Orthogonality and Eigenvectors

§1. Introduction

Recall:
1) P is unitary if $P^* = P^{-1}$.
2) The matrix of transition between orthonormal bases is unitary.
3) Matrices A and B are *unitary similar* if $B = P^{-1}AP$ with P unitary so A and B represent the same transformation with respect to different *orthonormal* bases.
We find unitary similar matrices which are of one or another special form.

§2. Schur Decomposition

Theorem (Schur decomposition) Given a square matrix A there is a unitary P with $\Delta = P^{-1}AP$ upper triangular. If A is real and has only real eigenvalues then P can be selected to be real.

Note Δ has the eigenvalues of A along its diagonal because Δ and A are similar and Δ has its eigenvalues on the diagonal.

Proof By induction assume the result is true for order $n-1$ matrices and let A be order n. Let $\mathbf{v}$ be an eigenvector with value λ. Normalize $\mathbf{v}$, that is, replace $\mathbf{v}$ by $\mathbf{v}/\|\mathbf{v}\|$ if necessary. Extend $\mathbf{v}$ to an orthonormal basis $\mathbf{v}$, $\mathbf{w}_1$, …, $\mathbf{w}_{n-1}$ using Gram-Schmidt. Let Q be the matrix of transition to this basis. So Q is unitary and $Q^{-1}AQ$ is of form

$$\begin{pmatrix} \lambda & * & \dots & * \\ 0 & & & \\ \vdots & & C & \\ 0 & & & \end{pmatrix}.$$

By induction there is a unitary V with $V^{-1}CV$ upper triangular. Consider

$$U = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & V & \\ 0 & & & \end{pmatrix}$$

which is also unitary. Then QU is unitary and $(QU)^{-1}A(QU) = U^{-1}(Q^{-1}AQ)U =$

$$= \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & V^{-1} & \\ 0 & & & \end{pmatrix} \begin{pmatrix} \lambda & * & \dots & * \\ 0 & & & \\ \vdots & & C & \\ 0 & & & \end{pmatrix} \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & V & \\ 0 & & & \end{pmatrix} = \begin{pmatrix} \lambda & * & \dots & * \\ 0 & & & \\ \vdots & & V^{-1}CV & \\ 0 & & & \end{pmatrix}$$

which is upper triangular. ∎

Remark The proof gives an algorithm for finding P.

§3. **Normal Matrices**

Lemma 1 Let A and B be unitary similar. Then A is normal iff B is normal.

Proof Let P be unitary with $B = P^*AP$. Then $B^* = P^*A^*P$ and $P^* = P^{-1}$. So $B^*B - BB^* = P^*A^*PP^*AP - P^*APP^*A^*P = P^*(A^*A - AA^*)P$. Thus $B^*B - BB^* = 0$ iff $A^*A - AA^* = 0$. ∎

Lemma 2 If A is normal then for each i, the i-th row of A has the same length as the i-th column of A.

Proof $\langle A\mathbf{e}_i, A\mathbf{e}_i\rangle = \langle \mathbf{e}_i, A^*A\mathbf{e}_i\rangle = \langle \mathbf{e}_i, AA^*\mathbf{e}_i\rangle = \langle A^*\mathbf{e}_i, A^*\mathbf{e}_i\rangle$. Thus $\|A\mathbf{e}_i\| = \|A^*\mathbf{e}_i\|$. The left side is the length of the i-th column of A. The right side is the length of the i-th column of A^*, that is, the i-th row of the conjugate of A; but the conjugate of a vector has the same length. ∎

The argument in the proof works for any vector $\mathbf{x}$ instead of $\mathbf{e}_i$, which proves the following lemma, whose converse is also true.

Lemma 3 If A is normal then $\|A\mathbf{x}\| = \|A^*\mathbf{x}\|$ for all $\mathbf{x}$.

Theorem (Unitary Similar Diagonalization) For a square matrix A, A is unitary similar to a diagonal matrix iff A is normal, that is, there is a unitary P with $P^{-1}AP = \Lambda$ diagonal iff $A^*A = AA^*$.

Proof $\Rightarrow$: Suppose $P^{-1}AP = \Lambda$ is diagonal. Diagonal matrices are normal since $\Lambda\Lambda^* = \begin{pmatrix} \lambda\bar{\lambda} & 0 \\ 0 & \ddots \end{pmatrix} = \Lambda^*\Lambda$. By Lemma 1, A is normal.

$\Leftarrow$: By the previous theorem (Schur), A is unitary similar to an upper triangular matrix Δ. We show Δ is diagonal. By Lemma 1, Δ is normal and by Lemma 2, the i-th row of Δ has the same length as its i-th column. If Δ is not diagonal, then select i so that the i-th row is the first row with a nonzero off-diagonal element:

$$\Delta = \begin{pmatrix} t_{11} & 0 & \dots & 0 & \dots & 0 \\ 0 & t_{22} & \dots & 0 & \dots & 0 \\ 0 & 0 & \dots & t_{ii} & * & * \\ 0 & 0 & \dots & 0 & * & * \\ \vdots & \vdots & \dots & \vdots & \vdots & \ddots \end{pmatrix}.$$

The length of the i-th column is $|t_{ii}|$ but the length of i-th row is greater than $|t_{ii}|$, a contradiction. So Δ can have no nonzero off-diagonal elements. ∎

Since the columns of P are eigenvectors of A, the next corollary follows immediately.

Corollary There is an *orthonormal* basis of eigenvectors of A iff A is normal.

Lemma Let A be normal. $A\mathbf{x} = \lambda\mathbf{x}$ iff $A^*\mathbf{x} = \bar{\lambda}\mathbf{x}$.

Proof $A\mathbf{x} = \lambda\mathbf{x}$ is equivalent to $\|(A - \lambda I)\mathbf{x}\| = 0$. It is easy to show $A - \lambda I$ is normal, so Lemma 3 shows that $\|(A - \lambda I)^*\mathbf{x}\| = \|(A^* - \bar{\lambda} I)\mathbf{x}\| = 0$ is equivalent. This in turn is equivalent to $A^*\mathbf{x} = \bar{\lambda}\mathbf{x}$. ∎

Proposition (Eigenspaces are Orthogonal) If A is normal then the eigenvectors corresponding to different eigenvalues are orthogonal.

Proof Suppose $A\mathbf{v} = \lambda\mathbf{v}$ and $A\mathbf{w} = \mu\mathbf{w}$, where $\lambda \neq \mu$. $\langle \mathbf{v}, A\mathbf{w}\rangle = \langle \mathbf{v}, \mu\mathbf{w}\rangle = \mu\langle \mathbf{v}, \mathbf{w}\rangle$. However, $\langle \mathbf{v}, A\mathbf{w}\rangle = \langle A^*\mathbf{v}, \mathbf{w}\rangle$ which by the lemma is $\langle \overline{\lambda}\mathbf{v}, \mathbf{w}\rangle = \lambda\langle \mathbf{v}, \mathbf{w}\rangle$. So $\mu\langle \mathbf{v}, \mathbf{w}\rangle = \lambda\langle \mathbf{v}, \mathbf{w}\rangle$ but since $\lambda \neq \mu$, $\langle \mathbf{v}, \mathbf{w}\rangle = 0$. ∎

Remark The above proposition provides the key to computing the matrix P in the Diagonalization Theorem. For each eigenspace, find a basis as usual. Orthonormalize the basis using Gram-Schmidt. By the proposition all these bases together form an orthonormal basis for the entire space. Examples will follow later (but not in these notes).

§4. **Special Cases**

Corollary If A is Hermitian ($A^* = A$), skew Hermitian ($A^* = -A$ or equivalently iA is Hermitian), or unitary ($A^* = A^{-1}$), then A is unitary similar to a diagonal matrix and A has an orthonormal basis of eigenvectors.

Proof These types of matrices are normal. Apply the previous theorem and corollary.

Proposition
If A is Hermitian then the eigenvalues of A are real.
If A is skew Hermitian then the eigenvalues of A are imaginary.
If A is unitary then the eigenvalues of A are of length 1.

Proof We note $A^* = A$ if A is Hermitian, $A^* = -A$ if A is skew Hermitian, and $A^* = A^{-1}$ if A is unitary. For convenience, combine these cases as $A^* = \pm A^{\pm 1}$.
Suppose $A\mathbf{x} = \lambda\mathbf{x}$ with $\mathbf{x} \neq \mathbf{0}$. Then $\langle \mathbf{x}, A\mathbf{x}\rangle = \langle \mathbf{x}, \lambda\mathbf{x}\rangle = \lambda\langle \mathbf{x}, \mathbf{x}\rangle$, whereas $\langle A^*\mathbf{x}, \mathbf{x}\rangle = \langle \pm A^{\pm 1}\mathbf{x}, \mathbf{x}\rangle = \langle \pm\lambda^{\pm 1}\mathbf{x}, \mathbf{x}\rangle = \pm\overline{\lambda}^{\pm 1}\langle \mathbf{x}, \mathbf{x}\rangle$.
$\langle \mathbf{x}, \mathbf{x}\rangle \neq \mathbf{0}$ since $\mathbf{x} \neq \mathbf{0}$, so canceling gives $\lambda = \pm\overline{\lambda}^{\pm 1}$ which shows λ is real, imaginary or of length one for the respective cases.

§5. **The Real Case**

A real matrix can have non-real eigenvalues (as we saw when we diagonalized a rotation matrix) since characteristic polynomials with real coefficients can have non-real roots. If A is Hermitian, which for a real matrix amounts to A is symmetric, then we saw above it has real eigenvalues.

Theorem (Orthogonal Similar Diagonalization) If A is *real symmetric* then A has an orthonormal basis of real eigenvectors and A is orthogonal similar to a real diagonal matrix $\Lambda = P^{-1}AP$ where $P^{-1} = P^T$.

Proof A is Hermitian so by the previous proposition, it has real eigenvalues. We would know A is unitary similar to a real diagonal matrix, but the unitary matrix need not be real in general. By the Schur Decomposition Theorem, $P^{-1}AP = \Delta$ for some *real* upper triangular matrix Δ and *real* unitary, that is, orthogonal matrix P. The argument of the last theorem shows Δ is diagonal. ∎

Remark The converse to this theorem holds: If A is real and orthogonal similar to a diagonal matrix, then A is real and symmetric.

Remark Since not all real matrices are symmetric, sometimes an artifice is used. For example, A^TA and AA^T are symmetric even when A is not.