VIGRE Algebra QR Review, October 21, 2002

Summary of Session

Concerning the sequence 10, 11, 12, 13, 14, 15, 16, 17, 20, 22, 24, 31, 100, 121, 10000: these are all the number 16, written in base 16, base 15, ... base 5, base 4, base 3, base 2.

Here's a new sequence to think about: what rule is being used to generate each row of numbers form the preceding row?

1
1 1
2 1
1 2 1 1
1 1 1 2 2 1
3 1 2 2 1 1
1 3 1 1 2 2 2 1
1 1 1 3 2 1 3 2 1 1
3 1 1 3 1 2 1 1 1 3 1 2 2 1
...

There was some discussion of why one can prove certain theorems about matrices by proving them instead about matrices of indeterminates over the integers. Suppose that $Y = (y_{ij})$ and $Z = (z_{ij})$ are two $n \times n$ matrices with indeterminate entries over $\mathbb{Z}$, and that we have already verified that YZ and ZY have the same characteristic polynomial. Let x be an additional indeterminate. Let R be a commutative ring with identity (take it to be a field if you like, but this is not necessary), and let $A = (a_{ij})$ and $B = \mathbb{Z}(b_{ij})$ be matrices over R. There is a homomorphism $D = [y_{ij}, z_{ij}][x] \to R[x]$ that maps y_{ij} to a_{ij} for all i, j, that maps z_{ij} to b_{ij} for all i, j, and that maps x to x. This homomorphism maps the characteristic polynomial of YZ to the characteristic polynomial of AB and the characteristic polynomial of ZY to that of BA. (Heuristically, the point is that in computing the characteristic polynomial of AB when the entries are not known, it comes out as a polynomial in the entries of A and B and the variable x. One is automatically manipulating those entries is if one were working in a polynomial ring.) It follows that AB and BA will have the same characteristic polynomial provided that YZ and ZY do. But one may consider the problem over the field $L = \mathbb{Q}(y_{ij}, z_{ij})$, which is the field of fractions of the integral domain $D = \mathbb{Z}[y_{ij}, z_{ij}]$. Working over the field does not affect whether the two characteristic polynomials are equal (the homomorphism $D[x] \to R[x]$ will not necessarily extend to a homomorphism defined on $L[x]$, but this does not matter). But now Y (and Z for that matter) are invertible, and so $YZ = Y(ZY)Y^{-1}$ and ZY are similar.

Many problems become more tractable when the objects being studied are placed in "general position" in a suitable sense. Desargue's theorem in geometry is an example: suppose that three lines meet at a point: let A, A' be two points lying on the first line, B, B' on the second, and C, C' on the third. Then the point X where line AB meets line

$A'B'$, the point Y where BC meets $B'C'$, and the point Z where CA meets $C'A'$ lie on a straight line. In the plane this is a bit mysterious, and not so easy to prove. But if one considers the case where the original lines are in 3-space and triangles ABC and $A'B'C'$ are in different planes, it becomes completely obvious: X, Y, and Z lie on the intersection of the planes. (The plane geometry case can then be proved by interpreting the picture as the projection on the plane of a picture in 3-space, or by a limit argument.)

Coming back to algebra, the Cayley-Hamilton theorem, which asserts that a square matrix satisfies its characteristic polynomial, is susceptible to the same kind of analysis as the proof above that the characteristic polynomials of AB and BA are the same. It suffices to consider a matrix (y_{ij}) of indeterminates over $\mathbb{Z}$. One can show that the eigenvalues of such a matrix are mutually distinct, and this means that the matrix, considered over an algebraic closure of $\mathbb{Q}(y_{ij})$, is diagonalizable. But the case of a diagonal matrix is easy.

The following problems are from the list for October 21 — I am including solutions for all of these problems (1. and 7. were not discussed).

1. Let T be a linear transformation of a vector space T such that T has minimal polynomial $x^2(x-1)^2$ and characteristic polynomial $x^3(x-1)^4$. What are the possible Jordan canonical forms (resp., rational canonical forms) for (the matrix of) T?

Solution. In the Jordan form, the largest block for the eigenvalue 0 must have size 2 (multiplicity in the minimal polynomial), and the block sizes add up to 3 (multiplicity in the characteristic polynomial). The only possibility is a block of size 2 and a block of size 1. The largest block for the eigenvalue 1 must have size 2, and the sum of the block sizes must be 4. There are two possibilities: 2 blocks of size 2 or 1 block of size 2 and 2 of size 1. Thus, there are only two possible Jordan forms. (I am not writing out all of the 4×4 matrices one gets: on the exam, it would be safest to do that.)

The possible rational forms correspond to sequences of monic polynomials such that each divides the next, the one of largest degree is the minimal polynomial, and the product is the characteristic polynomial. There are two possibilities: $x(x-1)^2$ and $x^2(x-1)^2$ (if there are to be just two) and $x-1$, $x(x-1)$, and $x^2(x-1)^2$ if there are three. The actual rational forms (not given explicitly here) are the direct sums of the companion matrices of the polynomials in the sequence.

2. (a) Find the Jordan canonical form of the complex matrix $A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix}$.

(b) Let F be any field. Suppose that a normal subgroup N of $\mathrm{GL}(2, F)$ contains one matrix of the form $\begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix}$ with $a \in F - \{0\}$. Prove that N contains every matrix of this form.

Solution. For part (a), $xI - A = \begin{pmatrix} x-1 & -1 & -1 \\ -1 & x-1 & -2 \\ 0 & 0 & x \end{pmatrix}$, and expanding the determinant by the third row yields $x((x-1)^2 - 1) = x^2(x-2)$. There will be one Jordan block of size one for the eigenvalue 2, and either two blocks of size 1 or one block of size 2 for the eigenvalue 0. If there are two blocks of size 1 for the eigenvalue 0, the matrix is similar to a diagonal matrix and the minimal polynomial would be $x(x-2)$. By calculation, $A^2 \neq 2A$ (the third entry of the first row of A^2 is 3), and so this is not true and the Jordan form has one block of size 1 for the eigenvalue 2 and one of size 2 for the eigenvalue 0. Specifically, one gets $\begin{pmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$, using lower triangular Jordan blocks.

For part (b), note that by an easy calculation, the characteristic polynomial for any of the specified matrices is $(x-1)^2$ while the minimal polynomial is also $(x-1)^2$ (because $a \neq 0$). Thus, there is only one block in the rational form, and all of these matrices are similar to the companion matrix of $(x-1)^2$. This means that they are all in the same conjugacy class. Since normal subgroups of $\mathrm{GL}(2, F)$ are closed under conjugation, a normal subgroup containing one of them contains them all.

3. Let A and B be two $n \times n$ matrices with complex entries.

(a) Show that for $\lambda \in \mathbb{C}$, λ is an eigenvalue of AB if and only if λ is an eigenvalue of BA.

(b) Show that if A or B is invertible, then AB and BA are conjugate.

(c) Give an example for $n = 2$ where AB and BA are not conjugate.

Solution. $AB = A(BA)A^{-1}$ and $BA = B(AB)B^{-1}$, provided that A is invertible in the first case and that B is invertible in the second case. This handles (b). The argument for showing that AB and BA have the same characteristic polynomial in general is give the discussion prior to the problems. It follows that they have the same eigenvalues. For part (c), let $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ and $B = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$. Then $AB = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, while $BA = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$. These are not similar: the zero matrix is similar only to itself.

4. Let A be an $n \times n$ matrix over a field k of characteristic 0, and suppose that for every positive integer t, the matrix A^t has trace 0. Prove that A is nilpotent.

Let $a_1, \ldots, a_n$ be the eigenvalues of A. We shall show that if $a_1^t + \cdots a_n^t = 0$, $1 \leq t \leq n$, then all the eigenvalues are 0, which implies that A is nilpotent. Renumber the eigenvalues so that all that are nonzero precede all the ones that are 0. If only h are nonzero with $h < n$ then $a_1, \ldots, a_h$ satisfy the corresponding condition, and we have decreased the size of the problem. Therefore, we may assume that all the eigenvalues are nonzero. Now, the characteristic polynomial has the form $A^n + c_1 A^{n-1} + \cdots + c_{n-1}A + c_n I$ where the c_i are, up to sign (which alternates) the elementary symmetric functions of $a_1, \ldots, a_n$: in particular, $c_n = \pm(a_1 \cdots a_n)$. Now the trace map from $n \times n$ matrices over k to k is k-linear. Since the characteristic polynomial vanishes, and the traces of $A, \ldots, A^n$ all vanish, we get that $c_n I$ has trace 0, and this is $c_n n$. Since the field has characteristic 0, $n \neq 0$, and c_n is not 0 either. This is a contradiction.

5. Let M be the ring of all $n \times n$ matrices over a field. The group G of invertible matrices acts on M by conjugation. Show that a conjugate of a nilpotent matrix is nilpotent. When $n = 6$ count the number of conjugacy classes of nilpotent matrices, and justify your answer.

Since conjugation is a ring automorphism, $(BAB^{-1})^s = B(A^s)B^{-1} = 0$, and so if $A^s = 0$ then $(BAB^{-1})^s = 0$. The Jordan form is determined by the sizes of the Jordan blocks for the unique eigenvalue 0, and so the number of conjugacy classes is the same as the number of ways to write 6 as a sum of positive integers, ignoring order. These are $6 = 5+1 = 4+2 = 4+1+1 = 3+3 = 3+2+1 = 3+1+1+1 = 2+2+2 = 2+2+1+1 = 2+1+1+1+1 = 1+1+1+1+1+1$, so there are 11 conjugacy classes.

6. Let A, B be two square $n \times n$ matrices over a field of characteristic 0 satisfying $AB - BA = A$. Show that $A^kB - BA^k = kA^k$ for every positive integer k. Deduce from this that A is a nilpotent matrix. (The char. 0 hypothesis was omitted on the original exam. As an extra problem, show that this is false in char. 2 for $n = 4$.)

Solution. For the first assertion use induction. The case $k = 1$ is given. Assuming that $A^kB - BA^k = kA^k$, multiply on the left by A to get $A^{k+1}B - ABA^k = kA^{k+1}$, and substitute $AB = BA + A$ in the second term on the left to get $A^{k+1}B - (BA + A)A^k = kA^{k+1}$, which simplifies easily to the required result.

For the second part we give two solutions. Since A^kB and BA^k have the same characteristic polynomial, they have the same trace, which shows that the trace of $A^kB - BA^k$ is 0, and so the trace of every kA^k is 0. Since the characteristic is 0, each k is nonzero, and so one gets that the trace of every A^k is 0. One may then proceed as problem 4. above.

The second method is to define a linear transformation T from the vector space M of $n \times n$ matrices over the field to M by the rule $T(C) = CB - BC$. The fact established in the first paragraph shows that if $A^k \neq 0$, then it is an eigenvector of T for the eigenvalue k. In characteristic 0, the integers k are all distinct. Since a matrix of size n cannot have more than n eigenvalues, we get a contradiction unless $A^k = 0$ for some $k \geq 1$.

In char. 2 there is a 2×2 counterexample with $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ and $B = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$. This can be checked by an easy calculation. A size 4 counterexample can be obtained by using the direct sum of two copies of A for one matrix and the direct sum of two copies of B for the other.

7. (Not a QR problem.) Let T be a linear transformation over $\mathbb{R}$ on a finite-dimensional vector space V. Show that there is either a one-dimensional or a two-dimensional subspace of V that is mapped into itself by T. (One suggestion: over $\mathbb{C}$, T has at least one eigenvalue, and a corresponding eigenvector.)

Solution. We work with a square matrix A for T, which will have real entries. Over the complex numbers there is at least one eigenvalue $a + bi$, with a, b real, and at least one (nonzero) eigenvector $u + vi$, where u, v are real vectors. If u and v are linearly dependent, say both multiples of a nonzero vector real vector w, then $u + vi$ can be written as a complex number times w, and it follows that w is a real eigenvector of A. The line it spans will be mapped into itself. If u and v are linearly independent over $\mathbb{R}$, then the plane they span is mapped into itself: since $A(u + vi) = (a + bi)(u + vi) = (au - bv) + (av + bu)i$ and is also $A(u) + A(v)i$, we have that $A(u) = au - bv$ and $Av = av + bu$.