

The Jordan normal form and the Euclidean algorithm

12 October, 2007 in [expository](#), [math.RA](#) | Tags: [Bezout identity](#), [Euclidean algorithm](#), [Jordan normal form](#), [nilpotent matrices](#), [principal ideal domain](#) | by [Terence Tao](#)

In one of my recent posts, I used the [Jordan normal form](#) for a matrix in order to justify a couple of arguments. As a student, I learned the derivation of this form twice: firstly (as an undergraduate) by using the [minimal polynomial](#), and secondly (as a graduate) by using the [structure theorem for finitely generated modules over a principal ideal domain](#). I found though that the former proof was too concrete and the latter proof too abstract, and so I never really got a good intuition on how the theorem really worked. So I went back and tried to synthesise a proof that I was happy with, by taking the best bits of both arguments that I knew. I ended up with something which wasn't too different from the standard proofs (relying primarily on the [\(extended\) Euclidean algorithm](#) and the [fundamental theorem of algebra](#)), but seems to get at the heart of the matter fairly quickly, so I thought I'd put it up on this blog anyway.

Before we begin, though, let us recall what the Jordan normal form theorem is. For this post, I'll take the perspective of abstract linear transformations rather than of concrete matrices. Let $T : V \rightarrow V$ be a linear transformation on a finite dimensional complex vector space V , with no preferred coordinate system. We are interested in asking what possible “kinds” of linear transformations V can support (more technically, we want to classify the conjugacy classes of $\text{End}(V)$, the ring of linear endomorphisms of V to itself). Here are some simple examples of linear transformations.

1. **The right shift.** Here, $V = \mathbb{R}^n$ is a standard vector space, and the *right shift* $U : V \rightarrow V$ is defined as $U(x_1, \dots, x_n) = (0, x_1, \dots, x_{n-1})$, thus all elements are shifted right by one position. (For instance, the 1-dimensional right shift is just the zero operator.)
2. **The right shift plus a constant.** Here we consider an operator $U + \lambda I$, where $U : V \rightarrow V$ is a right shift, I is the identity on V , and $\lambda \in \mathbb{C}$ is a complex number.
3. **Direct sums.** Given two linear transformations $T : V \rightarrow V$ and $S : W \rightarrow W$, we can form their direct sum $T \oplus S : V \oplus W \rightarrow V \oplus W$ by the formula $(T \oplus S)(v, w) := (Tv, Sw)$.

Our objective is then to prove the

Jordan normal form theorem. Every linear transformation $T : V \rightarrow V$ on a finite dimensional complex vector space V is similar to a direct sum of transformations, each of which is a right shift plus a constant.

(Of course, the same theorem also holds with left shifts instead of right shifts.)

– Reduction to the nilpotent case –

Recall that a linear transformation $T : V \rightarrow V$ is [nilpotent](#) if we have $T^m = 0$ for some positive integer m . For instance, every right shift operator is nilpotent, as is any direct sum of right shifts. In fact, these are essentially the only nilpotent transformations:

Nilpotent Jordan normal form theorem. Every nilpotent linear transformation $T : V \rightarrow V$ on a finite dimensional vector space is similar to a direct sum of right shifts.

We will prove this theorem later, but for now let us see how we can quickly deduce the full Jordan normal form theorem from this special case. The idea here is, of course, to split up the [minimal polynomial](#), but it turns out that we don't actually need the minimal polynomial *per se*; any polynomial that annihilates the transformation will do.

More precisely, let $T : V \rightarrow V$ be a linear transformation on a finite-dimensional complex vector space V . Then the powers I, T, T^2, T^3, \dots are all linear transformations on V . On the other hand, the space of all linear transformations on V is a finite-dimensional vector space. Thus there must be a non-trivial linear dependence between these powers. In other words, we have $P(T) = 0$ (or equivalently, $V = \ker(P(T))$) for some polynomial P with complex coefficients.

Now suppose that we can factor this polynomial P into two coprime factors of lower degree, $P = QR$. Using the [extended Euclidean algorithm](#) (or more precisely, [Bézout's identity](#)), we can find more polynomials A, B such that $AQ + BR = 1$. In particular,

$$A(T)Q(T) + B(T)R(T) = I. \quad (1)$$

The formula (1) has two important consequences. Firstly, it shows that $\ker(Q(T)) \cap \ker(R(T)) = \{0\}$, since if a vector v was in the kernel of both $Q(T)$ and $R(T)$, then by applying (1) to v we obtain $v=0$. Secondly, it shows that $\ker(Q(T)) + \ker(R(T)) = V$. Indeed, given any $v \in V$, we see from (1) that $v = R(T)B(T)v + Q(T)A(T)v$; since $Q(T)R(T) = R(T)Q(T) = P(T) = 0$ on V , we see that $R(T)B(T)v$ and $Q(T)A(T)v$ lie in $\ker(Q(T))$ and $\ker(R(T))$ respectively. Finally, since all polynomials in T commute with each other, the spaces $\ker(Q(T))$ and $\ker(R(T))$ are T -invariant.

Putting all this together, we see that the linear transformation T on $\ker(P(T))$ is similar to the direct sum of the restrictions of T to $\ker(Q(T))$ and $\ker(R(T))$ respectively. We can iterate this observation, reducing the degree of the polynomial P which annihilates T , until we reduce to the case in which this polynomial P cannot be split into coprime factors of lesser degree. But by the [fundamental theorem of algebra](#), this can only occur if P takes the form $P(t) = (t - \lambda)^m$ for some $\lambda \in \mathbb{C}$ and $m \geq 0$. In other words, we can reduce to the case when $(T - \lambda I)^m = 0$, or in other words T is equal to λI plus a nilpotent transformation. If we then subtract off the λI term, the claim now easily follows from the nilpotent Jordan normal form theorem.

[From a modern algebraic geometry perspective, all we have done here is split the spectrum of T (or of the ring generated by T) into connected components.]

It is interesting to see what happens when two eigenvalues get very close together. If one carefully inspects how the Euclidean algorithm works, one concludes that the coefficients of the polynomials $A(T)$ and $B(T)$ above become very large (one is trying to separate two polynomials $Q(T)$ and $R(T)$ that are only barely coprime to each other). Because of this, the Jordan decomposition becomes very unstable when eigenvalues begin to collide.

Because the fundamental theorem of algebra is used, it was necessary to work in an algebraically closed field such as the complex numbers \mathbb{C} . (Indeed, one can in fact *deduce* the fundamental theorem of algebra from the Jordan normal form theorem.) Over the reals, one picks up other “elliptic” components, such as 2×2 rotation matrices, which are not decomposable into translates of shift

operators.

Thus far, the decompositions have been canonical – the spaces one is decomposing into can be defined uniquely in terms of T (they are the kernels of the primary factors of the minimal polynomial). However, the further splitting of the nilpotent (or shifted nilpotent) operators into smaller components will be non-canonical, depending on an arbitrary choice of basis. [However, the multiplicities of each type of shift-plus-constant factor will remain canonical; this is easiest to see by inspecting the dimensions of the kernels of $(T - \lambda I)^m$ for various λ, m using a Jordan normal form.]

— Proof of the nilpotent case —

To prove the nilpotent Jordan normal form theorem, I would like to take a dynamical perspective, looking at orbits x, Tx, T^2x, \dots of T . (These orbits will be a cheap substitute for the concept of a Jordan chain.) Since T is nilpotent, every such orbit terminates in some finite time m_x , which I will call the *lifespan* of the orbit (i.e. m_x is the least integer such that $T^{m_x}x = 0$). We will call x the *initial point* of the orbit, and $T^{m_x-1}x$ the *final point*.

The elements of a finite orbit $x, Tx, \dots, T^{m_x-1}x$ are all linearly independent. This is best illustrated with an example. Suppose x has lifespan 3, thus $T^3x = 0$ and $T^2x \neq 0$. Suppose there linear dependence between x, Tx, T^2x , say $3x + 4Tx + 5T^2x = 0$. Applying T^2 we obtain $3T^2x = 0$, a contradiction. A similar argument works for any other linear dependence or for any other lifespan. (Note how we used the shift T here to eliminate all but the final point of the orbit; we shall use a similar trick with multiple orbits shortly.)

The vector space spanned by a finite orbit $x, Tx, \dots, T^{m_x-1}x$ is clearly T -invariant, and if we use this finite orbit as a basis for this space then the restriction of T to this space is just the right shift. Thus to prove the nilpotent Jordan normal form theorem, it suffices to show

Nilpotent Jordan normal form theorem, again. Let $T : V \rightarrow V$ be nilpotent. Then there exists a basis of V which is the concatenation of finite orbits $x, Tx, \dots, T^{m_x-1}x$.

We can prove this by the following argument (basically the same argument used to prove the Steinitz exchange lemma, but over the ring $\mathbb{C}[T]$ instead of \mathbb{C}). First observe that it is a triviality to obtain a concatenation of finite orbits which *span* V : just take any basis of V and look at all the finite orbits that they generate. Now all we need to do is to keep whittling down this over-determined set of vectors so that they span *and* are linearly independent, thus forming a basis.

Suppose instead that we had some collection of finite orbits $x_i, Tx_i, \dots, T^{m_{x_i}-1}x_i$, for $i = 1, \dots, r$ which spanned V , but which contained some non-trivial linear relation. To take a concrete example, suppose we had three orbits x, Tx, y, Ty, T^2y and z, Tz, T^2z, T^3z which had a non-trivial linear relation

$$3x + 4Tx + 5Ty + 6T^2y + 7T^2z = 0.$$

By applying some powers of T if necessary, and stopping just before everything vanishes, we may assume that our non-trivial linear relation only involves the final points $T^{m_{x_i}-1}x_i$ of our orbits. For instance, in the above case we can apply T once to obtain the non-trivial linear relation

$$3Tx + 5T^2y + 7T^3z = 0.$$

We then factor out as many powers of T as we can; in this case, we have

$$T(3x + 5Ty + 7T^2z) = 0.$$

The expression in parentheses is a linear combination of various elements of our putative basis, in this case x , Ty , and T^2z , with each orbit being represented at most once. At least one of these elements is an initial point (in this case x). We can then replace that element with the element in parentheses, thus shortening one of the orbits but keeping the span of the concatenated orbits unchanged. (In our example, we replace the orbit x, Tx of lifespan 2 with an orbit $3x + 5Ty + 7T^2z$ of lifespan 1.)

Iterating this procedure until no further linear dependences remain, we obtain the nilpotent Jordan normal form theorem, and thus the usual Jordan normal form.

Possibly related posts: (automatically generated)

- [Uniqueness of Jordan Normal Forms](#)
- [Jordan Normal Form](#)

15 comments

[Comments feed for this article](#)

[13 October, 2007 at 12:20 am](#)

Attila Smith



Dear Terence,

like you I was dissatisfied by both paths to Jordan's theorem, but I would never have had the guts to admit it, nor the talent to find a satisfying new way .

Your blog is truly liberating in showing how a great mathematician cares and thinks about relatively elementary mathematics, and then improves the traditional presentation.

I value this psychological effect of your blog as highly as its wonderful mathematical content.

Yours thankfully ,

Attila

[Reply](#)

[13 October, 2007 at 7:03 am](#)

im



bring back snap!

0 0 Rate This

[Reply](#)

[14 October, 2007 at 1:35 am](#)