Notes on Jordan Form
Northwestern University, Summer 2015

These notes are meant to clarify the notion of a Jordan form which the book talks about in the last section of Chapter 8. In particular, we give the actual definition of a Jordan form, which the book never quite gets to explicitly, although they do it in disguise!. We'll also talk about how to explicitly compute Jordan forms.

Throughout we work with a finite-dimensional complex vector space V, even though some of these concepts make sense for real vector spaces as well. (Essentially, whenever we have a real operator which has as many real eigenvalues as possible—meaning has $\dim V$ real eigenvalues when counted with multiplicities—what we say here applies as well.)

The Jordan Form of an Operator

Definition. A *Jordan block* of size k is a $k \times k$ matrix of the form

$$\begin{pmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ & & & & \lambda \end{pmatrix}$$

where the missing entries are all zero. In other words, a Jordan block is almost a multiple of the identity, except for 1's above the main diagonal.

Definition. A square matrix is said to be in *Jordan form* if it is block diagonal where each block is a Jordan block.

This is precisely the type of matrix described in the book near the top of page 186. With this terminology, Theorem 8.47 can then be restated as follows:

Theorem. *Any operator T on V can be represented by a matrix in Jordan form. This matrix is unique up to a rearrangement of the order of the Jordan blocks, and is called the Jordan form of T.*

A basis of V which puts $\mathcal{M}(T)$ in Jordan form is called a *Jordan basis* for T. This last section of Chapter 8 is all about proving the above theorem. The key step is Lemma 8.40, which shows that Jordan bases exist for nilpotent operators. The above theorem then follows by applying the lemma to the nilpotent operators which are the restrictions of $T - \lambda_i I$ to the generalized eigenspaces of T corresponding to the eigenvalues λ_i.

As you can see when reading this last section, the proof of this lemma is quite difficult. Indeed, most people probably never actually see a proof, which is okay because the most important thing is the existence of the Jordan form and what it tells you about the operator. This is what we will clarify in these notes.

If v is a nonzero generalized eigenvector of T corresponding to the eigenvalue λ, then there exists a smallest positive integer k so that

$$(T - \lambda I)^k v = 0.$$

The list $(v, (T - \lambda I)v, \ldots, (T - \lambda I)^{k-1}v)$ is linearly independent, as an argument similar to the one for Problem 4 on the midterm shows. (Check the midterm solutions to see this.)

Definition. For a nonzero generalized eigenvector v of T corresponding to an eigenvalue λ, the list

$$(v, (T - \lambda I)v, \ldots, (T - \lambda I)^{k-1}v),$$

where k is defined as above, is called the *Jordan chain* corresponding to v.

Note that for such a Jordan chain, the final vector $(T - \lambda I)^{k-1}v$ in the chain is an oridnary eigenvector of T since it is nonzero and applying $T - \lambda I$ to it gives zero.

A Jordan basis is then exactly a basis of V which is composed of Jordan chains. Lemma 8.40 (in particular part (a)) says that such a basis exists for nilpotent operators, which then implies that such a basis exists for any T as in Theorem 8.47. Each Jordan block in the Jordan form of T corresponds to exactly one such Jordan chain. Indeed, the point of Theorem 8.47 is that the matrix of an operator corresponding to a specific Jordan chain written in reverse order

$$((T - \lambda I)^{k-1}v, \ldots, (T - \lambda I)v, v)$$

is a Jordan block. This is how we get a matrix $\mathcal{M}(T)$ which is block diagonal and where each block is a Jordan block.

Let λ be a specific eigenvalue of T. Then we know there is a basis for the generalized eigenspace of T corresponding to λ consisting of Jordan chains. How many such Jordan chains are there? The final vectors in these Jordan chains are eigenvectors corresponding to λ, which in fact give a basis for the (orgindary) eigenspace corresponding to λ. (This is essentially what part (b) of Lemma 8.40 says.) In other words:

Proposition. *The number of Jordan blocks in the Jordan form of T corresponding to λ (or equivalently the number of Jordan chains corresponding to λ in a Jordan basis of V) is the dimension of the (usual) eigenspace of T corresponding to λ.*

The problem we are interested in now is finding the Jordan form of an operator. The characteristic polynomial tells us how many times a certain eigenvalue will appear in this Jordan form, and the dimension of each eigenspace tells us how many Jordan blocks there will be for a specific eigenvalue. What about the size of the blocks themselves? The important fact here is the following:

Proposition. *The size of the largest Jordan block corresponding to an eigenvalue λ of T is exactly the degree of the $(z - \lambda)$ term in the minimal polynomial of T.*

We have essentially already seen this in class: to kill off a Jordan block of size k, say of the form given in the first definition in these notes, we need the polynomial $(z - \lambda)^k$. Any smaller Jordan block with the same eigenvalue will also be killed off under this polynomial, and this implies the result of the proposition.

Let us summarize what we have learned so far about the relation between the Jordan form of an operator, its characteristic polynomial, and its minimal polynomial:

- the characteristic polynomial tells us the eigenvalues and the dimension of each generalized eigenspace, which is the number of times the eigenvalue appears along the diagonal of the Jordan form (also known as the "multiplicity" of λ),
- the dimension of each eigenspace tells us how many Jordan blocks corresponding to that eigenvalue there are in the Jordan form, and
- the exponents of the different terms in the minimal polynomial tell us the sizes of the largest Jordan blocks corresponding to each eigenvalue.

In the best situation, these facts give us all the information we need to explicitly find the Jordan form of an operator. If nothing else, these facts narrow down the possible Jordan forms an operator can have, and then we only need to do a bit of further computation to determine which of these possibilities is correct.

Examples of Jordan Forms

Example 1. For a first example, we determine the Jordan form of the operator represented by the matrix

$$A = \begin{pmatrix} 1 & -1 & 0 \\ -1 & 4 & -1 \\ -4 & 13 & -3 \end{pmatrix}.$$

We can view this as either an operator on $\mathbb{C}^3$ or $\mathbb{R}^3$, but it doesn't matter in this case since A has 3 real eigenvalues when counted with multiplicities, so everything we said before for operators on complex vector spaces will work here even when viewing A as an operator on $\mathbb{R}^3$.

The characteristic polynomial of A is

$$z(z-1)^2,$$

which can be found by computing $\det(A - zI)$. (As I said in class, I would likely just give you this polynomial in order to save time.) Thus, A has eigenvalues 0 with multiplicity 1 and 2 with multiplicity 2. Hence, the generalized eigenspace corresponding to 0 with just the ordinary eigenspace, so there will only be a single Jordan block corresponding to 0 in the Jordan form of A. Moreover, this block has size 1 since 1 is the exponent of z in the characteristic (and hence in the minimial as well) polynomial of A.

The only thing left to determine is the number of Jordan blocks corresponding to 1 and their sizes. We determine the dimension of the eigenspace corresponding to 1, which is the dimension of the null space of

$$A - I = \begin{pmatrix} 0 & -1 & 0 \\ -1 & 3 & -1 \\ -4 & 13 & -4 \end{pmatrix}.$$

Row-reducing gives

$$\begin{pmatrix} 0 & -1 & 0 \\ -1 & 3 & -1 \\ -4 & 13 & -4 \end{pmatrix} \to \begin{pmatrix} -1 & 3 & -1 \\ 0 & -1 & 0 \\ 0 & 1 & 0 \end{pmatrix} \to \begin{pmatrix} -1 & 3 & -1 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

which has a 1-dimensional kernel. Thus the dimension of the eigenspace corresponding to 1 is 1, meaning that there is only one Jordan block corresponding to 1 in the Jordan form of A. Since 1 must appear twice along the diagonal in the Jordan form, this single block must be of size 2. Thus the Jordan form of A is

$$\begin{pmatrix} \color{red}{0} & 0 & 0 \\ 0 & \color{blue}{1} & \color{blue}{1} \\ 0 & \color{blue}{0} & \color{blue}{1} \end{pmatrix},$$

where the colors highlight the two Jordan blocks. In this case the minimal polynomial is the same as the characteristic polynomial.

Example 2. Next we determine the Jordan form of

$$B = \begin{pmatrix} 5 & -1 & 0 & 0 \\ 9 & -1 & 0 & 0 \\ 0 & 0 & 7 & -2 \\ 0 & 0 & 12 & -3 \end{pmatrix}.$$

This has characteristic polynomial

$$(z-2)^2(z-3)(z-1),$$

so since all eigenvalues are real it again doesn't matter if we consider this to be an operator on $\mathbb{R}^4$ or $\mathbb{C}^4$. From the multiplicities we see that the generalized eigenspaces corresponding to 3 and to 1 are the orindary eigenspaces, so each of these give blocks of size 1 in the Jordan form.

The eigenspace corresponding to 2 is the null space of

$$B - 2I = \begin{pmatrix} 3 & -1 & 0 & 0 \\ 9 & -3 & 0 & 0 \\ 0 & 0 & 5 & -2 \\ 0 & 0 & 12 & -5 \end{pmatrix},$$

which row-reduces to

$$\begin{pmatrix} 3 & -1 & 0 & 0 \\ 9 & -3 & 0 & 0 \\ 0 & 0 & 5 & -2 \\ 0 & 0 & 12 & -5 \end{pmatrix} \to \begin{pmatrix} 3 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 5 & -2 \\ 0 & 0 & 0 & -1 \end{pmatrix}.$$

This has a 1-dimensional null space, so the eigenspace corresponding to 2 has dimension 1. Thus there is only one Jordan block corresponding to 2 in the Jordan form, so again it must be of size 2 since 2 has multiplicity 2. The Jordan form of B is thus:

$$\begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

and again the minimal polynomial of B is the same as the characteristic polynomial.

But let's go one step further in this case, and actually find a *Jordan basis* which puts B into this form. Recall that this should be a basis consisting of *Jordan chains*, where each chain corresponds to one Jordan block. For the blocks of size 1, the chains will be of length 1 and will each consist of a single eigenvector for the corresponding eigenvalue. You can check that

$$\begin{pmatrix} 0 \\ 0 \\ -1 \\ 2 \end{pmatrix} \text{ is an eigenvector of } B \text{ with eigenvalue 3 and}$$

$$\begin{pmatrix} 0 \\ 0 \\ -1 \\ 3 \end{pmatrix} \text{ is an eigenvector of } B \text{ with eigenvalue 1.}$$

These give one Jordan chain each. Going back to the row-reduction we did above when finding the dimension of the eigenspace corresponding to 2, we can compute that a basis for the eigenspace corresponding to 2 is given by

$$\begin{pmatrix}1\\3\\0\\0\end{pmatrix}.$$

The final Jordan chain we are looking for (there are only three Jordan chains since there are only three Jordan blocks in the Jordan form of B) must come from this eigenvector, and must be of the form

$$(v, (B-2I)v)$$

since the length has to be the size of the corresponding Jordan block. The final term here should be the ordinary eigenvector we found above, so we must "backtrack" and find a vector v such that

$$(B-2I)v = \begin{pmatrix}1\\3\\0\\0\end{pmatrix}.$$

Solving

$$\begin{pmatrix}3 & -1 & 0 & 0\\ 9 & -3 & 0 & 0\\ 0 & 0 & 5 & -2\\ 0 & 0 & 12 & -5\end{pmatrix} v = \begin{pmatrix}1\\3\\0\\0\end{pmatrix}$$

yields

$$v = \begin{pmatrix}1\\2\\0\\0\end{pmatrix}$$

as one possible solution. Thus

$$(v, (B-2I)v) = \left(\begin{pmatrix}1\\2\\0\\0\end{pmatrix}, \begin{pmatrix}1\\3\\0\\0\end{pmatrix}\right)$$

is a Jordan chain corresponding to the size 2 Jordan block in the Jordan form of B. Hence

$$\begin{pmatrix}1\\3\\0\\0\end{pmatrix}, \begin{pmatrix}1\\2\\0\\0\end{pmatrix}, \begin{pmatrix}0\\0\\-1\\2\end{pmatrix}, \begin{pmatrix}0\\0\\-1\\3\end{pmatrix}$$

is a Jordan basis corresponding to B, meaning that relative to this basis of $\mathbb{R}^4$ (or $\mathbb{C}^4$) the matrix of B is the Jordan form determined above.

Example 3. Finally we determine the Jordan form of

$$C = \begin{pmatrix}1 & -1 & -2 & 3\\ 0 & 0 & -2 & 3\\ 0 & 1 & 1 & -1\\ 0 & 0 & -1 & 2\end{pmatrix}.$$

This has characteristic polynomial

$$(z-1)^4,$$

so we can't immediately say anything about the Jordan form except for the fact that it can only have 1's down the diagonal, since this is the only eigenvalue of C. Next, we determine the dimension of the eigenspace corresponding to 1: $C - I$ reduces as

$$\begin{pmatrix} 0 & -1 & -2 & 3 \\ 0 & -1 & -2 & 3 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & -1 & 1 \end{pmatrix} \to \begin{pmatrix} 0 & -1 & -2 & 3 \\ 0 & 0 & -2 & 2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

which implies that this eigenspace is 2-dimensional. Hence there are two Jordan blocks corresponding to the eigenvalue 1 in the Jordan form. However, this alone does not give us enough information to fully determine the Jordan form since we could have two blocks of size 2 or one block of size 3 and the other of size 1; i.e. the Jordan form of C is either

$$\begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \text{ or } \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Correspondingly, the minimal polynomial of C is either $(z-1)^2$ in the first case or $(z-1)^3$ in the second.

To determine which it is, we must determine the lengths of the Jordan chains. We start with ordinary eigenvectors: a possible basis for the eigenspace corresponding to 1 is

$$\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix},$$

which we get by finding a basis for the null space of $C - I$ using the row-reduction above. Each of these will give rise to a Jordan chain. We start by trying to find v such that

$$(C-I)v = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

However, when attempting to solve this system of equations you end up with no solution, meaning that there is no such v. Thus, this specific eigenvector cannot occur at the end of a Jordan chain of length greater than 1, so it is its own Jordan chain. For the second eigenvector, we look for w such that

$$(C-I)w = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}.$$

Solving this system gives

$$w = \begin{pmatrix} 1 \\ 1 \\ -1 \\ 0 \end{pmatrix}$$

as one possible solution. This gives us the "previous" vector in a Jordan chain which has $\left(\begin{smallmatrix}1\\1\\1\\1\end{smallmatrix}\right)$ at the end, so this is a Jordan chain of length at least 2. To see if this is a Jordan chain of larger length, we next try to find a vector u such that

$$(C-I)u = \begin{pmatrix}1\\1\\-1\\0\end{pmatrix}.$$

Solving this gives

$$u = \begin{pmatrix}1\\-1\\0\\0\end{pmatrix}$$

as a solution, meaning we have found a Jordan chain of length at least 3:

$$(u, (C-I)u, (C-I)^2u) = \left(\begin{pmatrix}1\\-1\\0\\0\end{pmatrix}, \begin{pmatrix}1\\1\\-1\\0\end{pmatrix}, \begin{pmatrix}1\\1\\1\\1\end{pmatrix}\right).$$

We can stop here since we know there is not going to be a Jordan chain of length 4 since the Jordan form does not have a Jordan block of size 4. Thus since we have found a Jordan chain of length 3 and one of length 1, the Jordan form of C must have a Jordan block of size 3 and one of size 1, so it is

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

A Jordan basis is obtained by putting together the Jordan chain of length 3 together with the chain of length 1, and we can now see that this matrix has minimal polynomial $(z-1)^3$, where the exponent 3 corresponds to the size of the largest Jordan block.

You should notice that explicitly finding Jordan chains can be a lot of work, since it involves solving various linear equations. However, if all we are interested in is finding the Jordan form and not the explicit Jordan basis which gives this form, there is another way to proceed. Using the same example as above, we start with the fact that $\dim \operatorname{null}(C-I) = 2$, which is what tells us that there are two Jordan blocks and two Jordan chains corresponding to the eigenvalue 1. Let us denote the two eigenvectors we get from this by dots:

$$\begin{array}{c}\operatorname{null}(C-I)\\ \hline \bullet\\ \bullet\end{array}$$

Each of these dots represents the start of one Jordan chain.

The key observation is that the next vector in either one of these Jordan chains *must* come from $\operatorname{null}(C-I)^2$. Indeed, the next vector w in such a chain must have the property that $(C-I)w$ gives the first (i.e. previous) vector in the chain, which is an ordinary eigenvector, so $(C-I)^2w =$

$(C-I)[(C-I)w]=0$ since applying $C-I$ to an eigenvector gives zero. In our case,

$$(C-I)^2 = \begin{pmatrix} 0 & -1 & -1 & 2 \\ 0 & -1 & -1 & 2 \\ 0 & -1 & -1 & 2 \\ 0 & -1 & -1 & 2 \end{pmatrix},$$

which has a 3-dimensional null space. Two of these dimensions are already accounted for in $\dim \operatorname{null}(C-I)$, meaning that we only get one new generalized eigenvector at this stage. Again we denote this as a dot:

null$(C-I)$	null$(C-I)^2$
•	•
•	

Thus, so far we see that one Jordan chain has length at least 2, while the other has length 1.

The third vector in a Jordan chain must come from null$(C-I)^3$. In our case, $(C-I)^3$ turns out to be the zero matrix, so it has a 4-dimensional kernel. Three of these dimensions are already accounted for in lower powers of $C-I$, so we get one new generalized eigenvector at this stage;

null$(C-I)$	null$(C-I)^2$	null$(C-I)^3$
•	•	•
•		

Note that this generalized eigenvector can only add on to the existing chain of length at least 2, since to have something like

null$(C-I)$	null$(C-I)^2$	null$(C-I)^3$
•	•	
•	•	

we would have needed two new generalized eigenvectors at the null$(C-I)^2$ stage. Thus, we will have one Jordan chain of length 3 and one of length 1, agreeing with what we found previously. The point is that here we didn't need to find the explicit vectors in each Jordan chain, but rather were able to determine the lengths of each chain by comparing the dimensions of successive powers of $C-I$ to one another.

Applications of Jordan Forms

Finally, we say a word about the importance of the Jordan form and look at an application. We know that an operator on a finite-dimensional complex vector space can sometimes be represented by a diagonal matrix — i.e. exactly when the operator is diagonalizable. Note that for a diagonalizable operator, the Jordan form itself only consists of 1×1 Jordan blocks and is diagonal; this follows from the fact that for a diagonalizable operator, each generalized eigenvectors is a usual eigenvector. For a general operator, we can always at least represent it by an upper-triangular matrix. So the question is: what is the "best" possible matrix we can use to represent a general operator? The answer is the Jordan form, which is the "best" such matrix in the sense that it is very close to being diagonal except for a few ones above the main diagonal. The point is that this special form still allows us to do many of the nice things we can do with diagonal matrices.

To see an application, recall that for an operator T, we define the operator e^T by

$$e^T = I + \sum_{k=1}^{\infty} \frac{T^k}{k!}.$$

The motivation for this of course comes from the power series expansion of e^x. In general this may not be so easy to actually compute, but we will see the Jordan forms give us a nice way of describing e^T via a matrix.

If A is a square matrix, we have the same definition. If A is block diagonal:

$$A = \begin{pmatrix} A_1 & & \\ & \ddots & \\ & & A_m \end{pmatrix},$$

it turns out that A^k is also block-diagonal and

$$A^k = \begin{pmatrix} A_1^k & & \\ & \ddots & \\ & & A_m^k \end{pmatrix}.$$

This implies that

$$e^A = \begin{pmatrix} e^{A_1} & & \\ & \ddots & \\ & & e^{A_m} \end{pmatrix}.$$

So, if we can represent T by a block diagonal matrix, we can describe e^T by computing the exponential of each block. We know that if a block is diagonal, computing its exponential is easy. The point is that it is also easy for Jordan blocks!

For simplicity, consider the 2×2 Jordan block

$$J = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}.$$

Then

$$J^2 = \begin{pmatrix} \lambda^2 & 2\lambda \\ 0 & \lambda^2 \end{pmatrix}, \; J^3 = \begin{pmatrix} \lambda^3 & 3\lambda^2 \\ 0 & \lambda^3 \end{pmatrix},$$

and in general

$$J^k = \begin{pmatrix} \lambda^k & k\lambda^{k-1} \\ 0 & \lambda^k \end{pmatrix}.$$

Thus

$$\begin{aligned} e^J &= I + \sum_{k=1}^{\infty} \frac{J^k}{k!} \\ &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \sum_{k=1}^{\infty} \frac{1}{k!} \begin{pmatrix} \lambda^k & k\lambda^{k-1} \\ 0 & \lambda^k \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \sum_{k=1}^{\infty} \begin{pmatrix} \frac{\lambda^k}{k!} & \frac{\lambda^{k-1}}{(k-1)!} \\ 0 & \frac{\lambda^k}{k!} \end{pmatrix} \\ &= \begin{pmatrix} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} & \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} \\ 0 & \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} \end{pmatrix} \\ &= \begin{pmatrix} e^\lambda & e^\lambda \\ 0 & e^\lambda \end{pmatrix}, \end{aligned}$$

Thus we know exactly what the expoential of a 2×2 Jordan block is. Similarly, one can show that for a size k Jordan block, its exponential will be the upper-triangular matrix which as e^λ down the main diagonal, $e^\lambda/1!$ down the diagonal above that, $e^\lambda/2!$ down the diagonal above that, $e^\lambda/3!$ down the diagonal above that, and so on. Thus, we know exactly what the exponential of any Jordan block is, and hence what the exponential of any Jordan form is. So, we can always find a way to explicitly express e^T via a matrix representation.

As one final comment, here is another use of Jordan forms. We know that two matrices represent the same operator with respect to different bases if and only if they are similar. The still unanswered question is: Is there a quick way to determine if two matrices are similar? The answer is yes:

Theorem. *Two square complex matrices are similar if and only if they have the same Jordan form, up to a rearrangement of the Jordan blocks. In particular then, two square complex matrices represent the same operator if and only if they have the same Jordan form.*