Tom Leinster Notes PhL 12.19.04

These are excellent notes on the subject of Minimal Polynomials and Jordan Canonical Form.

Def: an **annihilating polynomial** $p(\alpha) = 0$. In this section, think of α = a square matrix A. So this polynomial is like the K(A) = 0 result of the Cayley-Hamilton theorem.

Def: a **minimal polynomial** $m(\alpha)$ is *the lowest degree poly* that does $m(\alpha) = 0$. Clearly, $m(\alpha) \mid p(\alpha)$, which notation means that m divides evenly into p, no remainder. I have then stated Prop 2.

Prop 1: every matrix α has exactly one minimal polynomial. Proof is sketchy, seems reasonable.

On page 2, he writes the $K(\lambda)$ but his notation is $\chi_\alpha(t)$, which is of course $\det(\alpha - tI)$, secular. The r's you see here are the algebraic multiplicities. The claim is that the minimal polynomial looks similar, but each exponent is something in the range (1, alg mult), or as he writes it $1 \le s_i \le r_i$.

I am unsure how you would show that $m(\alpha)$ really is an annihilator. You have to build it and see if it is zero or not. That is, is the matrix $m(\alpha)$ equal to the zero matrix? It certainly seems reasonable that such a thing might exist that is just a piece of $K(\alpha)$. [I now know how to find $m(\alpha)$ for a matrix α, so this is not as confusing as on first reading.]

If matrix is diagonal, then $m(\alpha) = K(\alpha)$. This is because all powers are 1 and you cannot go any lower.

Lemma 4 is Sylvester's triangle law of nullity, and I now see that **Ker** is the way some people say Nullspace. He actually proves this law right in situ, but I ignore the proof.

How does he use this Lemma on the bottom of page 3 ? There are three claims made at the bottom

(1) the first line says dim(V) = dim Nullspace of (K(A)). We know that K(A) = 0 from C-H, so we know that K(A)x = 0 for all x in V, so that certainly means that the dim of this nullspace is V, so OK.

(2) Here he is saying that dim(V) is less than the sum of the dims of the nullspaces of each factor, and this is just Sylvester generalized to many factors, so this is fine too. Think of the RHS as being dim N_1 + dim N_2 + etc = the sum of dims of a bunch of nullspaces. The idea here is that N(ABC...) cannot exceed the sum of the dimensions of the individual nullspaces.

(3) In the last line, you can just re-express this as dim ($N_1 \oplus N_2$ ...) where you make a new space N = $N_1 \oplus N_2$ ... by just superposing the other nullspaces. So dim(V) ≤ RHS of last line. I think in the case that each null space has dimension 1, which is the case for diagonalizable A, then the ≤ is really = because we are adding up N numbers that are all 1.

Still I am confused as to what the point is here. It is the sentence stated at the top, but I lose it.

On page 4 we learn this his word for "similarity" is "conjugates", this is the first time I have seen this usage. M&M called them similarity of collinearity. We are reminded here that nearly everything is conserved under a similarity, and it is just like a change of basis.

On page 5 we get the definition of a Jordan block and the idea that we are going to write A in block diagonal form which is a direct sum (I think I conjectured this direct sum notation early on when first reading M&M, and here it is in the flesh).

Page 6 then states the Big Theorem without proof:

Big Theorem: Any matrix A can be converted by a similarity to Jordan Canonical Form. In this form, you have the direct sum of blocks. Here are the key results:

(1) the eigenvalues are on the diagonal and each eigenvalue appears a number of times equal to its algebraic multiplicity. Thus must be the case since $\det(J - \lambda I) = K(\lambda)$ = the usual thing.

(2) A degenerate eigenvalue λ_1 can have several Jordan blocks of various sizes. The size of the largest block in this group tells you the exponent to use for the factor $(\lambda - \lambda_1)^s$ in the *minimal* polynomial.

(3) The number of Jordan blocks for a degenerate eigenvalue λ_1 is the geometric multiplicity of that λ_1.

On page 9 we have a 6x6 example, and he goes through all the above items for this example.

Warning is issued: even if two matrices how all invariants the same, they may not be similarity conjugates.

Now back to page 6-7.

(1) This Jordan thing is MUCH stronger than the Schur theorem which just says you can transform into upper triangular with some similarity.

(2) The block structure is akin to the notion of breaking a number into powers of primes.

Exercise 1: Here we look at just one Jordan block that is d x d for some λ. The key point I learn here is that this matrix only has one eigenvector! So it has geomult = 1 by its construction. That eigenvector is just (1,0,0...).

Exercise 2: Here we look at the direct sum of just two blocks C and D. Most things are obvious, but item b is where we have something new. You want to construct the minimal polynomial for $C \oplus D$. Suppose these two matrix C and D have the same eigenvalue. Suppose $KC(\lambda)$ contains $(\lambda - \lambda_1)^3$ and $KD(\lambda)$ has $(\lambda - \lambda_1)^2$. Well, we know the overall $K(\lambda)$ will have $(\lambda - \lambda_1)^5$, but what about the min pol $m(\lambda)$? Each of the above factors has to divide evenly into $m(\lambda)$ otherwise you can show that m is not a min pol. This is why we get the notion that $m = (\lambda - \lambda_1)^3$ because we can do both $(\lambda - \lambda_1)^3 / (\lambda - \lambda_1)^3$ and $(\lambda - \lambda_1)^3 / (\lambda - \lambda_1)^2$ with no remainder, so we have to take the MAX of the two exponents.

Exercise 3: Now we look at an arbitrary set of blocks all for the same λ. We know that each block has geo = 1, so if there are g blocks, then geo = g. So boom, we see why the geomult is the number of blocks in the big Jordan form result. Secondly, the power in the min poly is again going to be the max of all the powers of $(\lambda - \lambda_1)$ for each block to get that common divisor effect as above. Of course now $K(\lambda)$ has lots of factors, and this comment applies only to the factor for λ_1 and its Jordan blocks.

Exercise 4: Now we go out to the full matrix and state the final results. Results are the same, they just apply separately for each λ_i manifold.

My Comments:
(1) It would be nice to see a proof of the big theorem.
(2) I don't understand why the minimal polynomial is important.