MAS 5312 – Lecture for April 3, 2020

Richard Crew
Department of Mathematics,
University of Florida

Recap

Let A be an $n \times n$ matrix with entries in a field K, and denote by V_A the $K[X]$-module which as an abelian group (or even K-vector space) is V, and where $f(X) \in K[X]$ acts as the $n \times n$ matrix $f(A)$. We have shown that matrices A and B are similar if and only if $V_A \simeq V_B$ as $K[X]$-modules, and this happens if and only if V_A and V_B have the same invariant factors, up to multiplication by constants. We will agree to make all of our invariant factors *monic*; this means that A and B are similar if and only if they have the same invariant factors.

This amounts to a classification of the conjugacy classes of $GL_n(K)$. In fact if $d_1, \ldots, d_r$ are the invariant factors of A then $\prod_i d_i$ is a monic polynomial of degree n. I will show, conversely that if $d_1, \ldots, d_r$ are monic polynomials such that $d_1 | d_2 | \cdots | d_r$ and $\prod_i d_i$ has degree n, there is a conjugacy class in $GL_n(K)$ for which $d_1, \ldots, d_r$ are the invariant factors.

In fact this is not hard. Recall first that if $f \in K[X]$ is any polynomial of degree d, the ring $K[X]/(f)$ has dimension d as a K-vector space. We've seen this before: the division algorithm shows that any element of $K[X]/(f)$ has a unique representative of the form $h + (f)$ with $\deg(h) < \deg(f) = d$. It follows that the elements

$$1 + (f), X + (f), \ldots, X^{d-1} + (f)$$

are a basis $K[X]/(f)$ as a K-vector space.

Suppose now $f(X) = X^d + a_{d-1}X^{d-1} + \cdots + a_0$ is monic, and let's compute the matrix representation of the linear transformation "multiplication by X" with respect to this basis. It will be convenient to denote by x the image of X in $K[X]/(f)$, so that $x^i = X^i + (f)$.

Recall how we find these matrix representations: if $T : V \to V$ is any linear map and $v_1, \ldots, v_n$ is any basis of V, the matrix $A = (a_{ij})$ is determined by the equations $T(v_j) = \sum_i a_{ij} v_i$ for all j. The easy way to think about this is to write it using row vectors:

$$(T(v_1)\ T(v_2) \cdots T(v_n)) = (v_1\ v_2 \cdots v_n)A.$$

Let's apply this to the case of $V = K[X]/(f)$ with the basis $1, x, x^2, \ldots, x^{d-1}$, as before, and T is multiplication by x. We have

$$T(1) = x,\ T(x) = x^2, \ldots, T(x^{d-2}) = x^{d-1}$$

and since $f(X) = X^d + a_{d-1}X^{d-1} + \cdots + a_0$,

$$T(x^{d-1}) = x^d = -a_0 - a_1 x - \cdots - a_{d-1}X^{d-1}.$$

Thus with this particular basis, the matrix representing multiplication by x is

$$\mathcal{C}_f = \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 \\ \vdots & & & & & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 \\ -a_0 & -a_1 & -a_2 & -a_3 & \cdots & -a_{n-1} \end{pmatrix}.$$

We call this the *companion matrix of f*. Since it represents the linear transformation "multiplication by X (or x)" on $K[X]/(f)$, this matrix has exactly one invariant factor, namely f itself. In particular its characteristic polynomial is $(-1)^n f(X)$, a fact which is easily checked by explicit computation (give it a shot – use cofactor expansions).

Suppose now that $V_1, \ldots, V_r$ are K-vector spaces, $\dim_K(V_i) = n_i < \infty$ and for each i $A_i : V_i \to V_i$ is a linear map. If we set $V = \bigoplus_i V_i$ there is a linear map $A : V \to V$ whose restriction to each $V_i \subset V$ is A_i; explicitly

$$A(v_1, \ldots, v_r) = (A_1 v_1, \ldots, A_r v_r).$$

If B_i is a basis of V_i then A_i is represented by an $n_i \times n_i$ matrix with entries in K. Then $B = \bigcup_i B_i$ is a basis of V, and if we order B suitably the matrix representing A in the basis B is block-diagonal with diagonal entries $A_1, \ldots, A_r$ (this was a homework problem when $r = 2$, and the general case is no different).

Let's apply this to $V = \bigoplus_i K[X]/(d_i)$ with monic $d_i \in K[X]$, and A is multiplication by X. On each summand $K[X]/(d_i)$ the "standard" basis $1, x, \ldots, x^{\deg(d_i)-1}$ is given by the companion matrix $\mathcal{C}_{d_i}$. We conclude that for the union of these bases, ordered suitably, the matrix representing multiplication by X is block diagonal with entries $\mathcal{C}_{d_1}, \ldots, \mathcal{C}_{d_r}$.

To summarize: suppose $A : K^n \to K^n$ is a linear map, identified with a matrix A, and the invariant factors of V_K are $d_1, \ldots, d_r$ with each d_i monic. Then A is similar to the block-diagonal matrix with entries $\mathcal{C}_{d_i}$; this is the *rational canonical form* of A, and it is uniquely determined by A.

By considering the possible rational canonical forms of a nonsingular matrix of a given size, we can determine the conjugacy classes of $GL_n(K)$.

Suppose for example that $n = 2$. If $A \in GL_2(K)$ then $\Phi_A(X)$ has degree two, so there are two cases:

- A has a single invariant factor d_1 of degree two, or
- A has two invariant factors d_1, d_2 such that $d_1 d_2 = \Phi_A(X)$. In particular d_1 and d_2 have degree one, and since $d_1 | d_2$ we must have $d_2 = d_1$ since both are monic.

In the first case we have $d_1(X) = a + bX + X^2$, which has the companion matrix $\begin{pmatrix} 0 & 1 \\ -a & -b \end{pmatrix}$. So there is one conjugacy class of this sort for each $a, b \in K$ with $a \neq 0$. In the second case $d_1 = d_2 = X - a$. The companion matrices are just the 1×1 matrix with entry a, so these conjugacy classes are the classes of scalar matrices, and there is one for each nonzero $a \in K$. So basically, our construction separates out the central classes from the noncentral ones.

This works for any kind of matrix, not just invertible ones. Let's find all similarity classes of *nilpotent* matrices in $GL_n(K)$. If A is nilpotent, say A^N, the minimal polynomial of A must divide X^N, so it is a power of X. Then all the invariant factors of A are powers of X, say $(X^{n_1}, \ldots, X^{n_r})$ with $n_1 \leq n_2 \leq \cdots \leq n_r$, and since $\Phi_A(X)$ has degree n and is the product of the invariant factors, we must have $\sum_i n_i = n$. The companion matrix to a polynomial of the form X^k is a $k \times k$ matrix with 1s just above the diagonal and 0s elsewhere. It follows that a nilpotent matrix A is similar to a block diagonal matrix in which each block has this form; furthermore the sizes of the individual blocks determine the class of A and conversely.

As a last example lets find all conjugacy classes in $GL_5(\mathbb{Q})$ of matrices A whose characteristic polynomial is $(X-1)^3(X+1)^2$. The simplest way to find the possible invariant factors is to first find the possible elementary divisors. The product of all of the elementary divisors is characteristic polynomial, so the elementary divisors of the $(X-1)$-primary part of V_A are $((X-1),(X-1),(X-1))$, $((X-1),(X-1)^2)$ or $(X-1)^3$, and the ones for the $(X+1)$-primary part are $((X+1),(X+1))$ or $(X+1)^2$. Using our standard algorithm we find that the possible sets of invariant factors are

1. $(X-1)^3(X+1)^2 = X^5 - X^4 - 2X^3 + 2X^2 + X - 1$,
2. $(X-1,(X-1)^2(X+1)^2) = (X-1, X^4 - 2X^2 + 1)$,
3. $(X-1, X-1, (X-1)(X+1)^2) = (X-1, X-1, X^3+X^2-X-1)$,
4. $(X+1,(X-1)^3(X+1)) = (X+1, X^4 - 2X^3 + 2X - 1)$,
5. $((X-1)(X+1),(X-1)^2(X+1)) = (X^2-1, X^3-X^2-X+1)$,
6. $X-1,(X-1)(X+1),(X-1)(X+1) = (X-1, X^2-1, X^2-1)$.

We find that there are six such classes, with representatives

$$1: \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & -1 & -2 & 2 & 1 \end{pmatrix}, \quad 2: \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & -1 & 2 & 0 & -1 \end{pmatrix},$$

$$3: \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & -1 \end{pmatrix}, \quad 4: \begin{pmatrix} -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & -2 & 2 & 1 \end{pmatrix},$$

$$5: \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 1 & 1 \end{pmatrix} \text{ and finally } 6: \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

Similar games can be played with any polynomial as characteristic polynomial (assuming it has degree equal to the dimension of V).

As a last example, let's find all conjugacy classes in the group $G = GL_3(\mathbb{F}_2)$. We will eventually show that this is a simple group of order $168 = 2^3 \cdot 3 \cdot 7$ (the order comes from our standard formula for the order of $GL_n(F)$, F a finite field). Note that all the invariant factors of an element must have nonzero constant term, since the constant term of the characteristic polynomial is the determinant. Furthermore the sum of the degrees of the invariant factors must be 3.

Therefore the possible sets of invariant factors are as follows:

1. A single invariant factor d_1, which is either
   1.1 irreducible: $X^3 + X + 1$ or $X^3 + X^2 + 1$; or
   1.2 reducible: $(X+1)^3$ or $(X+1)(X^2+X+1)$, since $X+1$ and X^2+X+1 are the only irreducible polynomials of degrees 1 and 2 with nonzero constant term.
2. Two invariant factors $d_1|d_2$, and since $X+1$ is the only linear polynomial with nonzero constant term, they are $X+1, (X+1)^2$.
3. Three invariant factors $d_1|d_2|d_3$, which are all $X+1$.

We conclude that there are 5 conjugacy classes. We can find their orders by writing down the companion matrices and explicitly computing the order, but there is an easier way.

Observe first that

$$X^7 - 1 = (X - 1)(X^3 + X + 1)(X^3 + X^2 + 1)$$

in $\mathbb{F}_2[X]$, so the 2 classes in 1.1 have order 7 (they have order dividing 7 and aren't the identity). Representatives of these classes are

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

Similarly $X^4 - 1 = (X + 1)^4$ which shows that the class in 1.2 with invariant factor $(X + 1)^3$ has order 4 (it must have order dividing 4, and can't have order 1 or 2 since the minimal polynomial $(X + 1)^3$ doesn't divide $X^2 + 1$). A representative is

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

Likewise $X^3 - 1 = (X+1)(X^2+X+1)$, so the class in 1.2 with invariant factor $(X+1)(X^2+X+1)$ has order 3; a representative is

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

Finally the single class in case 2 has order 2, with representative

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

and the class in case 3 is the identity.

In short, the possible orders of elements of $GL_3(\mathbb{F}_2)$ are 1, 2, 3, 4 or 7; there are two classes of elements of order 7 and only one class in the other cases.