Unitary Diagonalization of Matrices

Here we take a different approach than Lang and diagonalize all matrices which can possibly be diagonalized using unitary matrices.

Theorem 1. *The product of two unitary matrices is unitary.*

Proof: Suppose Q and S are unitary, so $Q^{-1} = Q^*$ and $S^{-1} = S^*$. Then $(QS)^* = S^*Q^* = S^{-1}Q^{-1} = (QS)^{-1}$ so QS is unitary ∎

Theorem 2. *(Schur Lemma) If A is any square complex matrix then there is an upper triangular complex matrix U and a unitary matrix S so that $A = SUS^* = SUS^{-1}$.*

Proof: Let q_1 be an eigenvector of A, which we may suppose has unit length. By the Gram-Schmidt process we may choose q_i' so that $\{q_1, q_2', \ldots, q_n'\}$ is an orthonormal basis. Let $Q_0 = [q_1 q_2' \cdots q_n']$, then Q_0 is unitary and $Q_0^* A Q_0 = \begin{pmatrix} \lambda_1 & * \\ 0 & A_1 \end{pmatrix}$ for some $(n-1)\times(n-1)$ matrix A_1. Likewise, we may find a unitary $(n-1)\times(n-1)$ matrix Q_1 so that $Q_1^* A_1 Q_1 = \begin{bmatrix} \lambda_2 & * \\ 0 & A_2 \end{bmatrix}$. Then if $S_1 = Q_0 \begin{bmatrix} 1 & 0 \\ 0 & Q_1 \end{bmatrix}$ we have $S_1^* A S_1 = \begin{bmatrix} \lambda_1 & * & * \\ 0 & \lambda_2 & * \\ 0 & 0 & A_2 \end{bmatrix}$. Note that S_1 is unitary by Theorem 2. Now continue in this fashion, letting $S_k = S_{k-1} \begin{bmatrix} I_k & 0 \\ 0 & Q_k \end{bmatrix}$, and we see that $U = S_n^* A S_n$ is upper triangular. Letting $S = S_n$ we see that $A = SUS^*$. ∎

Finally we characterize which matrices can be diagonalized by a unitary matrix. We say a matrix A is **normal** if $AA^* = A^*A$.

Theorem 3. *A matrix A is diagonalizable with a unitary matrix if and only if A is normal. In other words:*
a) If A is normal there is a unitary matrix S so that S^*AS is diagonal.
b) If there is a unitary matrix S so that S^*AS is diagonal then A is normal.

Proof: Suppose A is normal. By Theorem 2 there is a unitary matrix S and an upper triangular U so that $A = SUS^*$. Then

$$UU^* = S^*AS(S^*AS)^* = S^*ASS^*A^*S = S^*AA^*S = S^*A^*AS = S^*A^*SS^*AS = U^*U$$

But if we let u_{ij} denote the ij-th entry of U then the upper left entry of U^*U is $u_{11}\overline{u_{11}} = |u_{11}|^2$ but the upper left entry of UU^* is

$$u_{11}\overline{u_{11}} + u_{12}\overline{u_{12}} + \cdots + u_{1n}\overline{u_{1n}} = |u_{11}|^2 + |u_{12}|^2 + \cdots + |u_{1n}|^2$$

Since this equals $|u_{11}|^2$ and all summands are nonnegative real numbers we must have $u_{12} = u_{13} = \cdots = u_{1n} = 0$. Similarly, looking at the second diagonal entry we see that $u_{2j} = 0$ for all $j > 2$. Continuing in this way we see that U must be diagonal. So we have shown that if A is normal, then it is diagonalizable with a unitary matrix.

Now suppose that A is any matrix so that there is a unitary matrix S so that $S^*AS = D$ is diagonal. Note $DD^* = D^*D$. Then

$$AA^* = SDS^*(SDS^*)^* = SDS^*SD^*S^* = SDD^*S^* = SD^*DS^* = SD^*S^*SDS^* = A^*A$$

Consequently, A is normal. ∎

Examples of normal matrices are Hermitian matrices ($A = A^*$), skew Hermitian matrices ($A = -A^*$) and unitary matrices ($A^* = A^{-1}$) so all such matrices are diagonalizable.

The Schur Lemma above needed to use a complex unitary matrix S. Note that A and U have the same characteristic polynomial and hence the diagonal entries of U are the eigenvalues of A. So if A is a real matrix and we want to find a real unitary matrix S so that ${}^tSAS$ is upper triangular, this is only possible if all the eigenvalues of A are real. But a real matrix A often has nonreal eigenvalues. The following real Schur Lemma shows what we can do in the real case.

But first a definition. For the purposes of these notes, we will say that a matrix U is 2×2 block upper triangular if the entries u_{ij} of U satisfy:

a) $u_{ij} = 0$ if $i > j + 1$.

b) If $u_{i+1,i} \neq 0$, then the adjacent entries below the diagonal are 0, i.e., $u_{i+2,i+1} = 0$ and $u_{i,i-1} = 0$.

In other words a matrix U is 2×2 block upper triangular if we can write U in block form as:

$$U = \begin{pmatrix} B_{11} & B_{12} & \cdots & B_{1,k-1} & B_{1k} \\ 0 & B_{22} & \cdots & B_{2,k-1} & B_{2k} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & B_{k-1,k-1} & B_{k-1,k} \\ 0 & 0 & \cdots & 0 & B_{kk} \end{pmatrix}$$

where each B_{ij} is a 2×2 or 1×1 matrix.

Theorem 4. *(real Schur Lemma) If A is any square real matrix then there is a 2×2 block upper triangular real matrix U and a real unitary matrix S so that $A = SU\,{}^tS = SUS^{-1}$. The eigenvalues of each 2×2 block on the diagonal are a complex conjugate pair of non real eigenvalues of A.*

Proof: The proof is similar to that of theorem 2. If A has a real eigenvalue then we let q_1 be a real eigenvector and proceed as in theorem 2. If A has no real eigenvalues, let $a + bi$ be an eigenvalue and let $u + iv$ be a nonzero eigenvector with u and v real. Then

$$Au + iAv = A(u + iv) = (a + bi)(u + iv) = au - bv + i(bu + av)$$

Equating real and imaginary parts, we see that $Au = au - bv$ and $Av = bu + av$. Note that $u \neq 0$ since if $u = 0$ we would have $0 = Au = au - bv = -bv$ so $v = 0$ since $b \neq 0$. Likewise $v \neq 0$. We know that v and u are linearly independent because if $v = cu$ then $Au = au - bcu = (a - bc)u$ so A would have a real eigenvalue $a - bc$. We now let $\{q_1, q_2\}$ be an orthonormal basis for the subspace generated by u and v, for example $q_1 = u/||u||$ and $q_2 = (v- < v, q_1 > q_1)/||v- < v, q_1 > q_1||$. Extend to an orthonormal basis $\{q_1, q_2, \ldots, q_n\}$ of $\mathbb{R}^n$. Let $Q_0 = [q_1 q_2 \cdots q_n]$, then Q_0 is real unitary and $Q_0^* A Q_0 = \begin{pmatrix} B_1 & C_1 \\ 0 & A_1 \end{pmatrix}$ for some 2×2 matrix B_1, some $(n-2) \times (n-2)$ matrix A_1, and some $2 \times (n-2)$ matrix C_1. Now continue as in the proof of theorem 2. It is not hard to see that the determinent of a block triangular matrix is the product of the determinents of the diagonal blocks, and hence the characteristic polynomial of U is the product of the characteristic polynomials of its diagonal blocks. Since U and A have the same characteristic polynomial, each eigenvalue of a diagonal block of U is an eigenvalue of A. Eigenvalues of real matrices come in complex conjugate pairs and the 2×2 blocks correspond to Note that if all the eigenvalues of A are real, then the resulting U will be upper triangular. ■

If A is symmetric then, as shown in Lang, all its eigenvalues are real so we get ${}^tSAS = U$ is upper triangular, so

$${}^tU = {}^tS\,{}^tA S = {}^tSAS = U$$

so U is diagonal.

If A is real unitary, then we saw that all eigenvalues λ have modulus 1, $\lambda\overline{\lambda} = 1$.

—More to come, but I'll post this incomplete version now