigenspaces of A. We define a diagonalizable linear operator $D : V \to V$ by requiring that $D = \lambda I$ when restricted to the generalized eigenspace V_λ.

17. If a linear operator $B : V \to V$ commutes with A, then it commutes with D.[12]

18. The operator $B = A - D$ is nilpotent and $BD = DB$.[13]

This proves existence of a Jordan-Chevalley decomposition, but we also need to show uniqueness. As such, suppose that we also have $A = B' + D'$, where D' is diagonalizable, B' is nilpotent and $B'D' = D'B'$.

19. The operators B, D, B' and D' all commute with each other.[14]

20. We have $B - B' = D' - D = 0$, so the decomposition is unique.[15]

This concludes the proof of the Jordan-Chevalley Decomposition Theorem!

[11] It suffices to show that $TP_i u = P_i T u$ when $u \in U_j$, for each $j = 1, \ldots, n$. If $u \in U_j$ and $i \neq j$, then $P_i u = 0$.
[12] If $AB = BA$, then each V_λ is B-invariant.
[13] Verify nilpotency on each V_λ individually.
[14] First show that D commutes with B' and D'.
[15] The only operator that is both diagonalizable and nilpotent is 0.