Lecture 3-25: The rational canonical (Frobenius normal) and Jordan forms

March 25, 2024

Last term Julia classified finitely generated modules over a PID. As she pointed out then, a rich class of examples of these arises from a finite-dimensional vector space V over a field k equipped with a linear transformation T from V to itself. Then V becomes a finitely generated module, with k acting as usual by scalar multiplication, while $xv = Tv$ for all $v \in V$.

By the elementary divisor version of this classification we can write V as a direct sum $\bigoplus_{i=1}^{m} k[x]/(q_i)$ of quotients of $k[x]$ by principal ideals (q_i) with each q_i of the form $p_i^{n_i}$, a power of a monic irreducible polynomial p_i over k (there is no room for any summands $k[x]$). Now given *any* monic polynomial $q = x^n + \sum_{i=0}^{n-1} q_i x^i \in k[x]$, there is an obvious basis for the quotient $W = k[x]/(q)$ consisting of the first n powers $1, x \ldots, x^{n-1}$ of x. Multiplication by x sends any power x^i to x^{i+1} if $i < n-1$, while $xx^{n-1} = x^n = -\sum_{i=0}^{n-1} q_i x^i$.

Accordingly the matrix of multiplication by x on W with respect to this basis is

$$\begin{pmatrix} 0 & 0 & 0 & \dots & -q_0 \\ 1 & 0 & 0 & \dots & -q_1 \\ 0 & 1 & 0 & \dots & -q_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & -q_{n-1} \end{pmatrix}$$

We denote this last matrix by $C(q)$ and call it the companion matrix of q (p. 475 of DF). Clearly it has minimal polynomial q as an $n \times n$ matrix over k; it provides a handy way to write down an explicit matrix with a given minimal polynomial. One checks directly that the characteristic polynomial of $C(q)$ is also $\pm q$. Combining the classification with the above choice of basis on each quotient $k[x]/(p_i^{n_i})$ above we get

Rational canonical (or Frobenius normal) form for matrices

Any square matrix M over k is similar to a block diagonal one for which the blocks are companion matrices $C(p_i^{n_i})$ of powers of irreducible polynomials p_i over k. Two such block diagonal matrices are similar if and only if one can be obtained from the other by reordering the blocks.

The minimal polynomial of the above block diagonal matrix is clearly the least common multiple of the powers $p_i^{n_i}$. As a consequence, given an irreducible polynomial p over k of degree d, the companion matrix $C(p^m)$ is up to similarity the only $md \times md$ matrix with minimal polynomial p^m; all other matrices of this size will have minimal polynomials either involving a factor other than p or equal to a lower power of p.

The version of this theorem in Dummit and Foote (Theorem 16, p. 477) uses the invariant factor decomposition rather than the elementary divisor decomposition, so that the block diagonal factors take the form $C(q_1), \ldots, C(q_m)$, where the q_i are monic polynomials with $q_1 | q_2 | \cdots | q_m$.

A further remarkable consequence is that any square matrix M over k is similar to its transpose M^t. Indeed, by putting M into rational canonical form we see that it is enough to prove this for companion matrices $C = C(p^m)$ of powers of irreducible polynomials p. But now a polynomial $q(C) = 0$ if and only if $q(C^t) = 0$, so C^t has the same minimal polynomial as C and must be similar to it. Note that any Math 208 student could understand the statement of this last result, but its proof is well beyond the scope of that course.

Now assume that k is algebraically closed, or more generally that all eigenvalues (roots of the characteristic polynomial) of a linear transformation T from k^n to itself lie in k. Then the only possible powers occurring in the rational canonical form of any matrix of T are powers $(x - a_i)^{n_i}$ of linear polynomials. In this case it is customary to use a different basis for $W = k[x]/((x - a_i)^{n_i}$, namely the one consisting of the powers $(x - a_i)^{n_i - 1} \ldots, (x - a_i)^0$ in reverse order. The matrix of multiplication by x on W now takes the form

$$\begin{pmatrix} a_i & 1 & 0 & \dots & 0 \\ 0 & a_i & 1 & \dots & 0 \\ 0 & 0 & a_i & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & a_i \end{pmatrix}$$

We call this matrix a Jordan block (with eigenvalue a_i).

Then we have

Jordan canonical form for matrices: Theorem 23, p. 493, Dummit and Foote

Any square matrix M over a field k containing all of its eigenvalues is similar to a block diagonal matrix with Jordan blocks corresponding to eigenvalues of M. Two such block diagonal matrices are similar if and only if one can be obtained from the other by reordering the blocks.

In particular, for any element $a \in k$ and a fixed size n, there are only finitely many similarity classes of $n \times n$ matrices with a as their only eigenvalue. The similarity class of any such matrix is characterized by the sizes of its Jordan blocks. These form an unordered set of positive integers whose sum is n (a partition of n). Partitions are well-studied objects in enumerative combinatorics and a great deal is known about them. Another consequence of the Jordan form is the following result: any $n \times n$ complex invertible matrix M has a logarithm, so that $M = e^L = \sum_{i=0}^{\infty} \frac{L^i}{i!}$ for some complex matrix L.

I will close by returning to the rational canonical form in a very special case. Let $q = p^n$ be a power of a prime p. You will show in homework this week that some nonzero x in the finite field F_q of order q has multiplicative order $q - 1$. Hence any such x generates all of F_q over its prime subfield F_p, so that its minimal polynomial f over F_p necessarily has degree n. Then the matrix of multiplication by x, regarded as a linear transformation of F_q, also has multiplicative order $q - 1$ and minimal polynomial f, whence its rational canonical form must be $C(f)$. Thus $C(f)$ and its transpose $C(f)^t$ are invertible $n \times n$ matrices over F_p having order $q - 1$.

In the same homework assignment, you will use $C(f)^t$ to construct a deBruijn sequence of length q; that is, a sequence $a_0, \ldots, a_{q-1}$ of elements of F_p such that its consecutive subsequences $a_i, \ldots, a_{i+n-1}$ of length n run through all sequences of n elements of F_p, each such sequence occurring exactly once. Here the index i runs from 0 to $q-1$ and addition of subscripts takes place modulo q (so that the sequence "wraps around").