Orthogonal and Unitary Diagonalization

November 22, 2022

Contents

$$A \text{ is symmetric : } A^T = A$$

$$A \text{ is Hermitian : } A^H = A$$

$$Q \text{ is orthogonal : } Q^T Q = I$$

$$U \text{ is unitary : } U^H U = I$$

Every real symetric matrix is orthogonally diagonalizable.

Every complex Hermitian matrix is unitarily diagonalizable.

1 Definitions

Suppose A is a square matrix with real(or complex) entries.

Definition The *transpose* A^T of A is a matrix B such that

$$b_{ij} = a_{ji}$$

for every $1 \leq i, j \leq n$. The *Hermtian* A^H of A is a matrix C such that

$$c_{ij} = \overline{a_{ji}}$$

for every $1 \leq i, j \leq n$. Note that $A^H = A^T$ if A is real.

Definition A is said to be *symmetric* if

$$A^T = A.$$

A is said to be *Hermitian* if

$$A^H = A.$$

Note that real symmetric matrices are Hermitian.

Definition Q is said to be *orthogonal* if

$$Q^T Q = I,$$

which is equivalent to saying that

$$QQ^T = I \quad \text{or} \quad Q^{-1} = Q^T.$$

U is said to be *unitary* if

$$U^H U = I,$$

which is equivalent to saying that

$$UU^H = I \quad \text{or} \quad U^{-1} = U^H.$$

Note that real orthogonal matrices are unitary.

Definition A is said to be *orthogonally diagonalizable* if there exist a orthogonal matrix Q and a diagonal matrix D such that

$$A = QDQ^T.$$

Definition A is said to be *unitarily diagonalizable* if there exist a unitary matrix U and a diagonal matrix D such that

$$A = UDU^H.$$

2 Matrices of Distinct Eigenvalues

For the properties below, we assume that A is an n by n Hermitian matrix and x is an n dimensional vector.

Property 1 $x^H Ax$ is real.

For example, consider the case when the dimension is 2. Set

$$x = \begin{bmatrix} u \\ v \end{bmatrix}, \quad A = \begin{bmatrix} a & b \\ \bar{b} & c \end{bmatrix}$$

where all the entries are complex numbers. Note first that since A is Hermitian, a and c are real. Then

$$x^H Ax = \begin{bmatrix} \bar{u} & \bar{v} \end{bmatrix} \begin{bmatrix} a & b \\ \bar{b} & c \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = a\bar{u}u + \bar{b}u\bar{v} + b\bar{u}v + c\bar{v}v$$

is real.

For the general proof, note that $x^H Ax$ is a 1×1 matrix. We have

$$\overline{(x^H Ax)} = (x^H Ax)^H = x^H Ax.$$

Property 2 Eigenvalues of A are real.

Suppose $Ax = \lambda x$. Multiplying x^H to the left on both sides yield

$$x^H Ax = x^H(\lambda x) = \lambda x^H x.$$

By the Property 1 and the fact that $x^H x \neq 0$, it follows that λ is real.

Property 3 If λ_i's are distinct, then x_i's are orthogonal.

Suppose $Ax_1 = \lambda_1 x_1$ and $Ax_2 = \lambda_2 x_2$ with $\lambda_1 \neq \lambda_2$. Then

$$\begin{aligned} \lambda_1 {x_1}^H x_2 &= (\lambda_1 x_1)^H x_2 = (Ax_1)^H x_2 = {x_1}^H Ax_2 \\ &= {x_1}^H(\lambda_2 x_2) = \lambda_2 {x_1}^H x_2 \end{aligned}$$

We used the property 2 that $\overline{\lambda_1} = \lambda_1$ in the first equality. It follows that ${x_1}^H x_2 = 0$, which means that

$$\langle x_1, x_2 \rangle = 0,$$

or that x_1 and x_2 are orthogonal.

Lemma(False) Suppose A is a real symmetric matrix. Then A has n distinct eigenvalues.

The roots of characteristic polynomals can be repeated. For example, $A = I$ has the only eigenvalue $\lambda = 1$, whose algebraic multiplicity is n. We regard $\lambda_1 = \lambda_2 = \cdots = \lambda_n = 1$. Any diagonal matrix with repeated diagonal entries can also be a counterexample : $A = \text{diag}\{1, 1, 4\}$ has three eigenvalues, $\lambda_1 = 1$, $\lambda_2 = 1$, $\lambda_3 = 4$.

Theorem Suppose A is an $n \times n$ real symmetric matrix with n distinct eigenvalues. Then A is orthogonally diagonalizable.

By the hypothesis, there exist $\lambda_1, \cdots, \lambda_n$ which are all distinct. And there correspond eigenvectors $x_1, \cdots, x_n$ such that

$$Ax_i = \lambda_i x_i \tag{$*$}$$

for each $i = 1, 2, \cdots, n$. By the property 3, x_i's are orthogonal. Moreover, we may assume that x_i are orthonormal. Let

$$Q = \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix}$$

be a $n \times n$ matrix. By orthonormality of x_i, Q is orthogonal. The above conditions $(*)$ reduces to

$$AQ = QD$$

where $D = \text{diag}\{\lambda_i\}$. A is orthogonally diagonalizable;

$$A = QDQ^T$$

3 Repeated Roots

Suppose that eigenvalues of A need not be distinct. That is, consider the case when A might have repeated roots.

Definition Suppose A is an $n \times n$ matrix. If M is another $n \times n$ matrix, A and $M^{-1}AM$ are said to be *similar*.

Remark 1 If we write $A \sim B$ for similarity, $\sim$ is an equivalance relation.

Remark 2 Suppose $A \sim B$ with $B = M^{-1}AM$. Then A and B have the same eigenvalues. And every eigenvector x of A corresponds to an eigenvector $M^{-1}x$ of B.

Note that $A - \lambda I$ and $B - \lambda I$ have the same determinants;

$$B - \lambda I = M^{-1}AM - \lambda I = M^{-1}(A - \lambda I)M$$
$$\det(B - \lambda I) = \det M^{-1} \det(A - \lambda I) \det M = \det(A - \lambda I)$$

Suppose $Ax = \lambda x$. Then $MBM^{-1}x = \lambda x$. It follows that $B(M^{-1}x) = \lambda(M^{-1}x)$.

Lemma(Schur) Suppose A is a complex square matrix. Then there exists a unitary matrix U such that

$$U^{-1}AU$$

is triangular.

Suppose A is a 4 by 4 matrix. A has at least one unit eigenvector x_1, which we place in the first column of U. By the Gram-Schmidt process, there exists a unitary U_1 such that

$$U_1{}^{-1}AU_1 = \begin{bmatrix} \lambda_1 & * & * & * \\ 0 & * & * & * \\ 0 & * & * & * \\ 0 & * & * & * \end{bmatrix}$$

Now consider the 3 by 3 submatrix in the lower right-hand corner. It has a unit eigenvector x_2, which becomes the first column of a unitary matrix M_2.

$$\text{Set } U_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & & & \\ 0 & & M_2 & \\ 0 & & & \end{bmatrix} \qquad \text{then} \qquad U_2{}^{-1}U_1{}^{-1}AU_1U_2 = \begin{bmatrix} \lambda_1 & * & * & * \\ 0 & \lambda_2 & * & * \\ 0 & 0 & * & * \\ 0 & 0 & * & * \end{bmatrix}$$

In a similar fashion,

$$U_3{}^{-1}U_2{}^{-1}U_1{}^{-1}AU_1U_2U_3 = \begin{bmatrix} \lambda_1 & * & * & * \\ 0 & \lambda_2 & * & * \\ 0 & 0 & \lambda_3 & * \\ 0 & 0 & 0 & * \end{bmatrix}$$

Thus, $U_1U_2U_3$ serves as U and the matrix on the right hand side is triangular.

Example The 4×4 matrix is too complicated for us to demonstrate the lemma as an example. Instead, consider the following 3×3 matrix;

$$A = \begin{bmatrix} 2 & 1 & -2 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

It has $\lambda = 1, -1, 2$ as eigenvalues. Choose $\lambda_1 = 1$. The corresponding eigenvector of length 1 is

$$x_1 = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \tag{1}$$

Since we have

$$Ax_1 = x_1, \tag{2}$$

construct a matrix U_1 which have x_1 as the first column. And impose U_1 to be unitary. We may set, for example,

$$U_1 = \begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & -\frac{2}{\sqrt{6}} \end{bmatrix}$$

Note that U_1 is not unique. From (2), we have

$$AU_1 = U_1 \begin{bmatrix} 1 & * & * \\ 0 & * & * \\ 0 & * & * \end{bmatrix} \tag{3}$$

Let the unkown matrix on the right hand side be B. We get

$$B = U_1{}^{-1} A U_1 = \begin{bmatrix} 1 & \frac{1}{\sqrt{6}} & \frac{3}{\sqrt{2}} \\ 0 & 0 & \sqrt{3} \\ 0 & \frac{2}{\sqrt{3}} & 1 \end{bmatrix}.$$

Let

$$\bar{B} = \begin{bmatrix} 0 & \sqrt{3} \\ \frac{2}{\sqrt{3}} & 1 \end{bmatrix}.$$

It has $\lambda = -1, 2$. Note that A and B[defined as in (7)] have exactly the same eigenvalues, which is trivial since A and B are similar(**Remark 2**). Choose $\lambda_2 = -1$. The corresponding unit eigenvector is

$$\overline{x_2} = \frac{1}{2} \begin{bmatrix} \sqrt{3} \\ -1 \end{bmatrix} \tag{4}$$

We have

$$\bar{B}\overline{x_2} = -\overline{x_2}. \tag{5}$$

Let $\overline{U_2}$ have $\overline{x_2}$ as the first column. Again, U_2 is unitary:

$$\overline{U_2} = \begin{bmatrix} \frac{\sqrt{3}}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{\sqrt{3}}{2} \end{bmatrix}.$$

(Here, $\overline{U_2}$ is not unique, but we only have two possibilities.) From (5), we have

$$\bar{B}\overline{U_2} = \overline{U_2} \begin{bmatrix} -1 & * \\ 0 & * \end{bmatrix}. \tag{6}$$

Let the unknown matrix on the right hand side be $\bar{C}$. We get

$$\bar{C} = \overline{U_2}^{-1} \bar{B} \overline{U_2} = \begin{bmatrix} -1 & \frac{1}{\sqrt{3}} \\ 0 & 2 \end{bmatrix}.$$

Let B, U_2 and C be 3×3 matrices and let x_2 be a column vector in $\mathbb{R}^3$ such that

$$B = \begin{bmatrix} 1 & 0 \\ 0 & \bar{B} \end{bmatrix}, \quad U_2 = \begin{bmatrix} 1 & 0 \\ 0 & \overline{U_2} \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 0 \\ 0 & \bar{C} \end{bmatrix}, \quad x_2 = \begin{bmatrix} 0 \\ \overline{x_2} \end{bmatrix}. \tag{7}$$

Then (6) reduces to

$$BU_2 = U_2 C. \tag{8}$$

By (3) and (8),

$$C = U_2{}^{-1} B U_2 = U_2{}^{-1} U_1{}^{-1} A U_1 U_2.$$

Note that C is a triangular matrix.

Theorem Suppose A is real symmetric(or complex Hermitian) matrix, Then A is orthogonally(unitarily) diagonalizable.

By Schur's lemma, there exists a unitary matrix U such that

$$U^{-1}AU = T$$

where T is a triangular matrix. Taking Hermtian on T yields

$$T^H = (U^H AU)^H = U^H AU = T.$$

Thus T is a diagonal matrix. Denote $T = D$, then

$$A = UDU^H.$$

and A is unitarily diagonalized.

References

[1] Gilbert Strang (4th edition) (2006) *Linear Algebra and its Applications* Belmont, CA : Thomson Brooks/Cole,

Section 5.5, 5.6