**The Block-Diagonal Form Theorem** PhL 12.10.04

This document contains notes which relate to pages 320 and 321 of M&M's book. Their statement of the theorem is the last equation on page 321. Their derivation is a bit sketchy, shall we say.

**Overview.** I now realize that M&M are really describing the Schur Decomposition, the idea that you can transform any matrix A to upper triangular form by a unitary similarity. This is not block diagonal form, and when you go the extra steps to block diagonal form, your overall similarity loses it unitarity.

Here is what is done below:

(1) Digression. We examine the geometric multiplicity of a few sample triangular blocks, and show that you generally do not have geomult = algemult. We show how rank and nullity work out correctly.

(2) We then consider now we might use (non-unitary) similarities to "clear out" regions of the fully triangular Schur matrix to get it into block diagonal form. This is a very messy discussion. I of course know this can be done from the Primary Decomposition Theorem which gives us a direct sum of the appropriate nullspaces, but here we are staying very mechanical.

(3) We state the Block Diagonal Form theorem. Note that Allen's notes have a similarly messy proof to mine.

(4) We apply my block-diagonalizing process to the simplest possible non-trivial case, a 3x3 triangular matrix where one eigenvalue is repeated twice. We show exactly in this example how we clear out the area we want to clear out, and we find that the similarity to do so is definitely non-unitary.

(5) We end this paper with some comments and a web quote of the Schur Decomposition and a few references to books.

**(1) Digression: A triangular eigenmanifold matrix generally does not have full geomult.**

As our first problem, consider a matrix which has one degenerate eigenvalue appearing N times. Suppose we apply the process that M&M describe in their chapter 10 section 10.15. We end up with our matrix A appearing as upper triangular as in 10-41 but this is our entire matrix A for our special case. We were able to produce this triangular matrix by applying a series of N successive rotation similarities to the original matrix A. The first rotation lined up the first eigenvector x with $e_1$, then the second rotation lined up the second eigenvector with $e_1$ in the N-1 space, and so on. The total rotation to get to the form 10-41 is the concatenation of all these rotations. Here is a 3x3 **example** where we assume we have gotten to the triangular form and $\lambda = 1$ and algemult = 3:

$$A = \begin{pmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{pmatrix} \qquad \text{and } A - \lambda E = \begin{pmatrix} 1-\lambda & a & b \\ 0 & 1-\lambda & c \\ 0 & 0 & 1-\lambda \end{pmatrix} \quad \text{secular says } (1-\lambda)^3 = 0 \text{ so } \lambda = 1 \text{ triple}$$

If you do the determinant by working down the first column, you see that the numbers a,b,c play no role in the secular equation. If we install the eigenvalue, we get

$$A - 1E = \begin{pmatrix} 0 & a & b \\ 0 & 0 & c \\ 0 & 0 & 0 \end{pmatrix}$$

If a and c are both nonzero, then we have rank 2, otherwise we have rank 1. In these two cases, we expect nullity = 1 or 2, so we expect either 1 or 2 eigenvectors. If b=0 as well, then we have rank 0 and we expect all 3 eigenvectors (which will be the trivial ones since then A = identity).

Now let's try to find the eigenvectors. Applying A to column vector (x,y,z) gives this requirement:

| | | |
|---|---|---|
| x = 1*(x + ay + bz) | => | ay + bz = 0 |
| y = 1*( y + cz) | => | cz = 0 |
| z = 1*( z) | => | nothing |

If a,b,c are all nonzero, we need z = 0 and y = 0 so our eigenray is (x,0,0) and we have only the eigenvector (1,0,0). So we ALWAYS will have this eigenvector regardless of the off-diagonal elements.
Now suppose c = 0 and a,b≠0. Then we can have z = -(a/b)y, and any x, so ray is (x,y,-(a/b)y). We can think of this situation as (1,0,0) and (0,1,-(a/b)), and we have two basis vectors. This agrees with the rank analysis which says if c = 0, then rank = 1 so nullity = 2.
Suppose b = 0 and a,c ≠ 0. Then we need z=0 still and y=0 too, so only have the (x,0,0) solution, again consistent with rank = 2 and nullity = 1.

So what can we see in the general NxN case of a degenerate eigenvalue? Let's just go up one more dimension and write things like so:

$$\begin{pmatrix} \lambda & a & b & c \\ 0 & \lambda & d & e \\ 0 & 0 & \lambda & f \\ 0 & 0 & 0 & \lambda \end{pmatrix}$$

The equations will look like this in general when applied to (x,y,z,t)

| | | |
|---|---|---|
| λx = (λx + ay + bz + ct) | => | ay + bz + ct = 0 |
| λy = ( λy + dz + et) | => | dz + et = 0 |
| λz = ( λz + ft) | => | ft = 0 |
| λt = ( λt) | => | nothing |

As before, the vector (x,0,0,0) is a solution. Other solutions depend on the values of the parameters.

So in the general case, we always have (1,0,0,0....) as a guaranteed eigenvector, but no other ones are guaranteed! In general, if we try to make a matrix U of the eigenvectors, we will have to either put in columns of 0's, or perhaps repeat earlier column good eigenvectors. In either case, detU = 0 and we cannot make $U^{-1}$ as before, so we cannot diagonalize A.

Basically, there is nothing we can do, we just have to accept the "box form". ( Well, we could continue on to the Jordan form. )

**(2) The Problem of "clearing out" the regions outside the diagonal blocks.**

Suppose we have a 6x6 matrix A which has the above Schur triangular form, but which has diagonal elements like this:

$$\begin{bmatrix} \lambda_1 & x & x & x & x & x \\ & \lambda_2 & x & x & x & x \\ & & \lambda_3 & x & x & x \\ & & & \lambda_4 & x & x \\ & & & & \lambda_4 & x \\ & & & & & \lambda_4 \end{bmatrix}$$

We got to the form shown here following the M&M procedure (which is the Schur procedure).

Now recall an earlier theorem which says that if you have an eigenvalue problem $Ax = \lambda x$ with all eigenvalues different, then the eigenvectors are linearly independent, and you can combine them to make a matrix X which fully diagonalizes your matrix A to $\Lambda$. We want now to apply this idea to the 4x4 matrix P shown in the box above. It is already "partially diagonalized" by the M&M procedure, but the theorem should still apply to this partially diagonalized version of A. However, the 4x4 matrix X which diagonalizes the box outlined above will NOT be unitary, see Clarification document.

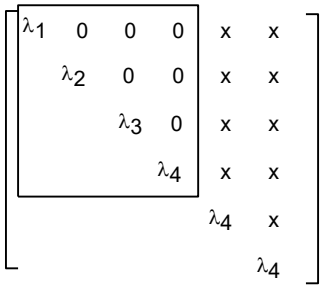
We got to the point shown above with nothing but rotations, so the triangular 4x4 matrix is I think connected to the original 4x4 matrix by a pure rotation. Now let's call the 4x4 matrix above by the name P for Partial diagonalized. Then we know that $P = R^{-1}AR$ where $R^{-1} = R^T$. If it happened that A were symmetric, then we can trivially show that $P^T = P$. This can only be so if ALL the off-diagonal elements are 0 in our 4x4. That is how symmetry of A makes it into our triangular world. In this case, the 4x4 would already be diagonal, and we would not have to do the transform we are about to do. So let's assume that A was not symmetric and proceed.

We have constructed our little 4x4 similarity X as described above. Since A is not symmetric, we know that this X will generally not be a rotation, it can have some "shear" component to it, stretching of some axes differently than others. If X were a pure rotation, then the TOTAL transformation for the 4x4 block would be XR = rotation. We know, however, that the TOTAL transform is only a rotation if the starting A is symmetric. We will see in our example below an explicit calculation of X and we will see that it is not a rotation. In any event, this 4x4 X similarity will take the 4x4 above to diagonal form, and of course the diagonal values stay just as they are, the upper triangle clears out.

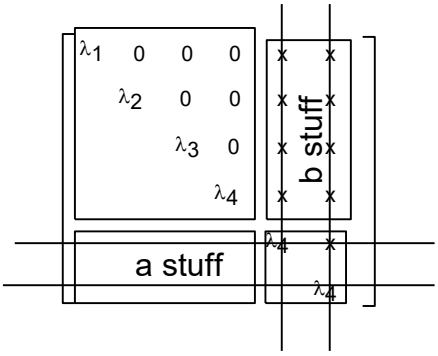
Suppose we then make a 6x6 matrix Q = diag(X,1). This is certainly invertible since $Q^{-1} = \text{diag}(X^{-1}, 1)$. Also I know that det(Q) = det(X). We apply this similarity Q to the *entire* 6x6 matrix, and what happens? Here is the matrix processing,

$$\begin{bmatrix} X^{-1} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} A & b \\ a & C \end{bmatrix} \begin{bmatrix} X & 0 \\ 0 & E \end{bmatrix} = \begin{bmatrix} X^{-1}AX & X^{-1}b \\ aX & C \end{bmatrix}$$

Our picture now looks like this,



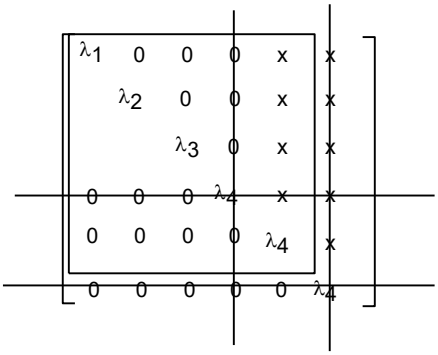

The similarity Q we did in the full space is "a similarity transformation that leaves the 5 and 6 axis unchanged." If you applied Q to a vector (x1,x2,x3,x4,x5,x6), you would get (x1',x2',x3',x4',x5,x6). You don't get this situation in 3-space, but very common for N-space. Here is a repeat of the picture above showing lines for the non-changing axes,



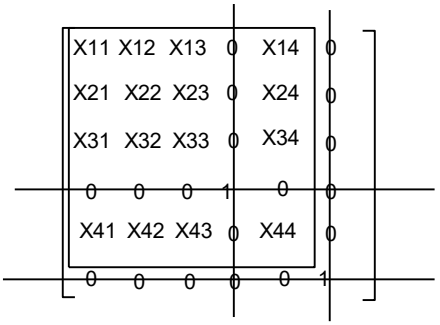

What did this similarity do to our A matrix? It diagonalizes the square section we want, it leaves the 5-6 part called C the same (lower right corner), the lower left portion "a" stays 0, but it "messes up" the upper right corner called b. I could have put x' instead of x in the above picture for the 8 x's up there in top right, since they got changed.

Notice carefully where the a and b portions are located in the above picture. The b stuff is on the V lines but not on the H lines, those 8 x's. The a stuff is on the H lines but not on the V lines. In the above picture we have all zeros in that lower left rectangle. Since this is the case, aX in the resulting matrix is all 0.

Now we want to apply a similarity to the above result which keeps the 4 and 6 axis unchanged. This is a little hard to draw, but let's give it a try: (note the 8 b-stuff items, and the 8 a-stuff items)




In the corresponding similarity matrix, you imagine $X_{ij}$ values sitting in the 16 locations in the large box excluding elements on the cross-out lines. We better go draw this matrix to avoid confusion:




This is just like our previous diag(X,E) transform, but it is much harder to visualize because the E is now split apart and so is the X part. There might be some cleverer way to handle this by swapping the last rows and columns, rotating with diag(X,E), then swapping back, but I don't know how to do that.

Now back to the previous picture, remember where to look for the a stuff and the b stuff. Notice that the b area now contains 7 zeros and a $\lambda_4$ whereas last time it was all zero. We need to see what happens to the a and b areas in the similarity.

The X similarity matrix we are looking to use has this form

$$X = \begin{bmatrix} 1 & 0 & 0 & X \\ 0 & 1 & 0 & X \\ 0 & 0 & 1 & X \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Remember that it contains the eigenvalues of our little 4x4 (non-crossed-out) eigenvalue problem, and for the first three eigenvectors we already have unit vectors. The fourth column is some unknown fourth eigenvector, but we do know that it is linearly independent of the first three columns because all four of our $\lambda$'s are different. We know it must have a non-vanishing 4th component to be linearly independent, so we normalize this eigenvector so that 4th component is 1.

Now notice this product: (L4 means $\lambda_4$)

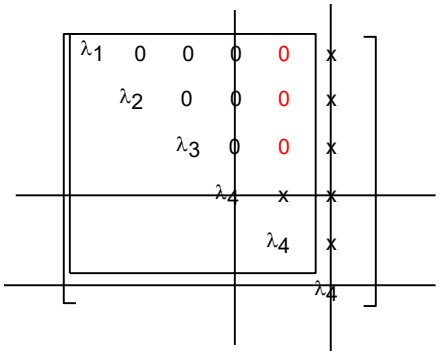
$$\begin{bmatrix} 0 & 0 & 0 & L4 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & X \\ 0 & 1 & 0 & X \\ 0 & 0 & 1 & X \\ 0 & 0 & 0 & 1 \end{bmatrix} = aX = \begin{bmatrix} 0 & 0 & 0 & L4 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

We had to do all this work to show that the lower part of our processed A matrix is not going to change!

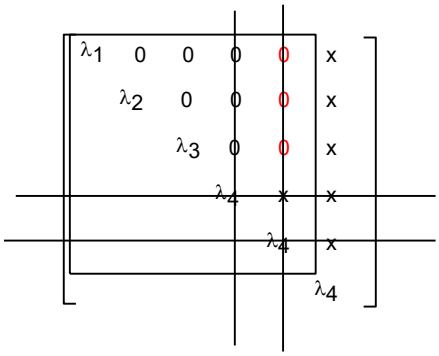
Now what about the $X^{-1}b$ piece? The b object contains the matrix elements on the vertical lines that are not on the horizontal lines, so this is a rectangle 4 high and 2 wide. When we multiply it by $X^{-1}$ based on X above, the result is that the left column is all zero, and the right column is all x.

The third thing we need to remember is that our rotation here is going to clear out the three x's between the vertical lines above the top horizontal line. Again, this is not generally a "rotation".

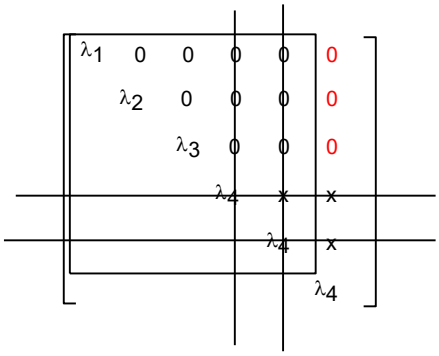
If we put all these results into our picture, here is what we have after this second X rotation:




As you might guess, the next step is to do a 4-5 rotation,




We have a very similar situation to last time. We construct a new X that will clear out the top three x's in the rightmost column. It has the same structure as our previous X, but likely a different fourth row. Our object b and a are similar to last time, though slightly different. We should carefully verify that they don't cause trouble, but I just assume this works because I am getting tired of all this. Then finally we do our new X rotation and clear out the top three x's and we get this result:




This is the result we have been seeking. We had to do three sequential similarities about different axis-pairs to get the upper right corner region "cleared out". We now have our matrix A in the famous "box diagonal form" which M&M claimed was possible. Of course now we can imagine that this all happened with a single similarity X which is the concatenation of the three X's we just described.

There are a few loose ends still.

One issue is that in the $\lambda_4$ subspace maybe we have less than three eigenvectors. If that is the case, then we just replicate existing ones to fill things out. We know for sure that there is always one eigenvector and in the 3x3 subspace it has unit vector (1,0,0).

The other issue is all the work we had to do above. We wanted to show that you could find similarities that would do the job. But we know that the eigenmanifolds are all disjoint, each one forming a closed subspace. And we know that "somehow" we can get to a basis where the manifolds are decoupled and the basis vectors for a given manifold take the simple form that none have non-zero components in any other manifolds. This is related to the idea of clearing out the far corners of the A matrix (even though this matrix is not quite a picture of the eigenvectors themselves). I don't have a good proof for how this is generally done, but I could look up the place that M&M quoted their results and probably find out. { Matthews notes add a lot to this story! }

I am reminded of the subject of angular momentum. We have a similar problem there, in that we start out with something like this: ( R is a rotation operator) (thinking j= 1 here)

$$R_{1m,\ 1m;} = \langle 1,m| R |1,m'\rangle$$

which is a 9x9 matrix of numbers. We know that we can get this thing into "block diagonal form" in a direct sum sense,

$$R = R^{(2)} \oplus R^{(1)} \oplus R^{(0)} \qquad\qquad 1 \otimes 1 = 2 \oplus 1 \oplus 0$$

where $R^{(2)}$ is a 5x5 matrix, $R^{(2)}$ is 3x3 and $R^{(0)}$ is 1x1. So if we drew the original 9x9 matrix, we could put these blocks on the diagonal. The "rotation" X that gets you go block-diagonal form is the Clebsch-Gordon coefficients. I think the very same thing is going on here.

So here is our final theorem on this subject:

**(3) The Block Diagonal Form Theorem**: Any matrix A with its spectrum of eigenvalues and algebraic multiplicities can be brought completely into block-diagonal form by some (generally non-unitary) NxN similarity transformation X. The eigenvalue spectrum can be obtained from the secular equation $\det(A - \lambda I) = 0$. All blocks have the upper triangular form shown in M&M 10-41 where the eigenvalue for the block appears on all the diagonal elements of the block. Obviously when algebraic multiplicity = 1, the block is 1x1 and consists of nothing but the eigenvalue. The blocks are of course all square, being m x m where m is the algebraic multiplicity of that eigenvalue. The issue of the geometric multiplicity within each block can be studied locally within that block, and as we know, it can range anywhere from 1 up to the algebraic multiplicity for that block.

When each eigenvalue occurs only once, we know that the (generally non-unitary) transformation X is equal to the matrix you get by assembling the eigenvectors of $Ax = \lambda x$. When there is degeneracy, we know it is certainly possible to *not have* N linearly independent column vector eigenvectors, so in such a case, X still exists, but it cannot possibly be the same matrix you get by assembling eigenvectors. We can still associate the non-degenerate columns of X with the singlet eigenvalues (they have simple column unit vectors matching the location of $\lambda$ in the matrix), but the columns which go with a degeneracy box are hard to interpret { they are generalized eigenvectors of the $\ker(p^b)$ space! }.

Finally, if matrix A is complex, I think everything still "goes through". You can still find "rotations" that rotate a complex vector to (1,0,0....) in the M&M program, but of course these are really complex rotations that are unitary transformations. In the end you will get some complex X that will put the matrix

A into the block-diagonal form we have been talking about here. I have not looked at this in detail, but I presume it is true. And of course if eigenvalues are singlets and A is Hermitian, then X is unitary.

**(4) A 3x3 Example where we explicitly construct similarities to "clear out" the "out of block" area.**

This is an example of the general procedure outlined above. We start with this matrix, which is about the simplest one you can think of that has $\lambda$ degeneracy and is more than 2x2:

$$A = \begin{pmatrix} \lambda_1 & a & b \\ 0 & \lambda_2 & c \\ 0 & 0 & \lambda_2 \end{pmatrix}$$

With an eye to perhaps clearing away the "a" element, we consider the upper left 2x2 matrix:

$$B = \begin{pmatrix} \lambda_1 & a \\ 0 & \lambda_2 \end{pmatrix} \qquad \text{eigenvectors are:} \qquad \begin{pmatrix} 1 \\ 0 \end{pmatrix} \text{ for } \lambda_1 \qquad \text{and} \begin{pmatrix} \alpha_a \\ 1 \end{pmatrix} \text{ for } \lambda_2 \quad \text{where } \alpha_a = a/(\lambda_2 - \lambda_1).$$

By doing a similarity with $S = \begin{pmatrix} 1 & \alpha_a \\ 0 & 1 \end{pmatrix}$ [ inverse $\begin{pmatrix} 1 & -\alpha_a \\ 0 & 1 \end{pmatrix}$ ] , we get $S^{-1} B S = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$, so we can successfully diagonalize this part of the matrix. Since B is not symmetrical, we know this is not really a "rotation", ie, S is not real orthogonal, though det=1. You can see by acting on (x,y) that the matrix S causes "shear" by treating axes differently. Let's show this here for the $\lambda_1$ eigenvalue:

$$\begin{pmatrix} \lambda_1 & a \\ 0 & \lambda_2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \lambda_1 \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \lambda_1 x + ay \\ \lambda_2 y \end{pmatrix} \quad = \text{looks like shear, not rotation}$$

Notice that we constructed S in the official fashion, by making its columns be the eigenvectors of our problem $BS = S\Lambda$, and we could do that since the eigenvalues are different.

We then extend these matrices to 3D and say (cA stands for a matrix to clear out "a")

$$cA = \begin{pmatrix} 1 & \alpha_a & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{and its inverse } cAI = \begin{pmatrix} 1 & -\alpha_a & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

When we compute A1 = cAI * A * cA, we get $A1 = \begin{pmatrix} \lambda_1 & 0 & b' \\ 0 & \lambda_2 & c \\ 0 & 0 & \lambda_2 \end{pmatrix}$ where $b' = b - c\alpha_a$. As expected, a is now cleared out, but b has been altered to b'. We next attempt to clear out b' using this matrix (cB means clear b')

$$cB = \begin{pmatrix} 1 & 0 & \alpha_{b'} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{and its inverse } cBI = \begin{pmatrix} 1 & 0 & -\alpha_{b'} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Then we find that A2 = cBI * A1 * cB $= \begin{pmatrix} \lambda_1 0 & 0 \\ 0 & \lambda_2 c \\ 0 & 0 \lambda_2 \end{pmatrix}$. We have checked this with Maple, see program stored in this local directory. We conclude then that the total similarity which clears out both a and b is this:

$$X = cA * cB = \begin{pmatrix} 1 & \alpha_a & \alpha_{b'} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad X^{-1} = \begin{pmatrix} 1 & -\alpha_a & -\alpha_{b'} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \text{ and Maple confirms that } A3 = X^{-1}AX = \begin{pmatrix} \lambda_1 0 & 0 \\ 0 & \lambda_2 c \\ 0 & 0 \lambda_2 \end{pmatrix}.$$

Notice that X has det=1, but $X^{-1} \neq X^T$, so this thing is not real orthogonal, that is, X is not formally a "rotation", it is just a similarity. The transformations that move you from the original A matrix (whatever it was) to the triangular form we show here for A *were* in fact all rotations (these were done in M&M) but the clearing out ones here are both "shear-like" and so the overall X is not a rotation. [ I think that if we had started with an original symmetric A, we would have both a = b = 0, so we would not be doing these shear transforms. ]

Recall now that A = $\begin{pmatrix} \lambda_1 a & b \\ 0 & \lambda_2 c \\ 0 & 0 \lambda_2 \end{pmatrix}$ and X = $\begin{pmatrix} 1 & \alpha_a & \alpha_{b'} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$. It would seem that the first column vector of X which is (1,0,0) goes with the singlet $\lambda_1$. What are the eigenvectors we expect for the lower 2x2 problem?

$\begin{pmatrix} \lambda_2 & c \\ 0 & \lambda_2 \end{pmatrix} \begin{pmatrix} y \\ z \end{pmatrix} = \lambda_2 \begin{pmatrix} y \\ z \end{pmatrix} = \begin{pmatrix} \lambda_2 y + cz \\ \lambda_2 z \end{pmatrix}$ . The upper equation says z = 0 so we get $\begin{pmatrix} y \\ z \end{pmatrix} = \begin{pmatrix} y \\ 0 \end{pmatrix} \sim \begin{pmatrix} 1 \\ 0 \end{pmatrix}$.

This is the ONLY eigenvector. In 3D you might imagine it to be $\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$, but neither the second nor the third column of X matches this vector.

So this is I think a good example of what happens when you have some degeneracy. The eigenvectors for the non-degenerate λ's will form the expected column vectors, but the columns for a multiple-λ "box" are something else altogether. In the end, we know X is invertible (detX = 1 at least in our example), so the columns are linearly independent, whatever they come out to be.

This example also shows an example of "clearing out the upper right corner" which for our small matrix is only the element b. We followed our general idea of doing two similarities to clear out the upper right stuff, and we got A to be A2 which is in standard diagonal box form!

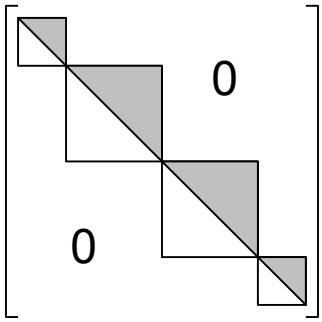
**(5) Comments.**

**Comment #1 on the Pivot and Twist Process: a fallacy**

This is done in M&M with a general matrix A. I wanted to argue that you get block diagonal form because the eigenvectors in different eigenmanifolds are orthogonal, and this would explain lots of the "zeros" in the block diagonal form of A. I now know that such eigenvectors are in general NOT orthogonal (unless A is Hermitian), so this is not the reason for those zeros. Moreover, this argument of orthogonal eigenvectors would put zeros into the column vector matrix X, not into the matrix A. In the Pivot and Twist process, it is true that the eigenvectors are progressively (*000) (**00) (***0), but these of course are not generally orthogonal.

**Comment #2 on the Pivot and Twist Process.**

Suppose you carry out the whole process and get to block diagonal form with some similarity W, so you have this picture of A:




It is true that each vertical section of the eigenvectors will map into itself, and this just says that each diagonal block represents a nullspace of $(A-\lambda I)x = 0$. This block form makes that idea transparent, and I think that is fast argument why this picture must be possible, see next comment. However, within each nullspace, you generally don't have full geometric dimensionality, than is, you generally have $m < k$, whereas the dimensions of the above boxes are the k values. For that very reason, in general you cannot think of the full space $E^N$ as really being partitioned into the nullspaces, though the above picture makes you think that is happening. { Matthews shows that the direct sum nullspaces are in fact those of $(A-\lambda I)^{bi}$, where $b_i$ is the min poly exponent, etc etc. } Within each block, you generally don't get a full basis (of true eigenvectors) for that section of the eigen column vector.

In the case that A is Hermitian, we do have $m = k$ and we do get a full basis. In this case, within each block, we can fully diagonalize to $\lambda E_k$ . In order to diagonalize, a block, you need a local full basis, otherwise you cannot construct your X of $AX = X\Lambda$. In the Hermitian case, not only can you diagonalize each eigenmanifold block, but the similarity that does it is unitary.

**Comment #3 on the Pivot and Twist Process: a simple argument for block-diagonal form.**

Here is a simple reason why the block-diagonal form is possible. We know we have a set of eigenmanifolds with geo dim = $m_i \le k_i$. These are the nullspaces of $(A-\lambda_i I)x=0$. Even though the $i^{th}$ nullspace only has $m_i$ linearly independent eigenvectors, we can artificially enlarge the dimension of the nullspace by adding linearly *dependent* vectors until we reach dimension $k_i$. We then choose a basis in the full space of dim = N in which vectors have the form x = ( *** ****** **** *****) where each grouping of coordinates is for a particular enlarged space of size $k_i$ . Thus, a vector in the third nullspace would have the form x = ( 000 000000 **** 00000) where $k_3 = 4$. In this choice of overall basis, matrix A must then have the block diagonal form shown above.

The M&M pivot and twist process is just a mechanical way to get to this form.

{ Well OK, I have left the above comments unchanged, but will comment here. We "artificially enlarge" the dimension of an eigenmanifold from $m_i$ to $k_i$ by adding "generalized eigenvectors". Then in each k-size block, we get a full set of basis vectors for $\mathrm{Ker}(p_i^{bi})$. In terms of this set of basis vectors, I think the above comment is OK. }

**********

**Some Web Excerpts.** This first one is pretty much our M&M Pivot and Twist idea where we end up with the overall triangular matrix T just as described in the words below { Yes, it is *exactly* that! } The snippet below does not talk about block-diagonal form, just the one big triangular matrix T. This reduction to T is with unitary U called Schur's Theorem, ( I think the date is 1909, see Top 10 paper).

It turns out that a **"normal" matrix** is one such that $AA^{\dagger} = A^{\dagger}A$, and that such matrices are those that can be diagonalized by a unitary transformation. A normal matrix can have complex eigenvalues. Hermitian matrices are a special case of a normal matrix that have real eigenvalues. I have not attempted to prove this, trying to finish M&M first.

Note that many people use the letter H to mean Hermitian conjugate, so yet another notation. I use the symbol $^{\dagger}$ which admittedly a typing pain, Stakgold uses * which is very confusing, so maybe H is good.

## Schur Decomposition

Every square matrix **A** is unitarily similar to an upper triangular matrix **T**: $\mathbf{A}=\mathbf{U}^{H}\mathbf{T}\mathbf{U}$.

- The main diagonal of **T** contains the eigenvalues of **A** repeated according to their algebraic multiplicities.
    - The eigenvalues may be chosen to occur in any order along the diagonal of **T** and for each possible order the matrix **U** is unique.
- **T** is diagonal iff **A** is normal. The sum of the squares of the absolute values of the off-diagonal elements of **T** is **A**'s *departure from normality*.

---

## Schur Decomposition, Real

Every square real matrix **A** is orthogonally similar to an upper block triangular matrix **T**: $\mathbf{A}=\mathbf{Q}^{T}\mathbf{T}\mathbf{Q}$ where each block of **T** is either a 1#1 matrix or a 2#2 matrix having complex conjugate eigenvalues.

- **T** is diagonal iff **A** is symmetric.

I think this must be what M&M are talking about, and I have heard the name before. They have in fact shown how to get the matrix A into upper triangular form, so they have already done Schur. My question is how they get it into "box form". I don't know any name yet for box form.

Definition: A **Hessenburg Matrix** is upper triangular + the first lower off-diagonal.

Here are a few references I found for matrix stuff.

- *A Survey of Matrix Theory and Matrix Inequalities* by M Marcus & H Minc, Prindle, Weber & Schmidt, 1964 / Dover, 1992
- *Applied Linear Algebra* by B. Noble and J.W.Daniel, Prentice-Hall, 1988
- *Finite Dimensional Vector Spaces* by P.R.Halmos, D Van Nostrand, 1958

- *Generalized Inverses* by A.Ben-Israel and T.N.E.Greville, Wiley1974
- *Matrix Computations* by G.H.Golub & C.F.Van Loan, John Hopkins University Press, 1983 ISBN 0-946536-00-7/05-8
- *Matrix Methods in Stability Theory* by S.Barnett and C.Storey, Nelson, 1970