Comments on Diagonalization PhL 12.18.04

The big theorem is really this one: It says that ANY square matrix A can be put into upper triangular form (Schur form) by some unitary transformation. A Hermitian matrix is a special case of a normal matrix that has real eigenvalues, so normal is the larger class, and for this class, the result is all the way to diagonal.

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 last Real thing is correct, but I don't know why that upper triangular simplifies so much in the real case.]

It turns out that a "normal" matrix is one such that AA† = A†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 † which admittedly a typing pain, Stakgold uses * which is very confusing, so maybe H is good.

Comments:

(1) In M&M, that huge Pivot and Rotate/Twist discussion describes how, in principle, you can take a general square matrix by similarity and get it into upper triangular form. It is unclear whether A is real or complex in this discussion, but we can see that each step is a rotation if things are real, so the total transform getting to upper triangular is a product of a finite number of rotations. The authors then do some arm-waving and continue to get the result into "block diagonal form". We know that if all eigenvalues are different (regardless of Hermiticity) we can get all the way to diagonal form, so the lack of diagonality comes from the degenerate eigenvalues.

(2) The M&M discussion is "in principle" only, because there is no prescription for finding the right rotation at each step. For example, on page 320 they write $B = X^{-1}AX$. They don't tell you how to find the matrix X, they only comment that it must have at least one column that is an eigenvector, and so on. Now suppose the first column of A is some numbers. I can find a rotation X to turn that column into x,0,0,0 so that if I write B = XA, the first column of B has this form. But I don't know how to do this if I have to have a similarity, $B = XAX^{-1}$, which does this action. We are sort of begging the question here. You really have to have a solution to the whole problem to make X. Similarly, I can find a Householder matrix such that B = HA will make the first column be x,0,0,0, but this column will not survive in B = HAH. The right side H will mess up that first column, generally speaking. (See Recipe notes on this.)

(3) In practice, then, even if you know the eigenvalues, you don't really know what the upper triangular matrix looks like, except of course you do know the diagonal elements of it.

(4) In practice, you generally just have a matrix A and you do NOT know the eigenvalues, that is part of what you want to find! In this case, you apply your finite number of operations (Householders, Givens rotations, similarized EROs, whatever) to get down to Hessenberg form, THEN you have to do the QR iteration. You cannot get out of doing the iteration in a numerical problem where all you are given is the matrix A.

(5) This is similar to saying that if you are given a polynomial and you want to find a root, you have to do Newton-Raphson iteration to get the answer. The only exception here would be if you just happen to be able to factor the polynomial and see some exact eigenvalues. A similar thing happens when you compute a square root. If the number they hand you happens to be 4, yes, you know the result is 2, but for arbitrary numbers, you have to iterate. This is some kind of principle. Something about all these problems is transcendental-like.