EE263 Notes PhL 1.7.05

This is a 10 page set of lecture notes from www.stanford.edu/class/ee263/jcf.pdf from Prof. Steve Boyd, see site http://www.stanford.edu/class/ee263/ for more stuff. This is a good place to go for lots of stuff. This was one of the first sets of notes I got for the Jordan stuff, and I did not understand them the first time, so now having done all the Matthews work, we can come back and look at what Boyd has to say.

Page 10-2 Here we get $J = T^{-1}AT$ and he shows the 1's of the J block on the upper off diagonal, whereas Matthews puts it on the lower. I am not much concerned about this transpose issue. He calls the Jordan block "upper bidiagonal". All comments on this page are now clear. When I first read it, I did not even know if the matrix J was the same dimension as matrix A. He refers to T as a similarity.

Why does he restrict to A having real matrix elements? I don't think this is necessary, but I agree that if you do, J can still be complex since eigenvalues can be complex.

Page 10-3. Comment that no one uses JCF for numerical work, only for exact work.

Then he writes the char poly for A with variable s instead of x or λ, and his n_i are my a_i, the char poly exponents which are of course the algemults.

Next, he comments that if all the $n_i = 1$, A can be diagonalized. I know this is true, A then has all eigenvalues different.

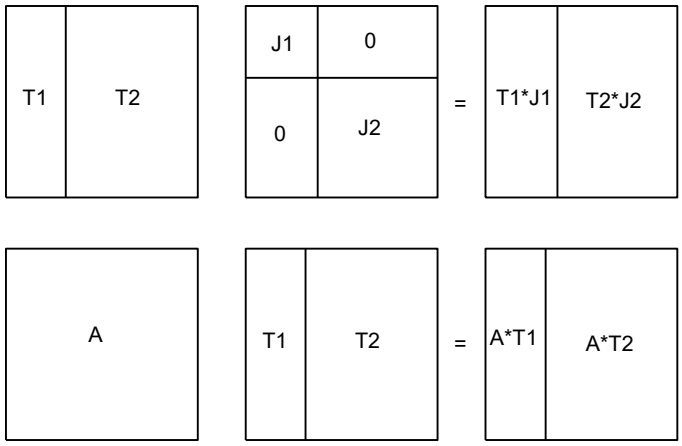
Next, dim Ker p = # Jordan blocks. I now understand this to be true since each Jordan block has geomult = 1, and dim Ker p = geomult of the entire A. What he really means is that dim Ker p_i is the number of Jordan blocks in the eigenmanifold $p_i^{a_i}$. If you then do this over all eigenmanifolds, then the total number of Jordan blocks in the final J is in fact Σ_i dim Ker p_i . So we tick him for vagueness on this point.

Now comes his famous "more generally" claim. He is talking about dim K p^k here for k = 1,2..., and these are the $\upsilon(p^k)$ of Matthews. Consider our Matthew's second example which we just wrote up with the Figure 1 Visio picture. In that case we had $\upsilon(p^3) = 6$ and $n_i = 6$ for our one eigenmanifold. So in this case, his equation is wrong. He says dim K $p^3 = \min(3,6) = 3$. Maybe on the left he means the sum over all such spaces, in which case his LHS = Σ_i dim Ker $p_i^k = \Sigma_i \upsilon(p_i^3)$ but this is nothing useful to anyone! So I think our guy has screwed up here. He comments about the "sizes of the Jordan blocks". I know these are the heights of the columns in the dot diagram. Well, it is true that we have $\upsilon(p^k)$ = dim Ker p^k and then from these numbers we can compute the $\upsilon_{k,p}$ numbers using the subtraction rule, and from those we can get the dot diagram, and from that we can get the sizes of the Jordan blocks. But we do not have the simple claim he makes that $\upsilon(p^k)$ = some expression. I cannot even ascribe any meaning to his summation notation. But the general idea is that from the $\upsilon(p^k)$ you can figure out the sizes of the Jordan blocks, that claim is true.

Next, he claims that $T\, p_i(A)\, T^{-1} = (J-\lambda_i)$ which I know is the companion matrix. He then shows that as you take higher powers of this matrix, it soon becomes all zeros. This is just the idea that the min poly of a C matrix (and of the J matrix) is the full dimension of the matrix. In his case, the min poly is of degree 3 and we can say $C^3 = 0$.

His last statement concerns an off diagonal power $(J_j - \lambda_i)^k$. I could verify this I suppose by brute force computation, but I don't see any use for the result, so I won't try to verify it.

Page 10-4 Now we are going to examine the matrix T which Matthews calls P. Instead of expressing T in columns, he puts it into groups of columns called T_1. Well, here is an interesting fact to keep in mind:



So if we are thinking about $A\,T = T\,[\,J_1 \oplus J_2\,]$, we can break this into two separate problems:

$AT_1 = T_1J_1$ and $AT_2 = T_2J_2$

so I agree with his general breakdown into $AT_i = T_iJ_i$. Then you take one of these groups of columns and say this:

$T_i = [\, v_{i1}\, v_{i2}\, \ldots\, v_{i\,ai}\,]$ where a_i = his n_i = the algemult of the J block i.

These v's are then single column vectors. We then get his nice result that the first column (only) of one a T_i group is a true eigenvector. The other columns have a slightly different equation and they are called the generalized eigenvectors. I have derived all this stuff. Now in Matthews, we got a similar situation, but there the generalized eigenvectors formed the cyclic basis of a T cyclic subspace for J.

Well, I think I can make the connection as follows. His general equation is this:

$(A - \lambda_i)\, v_{i\,j} = v_{i\,j-1}$ $\quad j > 1$ up to $j = a$ $\quad \Rightarrow p_i\, v_{i\,j} = v_{i\,j-1}$

so we have: $p_i\, v_{i\,2} = v_{i\,1}$, $p_i\, v_{i\,3} = v_{i\,2}$, $p_i\, v_{i\,4} = v_{i\,3}$ .... $p_i\, v_{i\,a} = v_{i\,a-1}$

which I can rewrite as (for the case a = 4)

$(p_i)^3\, v_{i\,4} = (p_i)^2\, v_{i\,3} = (p_i)^1\, v_{i\,2} = v_{i\,1}$

$(p_i)^2\, v_{i\,4} = (p_i)\, v_{i\,3} = v_{i\,2}$

$(p_i)^1\, v_{i\,4} = v_{i\,3}$

$(p_i)^0\, v_{i\,4} = v_{i\,4}$

So here we see the Matthews idea that you get the basis vectors of the J_4 space by applying powers of p to a generator vector which here is v_{14} , and the highest power $(p_i)^3\, v_{i\,4}$ then gives you the true eigenvector

$v_{i\ 1}$, and all the other powers give you those generalized eigenvectors, so we are on the same page! What you would really like is to first find v_{14}, then generate the various other guys by applying powers of p.

Page 10-5 Now suddenly we are looking at a Linear Differential System LDS $dX/dt = AX$ where A is a matrix. This was also touched upon in Matthews. Here Boyd suggests doing a change of basis as shown to get a much simpler problem in which the matrix J is used. Let's look more into this:

$$dX/dt = AX \Rightarrow T^{-1}dX/dt = T^{-1}A\,T\,T^{-1}X \Rightarrow dX'/dt = [J_1 \oplus J_2]X'$$

We can now break our tall vector X' into $X' = [.....]^t = [\ x_1'\ \ x_{2'}\ ]^t$ and the problem then decouples into this set of smaller problems:

$d\mathbf{x'}_i/dt = J_i\ \mathbf{x'}_i$ // just as he says.

Now, the i index tells you which Jordan block. Let's look just at a single Jordan block for some fixed i, and also let's get rid of the primes for now, knowing where we are. Then we have

$d\mathbf{x}/dt = J\ \mathbf{x}$

and if we assume Laplace $s = i\omega$ sine action, we can write this as

$J\ \mathbf{x} = s\ \mathbf{x}$

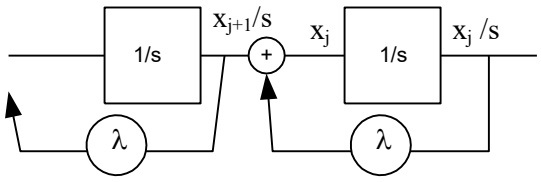
But remember that $J = C + \lambda I$ and that $C_{jk} = \delta_{j+1,k}$ so we can write our equation as

$$(C + \lambda I)\ \mathbf{x} = s\mathbf{x} \Rightarrow \Sigma_k(\delta_{j+1,k} + \lambda\delta_{j,k})\ x_k = sx_j$$

$$\Rightarrow x_{j+1} + \lambda\ x_j = s\ x_j$$

$$\Rightarrow x_{j+1}/s + \lambda\ x_j\ /s = x_j$$

and we then get this **Jordan chain** picture



$x_{j+1}/s + \lambda\ x_j\ /s = x_j$

where we have one "stage" for each component of the vector x. Just think of the 1/s box as an integration operator if you like. So within each eigenmanifold, you can solve your system in this manner, then combine all the solutions into the full X' vector, then transform back to the original space. Lot's of work.

Page 10-6 Resolvent, Expo of J. Now we look at the meaning of p^{-1} . This is a matrix. It is the inverse of p, and we find a simple closed form result for p^{-1}. When we write p^{-1} p = I, we have product of two triangulars is a triangular. He defines F_i to mean a diagonal of 1's and this allows the nice form shown for p^{-1}.

He next shows that you can compute $\exp(tJ_\lambda)$, and his result is like Matthew's result on page 83, except we have the transpose issue in the definition of J.

What is the "resolvent"? It is exactly the inverse of p, something I am tempted to call a propagator. But resolvent is a standard name for the inverse secular operator.

Page 10-7 Generalized modes. We return to our system of first order ODE's as shown. We can solve it as shown on page bottom using the expo stuff (as Matthews did). If we pick the special t=0 starting vector as shown at the top of the page, then for all later time, the vector x(t) is controlled only by one of the Jordan blocks of our matrix A. You think of this as if it were a vibrational normal mode. Then at the bottom you see that with a general starting vector x(0), you have activity in all the "modes".

I can see that this is a whole new subject that one could study, I never heard of it before.

Page 10-8 Here he simple states the C-H theorem.

Page 10-9,10 The corollary here is that or arbitrary powers of A, you have a basis that is a finite number of powers, as determined by C-H (or perhaps by the min poly, but that is not mentioned). He directly proves this little corollary based on the C-H which makes $\chi(A) = 0$. If A is invertible, the same claim of finite basis applies to negative powers of A.

Page 10-11. His proof of C-H. He first assumes A can be diagonalized into Λ, then C-H is quite easy to show since all matrices are diagonal. Each one has a 0 somewhere on the diagonal and the effect of all the zeros is to kill the result so $\chi(\Lambda) = 0$, so that means $\chi(A) = 0$. More generally, we can get A into J form. Then you can show that $\chi(J_i) = 0$ for each i as shown, and then the direct sum business tells you that the full $\chi(J) = 0$, and so $\chi(A) = 0$. This is a nice proof I think! But you have to use the Jordan form, so it depends where you are starting your circle of logic.

Comments: This is a nice little lecture I think, lots of good topics are reached. He presents a direct method for computing the matrix T. You cannot say that he "proves" the Jordon Form in the sense that Matthews does. He assumes it, and then computes T with that assumption. He then touches on these applications: system of first order DE's, resolvent, generalized modes, and the C-H idea and what it implies.

****

Comments on the EE263 Lecture 10 Document

These are my original notes on this lecture when I knew very little about the subject.

These notes add a small amount to my picture which is still very incomplete. Page 2 states the Jordan block form which I now see better. At first I thought maybe the T were non-square, but not so. So page 2 is just fine.

Page 10-3

Page 3 is full of mysteries. I agree with the fact that dim N (λI-A) = geometric multiplicity of the manifold associated with eigenvalue λ. And I know from the above notes that this equals the number of Jordan blocks for this λ (but I don't know why that should be so). The next line where author says "more generally" -- this means nothing to me at all. I am not grokking a major point here, and author is not helping me at all with words.

Now I do know from the TL notes above that in K(A) we have factors like $(\lambda I - A)^r$ where r is the algebraic multiplicity. And I know that there is a thing called "the minimal polynomial" which has the same factor but possibly raised to a lower power $(\lambda I - A)^s$. In the TL notes page 8 I agreed with the fact that this exponent s must be max of the sizes of all the Jordan blocks associated with this particular λ. But this does not seem to relate.

Recall from earlier work that Ker(A-Iλ) is the nullspace whose dimension m is the geomult of this eigenvalue λ, the number of different eigenvectors. We are somehow breaking this nullspace down into m distinct nullspaces, each of which has geomult =1. This is exactly TL's Exercise 3. And we also know that the sum of the dims of these subspaces is r, the algemult. So we are doing a factorization like so:

$$(\lambda I - A)^6 = (\lambda I - A)^1 (\lambda I - A)^2 (\lambda I - A)^3 \qquad 6 = \text{algemult} \quad 3 = \text{geomult}$$

and we associate each factor with one of these new nullspaces.

AHA...! (as one says). We know that the K(A) equation is invariant under similarity, that is very easy to show and I have done it. So consider the above product in the language of J

$$(\lambda I - J)^6 = (\lambda I - J)^1 (\lambda I - J)^2 (\lambda I - J)^3 \qquad 6 = \text{algemult} \quad 3 = \text{geomult}$$

Suppose for the moment that the entire J has only this one eigenvalue λ, but there are three Jordan blocks which we will call J = diag(J_1, J_2, J_3) which are 1,2 and 3 in dimension. In the above equation, I and J are both in block diagonal form, so this means we have a direct sum view and we can look only at a particular block, in which case we get this:

$$(\lambda I - J_i)^6 = (\lambda I - J_i)^1 (\lambda I - J_i)^2 (\lambda I - J_i)^3 \qquad 6 = \text{algemult} \quad 3 = \text{geomult} \quad i = 1,2 \text{ or } 3$$

Our char equation is just $(\lambda I - J_i)^6 = 0$ for our hypothetical matrix. But within each nullspace, we want to have $(\lambda I - J_i)^{dimNS} = 0$. This will then say that $(\lambda I - J_i)^{dimNS}\, x = 0$ for all x, so then

$$\dim\text{Ker}\,[\,(\lambda I - J_i)^{dimNS}\,] = \text{dimNS} \qquad \text{write as} \qquad \dim N\,(\lambda I - A)^k = k$$

Now our page 3 notes show that, with J_i defined as it is (the Jordan Block), we do in fact get $(\lambda I - J_i)^{dimNS}=0$. On each power, the ones slide up one diagonal, and then they slide off, just the way my little UH_n theorem says they should. So this at least motivates the form of the Jordan Block. On the right above, we have something that at least resembles what is on page 10-3 of the notes. Now we could never talk about dimNS being larger than the algemult of that λ.

Now in a more general matrix with several distinct λ_i , we do get the "off diagonal" powers as shown on the bottom of page 10-3. I have not checked the matrix on the right, but I can see it will be something that is non-zero.

Page 10-4

Now we get another small iota of insight. If we simply assume that some T exists to convert A to J, what can we say about this matrix T? Well, we certainly can say AT = TJ, the way we used to say $A\Lambda = T\Lambda$ where Λ was a diagonal matrix. Because J is in block diagonal form, we can write this as $AT_i = T_iJ_i$ where the J_i are the small square Jordon blocks, and the T_i are groups of column vectors of T which we can associate with that J_i . I have shown graphically how this works on that page. So we have a decoupling of the problem somewhat.

We then label the columns in each T_i . I like to say this as $T^{(i)} = [\ \mathbf{v_1^{(i)}}, \mathbf{v_2^{(i)}}, \mathbf{v_3^{(i)}}, \ldots \mathbf{v_N^{(i)}}\ ]$ We need to maintain some kind of label "i" to say which T_i set of columns we are talking about. But then I leave off the i label and just think about ONE of these T blocks. We then find these interesting facts:

(1) The first column $\mathbf{v_1^{(i)}}$ satisfies $A\ \mathbf{v_1^{(i)}} = \lambda_i\ \mathbf{v_1^{(i)}}$ and so it is a true eigenvector of λ_i. Remember that each Jordan block only has one true eigenvector, and this is it!!!

(2) The other columns satisfy equations like $A\ \mathbf{v_3^{(i)}} = \lambda_i\ \mathbf{v_3^{(i)}} + \mathbf{v_2^{(i)}}$. This is *close* to an eigenvector equation, *but* we have the extra term on the right which is the previous column! We know that these other columns cannot be true eigenvectors since each block has only one, so these other columns are called the generalized eigenvectors. They are some kind of dummies to fill out the T block. It is not clear to me looking at this equation $A\ \mathbf{v_3^{(i)}} = \lambda_i\ \mathbf{v_3^{(i)}} + \mathbf{v_2^{(i)}}$ whether or not there is a solution. I guess if there is not, that determines the width of the T block, which is to say, that determines the size of this Jordan block.

Page 10-8

I have to skip other pages because they are specialized to the application of this EE263 course at Stanford, but the last two pages 8 and 9 are good. He first makes a "statement" about what you mean by a poly with matrix coefficients, then he states the C-H theorem and gives a nice 2x2 example of it.

Page 10-9

The first part here is the "corollary" to C-H which says any power A^k can be written as a lincom of powers of A up to N-1, which I agree with. The second part is basically the same proof of "the corollary" that I put into my "matrix research" notes, but I think my details are much better, his is a little fudgy. Of course this proof depends on the C-H theorem.

Page 10-10.

This is a way to show that A^{-1} can also be expanded in the same lincom of powers of A, hence we can do any negative powers as claimed earlier if A is invertible.

Page 10-11

Now comes this guy's "proof" of the C-H theorem itself. It is all on the very bottom one line. It assumes that we can go to Jordan form, and then it uses the idea that the powers of the matrices vanish if you pick out the right one for each eigenvalue. Notice that the n_i here are NOT the algemults, they are the Jordan

block dimensions. So he uses Jordan decomposition to prove C-H, which is if nothing else interesting. Again, the fact that the secular equation is true in A or J is used.

I think I have now squeezed the most juice that I can out of each of these 10-page papers which I have printed and written on. But we need more info!