Jordan Canonical Form

Suppose A is a $n \times n$ matrix operating on $V = \mathbf{C}^n$.

First Reduction (to a repeated single eigenvalue).

Let

$$\phi(x) = det(x - A) \ = \ \prod_{i=1}^{r} (x - \lambda_i)^{e_i} \tag{1}$$

be the characteristic equation of A. Factor $\phi(x)$ into relatively prime factors

$$\phi(x) \ = \ p(x)\, q(x) \tag{2}$$

(if possible). By the euclidean algorithm, there exist polynomials $a(x)$ and $b(x)$ so that

$$a(x)\, p(x) \ + \ b(x)\, q(x) \ = \ 1. \tag{3}$$

Consider the subspaces

$$V_p \ = \ p(A)V \qquad \text{and} \qquad V_q \ = \ q(A)V. \tag{4}$$

Note that

i) Both V_p and V_q are invariant under A since

$$AV_p = Ap(A)V = p(A)AV \subset V_p.$$

ii)

$$V \ = \ V_p \ + \ V_q$$

since by (3),

$$v = a(A)p(A)v \ + \ b(A)q(A)v \ = \ v_p \ + \ v_q. \tag{5}$$

iii)

$$q(A)V_p \ = 0 \ = \ p(A)V_q \tag{6}$$

since by Cayley-Hamilton, A satisfies its characteristic equation (2).

Consequently, the representation in (5) is unique. Moreover,

iv)

$$V_p \cap V_q \;=\; \{0\}. \tag{7}$$

We thus say V is the *direct sum* of V_p with V_q and write

$$V \;=\; V_p \;\oplus\; V_q.$$

v) Choosing bases for V_p and V_q, we may move to this new combined basis for V so that A is now represented by the block diagonal

$$P^{-1}AP \;=\; \begin{pmatrix} A_p & 0 \\ 0 & A_q \end{pmatrix}, \tag{8}$$

vi) where

$$\phi(x) \;= \phi_{A_p}(x)\, \phi_{A_q}(x), \tag{9}$$

and where

$$\phi_{A_p}(x) \;=\; q(x) \quad \text{and} \quad \phi_{A_q}(x) \;=\; p(x). \tag{10}$$

Relation (9) is clear from the block structure. But (10) is not at all clear: Suppose $\phi_{A_p}(\lambda_p) = 0$. Then A_p, hence A, possesses a corresponding eigenvector v_p in V_p. But then by (6),

$$q(A)v_p \;=\; 0 \;=\; \prod_{q(\lambda_q)=0} (A - \lambda_q)v_p \;=\; \prod_{q(\lambda_q)=0} (\lambda_p - \lambda_q) \cdot v_p,$$

hence some $\lambda_q = \lambda_p$. In short, the eigenvalues of A_p are roots of $q(x) = 0$. By reversing roles and recalling that $p(x)$ and $q(x)$ are relatively prime we obtain (10).

This realizes our first goal:

First Reduction. An $n \times n$ matrix A on $V = \mathbf{C}^n$ can be brought to block diagonal form

$$P^{-1}AP \;=\; \begin{pmatrix} A_1 & 0 & \dots & 0 \\ 0 & A_2 & 0 & \dots \\ \vdots & \dots & \ddots & 0 \\ 0 & \dots & 0 & A_r \end{pmatrix}, \tag{11a}$$

where each block has but one distinct eigenvalue, i.e.,

$$\phi_{A_i}(x) \;=\; (x - \lambda_i)^{e_i}. \tag{11b}$$

Second Reduction (to Jordan blocks).

Suppose A is $n \times n$ operator on $V = \mathbf{C}^n$ with characteristic polynomial

$$\phi(x) \;=\; det(x - A) \;=\; (x - \lambda)^n. \tag{12}$$

Let $m \leq n$ be the least integer so that

$$(A - \lambda)^m V \;=\; 0. \tag{13}$$

(The polynomial $\mu(x) = (x - \lambda)^m$ is called the *minimal polynomial* for A.) Find a vector v_0 of V of maximal cyclic order, i.e., where

$$(A - \lambda)^m v_0 \;=\; 0, \tag{14a}$$

$$(A - \lambda)^{m-1} v_0 \;\neq\; 0, \tag{14b}$$

but where m is largest possible.

Consider the invariant subspace of A given by

$$V_0 \;=\; span\{v_0, Av_0, A^2v_0, \dots, A^{m-1}v_0\}. \tag{15}$$

Note that $\{v_0, Av_0, A^2v_0, \dots, A^{m-1}v_0\}$ form an independant set, for if not, by the division algorithm we see $(A - \lambda)^{m-1}v_0 \;=\; 0$. (Alternatively, apply the operator $A - \lambda$ repeatedly to any relation among between this putative basis.)

Note also that this tower of 'pseudoeigenvectors' $v_i = (A-\lambda)^{i-1}v_0$ over the eigenvector v_{m-1} of A satisfies

$$Av_i = (A-\lambda)v_i + \lambda v_i = v_{i-1} + \lambda v_i. \tag{16}$$

Thus cutting A back to V_0 with this tower as basis yields the representation

$$J_i = \begin{pmatrix} \lambda & 0 & 0 & \cdots & & 0 \\ 1 & \lambda & 0 & \cdots & \cdots & 0 \\ 0 & 1 & \lambda & 0 & \cdots & \vdots \\ \vdots & 0 & 1 & \lambda & 0 & \cdots \\ \vdots & & & 1 & \ddots & \\ \vdots & & & 0 & & \\ \vdots & & & \vdots & & \\ 0 & \cdots & & 0 & 1 & \lambda \end{pmatrix}, \tag{17}$$

called a *Jordan block.*

It remains to show that either $V_0 = V$ or that V_0 *decomposes* V, i.e., that there exists a second invariant subspace V_1 with $V = V_0 \oplus V_1$. We can then continue with the above proceedure on V_1.

There are two cases: Either V_0 contains the eigenspace of λ or there exists an eigenvector u not in V_0. **In the first case,** the null space of $A-\lambda$ has rank 1, hence its range has rank $n-1$. In fact, repeatedly applying $A-\lambda$ yields the descending sequence of subspaces

$$V \supset (A-\lambda)V \supset (A-\lambda)^2V \supset \cdots \supset (A-\lambda)^{m-1}V \supset (A-\lambda)^mV = 0, \tag{18}$$

where at each step the rank can decrease by at most 1 since the eigenspace of A has rank 1. Thus $m = n$ and $V_0 = V$.

In the second case, suppose u is an eigenvector not in V_0 and let $U = span\{u\}$. Consider the *factor space* V/U, that is, consider the collection of residue classes (blocks) $v + U$. As we have seen before, these blocks form a disjoint covering of V. Moreover, these blocks inherit the operations of addition and scalar multiplication from their representatives making V/U into a vector space in its own right. The map

$$\pi : V \longrightarrow V/U$$

given by

$$\pi(v) \;=\; v + U$$

is a linear transformation with kernel U and hence V/U has dimension $n-1$. Note that A induces naturally an operator $\tilde{A}$ on the factor space $\tilde{V} = V/U$ by the rule $\tilde{A}\pi(v) = Av$.

Note also that because V_0 meets U only at 0, the image $\tilde{V_0}$ of V_0 under π is identical algebraically (isomorphic) and of the least dimension m such that $(\tilde{A} - \lambda)^m \tilde{V} = 0$, but it now lies in a vector space V/U of smaller dimension. Hence we may proceed by induction: The image $\pi(V_0)$ of V_0 under this algebraic projection π decomposes V/U into $\tilde{A}$-invariant subspaces,

$$V/U \;=\; \pi(V_0) \;\oplus\; \tilde{V_1}.$$

But then pulling back, we have the decomposition

$$V \;=\; V_0 \;\oplus\; V_1$$

into invariant subspaces where $V_1 = \pi^{-1}(\tilde{V_1})$.

Jordan Canonical Form.[1] An operator A on $V = \mathbf{C}^n$ can be brought to the form

$$P^{-1}AP \;=\; \begin{pmatrix} J_1 & 0 & \dots & 0 \\ 0 & J_2 & 0 & \dots \\ \vdots & \dots & \ddots & 0 \\ 0 & \dots & 0 & J_r \end{pmatrix}, \tag{18}$$

where each block J_i is a Jordan block (17).

[1]The above proof scheme first appeared in the *American Mathematical Monthly,* Vol. 91, No. 1, January 1984 — “A simple proof of the fundamental theorem on finite abelian groups,” (C. R. MacCluer).

Suggestion. When the space is factored into these cyclic invariant subspaces

$$V = V_1 \oplus V_2 \oplus \cdots \oplus V_s,$$

renumber so that the minimal polynomials of each subspace are nonincreasing in degree.

So if say A is 6×6, then the possible factorizations of V into the direct sum of invariant subspaces would have minimal polynomials of one of the following patterns:

$$(x-\lambda)^6,$$
$$(x-\lambda)^5,\ x-\lambda,$$
$$(x-\lambda)^4,\ x-\lambda,\ x-\lambda,$$
$$(x-\lambda)^4,\ (x-\lambda)^2,$$
$$(x-\lambda)^3,\ x-\lambda,\ x-\lambda,\ x-\lambda,$$
$$(x-\lambda)^3,\ (x-\lambda)^2,\ x-\lambda,$$
$$(x-\lambda)^3,\ (x-\lambda)^3,$$

and so on, each with a distinct Jordan form.