M&M Support Document for Chapter 10 PhL 12.8.04

I don't want to clutter up the already long notes on Chapter 10 with trial balloon ideas, so put those here.

The stuff in this section was written as I was learning things and I maintain it just for the record. Important theorems are all in other documents.

****

Projection Theorem S1: Suppose space A of dimension N has a subspace S of dimension M. Then any vector in A can be decomposed as v = s + r where s is in S, and r is in $S^{\perp}$. We can write $A = S + S^{\perp}$. We know that r•s = 0. We could define projections such that P_S v = s and $P_{S\perp}$ v = r and of course $P_S + P_{S\perp} = I$. But in general $S^{\perp}$ will not also be a subspace.

Example: Let $A = E^3$ and let $S = E^2$. Then write (x,y,z) = (x,y,0) + (0,0,z) = s + r. Note that s•r = 0.

Example: Let S be a subspace in E^3 with M = 2 where A**s** = λ**s** for some λ. This would say that the geometric multiplicity of S is 2. We can find some basis vectors call then s_1 and s_2 in S.

Now our theorem says that if **v** is in E^3 , then we can write **v** = **s** + **r** where **r•s** = 0.

Projection Theorem S2 (Generalization). For a general matrix A that has several subspaces (such as nullspaces of A in this case, Ax = 0) and a residual, you could perhaps say that $B = S_1 + S_2 + S_3 + R$ where we have three subspaces and whatever is left over. We could then arrange our vectors in this space to match and have the form $x = s_1 + s_2 + s_3 + r$. We could make a projection operator into each chunk. Notice that in general, R is not likely to be a subspace because elements in it might map into a combination of spaces under the action of (A-λI)x, but we could still wrote projections for each section. When B is Hermitian and A = (B - λI), the nullspaces of A are the eigenmanifolds of A and there is no residual piece R.

****

Diagonalization of A when all eigenvalues are different.

Now, suppose we have our usual eigenvalue problem $Au_i = \lambda_i u_i$ where all eigenvalues are different. We can construct matrix U from the column vectors u_i (the eigenvectors) and we know from a theorem proven elsewhere that these column vectors are linearly independent, so we know that detU≠0 so U is invertible. And we can write the eigenvalue problem as AU = UΛ where Λ =diag(λ_i). Putting these two facts together allows us to write $U^{-1}A\,U = \Lambda$. We now have <u>several conclusions to draw from this result</u>:

(a) This shows that any matrix A with all-different eigenvalues can be diagonalized by a similarity with the matrix U of its eigenvectors, and this is true regardless of the normalization or orthogonality of the eigenvectors which make up U. All we needed was the fact that the columns of U are linearly independent so that U^{-1} exists.

(b) the matrix U which solves $U^{-1}A\,U = \Lambda$ is not unique. We can change our initial problem from

$AU = U\Lambda$ to $AU' = U'\Lambda$ by an arbitrary rescaling of each separate eigenvector u_i column in U. Then we end up with $U'^{-1}A\,U' = \Lambda$ and $U \neq U'$. However, if we insist that the u_i are normalized, then U is unique.

(c) In general, U is not going to be real orthogonal, because although we can normalize the columns of U, they are not in general orthogonal, so U is not a rotation (even if we scale things to get detU=1). Nevertheless, any transformation of the form $U^{-1}A\,U = A'$ causes $\det(A-\lambda I) = \det(A' - \lambda I)$ and therefore preserves eigenvalues. $U^{-1}A\,U = A'$ is called a similarity without regard to the nature of U.

(d) However, if A has all-different eigenvalues, is symmetric, and the u_i are normalized, then U is real orthogonal! First, we have $A^T = A$. Since $\Lambda^T = \Lambda$ as well, we can transpose our equation $U^{-1}A\,U = \Lambda$ to get $U^T\,A\,(U^T)^{-1} = \Lambda$. Since U is unique in this case, we conclude that $U^{-1} = U^T$ or $U^TU = 1$. This tells us that U is real orthogonal, and also that det U = 1. So in this situation, U really is a rotation in E^N. Another way to understand U is that its columns are normalized, but also orthogonal due to different eigenvalues.

It is the *symmetric* part that forces columns to not just be linearly independent, but to also be orthogonal, as we shall see later. [It turns out this is the case even if eigenvalues are not different for a symmetric matrix.]

(e) We can interpret our result $U^{-1}A\,U = \Lambda$ in the following manner. We have transformed A into a new matrix $A' = \Lambda$ by a certain similarity U. The transformed eigenvalue equation is $A'\,I = I\,\Lambda$. This equation is obviously true since $[\Lambda,I] = 0$, we just want to put it into "standard form". Now we can interpret the columns of the identity matrix I to be the eigenvectors of this transformed problem. So the point is that we can always start with any matrix A with $AU = U\Lambda$ having all-different eigenvalues and eigenvectors u_i, and we can transform this into another problem $A'V = V\Lambda$ where $A' = \Lambda$ and the eigenvectors in V are just the usual unit vectors in E^N.

(f) Suppose we start with the problem $A'V = V\Lambda$ where $A' = \Lambda$. We of course trivially know the eigenvalues and eigenvectors of this problem (all λ are different). Then for each invertible matrix U that you can make, you can define an equivalent eigenvalue problem $AU = U\Lambda$. Any NxN matrix U with rank N will do. There are a very large number of such matrices U! For example, if you select $N^2 - 1$ arbitrary but non-zero numbers for the matrix except for the lower right corner entry, then you make sure that you select that final value so that $\det U \neq 0$, you have a viable U. So in some sense there are $N^2 - 1$ free parameters for you to play with on constructing U.

****

A simple example where alg ≠ geo multiplicity.

Question: what is happening when we have a matrix with geometric multiplicity ≠ algebraic multiplicity?

First, consider the problem where A = I,

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

The secular equation give us $\lambda = 1$ two times, so alg = 2. Trial eigenvalue is (x,y), then eigenvalue equation says that (x,y) = (x,y). We may choose as our basis vecs (1,0) and (0,1) and then geo = alg = 2.

Now let's look at the example of Stakgold on page 158 which is this simple "shear" matrix:

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

The secular is $(1-\lambda)^2 = 0$ so $\lambda=1$ with alg= 2. The eigenvalue equation with $\lambda=1$ says that (x+y, y) = (x,y). But this implies that y = 0, so eigenvector must have the form (1,0) so geo = 1 !!! Very good. If you try the vector (0,1) then $Ax = \lambda x$ says (1,1) = (0,1) which is just not true. So in this example, (1,0) is an eigenvector. The vector (0,1) is certainly "something you can write", but it is not an eigenvector.

Obviously this exact situation could happen within an eigenmanifold of a larger matrix.

****

Projection Operator Example. Now consider example 4 on Stak page 158. Let S be a manifold of dim k within space H of dim n. For any h in H, imagine the projection operator P which creates s = Ph where s is in S. These things are all matrices!

[To be specific, suppose we associate the first k dimensions of E^n with the space S. Then we can write the projection matrix as P = diag(I, 0) where I is kxk and 0 is
n-k x n-k. The secular equation for this P is $(1-\lambda)^k \lambda^{n-k} = 0$, so we see at once that eigenvalue 0 has
alg = n-k and eigenvalue 1 has alg = k.]

Clearly, we can say that Ps = s for any s in manifold S, so for such vectors s, eigenvalue of P is 1. P is just the identity for such s, and we know that alg =k and also geo = k. Now what about q in $S^{\perp}$? We know that for such q, we have Pq = 0. The dim of the space of such q is n-k, so geo = n-k. To find its alg, I rely on the bracketed comments above, it is also n-k.

Question: Can P have some other matrix form?

(1) If we assume the basis of unit vectors in each dimension, then any vector of the form {*,*,*...0,0,0..} is in S, and {0,0,0....*,*,*,*...} is in $S^{\perp}$. In this case, P must have the form claimed. You could shuffle the order of the unit vectors and cause a trivial change in P by so doing.

(2) What happens if you now apply a similarity transformation to $Pe_i = \lambda_i e_i$ where we are starting in the form described above? Let's apply some arbitrary transformation R with $\det R \neq 0$ to the basis vectors and we get v with $x_i' = Re_i$ and $P' = (RPR^{-1})$. Now let's look at P' in more detail. Assume these forms for P, R and R^{-1}:

$$R^{-1} = \left[\begin{array}{c|c} A & b \\ \hline c & D \end{array}\right] \qquad R = \left[\begin{array}{c|c} A' & b' \\ \hline c' & D' \end{array}\right] \qquad P = \left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & 0 \end{array}\right]$$

Then we can calculate P' to be [cap letters are square matrices, lower case are not]

$$P' = \left[\begin{array}{c|c} A'A & A'b \\ \hline c'A & c'b \end{array}\right] \qquad P' = \left[\begin{array}{c|c} A^TA & A^Tb \\ \hline b^TA & b^Tb \end{array}\right]$$

which is pretty ugly looking!

Suppose we now restrict our attention to transformations R which are rotations. By the word "rotation" we mean R that does not change the length of a vector, and we know that this means that $R^TR = 1$ or that $R^{-1} = R^T$ (so R is real orthogonal). This also means that $\det(R) = \pm 1$, and we usually think of the +1 as the pure rotations without inversions. In this case, in terms of the above matrices we know that $A' = A^T$ and $D' = D^T$ and $b' = c^T$ and $c' = b^T$. In this case we get the form shown on the right above. Still, it is quite a mess.

More generally, when $R^TR = 1$, the size of any dot product is preserved as well. So this means that a rotation preserves the lengths of vectors and also preserves "angles" in the sense that $\cos\theta = a\bullet b/ab$. In particular, the orthogonality of vectors is preserved under a true rotation.

Therefore, when we have R = rotation, then we have moved from our original unit vector basis to some other basis that is still orthonormal.

But when we have R = some other transformation with $\det R \neq 0$, then the new basis vectors are no longer normalized, nor are they orthogonal. We do have the equation $P'x_i' = \lambda_i x_i'$, and $P'X' = X'\Lambda$. The new eigenvectors are, however, still linearly independent, so we do still have a basis! Here is the proof that this is so:

Theorem S3: If you map a set of N vectors $\mathbf{x_i}$ to a new set $\mathbf{x_i'}$ using $R\mathbf{x_i} = \mathbf{x_i'}$ where $\det R \neq 0$, then if the original vectors were linearly independent, then so are the final vectors.

Proof: Line up your N vectors $\mathbf{x_i}$ as columns of a matrix X. Similarly, form X'. Then RX = X'. This tells us that det(R) det(X) = det(X'). If the x_i vectors are linearly independent, we know by Theorem 9.7 that $\det(X) \neq 0$. Since $\det(R) \neq 0$, this tells us that $\det(X') \neq 0$, and that means that the $\mathbf{x_i'}$ are linearly independent by that same theorem.

And here is another theorem just along the way that I proved but don't need:

Theorem S4: If $\det R \neq 0$, then $R\mathbf{e_1} = \mathbf{0}$ is not possible.

Proof: To say $R\mathbf{e_1} = \mathbf{0}$ is to say that the leftmost column of R is all zeros, and then det R = 0.

So my conclusion is that if you allow an *arbitrary* transformation R with the only restriction being that $\det R \neq 0$, you will map your nice orthonormal unit vector basis into some other basis where the vectors are still linearly independent, but the lengths have changed and they are no longer orthogonal. We have proved the following theorem.

Theorem S5: Consider an NxN eigenvalue problem $A\mathbf{x_i} = \lambda_i\mathbf{x_i}$ in which it happens that the N eigenvectors form an orthonormal basis. If you transform this problem in the usual similarity way to $A'\mathbf{x_i}' = \lambda_i\mathbf{x_i}'$ by a transformation $\mathbf{x_i}' = R\mathbf{x_i}$ with the only restriction being that $\det R \neq 0$, then in the new problem, the vectors $\mathbf{x_i}'$ still form a basis, but they will in general be neither normalized nor orthogonal. If it happens that R is a real orthogonal transformation (a "rotation"), then the new basis will still be orthonormal.

****

Question: In the world of 2x2 matrices, can you construct a matrix such that there are *no* non-zero eigenvectors?

[Here is a "modern" answer. Any matrix has at least 1 eigenvalue and 1 eigenvector, so no.]

Answer: [a geometric answer] No, and here is why (it is a long answer). [First of all, if λ is not an eigenvalue, the secular equation deal tells you that $(A-\lambda)$ is invertible, and then only x=0 solves $Ax = \lambda x$. Fine.] If you write out your two equations in two unknowns for the matrix (a,b,c,d)

$ax + by = \lambda x$ secular: $(a-\lambda)(d-\lambda) - bc = 0$

$cx + dy = \lambda y$ $\Rightarrow \lambda^2 - (a+d)\lambda + \det A = 0$

$\Rightarrow \lambda = -(a+d)/2 \pm \text{sqrt}((a+d)^2 - 4\det A)/2$

You have two lines passing through the origin. If λ is an eigenvalue, then these two lines have the same slope and are then the same line. That is, if you write down the slope of the two lines shown above and set them equal to each other, you get the secular equation. For the first equation, $(a-\lambda)x + by = 0$, $m = (\lambda-a)/b$, and for the second $m' = -c/(d-\lambda)$. So, if you pick one of your eigenvalues λ, then $m = (\lambda-a)/b$ is the slope that both lines have, and therefore each line defines the same eigenvector. To summarize, for each eigenvalue like λ_1, you get two lines that are the same line.

Now, the secular equation for 2x2 has 2 eigenvalues $\lambda 1$ and $\lambda 2$.

If they are **different**, you can see that the two lines have different slope since $m = (\lambda-a)/b$ (if $b \neq 0$) so you know that you will have two different line solutions, and then two different eigenvector rays. In the case that b=0, this is still true, but it takes a bit of work to show. If b=0, the secular equation says $(a-\lambda)(d-\lambda) = 0$ so $\lambda_1 = a$ and $\lambda_2 = d$. So let's assume then that $a \neq d$ so we stay in our case of interest. For $\lambda_1 = a$, the first equation above says nothing, and the second is a line of slope $-c/(d-\lambda) = -c/(d-a)$ and that defines that eigenray. For $\lambda_2 = d$, the first equation says $ax = dx$. But since $a \neq d$, the only solution is $x = 0$. The second equation then says $cx=0$ which is consistent with $x = 0$. So the ray here is going to be (0,y) which has ∞ slope. To summarize, for $\lambda_1 = a$ line has slope $-c/(d-a)$ and for $\lambda_2 = d$ line has slope ∞ . These slopes are always different, so there are always two eigenrays when eigenvalues are different. This is consistent with our earlier theorem which said that the eigenvectors of a matrix with all different eigenvalues are linearly independent. So in our general quest for problems of the form $Ax = \lambda x$ which have no non-trivial solutions, we need not consider 2x2 matrices with different eigenvalues, because those always have two non-trivial solutions.

This then leaves us thinking about 2x2 matrices with the **same** eigenvalues. In the matrix (a,b,c,d) this occurs when $c = (a-d)^2/(-4b)$ and the eigenvalues are $\lambda 1 = \lambda 2 = (a+d)/2$. In this case, our two "different lines" have slope $m = (\lambda-a)/b$, and they are really the same line, so there is only one eigenray and it is given by $ax + by = [(a+d)/2]x$ which says $[(a-d)/2]x + by = 0$, so the ray is $(1, -(a-d)/2b)$.

An example of this case was the sheer matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. The eigenray is (1,0).

So we have exhausted all the cases I think and have this result:

Theorem S6: Consider the world of 2x2 matrices. If the two eigenvalues are different, there are two distinct non-zero eigenvectors, and geometric multiplicity = algebraic multiplicity = 1 for each eigenmanifold. If the two eigenvalues are the same, there is exactly one non-zero eigenvector, so geometric multiplicity = 1 and algebraic multiplicity = 2. There are no 2x2 matrices which have geometric multiplicity = 0.

Proof: We just showed this above, but here is a more general answer which assumes more knowledge. Such a matrix has N = 2, but the matrix $(A-\lambda I)$ has rank 1 unless the matrix is all zeros. This means there will be k = N - rank = 2-1 = 1 independent eigenvector. If $A = \lambda E$, then get rank 0 and 2 eigenvectors. This is the only possibility.

More generally, for an NxN matrix, only the matrix $A = \lambda E$ will have N eigenvectors. All other matrices will have a number somewhere between 1 and N-1. See matrix theorems doc.

****

Now let's redo the above 2x2 stuff in a slightly different language. We have

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \qquad \text{and } A - \lambda I = \begin{pmatrix} a-\lambda & b \\ c & d-\lambda \end{pmatrix}$$

By setting $\det(A-\lambda I)=0$, we are forcing the rows of $(A-\lambda)$ to be linearly dependent. That means that the set of equations we get:

$$(a-\lambda)x + by = 0$$
$$cx + (d-\lambda)y = 0$$

are not independent. The second one is a multiple of the first one! So the eigenray for this λ is given by the equation $(a-\lambda)x + by = 0$. If we have the λ's different, $\lambda 1 \neq \lambda 2$, then these eigenrays are clearly different, and we have our two solutions. If the λ's are the same, then we have only one eigenray. This agrees with the conclusions obtained earlier.

****

More geometric interpretation: in 3x3 each equation is a plane, etc.

Comment: other methods let us avoid thinking about visualizing things geometrically and are therefore "better" especially when N > 3. But here we think in geometry one more time.

Now let's ascend to the 3x3 world but in the language of the last section. We have

$$A = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} \qquad \text{and } A - \lambda I = \begin{pmatrix} a-\lambda & b & c \\ d & e-\lambda & f \\ g & h & i-\lambda \end{pmatrix} \qquad (A-\lambda I)\mathbf{x} = 0$$

For eigenvalues, we know that at least one of the three rows in A-λI can be written as a sum of the other two. This is so because det = 0 implies linear dependence of rows (see theorem elsewhere). We can think of each row as determining the normal vector of a plane in E^3 . For example, the top row is
$\mathbf{n_1}$ = (a-λ, b, c).

But suppose $\mathbf{n_1}$ = (0,0,0). In this case, the row corresponds to the equation 0 = 0, so there is no plane, there is no equation. So this could happen with one or more rows.

For the moment, assume that none of the three $\mathbf{n_i}$ are (0,0,0). The three rows imply three planes due to (A-λI)x = 0. If λ is not an eigenvalue, then the three normal vectors are linearly independent, and that means that the only solution is **x** = 0. If λ is an eigenvalue, then we know that one of the 3 normals is a lincom of the others. In this case, we can write the last row as a linear combination of the first two rows, so we really only have 2 equations in 3 unknowns, and each equation implies a plane. We have studied this little problem in a separate document, and the conclusions are simple: if the two equations are linearly independent, (the normals don't line up with each other), then geo = 1 and we have a line which is the intersection of the two planes. But if the two equations are multiples, then we really only have 1 equation in 3 unknowns, and in this case geo = 2.

Now suppose $\mathbf{n_1}$ = (0,0,0) but the other two are not null. We then have 2 equations in 3 unknowns, and the conclusions are exactly those of the last paragraph, and geo = 1 or 2.

Suppose that two of the n_i vanish. Then we have 1 equation in 3 unknowns which is one plane, and in this case geo = 2.

Let's test this idea with a simple example which has secular

$$A = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \text{ and } A - \lambda I = \begin{pmatrix} 1-\lambda & 1 & 0 \\ 0 & 1-\lambda & 0 \\ 0 & 0 & 1-\lambda \end{pmatrix} \text{ secular says } (1-\lambda)^3 = 0 \text{ so } \lambda = 1 \text{ triple}$$

We have $A - 1*I = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$. Here we have only 1 equation in 3 unknowns (the top one), so we expect to have geo = 2. I seem to find this general solution

$$\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} A \\ 0 \\ C \end{pmatrix} = 1 * \begin{pmatrix} A \\ 0 \\ C \end{pmatrix}$$

which does indeed have geo = 2.

I think I finally realize the rule I have been looking for. Suppose you have an NxN matrix. You are then interested in the nullspace of the equation (A - λI)x = 0. When A has real numbers only, Stakgold page 154 tells us that the

$$\text{nullity}(A - \lambda I) = N - \text{rank}(A^T - \lambda I).$$

The nullity is the dimensionality of the nullspace of (A - λI)x = 0. This nullspace is a linear manifold or subspace of the E^N space. This nullity IS the geometric multiplicity for the eigenvalue λ.

(1) Examine this rule in our 2x2 matrix world. We had $A - \lambda I = \begin{pmatrix} a-\lambda & b \\ c & d-\lambda \end{pmatrix}$. We know it is not rank=2 because λ is an eigenvalue. Unless all four entries in this matrix vanish, we have rank = 1 for $A^T - \lambda I$. Our prediction would then be that nullity = 2 - 1 = 1, so that geo = 1. This agrees with work done above.

Now, how might we make all four elements vanish? First, we need b = c = 0. The secular then says λ = a and λ = d. In either of these cases, one of the 4 elements in $A - \lambda I$ is nonzero, unless a = d. In this case, A = multiple of identity. We then have $A - \lambda I = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ and rank = 0. We then expect to have nullity = 2. Look at $A = \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix}$. The eigenvectors are (a,0) and (0,a) so in fact nullity = 2 as predicted.

(2) Examine this rule in the 3x3 world. In general, the matrix $A - \lambda I$ is going to have rank = 0,1 or 2. This means the nullity will be 3, 2 or 1. The only way we can have nullity = 3 is if all those off diagonals are 0 and then we have $A - \lambda I = \begin{pmatrix} a-\lambda & 0 & 0 \\ 0 & e-\lambda & 0 \\ 0 & 0 & i-\lambda \end{pmatrix}$. Now we need all three diagonals to vanish at once, and that means a = e = i, and again we have a multiple of the identity and nullity = 3 is correct.

How about our explicit example? We repeat from above,

$A = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ and $A - \lambda I = \begin{pmatrix} 1-\lambda & 1 & 0 \\ 0 & 1-\lambda & 0 \\ 0 & 0 & 1-\lambda \end{pmatrix}$ secular says $(1-\lambda)^3 = 0$ so $\lambda = 1$ triple

The rank of $A^T-\lambda I$ is 1 due to that one off diagonal element, so rank = 1 and nullity = 2. This agrees with our solution which was written above as (A,0,C).

Theorem S7: Consider the problem $Ax_i = \lambda_i x_i$. Find the eigenvalues as usual and their algebraic multiplicities. Within each eigenmanifold, here is how we can compute the geometric multiplicity for that manifold

geometric multiplicity = nullity(Ax - λI) = N - rank(Ax - λI)

Another example:

$A = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$ and $A - \lambda I = \begin{pmatrix} 1-\lambda & 1 & 0 \\ 0 & 1-\lambda & 0 \\ 0 & 0 & 2-\lambda \end{pmatrix}$ secular says $(1-\lambda)^2(2-\lambda) = 0$

so $\lambda = 1$ is a double, and $\lambda = 2$ is a single. For $\lambda = 1$ we have $A - \lambda I = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ which has rank 2. This is because the subdet $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ has det = 1. Therefore we expect geo multiplicity = 3 - 2 = 1. I find the solution to be (x,0,0) which means geo = 1.

****

Theorem. Here is how you find a congruence Y which is real orthogonal which diagonalizes A.

First, find Q such that $Q^TAQ = I$. [This is only doable if $\det A \neq 0$, by the way.] Then apply GS to this Q as in my notes on the subject to get $X = QT$. Here T is an upper triangular matrix (but that is only for the particular way I did GS), X is the new matrix which is orthonormalized by GS, hence X is real orthogonal. This did require that $\det Q \neq 0$ which I think is not a problem. Then we have that

$$X^TAX = (QT)^TA(QT) = T^TT$$

which is *not* in general diagonal. However, T^TT is symmetric (Hermitian) and can therefore be diagonalized by a real orthogonal R, so we then have

$$R^TX^TAXR = R^T(T^TT)R \Rightarrow (XR)^TA(XR) = D$$

But $R' = XR$ is a real orthogonal matrix, and we know such a thing preserves eigenvalues, so it must be that $D = \Lambda$ and we end up with

$$R'^TAR' = \Lambda$$

I am just trying to justify the words on M&M here on page 325:

Once you have Q with Q^TAQ, "the elements can then be orthonormalized and the resulting matrix X is both congruent and collineatory(similarity)"

I know that of the many congruences Q, for sure you can find one that is also real orthogonal and which then takes you to Λ. I know this because A is symmetric. So in effect they are right.