Chapter 9

The Diagonalization Theorems

Let V be a finite dimensional vector space and $T : V \to V$ be a linear transformation. One of the most basic questions one can ask about T is whether it is semi-simple, that is, whether T admits an eigenbasis. In matrix terms, this is equivalent to asking if T can be represented by a diagonal matrix. The purpose of this chapter is to study this question. Our main goal is to prove the Principal Axis (or Spectral) Theorem. After that, we will classify the unitarily diagonalizable matrices, that is the complex matrices of the form UDU^{-1}, where U is unitary and D is diagonal.

We will begin by considering the Principal Axis Theorem in the real case. This is the fundamental result that says every symmetric matrix admits an orthonormal eigenbasis. The complex version of this fact says that every Hermitian matrix admits a Hermitian orthonormal eigenbasis. This result is indispensable in the study of quantum theory. In fact, many of the basic problems in mathematics and the physical sciences involve symmetric or Hermitian matrices. As we will also point out, the Principal Axis Theorem can be stated in general terms by saying that every self adjoint linear transformation $T : V \to V$ on a finite dimensional inner product space V over $\mathbb{R}$ or $\mathbb{C}$ admits an orthonormal or Hermitian orthonormal eigenbasis. In particular, self adjoint maps are semi-simple.

9.1 The Real Case

We will now prove

Theorem 9.1. *Let $A \in \mathbb{R}^{n\times n}$ be symmetric. Then all eigenvalues of A are real, and there exists an orthonormal basis of $\mathbb{R}^n$ consisting of eigenvectors of A. Consequently, there exists an orthogonal matrix Q such that*

$$A = QDQ^{-1} = QDQ^T,$$

where $D \in \mathbb{R}^{n\times n}$ is diagonal.

There is an obvious conversely of the Principal Axis Theorem. If $A = QDQ^{-1}$, where Q is orthogonal and D is diagonal, then A is symmetric. This is rather obvious since any matrix of the form CDC^T is symmetric, and $Q^{-1} = Q^T$ for all $Q \in O(n, \mathbb{R})$.

9.1.1 Basic Properties of Symmetric Matrices

The first problem is to understand the geometric significance of the condition $a_{ij} = a_{ji}$ which defines a symmetric matrix. It turns out that this property implies several key geometric facts. The first is that every eigenvalue of a symmetric matrix is real, and the second is that two eigenvectors which correspond to different eigenvalues are orthogonal. In fact, these two facts are all that are needed for our first proof of the Principal Axis Theorem. We will give a second proof which gives a more complete understanding of the geometric principles behind the result.

We will begin by formulating the condition $a_{ij} = a_{ji}$ in a more useful form.

Proposition 9.2. *A matrix $A \in \mathbb{R}^{n\times n}$ is symmetric if and only if*

$$\mathbf{v}^T A\mathbf{w} = \mathbf{w}^T A\mathbf{v}$$

for all $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$.

Proof. To see this, notice that since $\mathbf{v}^T A\mathbf{w}$ is a scalar, it equals its own transpose. Thus

$$\mathbf{v}^T A\mathbf{w} = (\mathbf{v}^T A\mathbf{w})^T = \mathbf{w}^T A^T \mathbf{v}.$$

So if $A = A^T$, then

$$\mathbf{v}^T A\mathbf{w} = \mathbf{w}^T A\mathbf{v}.$$

For the converse, use the fact that

$$a_{ij} = \mathbf{e}_i^T A\mathbf{e}_j,$$

so if $\mathbf{e}_i^T A\mathbf{e}_j = \mathbf{e}_j^T A\mathbf{e}_i$, then $a_{ij} = a_{ji}$. □

The linear transformation T_A defined by a symmetric matrix $A \in \mathbb{R}^{n\times n}$ is called *self adjoint*. Thus $T_A : \mathbb{R}^n \to \mathbb{R}^n$ is self adjoint if and only if $T_A(\mathbf{v}) \cdot \mathbf{w} = \mathbf{v} \cdot T_A(\mathbf{w})$ for all $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$.

We now establish the two basic properties mentioned above. We first show

Proposition 9.3. *Eigenvectors of a real symmetric $A \in \mathbb{R}^{n\times n}$ corresponding to different eigenvalues are orthogonal.*

Proof. Let $(\lambda, \mathbf{u})$ and $(\mu, \mathbf{v})$ be eigenpairs such that $\lambda \neq \mu$. Then

$$\mathbf{u}^T A\mathbf{v} = \mathbf{u}^T \mu\mathbf{v} = \mu\mathbf{u}^T\mathbf{v}$$

while

$$\mathbf{v}^T A\mathbf{u} = \mathbf{v}^T \lambda\mathbf{u} = \lambda\mathbf{v}^T\mathbf{u}.$$

Since A is symmetric, Proposition 9.2 says that $\mathbf{u}^T A\mathbf{v} = \mathbf{v}^T A\mathbf{u}$. Hence $\lambda\mathbf{u}^T\mathbf{v} = \mu\mathbf{v}^T\mathbf{u}$. But $\mathbf{u}^T\mathbf{v} = \mathbf{v}^T\mathbf{u}$, so $(\lambda - \mu)\mathbf{u}^T\mathbf{v} = 0$. Since $\lambda \neq \mu$, we infer $\mathbf{u}^T\mathbf{v} = 0$, which finishes the proof. □

We next show that the second property.

Proposition 9.4. *All eigenvalues of a real symmetric matrix are real.*

We will first establish a general fact that will also be used in the Hermitian case.

Lemma 9.5. *Suppose $A \in \mathbb{R}^{n\times n}$ is symmetric. Then*

$$\overline{\mathbf{v}}^T A\mathbf{v} \in \mathbb{R}$$

for all $\mathbf{v} \in \mathbb{C}^n$.

Proof. Since $\overline{\alpha + \beta} = \overline{\alpha} + \overline{\beta}$ and $\overline{\alpha\beta} = \overline{\alpha}\,\overline{\beta}$ for all $\alpha, \beta \in \mathbb{C}$, we easily see that

$$\begin{aligned}\overline{\overline{\mathbf{v}}^T A\mathbf{v}} &= \mathbf{v}^T\overline{A}\overline{\mathbf{v}} \\ &= \mathbf{v}^T A\overline{\mathbf{v}} \\ &= \left(\mathbf{v}^T A\overline{\mathbf{v}}\right)^T \\ &= \overline{\mathbf{v}}^T A^T\mathbf{v},\end{aligned}$$

As $\alpha \in \mathbb{C}$ is real if and only if $\overline{\alpha} = \alpha$, we have the result. □

To prove the Proposition, let $A \in \mathbb{R}^{n\times n}$ be symmetric. Since the characteristic polynomial of A is a real polynomial of degree n, the Fundamental

Theorem of Algebra implies it has n roots in $\mathbb{C}$. Suppose that $\lambda \in \mathbb{C}$ is a root. It follows that there exists a $\mathbf{v} \neq \mathbf{0}$ in $\mathbb{C}^n$ so that $A\mathbf{v} = \lambda\mathbf{v}$. Hence

$$\overline{\mathbf{v}}^T A\mathbf{v} = \overline{\mathbf{v}}^T \lambda\mathbf{v} = \lambda\overline{\mathbf{v}}^T\mathbf{v}.$$

We may obviously assume $\lambda \neq 0$, so the right hand side is nonzero. Indeed, if $\mathbf{v} = (v_1, v_2, \dots v_n)^T \neq \mathbf{0}$, then

$$\overline{\mathbf{v}}^T\mathbf{v} = \sum_{i=1}^{n} \overline{v_i}v_i = \sum_{i=1}^{n} |v_i|^2 > 0.$$

Since $\overline{\mathbf{v}}^T A\mathbf{v} \in \mathbb{R}$, λ is a quotient of two real numbers, so $\lambda \in \mathbb{R}$. Thus all eigenvaluesof A are real, which completes the proof of the Proposition. $\square$

9.1.2 Some Examples

Example 9.1. Let H denote a 2×2 reflection matrix. Then H has eigenvalues ± 1. Either unit vector $\mathbf{u}$ on the reflecting line together with either unit vector $\mathbf{v}$ orthogonal to the reflecting line form an orthonormal eigenbasis of $\mathbb{R}^2$ for H. Thus $Q = (\mathbf{u}\ \mathbf{v})$ is orthogonal and

$$H = Q\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} Q^{-1} = Q\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} Q^T.$$

Note that there are only four possible choices for Q. All 2×2 reflection matrices are similar to $\mathrm{diag}[1, -1]$. The only thing that can vary is Q.

Here is another example.

Example 9.2. Let B be 4×4 the all ones matrix. The rank of B is one, so 0 is an eigenvalue and $\mathcal{N}(B) = E_0$ has dimension three. In fact $E_0 = (\mathbb{R}(1,1,1,1)^T)^{\perp}$. Another eigenvalue is 4. Indeed, $(1,1,1,1)^T \in E_4$, so we know there exists an eigenbasissince $\dim E_0 + \dim E_4 = 4$. To produce an orthonormal basis, we simply need to find an orthonormal basis of $(\mathbb{R}(1,1,1,1)^T)^{\perp}$. We will do this by inspection rather than Gram-Schmidt, since it is easy to find vectors orthogonal to $(1,1,1,1)^T$. In fact, $\mathbf{v}_1 = (1,-1,0,0)^T$, $\mathbf{v}_2 = (0,0,1,-1)^T$, and $\mathbf{v}_3 = (1,1,-1,-1)^T$ give an orthonormal basis after we normalize. We know that our fourth eigenvector, $\mathbf{v}_4 = (1,1,1,1)^T$, is orthogonal to E_0, so we can for example express B as QDQ^T where $Q = \begin{pmatrix} \frac{\mathbf{v}_1}{\sqrt{2}} & \frac{\mathbf{v}_2}{\sqrt{2}} & \frac{\mathbf{v}_3}{2} & \frac{\mathbf{v}_4}{2} \end{pmatrix}$ and

$$D = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 4 \end{pmatrix}.$$

9.1.3 The First Proof

Since a symmetric matrix $A \in \mathbb{R}^{n\times n}$ has n real eigenvalues and eigenvectors corresponding to different eigenvalues are orthogonal, there is nothing to prove when all the eigenvalues are distinct. The difficulty is that if A has repeated eigenvalues, say $\lambda_1, \dots \lambda_m$, then one has to show

$$\sum_{i=1}^{m} \dim E_{\lambda_i} = n.$$

In our first proof, we avoid this difficulty completely. The only facts we need are the Gram-Schmidt property and the group theoretic property that the product of any two orthogonal matrices is orthogonal.

To keep the notation simple and since we will also give a second proof, we will only do the 3×3 case. In fact, this case actually involves all the essential ideas. Let A be real 3×3 symmetric, and begin by choosing an eigenpair $(\lambda_1, \mathbf{u}_1)$ where $\mathbf{u}_1 \in \mathbb{R}^3$ is a unit vector. By the Gram-Schmidt process, we can include $\mathbf{u}_1$ in an orthonormal basis $\mathbf{u}_1$, $\mathbf{u}_2$, $\mathbf{u}_3$ of $\mathbb{R}^3$. Let $Q_1 = (\mathbf{u}_1\ \mathbf{u}_2\ \mathbf{u}_3)$. Then Q_1 is orthogonal and

$$AQ_1 = (A\mathbf{u}_1\ A\mathbf{u}_2\ A\mathbf{u}_3) = (\lambda_1\mathbf{u}_1\ A\mathbf{u}_2\ A\mathbf{u}_3).$$

Now

$$Q_1^T AQ_1 = \begin{pmatrix} \mathbf{u}_1^T \\ \mathbf{u}_2^T \\ \mathbf{u}_3^T \end{pmatrix} (\lambda_1\mathbf{u}_1\ A\mathbf{u}_2\ A\mathbf{u}_3) = \begin{pmatrix} \lambda_1\mathbf{u}_1^T\mathbf{u}_1 & * & * \\ \lambda_1\mathbf{u}_2^T\mathbf{u}_1 & * & * \\ \lambda_1\mathbf{u}_3^T\mathbf{u}_1 & * & * \end{pmatrix}.$$

But since $Q_1^T AQ_1$ is symmetric (since A is), and since $\mathbf{u}_1$, $\mathbf{u}_2$, $\mathbf{u}_3$ are orthonormal, we see that

$$Q_1^T AQ_1 = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & * & * \\ 0 & * & * \end{pmatrix}.$$

Obviously the 2×2 matrix in the lower right hand corner of A is symmetric. Calling this matrix B, we can find, by repeating the construction just given, a 2×2 orthogonal matrix Q' so that

$$Q'^T BQ' = \begin{pmatrix} \lambda_2 & 0 \\ 0 & \lambda_3 \end{pmatrix}.$$

Putting $Q' = \begin{pmatrix} r & s \\ t & u \end{pmatrix}$, it follows that

$$Q_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & r & s \\ 0 & t & u \end{pmatrix}$$

is orthogonal, and in addition

$$Q_2^T Q_1^T A Q_1 Q_2 = Q_2^T \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & * & * \\ 0 & * & * \end{pmatrix} Q_2 = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}.$$

But Q_1Q_2 is orthogonal, and so $(Q_1Q_2)^{-1} = Q_2^{-1}Q_1^{-1} = Q_2^T Q_1^T$. Therefore, putting $Q = Q_1Q_2$ and $D = \text{diag}(\lambda_1, \lambda_2, \lambda_3)$, we get $A = QDQ^{-1} = QAQ^T$. Therefore A has been orthogonally diagonalized and the first proof is done. □

Note that by using mathematical induction, we can extend the above proof to the general case. The drawback of the above technique is that it requires a repeated application of the Gram-Schmidt process, which is not very illuminating. In the next section, we will give the second proof.

9.1.4 A Geometric Proof

Let $A \in \mathbb{R}^{n\times n}$ be symmetric. We first prove the following geometric property of symmetric matrices:

Proposition 9.6. *If $A \in \mathbb{R}^{n\times n}$ is symmetric and W is a subspace of $\mathbb{R}^n$ such that $T_A(W) \subset W$, then $T_A(W^\perp) \subset W^\perp$ too.*

Proof. Let $\mathbf{x} \in W$ and $\mathbf{y} \in W^\perp$. Since $\mathbf{x}^T A\mathbf{y} = \mathbf{y}^T A\mathbf{x}$, it follows that if $A\mathbf{x} \cdot \mathbf{y} = 0$, then $\mathbf{x} \cdot A\mathbf{y} = 0$. It follows that $A\mathbf{y} \in W^\perp$, so the Proposition is proved. □

We also need

Proposition 9.7. *If $A \in \mathbb{R}^{n\times n}$ is symmetric and W is a nonzero subspace of $\mathbb{R}^n$ with the property that $T_A(W) \subset W$, then W contains an eigenvector of A.*

Proof. Pick an orthonormal basis $\mathcal{Q} = \{\mathbf{u}_1, \dots, \mathbf{u}_m\}$ of W. As $T_A(W) \subset W$, there exist scalars r_{ij} $(1 \le i, j \le m)$, such that

$$T_A(\mathbf{u}_j) = A\mathbf{u}_j = \sum_{i=1}^{m} r_{ij}\mathbf{u}_i.$$

This defines an $m \times m$ matrix R, which I claim is symmetric. Indeed, since T_A is self adjoint,

$$r_{ij} = \mathbf{u}_i \cdot A\mathbf{u}_j = A\mathbf{u}_i \cdot \mathbf{u}_j = r_{ji}.$$

Let $(\lambda, (x_1, \dots, x_m)^T)$ be an eigenpair for R. Putting $\mathbf{w} := \sum_{j=1}^m x_j \mathbf{u}_j$, I claim that $(\lambda, \mathbf{w})$ is an eigenpair for A. In fact,

$$\begin{aligned}
T_A(\mathbf{w}) &= \sum_{j=1}^{m} x_j T_A(\mathbf{u}_j) \\
&= \sum_{j=1}^{m} x_j (\sum_{i=1}^{m} r_{ij}\mathbf{u}_i) \\
&= \sum_{i=1}^{m} (\sum_{j=1}^{m} r_{ij} x_j)\mathbf{u}_i \\
&= \sum_{i=1}^{m} \lambda x_i \mathbf{u}_i \\
&= \lambda \mathbf{w}.
\end{aligned}$$

This finishes the proof of the Proposition. □

The proof of the Principal Axis Theorem now goes as follows. Starting with an eigenpair $(\lambda_1, \mathbf{w}_1)$, put $W_1 = (\mathbb{R}\mathbf{w}_1)$. Then $T_A(W_1) \subset W_1$, so $T_A(W_1^\perp) \subset W_1^\perp$. Now either $W_1^\perp$ contains an eigenvector or $n = 1$ and there is nothing to show. Suppose $W_1^\perp$ contains an eigenvector $\mathbf{w}_2$. Then either $(\mathbb{R}\mathbf{w}_1 + \mathbb{R}\mathbf{w}_2)^\perp$ contains an eigenvector $\mathbf{w}_3$, or we are through. Continuing in this manner, we obtain a sequence of orthogonal eigenvectors $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_n$. This implies there exists an orthonormal eigenbasisfor A, so the proof is completed.

9.1.5 A Projection Formula for Symmetric Matrices

Sometimes it's useful to express the Principal Axis Theorem as a projection formula for symmetric matrices. Let A be symmetric, let $\mathbf{u}_1, \dots, \mathbf{u}_n$ be an orthonormal eigenbasis of $\mathbb{R}^n$ for A, and suppose $(\lambda_i, \mathbf{u}_i)$ is an eigenpair. Suppose $\mathbf{x} \in \mathbb{R}^n$. By the projection formula of Chapter 8,

$$\mathbf{x} = (\mathbf{u}_1^T\mathbf{x})\mathbf{u}_1 + \cdots + (\mathbf{u}_n^T\mathbf{x})\mathbf{u}_n,$$

hence

$$A\mathbf{x} = \lambda_1(\mathbf{u}_1^T\mathbf{x})\mathbf{u}_1 + \cdots + \lambda_n(\mathbf{u}_n^T\mathbf{x})\mathbf{u}_n.$$

This amounts to writing

$$A = \lambda_1 \mathbf{u}_1 \mathbf{u}_1^T + \cdots + \lambda_n \mathbf{u}_n \mathbf{u}_n^T. \tag{9.1}$$

Recall that $\mathbf{u}_i \mathbf{u}_i^T$ is the matrix of the projection of $\mathbb{R}^n$ onto the line $\mathbb{R}\mathbf{u}_i$, so (9.1) expresses A as a sum of orthogonal projections.

Exercises

Exercise 9.1. Orthogonally diagonalize the following matrices:

$$\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 1 & 3 \\ 1 & 3 & 1 \\ 3 & 1 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}.$$

I claim that you can diagonalize the first and second matrices, and a good deal (if not all) of the third, without pencil and paper.

Exercise 9.2. Prove the Principal Axis Theoremfor and 2×2 real matrix $\begin{pmatrix} a & b \\ b & c \end{pmatrix}$ directly.

Exercise 9.3. Prove Proposition9.6.

Exercise 9.4. Answer either T or F. If T, give a brief reason. If F, give a counter example.

(a) The sum and product of two symmetric matrices is symmetric.

(b) For any real matrix A, the eigenvalues of A^TA are all real.

(c) For A as in (b), the eigenvalues of A^TA are all non negative.

(d) If two symmetric matrices A and B have the same eigenvalues, counting multiplicities, then A and B are orthogonally similar (that is, $A = QBQ^T$ where Q is orthogonal).

Exercise 9.5. Recall that two matrices A and B which have a common eigenbasis commute. Conclude that if A and B have a common eigenbasis and are symmetric, then AB is symmetric.

Exercise 9.6. Describe the orthogonal diagonalization of a reflection matrix.

Exercise 9.7. Let W be a hyperplane in $\mathbb{R}^n$, and let H be the reflection through W.

(a) Express H in terms of P_W and $P_{W^\perp}$.

(b) Show that $P_W P_{W^\perp} = P_{W^\perp} P_W$.

(c) Simultaneously orthogonally diagonalize P_W and $P_{W^\perp}$.

Exercise 9.8. * Diagonalize

$$\begin{pmatrix} a & b & c \\ b & c & a \\ c & a & b \end{pmatrix},$$

where a, b, c are all real. (Note that the second matrix in Problem 1 is of this type. What does the fact that the trace is an eigenvalue say?)

Exercise 9.9. * Diagonalize

$$A = \begin{pmatrix} aa & ab & ac & ad \\ ba & bb & bc & bd \\ ca & cb & cc & cd \\ da & db & dc & dd \end{pmatrix},$$

where a, b, c, d are arbitrary real numbers. (Note: thimk!)

Exercise 9.10. Prove that a real symmetric matrix A whose only eigenvalues are ± 1 is orthogonal.

Exercise 9.11. Suppose $A \in \mathbb{R}^{n\times n}$ is symmetric. Show the following.

(i) $\mathcal{N}(A)^{\perp} = \mathrm{Im}(A)$.

(ii) $\mathrm{Im}(A)^{\perp} = \mathcal{N}(A)$.

(iii) $\mathrm{col}(A) \cap \mathcal{N}(A) = \{\mathbf{0}\}$.

(iv) Conclude form (iii) that if $A^k = O$ for some $k > 0$, then $A = O$.

Exercise 9.12. Give a proof of the Principal Axis Theorem from first principles in the 2×2 case.

Exercise 9.13. Show that two symmetric matrices A and B that have the same characteristic polynomial are orthogonally similar. That is, $A = QBQ^{-1}$ for some orthogonal matrix Q.

Exercise 9.14. Let $A \in \mathbb{R}^{n\times n}$ be symmetric, and let λ_m and λ_M be its minimum and maximum eigenvalues respectively.

(a) Show that for every $\mathbf{x} \in \mathbb{R}^n$, we have

$$\lambda_m \mathbf{x}^T\mathbf{x} \leq \mathbf{x}^T A\mathbf{x} \leq \lambda_M \mathbf{x}^T\mathbf{x}.$$

(b) Use this inequality to find the maximum and minimum values of $|A\mathbf{x}|$ on the ball $|\mathbf{x}| \leq 1$.

Exercise 9.15. Prove that an element $Q \in \mathbb{R}^{n\times n}$ is a reflection if and only if Q is symmetric and $\det(Q) = -1$.

9.2 The Principal Axis Theorem for Hermitian Matrices

The purpose of this section is to extend the Principal Axis Theorem to the complex case.

9.2.1 Hermitian Matrices

A linear map $T : \mathbb{C}^n \to \mathbb{C}^n$ is self adjoint if and only if

$$T(\mathbf{w}) \cdot \mathbf{z} = \mathbf{w} \cdot T(\mathbf{z})$$

for all $\mathbf{w}, \mathbf{z} \in \mathbb{C}^n$. Recall that in the $\mathbb{C}^n$ setting, the inner productis the Hermitianinner product

$$\mathbf{w} \cdot \mathbf{z} := \overline{w_1} z_1 + \overline{w_2} z_2 + \cdots + \overline{w_n} z_n.$$

Thus T is self adjoint if and only if its matrix A satisfies $\overline{\mathbf{w}}^T \overline{A}^T \mathbf{z} = \overline{\mathbf{w}}^T A\mathbf{z}$. This is equivalent to saying that

$$\overline{A}^T = A.$$

Definition 9.1. Let $A = (\alpha_{ij}) \in \mathbb{C}^{m \times n}$. Then the *Hermitian transpose* of $A = (\alpha_{ij})$ is the matrix

$$A^H := \overline{A}^T,$$

where $\overline{A}$ is the matrix $(\overline{\alpha_{ij}})$ obtained by conjugating the entries of A. We say that A is *Hermitian* if and only if $A^H = A$.

Thus the linear transformation associated to $A \in \mathbb{C}^{n \times n}$ is self adjoint A is Hermitian.

Example 9.3. For example,

$$\begin{pmatrix} 1 & 1+i & -i \\ 1-i & 3 & 2 \\ i & 2 & 0 \end{pmatrix}$$

is Hermitian.

Clearly, the real Hermitian matrices are exactly the symmetric matrices. By repeating the proof of Proposition 9.4, we get the fact that all eigenvaluesof a Hermitian matrix are real.

9.2.2 The Main Result

Recall that a basis of $\mathbb{C}^n$ which is orthonormal for the Hermitian inner product is called a Hermitian basis. If $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$ is Hermitian, the matrix $U = (\mathbf{u}_1 \ \mathbf{u}_2 \ \ldots \ \mathbf{u}_n)$ is unitary, that is $U^{-1} = U^H$.

Recall that eigenvaluesof Hermitian matrices are real.

Theorem 9.8. *Every self adjoint linear map $T : \mathbb{C}^n \to \mathbb{C}^n$ admits a Hermitian eigenbasis. Equivalently, if $K \in \mathbb{C}^{n\times n}$ is Hermitian, then there exists a unitary matrix U and a real diagonal matrix D such that $K = UDU^{-1} = UDU^H$.*

The proof is a carbon copy of the proof in the real symmetric case, so we won't need to repeat it. Note however that in the complex case, the so called *principal axes* are actually one dimensional complex subspaces of $\mathbb{C}^n$. Hence the principal axes are actually real two planes (an $\mathbb{R}^2$) instead of lines as in the real case.

Exercises

Exercise 9.16. Find the eigen-values of $K = \begin{pmatrix} 2 & 3+4i \\ 3-4i & -2 \end{pmatrix}$ and diagonalize K.

Exercise 9.17. Unitarily diagonalize $R_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$.

Exercise 9.18. Show that the trace and determinant of a Hermitian matrix are real. In fact, show that the characteristic polynomial of a Hermitian matrix has real coefficients.

Exercise 9.19. Prove that the Hermitian matrices are exactly the complex matrices with real eigen-values that can be diagonalized using a unitary matrix.

Exercise 9.20. Show that $U(n, \mathbb{C})$ is a matrix group. Can you find a general description for $U(2, \mathbb{C})$?

Exercise 9.21. Show that two unit vectors in $\mathbb{C}^n$ coincide if and only if their dot product is 1.

Exercise 9.22. Give a description of the set of all 1×1 unitary matrices. That is, describe $U(1, \mathbb{C})$.

Exercise 9.23. Consider a 2×2 unitary matrix U such that one of U's columns is in $\mathbb{R}^2$. Is U orthogonal?

Exercise 9.24. Prove assertions (i)-(iii) in Proposition**??**.

Exercise 9.25. Supppose W is a complex subspace of $\mathbb{C}^n$. Show that the projection P_W is Hermitian.

Exercise 9.26. How does one adjust the formula $P_W = A(AA^T)^{-1}A^T$ to get the formula for the projection of a complex subspace W of $\mathbb{C}^n$?

Exercise 9.27. Give a direct proof of the Principal Axis Theorem in the 2×2 Hermitian case.

9.3 Self Adjoint Operators

The purpose of this section is to formulate the Principal Axis Theorem for an arbitrary finite dimensional inner product space V. In order to do this, we have to make some preliminary comments about this class of spaces.

9.3.1 The Definition

In the previous section, we defined the notion of a self adjoint linear map $T : \mathbb{R}^n \to \mathbb{R}^n$. The notion of a self adjoint operator on an arbitrary inner product space is exactly the same. We will treat this in a slightly more general way, however. First we make the following definition.

Definition 9.2. Let V be a real inner product space and suppose $T : V \to V$ is linear. Define the *adjoint* of T to be the map $T^* : V \to V$ determined by the condition that

$$(T^*(\mathbf{x}), \mathbf{y}) = (\mathbf{x}, T(\mathbf{y}))$$

for all $\mathbf{x}, \mathbf{y} \in V$. Then we say that T is *self adjoint* if and only if $T = T^*$.

Proposition 9.9. *Let V be a real inner product space and suppose $T : V \to V$ is linear. Then the adjoint $T^* : V \to V$ is also a well defined linear transformation. If V is finite dimensional, then T is self adjoint if and only if for every orthonormal basis $\mathcal{Q}$ of V, the matrix $\mathcal{M}^{\mathcal{Q}}_{\mathcal{Q}}(T)$ is symmetric. More generally, the matrix $\mathcal{M}^{\mathcal{Q}}_{\mathcal{Q}}(T^*)$ is $\mathcal{M}^{\mathcal{Q}}_{\mathcal{Q}}(T)^T$.*

Proof. The proof is left as an exercise. □

Hence a symmetric matrix is a self adjoint linear transformation from $\mathbb{R}^n$ to itself and conversely. Therefore the eigenvalue problem for self adjoint maps on a finite dimensional inner product space reduces to the eigenvalue problem for symmetric matrices on $\mathbb{R}^n$.

Here is a familiar example.

Example 9.4. Let W be a subspace of $\mathbb{R}^n$. Then the projection P_W is self adjoint. In fact, we know that its matrix with respect to the standard basis has the form $C(CC^T)^{-1}C^T$, which is clearly symmetric. Another way to see the self adjointness is to choose an orthonormal basis $\mathbf{u}_1, \ldots, \mathbf{u}_n$ of $\mathbb{R}^n$ so that $\mathbf{u}_1, \ldots, \mathbf{u}_m$ span W. Then, by the projection formula, $P_W(\mathbf{x}) = \sum_{i=1}^{k}(\mathbf{x} \cdot \mathbf{u}_i)\mathbf{u}_i$. It follows easily that $P_W(\mathbf{u}_i) \cdot \mathbf{u}_j = \mathbf{u}_i \cdot P_W(\mathbf{u}_j)$ for all indices i and j. Hence P_W is self adjoint.

To summarize the Principal Axis Theorem for self adjoint operators, we state

Theorem 9.10. *Let V be a finite dimensional inner product space, and let $T : V \to V$ be self adjoint. Then there exists an orthonormal eigenbasis $\mathcal{Q}$ of V consisting of eigenvectors of T. Thus T is semi-simple, and the matrix $\mathcal{M}^{\mathcal{Q}}_{\mathcal{Q}}(T)$ is diagonal.*

Proof. Let $\mathcal{B} = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ denote an orthonormal basis of V. The map $\Phi : V \to \mathbb{R}^n$ defined by $\Phi(\mathbf{u}_i) = \mathbf{e}_i$ is an isometry (see Proposition8.14). Now $S = \Phi T \Phi^{-1}$ is a self adjoint map of $\mathbb{R}^n$ (check), hence S has an orthonormal eigenbasis $\mathbf{x}_1, \dots, \mathbf{x}_n \in \mathbb{R}^n$. Since Φ is an isometry, $\mathbf{v}_1 = \Phi^{-1}(\mathbf{x}_1), \dots, \mathbf{v}_n = \Phi^{-1}(\mathbf{x}_n)$ form an orthonormal basis of V. Moreover, the $\mathbf{v}_i$ are eigenvectors of T. For, if $S(\mathbf{x}_i) = \lambda_i \mathbf{x}_i$, then

$$T(\mathbf{v}_i) = \Phi^{-1} S \Phi(\mathbf{v}_i) = \Phi^{-1} S(\mathbf{x}_i) = \Phi^{-1}(\lambda_i \mathbf{x}_i) = \lambda_i \Phi^{-1}(\mathbf{x}_i) = \lambda_i \mathbf{v}_i.$$

Thus T admits an orthonormal eigenbasis, as claimed. □

9.3.2 An Infinite Dimensional Self Adjoint Operator

We now give an example of a self adjoint operator (or linear transformation) in the infinite dimensional setting. As mentioned in the introduction, self adjoint operators are frequently encountered in mathematical, as well as physical, problems.

We will consider a certain subspace of function space $C[a, b]$ of all continuous functions $f : [a, b] \to \mathbb{R}$ with the usual inner product

$$(f, g) = \int_a^b f(x)g(x)\mathrm{d}x.$$

The condition for a linear transformation $T : C[a, b] \to C[a, b]$ to be self adjoint is that satisfies the condition $(Tf, g) = (f, Tg)$ for all f, g, that is

$$\int_a^b T(f)(x)g(x)\mathrm{d}x = \int_a^b f(x)T(g)(x)\mathrm{d}x.$$

Now let $[a, b] = [0, 2\pi]$, and let $\mathcal{P}$ (for periodic) denote the subspace of $C[0, 2\pi]$ consisting of all functions f which have derivatives of all orders on $[0, 2\pi]$ and satisfy the further condition that

$$f^{(i)}(0) = f^{(i)}(2\pi) \quad \text{if} \quad i = 0, 1, 2, \dots,$$

where $f^{(i)}$ denotes the ith derivative of f. Among the functions in $\mathcal{P}$ are the trigonometric functions $\cos \lambda x$ and $\sin \lambda x$ for all $\lambda \in \mathbb{R}$. We will show below that these functions are linearly independent if $\lambda > 0$, so $\mathcal{P}$ is an infinite dimensional space.

We next give an example of a self adjoint operator on $\mathcal{P}$. Thus symmetric matrices can have infinite dimensional analogues. By the definition of $\mathcal{P}$, it is clear that if $f \in \mathcal{P}$, then $f^{(i)} \in \mathcal{P}$ for all $i \geq 1$. Hence the derivative operator $D(f) = f'$ defines a linear transformation $D : \mathcal{P} \to \mathcal{P}$. I claim the second derivative $D^2(f) = f''$ is self adjoint. To prove this, we have to show $(D^2(f), g) = (f, D^2(g))$ for all $f, g \in \mathcal{P}$. This follows from integration by parts. For we have

$$\begin{aligned}(D^2(f), g) &= \int_0^{2\pi} f''(t)g(t)\mathrm{d}t \\ &= f'(2\pi)g(2\pi) - f'(0)g(0) - \int_0^{2\pi} f'(t)g'(t)\mathrm{d}t.\end{aligned}$$

But by the definition of $\mathcal{P}$, $f'(2\pi)g(2\pi) - f'(0)g(0) = 0$, so

$$(D^2(f), g) = -\int_0^{2\pi} f'(t)g'(t)\mathrm{d}t.$$

Since this expression for $(D^2(f), g)$ is symmetric in f and g, it follows that

$$(D^2(f), g) = (f, D^2(g)),$$

so D^2 is self adjoint, as claimed.

We can now ask for the eigenvalues and corresponding eigenfunctions of D^2. There is no general method for finding the eigenvalues of a linear operator on an infinite dimensional space, but one can easily see that the trig functions $\cos \lambda x$ and $\sin \lambda x$ are eigenfunctions for $-\lambda^2$ if $\lambda \neq 0$. Now there is a general theorem in differential equations that asserts that if $\mu > 0$, then any solution of the equation

$$D^2(f) + \mu f = 0$$

has the form $f = a \cos \sqrt{\mu} x + b \sin \sqrt{\mu} x$ for some $a, b \in \mathbb{R}$. Moreover, $\lambda = 0$ is an eigenvalue for eigenfunction $1 \in \mathcal{P}$. Note that although $D^2(x) = 0$, x is not an eigenfunction since $x \notin \mathcal{P}$.

To summarize, D^2 is a self adjoint operator on $\mathcal{P}$ such that every non positive real number is an ev. The corresponding eigenspaces are $E_0 = \mathbb{R}$ and $E_{-\lambda} = \mathbb{R} \cos \sqrt{\lambda} x + \mathbb{R} \sin \sqrt{\lambda} x$ if $\lambda > 0$. We can also draw some other consequences. For any positive $\lambda_1, \dots \lambda_k$ and any $f_i \in E_{-\lambda_i}$, $f_1, \dots, f_k$ are linearly independent. Therefore, the dimension of $\mathcal{P}$ cannot be finite, i.e. $\mathcal{P}$ is infinite dimensional.

Recall that distinct eigenvalues of a symmetric matrix have orthogonal eigenspaces. Thus distinct eigenvalues of a self adjoint linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ have orthogonal eigenspaces. The proof of this goes over unchanged to $\mathcal{P}$, so if $\lambda, \mu > 0$ and $\lambda \neq \mu$, then

$$\int_0^{2\pi} f_\lambda(t) f_\mu(t) \mathrm{d}t = 0,$$

where f_λ and f_μ are any eigenfunctions for $-\lambda$ and $-\mu$ respectively. In particular,

$$\int_0^{2\pi} \sin\sqrt{\lambda}t \sin\sqrt{\mu}t \mathrm{d}t = 0,$$

with corresponding identities for the other pairs of eigenfunctions f_λ and f_μ. In addition, $\cos\sqrt{\lambda}x$ and $\sin\sqrt{\lambda}x$ are also orthogonal.

The next step is to normalize the eigenfunctions to obtain an orthonormal set. Clearly if $\lambda \neq 0$, $|f_\lambda|^2 = (f_\lambda, f_\lambda) = \pi$, while $|f_0|^2 = 2\pi$. Hence the functions

$$\frac{1}{\sqrt{2\pi}}, \quad \frac{1}{\sqrt{\pi}}\cos\sqrt{\lambda}x, \quad \frac{1}{\pi}\sin\sqrt{\lambda}x,$$

where $\lambda > 0$ are a family of ON functions in $\mathcal{P}$. It turns out that one usually considers only the eigenfunctions where λ is a positive integer. The *Fourier series* of a function $f \in C[0, 2\pi]$ such that $f(0) = f(2\pi)$ is the infinite series development

$$f(x) \approx \frac{1}{\pi}\sum_{m=1}^{\infty} a_m \cos mx + \frac{1}{\pi}\sum_{m=1}^{\infty} b_m \sin mx,$$

where a_m and b_m are the Fourier coefficients encountered in §33. In particular,

$$a_m = \frac{1}{\sqrt{\pi}}\int_0^{2\pi} f(t)\cos mt \mathrm{d}t$$

and

$$b_m = \frac{1}{\sqrt{\pi}}\int_0^{2\pi} f(t)\sin mt \mathrm{d}t.$$

For a precise interpretation of the meaning $\approx$, we refer to a text on Fourier series. The upshot of this example is that Fourier series are an important tool in partial differential equations, mathematical physics and many other areas.

Exercises

Exercise 9.28. Show that if V is a finite dimensional inner product space, then $T \in L(V)$ is self adjoint if and only if for every orthonormal basis $\mathbf{u}_1, \dots, \mathbf{u}_n$ OF V, $(T(\mathbf{u}_i), \mathbf{u}_j) = (\mathbf{u}_i, T(\mathbf{u}_j))$ for all indices i and j.

Exercise 9.29. Let U and V be inner product spaces of the same dimension. Show that a linear transformation $\Phi : U \to V$ is an isometry if and only if Φ carries every orthonormal basis of U onto an orthonormal basis of V.

Exercise 9.30. Give an example of a linear transformation$\Phi : \mathbb{R}^2 \to \mathbb{R}^2$ that isn't an isometry.

Exercise 9.31. Show that the matrix of an isometry $\Phi : U \to U$ with respect to and orthonormal basis is orthogonal. Conversely show that given an orthonormal basis, any orthogonal matrix defines an isometry form U to itself.

Exercise 9.32. Give the proof of Proposition9.9 .

Exercise 9.33. Describe all isometries $\Phi : \mathbb{R}^2 \to \mathbb{R}^2$.

Exercise 9.34. Prove Proposition9.9.

9.4 Normal Matrices and Schur's Theorem

The result that any Hermitian matrix K can be expressed in the form $K = UDU^H$, where D is real diagonal and U unitary, suggests that we can ask which other matrices $A \in \mathbb{C}^{n\times n}$ can be unitarily diagonalized. To answer leads us to a beautiful class of matrices.

9.4.1 Normal matrices

Theorem 9.11. *An $n \times n$ matrix A over $\mathbb{C}$ is unitarily diagonalizable if and only if*

$$AA^H = A^H A. \tag{9.2}$$

Definition 9.3. A matrix $A \in \mathbb{C}^{n\times n}$ for which (9.2) holds is said to be *normal.*

The only if part of the above theorem is straightforward, so we'll omit the proof. The if statement will follow from Schur's Theorem, proved below.

Clearly Hermitian matrices are normal. We also obtain more classes of normal matrices by putting various conditions on D. One of the most interesting is given in the following

Example 9.5. Suppose the diagonal of D is pure imaginary. Then $N = UDU^H$ satisfies $N^H = UD^HU^H = -UDU^H = -N$. A matrix S such that $S^H = -S$ is called *skew Hermitian.* Skew Hermitian matrices are clearly normal, and writing $N = UDU^H$, the condition $N^H = -N$ obviously implies $D^H = -D$, i.e. the diagonal of D to be pure imaginary. Therefore, a matrix N is skew Hermitian if and only if iN is Hermitian.

Example 9.6. A real skew Hermitian matrix is called *skew symmetric.* In other words, a real matrix S is skew symmetric if $S^T = -S$. For example, let

$$S = \begin{pmatrix} 0 & 1 & 2 \\ -1 & 0 & 2 \\ -2 & -2 & 0 \end{pmatrix}.$$

The determinant of a skew symmetric matrix of odd order is 0 (see Exercise 9.35 below). The trace is obviously also 0, since all diagonal entries of a skew symmetric matrix are 0. Since S is 3×3, t its characteristic polynomial is determined by the sum $\sigma_2(S)$ of the principal 2×2 minors of S. Here, $\sigma_2(S) = 9$, so the characteristic polynomial of S up to sign is $\lambda^3 - 9\lambda$. Thus the eigenvalues of S are $0, \pm 3i$.

Since the characteristic polynomial of a skew symmetric matrix S is real, the nonzero eigenvalues of S are pure imaginary and they occur in conjugate pairs. Hence the only possible real eigenvalue is 0. Recall that a polynomial $p(x)$ is called even if $p(-x) = p(x)$ and odd if $p(-x) = -p(x)$. Only even powers of x occur in an even polynomial, and only odd powers occur in an odd one.

Proposition 9.12. *Let A be $n \times n$ and skew symmetric. Then the characteristic polynomial is of A is even or odd according to whether n is even or odd.*

Proof. Since the characteristic polynomial is real, if n is even, the eigenvalues occur in pairs $\mu \neq \overline{\mu}$. Thus the characteristic polynomial $p_A(\lambda)$ factors into products of the form $\lambda^2 - |\mu|^2$, $p_A(\lambda)$ involves only even powers. If n is odd, then the characteristic polynomial has a real root μ, which has to be 0 since 0 is the only pure imaginary real number. Hence $p_A(\lambda) = \lambda q_A(\lambda)$, where q_A is even, which proves the result. □

Example 9.7. Let $A = UDU^H$, where every diagonal entry of D is a unit complex number. Then D is unitary, hence so is A. Conversely, every unitary matrix is normal and the eigenvalues of a unitary matrix have modulus one (see Exercise 9.37), so every unitary matrix has this form. For example, the skew symmetric matrix

$$U = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

is orthogonal. U has eigenvalues $\pm i$, and we can easily compute that $E_i = \mathbb{C}(1, -i)^T$ and $E_{-i} = \mathbb{C}(1, i)^T$. Thus

$$U = U_1 D U_1^H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix} \begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -i \\ 1 & i \end{pmatrix}.$$

As a complex linear transformation of $\mathbb{C}^2$, the way U acts can be interpreted geometrically as follows. U rotates vectors on the principal axis $\mathbb{C}(1, i)^T$ spanned by $(1, i)^T$ through $\frac{\pi}{2}$ and rotates vectors on the orthogonal principal axis spanned by $(1, -i)^T$ by $-\frac{\pi}{2}$. Note, $U = R_{\pi/2}$ considered as a transformation on $\mathbb{C}^2$.

The abstract formulation of the notion of a normal matrix of course uses the notion of the adjoint of a linear transformation.

Definition 9.4. Let V be a Hermitian inner product space with inner product $(\ ,\ )$, and let $T : V \to V$ be $\mathbb{C}$-linear. Then T is said to be *normal* if and only if $TT^* = T^*T$, where T^* is the adjoint of T.

We leave it to the reader to formulate the appropriate statement of Theorem 9.11 for a normal operator T.

9.4.2 Schur's Theorem

The Theorem on normal matrices, Theorem 9.11, is a consequence of a very useful general result known as Schur's Theorem.

Theorem 9.13. *Let A be any $n \times n$ complex matrix. Then there exists an $n \times n$ unitary matrix U and an upper triangular T so that $A = UTU^{-1}$.*

Schur's Theorem can also be formulated abstractly as follows:

Theorem 9.14. *If V is a finite dimensional $\mathbb{C}$-vector space and $T : V \to V$ is linear over $\mathbb{C}$, then there exists a Hermitian orthonormal basis $\mathcal{U}$ of V for which the matrix $\mathcal{M}^{\mathcal{U}}_{\mathcal{U}}(T)$ of T is upper triangular.*

We will leave the proof Theorem 9.13 as an exercise. The idea is to apply the same method used in the first proof of the Principal Axis Theorem. The only essential facts are that A has an eigenpair $(\lambda_1, \mathbf{u}_1)$, where $\mathbf{u}_1$ can be included in a Hermitian orthonormal basis of $\mathbb{C}^n$, and the product of two unitary matrices is unitary. The reader is encouraged to write out a complete proof using induction on n.

9.4.3 Proof of Theorem 9.11

We will now finish this section by proving Theorem 9.11. Let A be normal. By Schur's Theorem, we may write $A = UTU^H$, where U is unitary and T is upper triangular. We claim that T is in fact diagonal. To see this, note that since $A^HA = AA^H$, it follows that $TT^H = T^HT$ (why?). Hence we need to show that an upper triangular normal matrix is diagonal. The key is to compare the diagonal entries of TT^H and TT^H. Let t_{ii} be the ith diagonal entry of T, and let $\mathbf{a}_i$ denote its ith row. Now the diagonal entries of TT^H are $|\mathbf{a}_1|^2, |\mathbf{a}_2|^2, \dots |\mathbf{a}_n|^2$. On the other hand, the diagonal entries of T^HT are $|t_{11}|^2, |t_{22}|^2, \dots, |t_{nn}|^2$. It follows that $|\mathbf{a}_i|^2 = |t_{ii}|^2$ for each i, and consequently T has to be diagonal. Therefore A is unitarily diagonalizable, and the proof is complete. $\square$

Exercises

Exercise 9.35. Unitarily diagonalize the skew symmetric matrix of Example 9.6.

Exercise 9.36. Let S be a skew Hermitian $n \times n$ matrix. Show the following:

(a) Every diagonal entry of S is pure imaginary.

(b) All eigenvalues of S are pure imaginary.

(c) If n is odd, then $|S|$ is pure imaginary, and if n is even, then $|S|$ is real.

(d) If S is skew symmetric, then $|S| = 0$ if n is odd, and $|S| \geq 0$ if n is even.

Exercise 9.37. Let U be any unitary matrix. Show that

(a) $|U|$ has modulus 1.

(b) Every eigenvalue of U also has modulus 1.

(c) Show that U is normal.

Exercise 9.38. Are all complex matrices normal? (Sorry)

Exercise 9.39. Formulate the appropriate statement of Theorem 9.11 for a normal operator T.

Exercise 9.40. The Principle of Mathematical Induction says that if a S(n) is statement about every positive integer n, then S(n) is true for all positive integers n provided:

(a) S(1) is true, and

(b) the truth of S($n-1$) implies the truth of S(n).

Give another proof of Schur's Theorem using induction. That is, if the theorem is true for B when B is $(n-1) \times (n-1)$, show that it immediately follows for A. (Don't forget the 1×1 case.)

9.5 Summary

The goal of this chapter was to prove the basic diagonalization theorems for real and complex matrices. This result, which is called the Principal Axis Theorem, is one of the most celebrated results in matrix theory. To be specific, we showed that every symmetric matrix A over $\mathbb{R}$ admits an orthonormal basis which consists of eigenvectors. More generally, every complex Hermitian matrix H admits a Hermitian orthonormal basisof eigenvectors. In particular, there is an orthogonal matrix Q and a unitary matrix U such that $A = ADQ^T$ and $H = UD\overline{U}^T$. In both cases, D is a real diagonal matrix, which implies that the eigenvalues of a Hermitian matrix are always real. In the real case, the general formulation of the Principal Axis Theorem says that every self adjoint linear operator on a finite dimensional inner product space V is semi-simple and, in fact, admits an orthonormal basisof eigenvectors.

The condition $H = UD\overline{U}^T$ with U unitary characterizes Hermitian matrices. More generally, we showed that the condition $N = UDU\overline{U}^T$ characterizes the class of complex matrices which are said to be normal. (A matrix $N \in \mathbb{C}^{n\times n}$ is normal if $N\overline{N}^T\overline{N}^T N$.) Among the normal matrices are skew symmetric real matrices and skew Hermitian complex matrices.