Jordan Canonical Forms

December 6, 2006

1 Introduction

We know that not every $n \times n$ matrix A can be diagonalized. However, it turns out that we can always put matrices A into something called *Jordan Canonical Form*, which means that A can be written as

$$A = B^{-1} \begin{bmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_k \end{bmatrix} B,$$

where the J_i are certain block matrices of the form

$$J_i = [\lambda], \text{ or } \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}, \text{ or } \begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{bmatrix}, \text{ or ...}$$

Here, λ is an eigenvalue of A.

In this note we will only be concerned with how to compute the Jordan blocks J_i, as well as how to apply it. Thus, we will not be concerned with *proving* that there is always a Jordan decomposition.

2 Determining what the J_i blocks look like

Fix an eigenvalue λ. To determine the size of the Jordan blocks J_i that are associated to λ, it turns out that all we need to know are the numbers

$$\text{nullity}(A - \lambda I),\ \text{nullity}((A - \lambda I)^2),\ \text{nullity}((A - \lambda I)^3),\ \ldots$$

Moreover, we have:

Key Facts.

- nullity$(A - \lambda I)$ is the number of Jordan blocks J_i associated to λ.

- The differences nullity$((A - \lambda I)^j)$ − nullity$((A - \lambda I)^{j-1})$ is the number of Jordan blocks associated to λ that are of size at least $j \times j$.

These claims are easy to prove, so let us see why they hold.

2.1 On the nullity of $A - \lambda I$

First, we note that if the block J_i is $n_i \times n_i$, then one can easily check that

$$A - \lambda I \;=\; B^{-1} \begin{bmatrix} J_1 - \lambda I_{n_1} & 0 & \cdots & 0 \\ 0 & J_2 - \lambda I_{n_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & J_k - \lambda I_{n_k} \end{bmatrix} B.$$

It is easy to see that the jth power of this matrix is

$$(A - \lambda I)^j \;=\; B^{-1} \begin{bmatrix} (J_1 - \lambda I_{n_1})^j & 0 & \cdots & 0 \\ 0 & (J_2 - \lambda I_{n_2})^j & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & (J_k - \lambda I_{n_k})^j \end{bmatrix} B.$$

Since B is invertible, the rank (and nullity) of $(A - \lambda I)^j$ is the same as the rank (and nullity) of the matrix [1]

$$\begin{bmatrix} (J_1 - \lambda I_{n_1})^j & 0 & \cdots & 0 \\ 0 & (J_2 - \lambda I_{n_2})^j & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & (J_k - \lambda I_{n_k})^j \end{bmatrix}$$

[1]This follows from the fact that if U and V are both $n \times n$ matrices, such that U is invertible, then rank(UV) = rank(VU) = rank(V).

Since the different blocks $(J_i - \lambda I_{n_i})^j$ lie in different rows and columns, the rank (and nullity) of this block diagonal matrix equals the sum of the ranks (and nullities) of the individual blocks.[2]

The only blocks that could possibly contribute to the nullity (when we sum up the nullities of the $(J_i - \lambda I_{n_i})^j$ blocks) are those whose eigenvalues equal λ, because otherwise $(J_i - \lambda I_{n_i})^j$ is an $n_i \times n_i$ upper triangular matrix whose diagonal contains non-zero entries, making it intertible.

We now know that to compute our nullities, we only need to focus on blocks corresponding to the same eigenvalue λ. So, let us assume that we have reordered the Jordan blocks $J_1, ..., J_k$ so that the blocks corresponding to λ are $J_1, ..., J_t$. Then, a typical $J_i - \lambda I_{n_i}$, $i = 1, ..., t$, might look like

$$[0], \text{ or } \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \text{ or } \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \text{ or } \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \text{ or } ...$$

Let us denote this block by

$$H_i \; := \; J_i - \lambda I_{n_i}.$$

We note that the nullity of this block H_i is 1 (its rank is $n_i - 1$), no matter what n_i happens to be. So, the sum of the nullities of $H_1, ..., H_t$ is t, which therefore proves

> The nullity of $A - \lambda I$ is the number t of Jordan blocks associated to the eigenvalue λ.

2.2 On the nullity$((A - \lambda I)^j) - $nullity$((A - \lambda I)^{j-1})$

So we know how many Jordan blocks there are, but we would like to also determine their various sizes. In order to understand how to do this, we need to understand what the various powers of the blocks H_i look like. Suppose that the first through fourth powers of H_i are

$$\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

[2] This is not completely obvious, but is not difficult to see if you work out a few examples.

Once we have a 0 block, all higher powers will also be a zero 0. Thus, at some point, all higher powers of H_i will be a zero block. In our example we have that there are j rows that equal 0 if the power j is $1, 2, 3$, or 4, and there are four 0 rows for $j = 5$ or higher.

By studying this example, one should be convinced that the following is true:

$$\text{nullity}(H_i^j) \;=\; \begin{cases} j, & \text{if } n_i \geq j; \\ n_i, & \text{if } n_i < j. \end{cases}$$

So,

$$\text{nullity}(H_i^j) - \text{nullity}(H_i^{j-1}) \;=\; \begin{cases} 1, & \text{if } n_i \geq j; \\ 0, & \text{if } n_i < j. \end{cases}$$

If we sum this up over all the blocks H_i, we get a sum of 1's when $n_i \geq j$, which means that:

> $\text{nullity}((A - \lambda I)^j) - \text{nullity}((A - \lambda I)^{j-1})$ equals the number of blocks of size at least $j \times j$ corresponding to the eigenvalue λ, as claimed.

3 An example

Suppose that

$$A \;=\; \begin{bmatrix} 0 & 1 & 0 & 0 \\ -3 & 4 & 0 & 0 \\ 2 & -1 & 2 & 0 \\ -1 & 1 & 1 & 2 \end{bmatrix}.$$

What does the Jordan Canonical form look like (i.e. find the Jordan blocks) ?

First, we will need to compute the characteristic polynomial of A, to find the eigenvalues. A routine calculation reveals that

$$\det(A - \lambda I) \;=\; (\lambda - 2)^4.$$

So, $\lambda = 2$ is the only eigenvalue.

There are lots of possibilities for the Jordan blocks, then, and here they all are:

$$\begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}, \begin{bmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}, \begin{bmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}, \begin{bmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{bmatrix},$$

and

$$\begin{bmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{bmatrix}.$$

In general, we have that

Claim. The number of possible sizes of the Jordan blocks of an $n \times n$ matrix having a single eigenvalue λ is the number of integer partitions of n, denoted by $p(n)$. It turns out that

$$\begin{aligned} p(1) &= 1 \\ p(2) &= 2 \\ p(3) &= 3 \\ p(4) &= 5 \\ p(5) &= 7 \\ p(6) &= 11 \\ p(7) &= 15 \\ p(8) &= 22 \\ p(9) &= 30 \\ p(10) &= 42. \end{aligned}$$

In order to reduce the possibilities, we will need to first compute the number of Jordan blocks by row reducing $A - 2I$: When we do this, we get

$$A - 2I = \begin{bmatrix} -2 & 1 & 0 & 0 \\ -3 & 2 & 1 & 0 \\ 2 & -1 & 0 & 0 \\ -1 & 1 & 1 & 0 \end{bmatrix} \to \begin{bmatrix} -1 & 1 & 1 & 0 \\ -3 & 2 & 1 & 0 \\ 2 & -1 & 0 & 0 \\ -2 & 1 & 0 & 0 \end{bmatrix}$$

$$\rightarrow \begin{bmatrix} -1 & 1 & 1 & 0 \\ 0 & -1 & -2 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & -1 & -2 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} -1 & 1 & 1 & 0 \\ 0 & -1 & -2 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

The nullity of this matrix is 2; and so, we know that we have two Jordan blocks. Thus, we either have

$$\begin{bmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{bmatrix}.$$

To determine which it is, we must compute the nullity of $(A-2I)^2$: First,

$$(A - 2I)^2 \;=\; \begin{bmatrix} 1 & 0 & 1 & 0 \\ 2 & 0 & 2 & 0 \\ -1 & 0 & -1 & 0 \\ 1 & 0 & 1 & 0 \end{bmatrix}.$$

The nullity is clearly 3, because the first column is a basis for the column space. Thus,

$$\text{nullity}((A - 2I)^2) - \text{nullity}(A - 2I) \;=\; 3 - 2 \;=\; 1.$$

which means that there is exactly 1 matrix having size at least 2×2. Thus,

$$A \;=\; B^{-1} \begin{bmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix} B.$$

4 An application

Just knowing the general shape of the Jordan form is enough to prove some very nice results – that is, in a lot of cases you don't really need the matrix B.

Here is one example. Suppose we have a sequence $x_0, x_1, ...$ defined by the recurrence relation

$$x_n \;=\; c_1 x_{n-1} + \cdots + c_k x_{n-k}.$$

Of course since $x_{-1}, x_{-2}, \ldots$ are not defined, we need to define $x_0, \ldots, x_{k-1}$ to be certain values, in order to compute terms of the sequence. This constitutes some "initial conditions". Now, with this in mind, just like with the Fibonacci numbers we worked with earlier in the course, we will have that there is a corresponding matrix equation, and in our case it is:

$$\begin{bmatrix} c_1 & c_2 & \cdots & c_{k-1} & c_k \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix} \begin{bmatrix} x_{n-1} \\ x_{n-2} \\ \vdots \\ x_{n-k} \end{bmatrix} = \begin{bmatrix} x_n \\ x_{n-1} \\ \vdots \\ x_{n-k+1} \end{bmatrix}.$$

So, if we let M be this matrix, then we have that

$$M^n \begin{bmatrix} x_{k-1} \\ x_{k-2} \\ \vdots \\ x_0 \end{bmatrix} = \begin{bmatrix} x_{k-1+n} \\ x_{k-2+n} \\ \vdots \\ x_n \end{bmatrix}.$$

Then, we have that:

x_n is some linear combination of the entries of M^n.

It is easy to check that

$$M^n = B^{-1} \begin{bmatrix} J_1^n & & & \\ & J_2^n & & \\ & & \ddots & J_t^n \end{bmatrix} B,$$

where J_i is the ith Jordan block in the Jordan Canonical Form associated to the matrix M.

It is a simple matter to check that the entries of J_i^n all are of the form $p_i(n)\lambda^n$, where $p_i(x)$ is a certain polynomial of degree at most $n_i - 1$. In fact, far more is true: If the Jordan block J_i is

$$[\lambda], \text{ or } \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}, \text{ or } \begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{bmatrix}, \text{ or } \ldots$$

then the nth powers of this block is

$$[\lambda^n],\ \text{or}\ \begin{bmatrix} \lambda^n & \binom{n}{1}\lambda^{n-1} \\ 0 & \lambda^n \end{bmatrix},\ \text{or}\ \begin{bmatrix} \lambda^n & \binom{n}{1}\lambda^{n-1} & \binom{n}{2}\lambda^{n-2} \\ 0 & \lambda^n & \binom{n}{1}\lambda^{n-1} \\ 0 & 0 & \lambda^n \end{bmatrix},\ \text{or ...}$$

5 An example

Linear recurrence sequences abound, and turn up in the most unlikely of contexts. One example is the sequence

$$x_0 = 0,\ x_1 = 1,\ x_2 = 1+2,\ x_3 = 1+2+3,$$

and in general

$$x_n \ = \ 1 + 2 + \cdots + n.$$

It has been known for a very long time that

$$x_n \ = \ \frac{n(n+1)}{2}.$$

If you didn't know this, how would you prove it? For that matter, how would you prove that $1^k + \cdots + n^k$ is a certain $(k+1)$st degree polynomial in n ? Well, there are lots of ways to prove these things, but here I will explain how to do it using Jordan Canonical Forms.

First, let us see that the above sequence is a linear recurrence sequence: We have that

$$x_n - x_{n-1} \ = \ n;$$

and so,

$$(x_n - x_{n-1}) - (x_{n-1} - x_{n-2}) \ = \ n - (n-1) \ = \ 1;$$

and continuing in this vein we find that

$$x_n - 3x_{n-1} + 3x_{n-2} - x_{n-3} \ = \ 0.$$

In other words,

$$x_n \ = \ 3x_{n-1} - 3x_{n-2} + x_{n-3}.$$

It follows, from what we worked out in the previous section, that

$$\begin{bmatrix} 3 & -3 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}^n \begin{bmatrix} x_2 \\ x_1 \\ x_0 \end{bmatrix} = \begin{bmatrix} x_{n+2} \\ x_{n+1} \\ x_n \end{bmatrix}.$$

The characteristic polynomial of this matrix A (without the power n) is

$$f(\lambda) = (1-\lambda)^3.$$

So, $\lambda = 1$ is the only eigenvalue; moreover, one easily sees that $A - I$ has rank 2, meaning that the nullity is 1, and therefore the Jordan form involves just one large 3×3 block. It follows that the entries of A^n are polynomials of degree at most 2 in n; and therefore,

$$x_n = g(n), \text{ where } \deg(g) \leq 2.$$

Testing with a few small values of n, one finds that $g(n) = n(n+1)/2$, as claimed.