Jordan Normal Form

Linden Disney-Hogg

June 2020

1 Introduction

The purpose of this document is to ask about the density of the Jordan normal form of matrices in the space of matrices. This is motivated by the following result:

Theorem 1.1. *The subset of diagonalisable matrices is dense in $M_n(\mathbb{C})$.*

Before proving this, we note that I have not told you what topology to take on $M_n(\mathbb{C})$. We take the metric topology induced by having a norm $||\cdot||$. That this is well defined requires a bit of theory[1]

Definition 1.2. *Let V be a real or complex vector space. Two norms $||\cdot||_a, ||\cdot||_b$ are* ***equivalent*** *if $\exists C_1, C_2 \in \mathbb{R}_{>0}$ s.t.*

$$\forall v \in V, \begin{cases} ||v||_a \leq C_1 ||v||_b \\ ||v||_b \leq C_2 ||v||_a \end{cases}$$

Lemma 1.3. *Equivalence of norms is an equivalence relation.*

Lemma 1.4. *Let V be a f.d. vector space. Any norm is continuous with respect to the l_1-norm*

Proof. Taking a basis $\{e_i\}_{i=1}^n$. We can calculate

$$\begin{aligned} \forall v \in V,\, ||v|| &= ||v_1 e_1 + \cdots + v_n e_n|| \\ &\leq ||e_1|| \, |v_1| + \cdots + ||e_n|| \, |v_n| \\ &\leq C ||v||_1 \text{ where } C = \max ||e_i|| \\ \Rightarrow |||v|| - ||w||| &\leq ||v - w|| \leq C ||v - w||_1 \end{aligned}$$

□

This has the corollary

Theorem 1.5. *If V is a real or complex finite-dimensional vector space, then any two norms on V are equivalent*

[1] I will assume that you know what a norm is, what the l_p-norm is, and what the triangle inequality is.

Proof. It is sufficient to show that any norm is equivalent to l_1. Consider $S_1 = \{v \in V \mid ||v||_1 = 1\} \subset V$ and the image of the norm $||\cdot||$ acting on S_1. As S_1 is compact and $||\cdot||$ is continuous we must have that the image is comapct in $\mathbb{R}$, so has a maximmum and minimun obtained C_1, C_2. Hence

$$\begin{aligned}\forall v \in V,\, C_2 \leq \left|\left|\frac{v}{||v||_1}\right|\right| &\leq C_1 \\ \Rightarrow C_2 \, ||v||_1 \leq ||v|| &\leq C_1 \, ||v||_1\end{aligned}$$

and so we are done. □

Now we can get to the proof of the theorem we care about

Proof of Theorem 1.1. The theorem is trivial is $n = 1$ so we assume $n \geq 2$.
Let us call X the subset of diagonalisable matrices. Note we have $Y \subset X$ where Y is the set of matrices with distinct eigenvalues. It will suffice to show $Y \subset M_n(\mathbb{C})$ is dense.
We now construct the map

$$\begin{aligned}p : M_n(\mathbb{C}) &\to P_n, \\ A &\mapsto p_A\end{aligned}$$

from matrices into polynomials of degree n where p_A is the characteristic polynomial defined by

$$p_A(t) = \det(tI - A)$$

This map is continuous (viewing P_n inside the vector space of polynomials of degree $\leq n$, and inheriting the metric). We also have the continuous map $\Delta : P_n \to \mathbb{R}$ given by the discriminant. Moreover, the composed map $\Delta \circ p : M_n(\mathbb{C}) \to \mathbb{R}$ is a non-constant polynomial in the components, and so cannot be locally constant. Recall that the discriminant of a polynomial is 0 iff it has a repeated root, and a repeated root in a characteristic polynomials corresponds to a an eigenvalue with algebraic multiplicity > 1 of the original matrix. Hence we have

$$\forall A \in M_n(\mathbb{C}) \setminus Y,\, (\Delta \circ p)(A) = 0$$

and if $U \ni A$ is an open neighbourhood

$$\exists d \in (\Delta \circ p)(U),\, d \neq 0$$

We thus know any neighbourhood of A intersects Y non-trivially, and so Y is dense. □

2 Jordan normal form

We now want to see if we can extend this idea to ask about when we know an eigenvalue has algebraic multiplicity ≥ 1, whether there is a subset of these matrices that is dense in the matrices with a given multiplicity.

Theorem 2.1. *The Jordan normal form of a matrix exists and is unique up to reordering of jordan blocks.*

We can define an equivalence relation on Jordan form matrices if they are the same up to a reordering of block. We can then fix notation to let $J : M_n(\mathbb{C}) \to M_n(\mathbb{C})$ be the map that sends a matrix to it's Jordan normal form equivalence class. That such a map exists and is well defined is a result of the above theorem. We can then biject between the image of J and the cosets of $M_n(\mathbb{C})$ quotiented by $GL_n(\mathbb{C})$ acting by conjugation.
We will also fix notation as such:

Definition 2.2. *The* ***Jordan block of size m*** *for $m \geq 2$ is the $m \times m$ matrix given by*

$$J_m(\lambda) = \lambda I_{m\times m} + E_{12} + \cdots + E_{m-1,m}$$

where E_{ij} is the standard basis of $M_m(\mathbb{R})$

A definition we will also want that will be useful is

Definition 2.3. *The* ***centraliser*** *of a matrix $A \in M_n(\mathbb{C})$ is*

$$C(A) = \{B \in M_n(\mathbb{C}) \mid [A, B] = 0\}$$

Lemma 2.4. *$B \in C(A) \cap GL_n(\mathbb{C}) \Leftrightarrow B^{-1}AB = A$*

Lemma 2.5. *$C(A) \cap GL_n(\mathbb{C}) \leq GL_n(\mathbb{C})$*

I will often consider the restriction of the centraliser to the invertible elements, as this will be important from here on in with regards to Jordan normal form.

2.1 $n = 2$

Let us start with an easy question we can make traction on, the case of 2×2 matrices. We can split the Jordan normal form of such matrices into three groups

A	p_A	m_A
$\left(\begin{smallmatrix}\lambda & 0\\ 0 & \lambda\end{smallmatrix}\right)$	$(t-\lambda)^2$	$(t-\lambda)$
$\left(\begin{smallmatrix}\lambda & 1\\ 0 & \lambda\end{smallmatrix}\right)$	$(t-\lambda)^2$	$(t-\lambda)^2$
$\left(\begin{smallmatrix}\lambda & 0\\ 0 & \mu\end{smallmatrix}\right)$	$(t-\lambda)(t-\mu)$	$(t-\lambda)(t-\mu)$

Here we have also shown the characteristic polynomial p_A and the minimal polynomial m_A. We see what we know to be true, that the characteristic polynomial is not sufficient to distinguish between all cases of Jordan normal form

Remark. *In the 2×2 case the minimal polynomial along is sufficient to characterise, that is there is a bijection between $m_A \leftrightarrow J(A)$. In the 3×3 case there is a bijection $(m_A, p_A) \leftrightarrow J(A)$. In higher cases you need more information that the minimal and characteristic polynomial.*

In this case we see that, if we fix the diagonal to be λ, λ, then there is only one matrix with the Jordan normal form $\left(\begin{smallmatrix}\lambda & 0\\ 0 & \lambda\end{smallmatrix}\right)$ (namely $\left(\begin{smallmatrix}\lambda & 0\\ 0 & \lambda\end{smallmatrix}\right)$ itself, as it is invariant under conjugation). Contrastingly, the only matrices which preserve $\left(\begin{smallmatrix}\lambda & 1\\ 0 & \lambda\end{smallmatrix}\right)$ under conjugation are those that commute with it. Let's calculate these conditions:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix} = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

$$\Rightarrow \begin{pmatrix} \lambda a & a + \lambda b \\ \lambda c & c + \lambda d \end{pmatrix} = \begin{pmatrix} c + \lambda a & d + \lambda b \\ \lambda c & \lambda d \end{pmatrix}$$

$$\Rightarrow c = 0,\, d = a$$

The condition that the matrix $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ is invertible then gives

$$C\left(\left(\begin{smallmatrix} \lambda & 1 \\ 0 & \lambda \end{smallmatrix}\right)\right) = \{\left(\begin{smallmatrix} a & b \\ 0 & a \end{smallmatrix}\right) \,|\, a, b \in \mathbb{C},\, a \neq 0\}$$

Note here that the centraliser of the Jordan block is independent of the value on the diagonal. This is in fact general

Proposition 2.6. $C(J_m(\lambda)) = C(J_m(0))$

In order to think about matrices with a fixed set of eigenvalues, we need to understand isospectral deformations.

3 What Harry Says

3.0.1 Some Linear Algebra

Its useful to gather some properties of commutators of matrices in Jordan form. Let $\tilde{L} = \mathrm{diag}(J_{\mu_1}(\lambda_1), \ldots, J_{\mu_t}(\lambda_t))$ be written in Jordan form. Here $J_{\mu_i}(\lambda_1)$ is a $\mu_i \times \mu_i$ Jordan block with diagonal λ_i and 1's just above the diagonal and we are not assuming the λ_i distinct. In considering the matrix $[\tilde{L}, \tilde{M}]$ we will write $\tilde{M}$ as $\mu_i \times \mu_j$ blocks and denote by M' the i-th diagonal block of $\tilde{M}$. We wish to solve

$$[\tilde{L}, \tilde{M}] = 0. \tag{3.0.1}$$

First, we see for the i-th diagonal block $[J_{\mu_1}(\lambda_1), M'] = 0$. Solving this we see M' is upper triangular with constant value along the $M'_{i,i+k}$ diagonals ($k = 0, \ldots, n-1$); for example with $\mu_i = 3$,

$$M' = \begin{pmatrix} x & y & z \\ 0 & x & y \\ 0 & 0 & x \end{pmatrix}.$$

For the off-diagonal blocks of $\tilde{M}$ consider $J_{\mu_i}(\lambda_i)B - BJ_{\mu_j}(\lambda_j) = 0$, where B is a $\mu_i \times \mu_j$ matrix. There are two cases. If $\lambda_i \neq \lambda_j$ then starting with the μ_i-th row and 1-st column we find $B_{\mu_i 1}(\lambda_i - \lambda_j) = 0$ and so $B_{\mu_i 1} = 0$; recursively substituting we find $B = 0$. If $\lambda_i = \lambda_j$ we get a block of forms

$$\begin{pmatrix} 0 & 0 & x & y & z \\ 0 & 0 & 0 & x & y \\ 0 & 0 & 0 & 0 & x \end{pmatrix}, \quad \begin{pmatrix} x & y & z \\ 0 & x & y \\ 0 & 0 & x \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

depending on whether $\mu_i < \mu_j$ or $\mu_j < \mu_i$ respectively. In either case we find this space has dimension $\min(\mu_i, \mu_j)$. Note that when we have the m Jordan blocks of size 1 associated with the same eigenvalue these combine to give a $m \times m$ matrix.
Example:

$$
\begin{pmatrix}
\lambda & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & \lambda & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & \lambda & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & \lambda & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & \lambda & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & \lambda & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & \lambda & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \lambda & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \lambda
\end{pmatrix}, \qquad
\begin{pmatrix}
m_{1,1} & m_{1,2} & 0 & 0 & a & 0 & 0 & 0 & e \\
m_{2,1} & m_{2,2} & 0 & 0 & b & 0 & 0 & 0 & f \\
c & d & x_1 & x_2 & x_3 & 0 & \alpha & \beta & \gamma \\
0 & 0 & 0 & x_1 & x_2 & 0 & 0 & \alpha & \beta \\
0 & 0 & 0 & 0 & x_1 & 0 & 0 & 0 & \alpha \\
g & h & \rho & \sigma & \tau & y_1 & y_2 & y_3 & y_4 \\
0 & 0 & 0 & \rho & \sigma & 0 & y_1 & y_2 & y_3 \\
0 & 0 & 0 & 0 & \rho & 0 & 0 & y_1 & y_2 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & y_1
\end{pmatrix}.
$$

First recall that a matrix A is similar to the matrix J if and only if there exists an invertible matrix T such that $A = TJT^{-1}$. Then if $A = TJT^{-1} = T_1JT_1^{-1}$ we have $S := T_1^{-1}T$ satisfies $SJ = JS$, giving

Lemma 3.1. *If $A = TJT^{-1}$ then $A = T_1JT_1^{-1}$ if and only if $S := T_1^{-1}T$ and*

$$SJ = JS. \tag{3.0.2}$$

Further,

$$\dim\{A \,|\, A = TJT^{-1}\} = \dim\{T \,|\, TT^{-1} = 1\} - \dim\{S \,|\, SS^{-1} = 1,\ SJ = JS\}. \tag{3.0.3}$$

This may be used to calculate the dimensions of various standard classes of matrices.

Let us work over $\mathbb{C}$ for simplicity. Here the dimension of the space of invertible matrices is $n^2 = \dim\{T \,|\, T \in M_n(\mathbb{C}),\ TT^{-1} = 1\}$ and the dimension of the set of matrices with distinct eigenvalues of specified multiplicities μ_i $(i = 1, \dots, t)$ has dimension t.

Lemma 3.2. *The set of diagonalizable matrices in $M_n(\mathbb{C})$ with (specified) distinct eigenvalues $(\lambda_1, \dots, \lambda_t)$ of multiplicities μ_i $(i = 1, \dots, t)$ has (complex) dimension*

$$n^2 - \sum_{i=1}^{t} \mu_i^2.$$

The dimension of the set of matrices with distinct eigenvalues of specified multiplicities μ_i $(i = 1, \dots, t)$ has (complex) dimension

$$n^2 - \sum_{i=1}^{t} (\mu_i^2 - 1).$$

We have seen when discussing matrices that commute with a matrices in Jordan form that if B is a $\mu_i \times \mu_j$ matrix such that $J_{\mu_i}(\lambda_i)B - BJ_{\mu_j}(\lambda_j) = 0$, then $B = 0$ if $\lambda_i \neq \lambda_j$. Thus when counting the dimension of matrices commuting with a matrix in Jordan form we can restrict attention to those matrices with a common eigenvalue, and then sum over the distinct eigenvalues. Let

$$J = \operatorname{diag}(J_{\mu_1}(\lambda), \dots, J_{\mu_k}(\lambda)), \quad \mu_1 \geq \mu_2 \geq \dots \geq \mu_k,$$

have k Jordan blocks with common eigenvalue λ. Then we have

$$\dim\{S \mid SS^{-1} = 1,\ SJ = JS\} = \sum_{i,j=1}^{k} \min(\mu_i, \mu_j) = \sum_{j=1}^{k} (2j-1)\mu_j.$$

Now suppose that a matrix has Jordan form J with p distinct eigenvalues λ_a of multiplicities n_a $(a = 1, \ldots, p)$ each of which is formed of formed from N_a blocks of size μ_{aj},

$$J = \mathrm{diag}(J_{\mu_{a1}}(\lambda_a), \ldots, J_{\mu_{aN_a}}(\lambda_a)), \quad \mu_{a1} \geq \mu_{a2} \geq \ldots \geq \mu_{aN_a}, \quad \sum_{i=1}^{N_a} \mu_{ai} = n_a. \tag{3.0.4}$$

Set $N^* = \max_a n_a$ and $\mu_{al} = 0$ for $l > N_a$. Then

$$\dim\{S \mid SS^{-1} = 1,\ SJ = JS\} = \sum_{a=1}^{p} \sum_{j=1}^{N_a} (2j-1)\mu_{aj} = \sum_{j=1}^{N^*} (2j-1) \sum_{a=1}^{p} \mu_{aj}$$

Lemma 3.3. *The set of complex $n \times n$ matrices A with Jordan form (3.0.4) and prescribed (distinct) eigenvalues λ_a is*

$$\dim\{A \mid p, \lambda_a N_a, \mu_{aj}\} = n^2 - \sum_{a=1}^{p} \sum_{j=1}^{N_a} (2j-1)\mu_{aj} = n^2 - \sum_{j=1}^{N^*} (2j-1) \sum_{a=1}^{p} \mu_{aj}. \tag{3.0.5}$$

The complex dimension of matrices A with these properties and any (distinct) eigenvalues λ_a is

$$\dim\{A \mid p, N_a, \mu_{aj}\} = n^2 - \sum_{a=1}^{p} \sum_{j=1}^{N_a} (2j-1)\mu_{aj} + p = n^2 - \sum_{j=1}^{N^*} (2j-1) \sum_{a=1}^{p} \mu_{aj} + p. \tag{3.0.6}$$

(a) When all the eigenvalues of A are simple and distinct then $p = n$, $n_a = N_a = \mu_{1j} = 1$,

$$\dim\{A \mid p = n,\ n_1 = N_1 = 1,\ \mu_{11} = 1\} = n^2,$$

(b) When A has one Jordan block with n-fold eigenvalue then $p = 1$, $n_1 = n$, $N_1 = 1$, $\mu_{11} = n$,

$$\dim\{A \mid p = 1,\ n_1 = n,\ N_1 = 1,\ \mu_{11} = n\} = n^2 - n + 1,$$

(c) When A has one Jordan block with n-fold eigenvalue and n blocks then $p = 1$, $n_1 = n$, $N_1 = n$, $\mu_{11} = 1$,

$$\dim\{A \mid p = 1,\ n_1 = n,\ N_1 = 1,\ \mu_{11} = 1\} = 1.$$

(d) When A has $n-1$ distinct eigenvalues and one nontrivial Jordan block with 2-fold eigenvalue then $p = n-1$, $n_1 = 2$, $N_1 = 1$, $\mu_{11} = 2$, $1 = n_j = N_j = \mu_{j1}$ $(j > 1)$,

$$\dim\{A \mid p = n-1,\ n_1 = 2,\ N_1 = 1,\ \mu_{11} = 2,\ 1 = n_j = N_j = \mu_{j1}\ (j > 1)\} = n^2 - 1,$$

(e) When A is diagonalizable with $n-1$ distinct eigenvalues $p = n-1$, $n_1 = 2$, $N_1 = 2$, $\mu_{11} = \mu_{12} = 1$, $1 = n_j = N_j = \mu_{j1}$ $(j > 1)$,

$$\dim\{A \mid p = n-1,\ n_1 = 2,\ N_1 = 2,\ \mu_{11} = \mu_{12} = 1,\ 1 = n_j = N_j = \mu_{j1}\ (j > 1)\} = n^2 - 3.$$

We see that when two eigenvalues coincide that the generic situation is for a Jordan block of codimension 1; the diagonalizable case is of codimension 3.

Example: Consider the matrix

$$\begin{pmatrix} 0 & 1-\epsilon^2 \\ -b^2 & 2b \end{pmatrix} = \Lambda . \begin{pmatrix} b+\epsilon b & 0 \\ 0 & b-\epsilon b \end{pmatrix} \Lambda^{-1}, \quad \Lambda = \begin{pmatrix} 1-\epsilon & 1+\epsilon \\ b & b \end{pmatrix}.$$

As $\epsilon \to 0$ this has Jordan form

$$\begin{pmatrix} b & 1 \\ 0 & b \end{pmatrix}.$$

Example: For 3×3 matrices we have for distinct λ_i that

(i) $\dim \operatorname{diag}(\lambda_1, \lambda_2, \lambda_3) = 9$,

(ii) $\dim \operatorname{diag}(J_2(\lambda_1), \lambda_2) = 8$,

(iii) $\dim \operatorname{diag}(J_3(\lambda_1)) = 7$,

(iv) $\dim \operatorname{diag}(\lambda_1, \lambda_1, \lambda_2) = 6$,

(v) $\dim \operatorname{diag}(\lambda_1, \lambda_1, \lambda_1) = 1$.

References