Jordan Canonical Form

S. Kumaresan
School of Math. and Stat.
University of Hyderabad
Hyderabad 500046
kumaresa@gmail.com

We shall assume that V is a finite dimensional vector space over an algebraically closed field. We start with an easy observation.

Lemma 1. *Let $v_1, v_2, \ldots, v_m$ be linearly independent in a vector space. Assume that $v_i \in \operatorname{Span}\{w_1, \ldots, w_n\}$. Then $n \geq m$.* □

Proposition 2. *Let $T \colon V \to V$ be a linear map. Then there exists a a polynomial $p(X) := \sum_{k=1}^{d} a_k X^k$ such that $p(T) := \sum_k a_k T^k = 0$.*

Proof. Fix a basis $v_1, \ldots, v_n$ of V. For each i with $1 \leq i \leq n$, we can find scalars $a_0, a_1, \ldots, a_n$ such that $a_0 v_1 + a_1 T v + \cdots + a_n T^n v = 0$. If we let $p_i(X) := \sum_i a_i X^i$, then $p_i(T) v_i = 0$. If we let $p = p_1 \cdots p_n$ be the product of these polynomials, then $p(T) v_i = 0$ for each i and hence $p(T) v = 0$ for all $v \in V$. □

Proposition 3. *Let $T \colon V \to V$ be linear. Let $N(T)$ and $R(T)$ denote the kernel and the range of T. Then they are invariant under T. If $N(T) \cap R(T) = \{0\}$, then $V = N(T) \oplus R(T)$.*

Proof. A trivial application of the rank-nullity theorem. □

Theorem 4. *Any linear map on a finite dimensional vector space over an algebraically closed field has an eigen vector.*

Proof. Let $T \colon V \to V$ be one such. By Proposition 2, there exists a polynomial $p(X)$ with coefficients in the field such that $p(T) = 0$. Since the filed is algebraically close, we can write $p(X) = c(X - \alpha_1) \cdots (X - \alpha_k)$. Let $v \in V$ be any nonzero vector. Let i be the least integer such that $c(T - \alpha_i I) \cdots (X - \alpha_i I) v = 0$. If $i = 1$, then v is an eigen vector with eigen value α_1, otherwise, $w := (T - \alpha_{i-1} I) \cdots (T - \alpha_1 I) v$ is an eigen vector with eigen value α_i. □

Proposition 5. *Let λ be an eigenvalue of a linear map $T \colon V \to V$. Let the generalize eigensubspace of λ be defined by*

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

Then there exists $r \geq 1$ such that $V(\lambda) = N((T - \lambda I)^r)$ and we have

$$V(\lambda) = N((T - \lambda I)^r) \text{ and } V = N((T - \lambda I)^r) \oplus R((T - \lambda I)^r).$$

The summands are invariant under T.

Proof. We show first that $V(\lambda)$ is a vector subspace. It is closed under scalar multiplication. If $v_i \in V(\lambda)$ such that $(T-\lambda I)^{k_i} v_i = 0$, then $(T-\lambda I)^k(v_1+v_2) = 0$ for $k = \max\{k_1, k_2\}$. Since $V(\lambda) \subset V$, it is finite dimensional. If we fix a basis of $V(\lambda)$, and we take r to be the maximum of the k's corresponding to these finite number of elements, it is clear that $V(\lambda) = N((T-\lambda I)^r)$.

We claim that the intersection of the kernel and the range of $T-\lambda I)^r$ is $\{0\}$. Let w be in their intersection. Then there exists $v \in V$ such that $w = (T-\lambda I)^r v$. Since w lies in the kernel of $(T-\lambda I)^r$, it follows that $(T-\lambda I)^r w = (T-\lambda I)^{2r} v = 0$. It follows that $v \in V(\lambda)$, that is in the kernel of $(T-\lambda I)^r$. But then $w = (T-\lambda I)^r v = 0$. The claim follows. By the rank-nullity theorem, we have the direct sum decomposition as in the theorem.

Each of the summands is invariant under $(T-\lambda I)$ as well as under λI and hence under $T = (T-\lambda I) + \lambda I$. □

Theorem 6. *Let $T\colon V \to V$ be a linear map on a finite dimensional vector space over an algebraically closed field. Then V is the direct sum of generalized eigen spaces of T.*

Proof. The result follows by induction (on the number of eigenvalues of T) using the last proposition. □

Definition 7. A linear map $T\colon V \to V$ is said to be *nilpotent* of index r if $T^r = 0$ but $T^{r-1} \neq 0$.

Lemma 8. *Let A be nilpotent of index r. Then we have a chain of strict inclusions*

$$N(A) \subset N(A^2) \subset \cdots \subset N(A^r) = V.$$

Proof. Inclusions are obvious. By the definition of index there exists $v \in V$ such that $A^{r-1}v \neq 0$. Then the vector $A^{r-i}v \in N(A^i)$ but not in $N(A^{i-1})$. □

Definition 9. We say that the vectors $v_1, \ldots, v_k$ is *independent of* a subspace $W \subset V$ if whenever a linear combination $a_1v_1 + \cdots + a_kv_k \in W$, it follows that $a_i =$ for $1 \leq i \leq k$.

Theorem 10 (Jordan canonical Form for Nilpotent Maps)**.** *Let $A\colon V \to V$ be nilpotent of index r. Let $S \subset V$ be linearly independent of A^{r-1}. Then*

(a) *there exist a number m and vectors $v_1, \ldots, v_m$ such that*

(i) *the nonzero vectors of the form $A^j v_i$, for $j \geq 0$ and $1 \leq i \leq m$ from a basis of V and*

(ii) *S is a subset of this basis.*

(b) *For $1 \leq i \leq m$, let r_i be the least integer such that $A^{r_i}v_i = 0$. Let V_i be the subspace spanned by the linearly independent set $\{A^j v_i : 0 \leq i \leq r_i - 1\}$. Then*

$$V = V_1 \oplus \cdots \oplus V_m. \tag{1}$$

(i) *Each V_i is invariant under A and*

(ii) *A is nilpotent of index r_i on V_i.*

(c) *Let $\varphi(i)$ be the the number of subspaces in the decomposition* (1) *of dimension at least i. Then*

$$\dim N(A^i) - \dim(A^{i-1}) = \varphi(i).$$

(d) The number of subspaces in (1) *of any given dimension is uniquely determined by A.*

Proof. We prove the statements in (a) by induction on r. When $r = 1$, we have $A = 0$. The result is trivial in this case. Assume that the result holds for $r - 1$ and consider a nilpotent map of index r.

Given a subset S linearly independent of $N(A^{r-1})$, extend it a maximal such set $w_1, \ldots, w_k$. Then they along with any basis of $N(A^{r-1})$ is a basis of V.

We claim that the vectors Aw_i, $1 \le i \le k$, lie in $N(A^{r-1})$ and are linearly independent of $N(A^{r-2})$. For, if $a_1 A w_1 + \cdots + a_k A w_k \in N(A^{r-2})$, then, $a_1 w_1 + \cdots + a_k w_k \in N(A^{r-1})$. By hypothesis on w_i, it follows that $a_i = 0$ for each i.

Now, A restricted to $N(A^{r-1})$ is nilpotent of degree $r - 1$. By induction hypothesis, there are vectors $v_1, \ldots, v_m$ including Aw_i's such that the nonzero vectors of the form $A^j v_i$, $1 \le i \le m$ form a basis for $N(A^{r-1})$. We adjoin the vectors $w_1, \ldots, w_k$ to them.

The rest (b) and (c) follow easily. □

Example 11. Let A be nilpotent of degree 5 on a vector space of dimension 22. The bottom i rows are $N(A^i)$.

V_1	V_2	V_3	V_4	V_5	V_6
v_1	v_2				
Av_1	Av_2	v_3	v_4		
A^2v_1	A^2v_2	Av_3	Av_4	v_5	
A^3v_1	A^3v_2	A^2v_3	A^2v_4	Av_5	
A^4v_1	A^4v_1	A^3v_3	A^3v_4	A^2v_5	v_6

We let $J_m(\lambda)$ denote the Jordan block matrix

$$\begin{pmatrix} \lambda & 0 & 0 & 0 & \ldots & 0 \\ 1 & \lambda & 0 & 0 & \ldots & 0 \\ 0 & 1 & \lambda & 0 & \ldots & 0 \\ 0 & 0 & \ddots & \ddots & & \\ 0 & 0 & 0 & 1 & \lambda & 0 \\ 0 & 0 & 0 & 0 & 1 & \lambda \end{pmatrix}.$$

In the decomposition of the theorem, the matrix of A restricted to V_i with respect to the specified basis, is $J_{r_i}(0)$.

Theorem 12. *Let $T\colon V \to V$ be a linear map on a finite dimensional vector space over an algebraically closed field. Then there exists a Jordan basis for T on V so that the matrix representation A of T with respect to this basis is of the form*

$$A = \operatorname{diag}\left(J_{r_1}(\lambda_1), \ldots, J_{r_k}(\lambda_k)\right).$$

Proof. Immediate from Theorems 6 and 10. □

Remark 13. The above proof is valid as long as there exists a polynomial p such that $p(A) = 0$ and p can be factored into linear polynomials over the field of definition of the vector space.