Chapter 10

Spectral Theorems in Euclidean and Hermitian Spaces

10.1 Normal Linear Maps

Let E be a real Euclidean space (or a complex Hermitian space) with inner product $u, v \mapsto \langle u, v \rangle$.

In the real Euclidean case, recall that $\langle -, - \rangle$ is bilinear, symmetric and positive definite (i.e., $\langle u, u \rangle > 0$ for all $u \neq 0$).

In the complex Hermitian case, recall that $\langle -, - \rangle$ is sesquilinear, which means that it linear in the first argument, semilinear in the second argument (i.e., $\langle u, \mu v \rangle = \overline{\mu} \langle u, v \rangle$), $\langle v, u \rangle = \overline{\langle u, v \rangle}$, and positive definite (as above).

In both cases we let $\|u\| = \sqrt{\langle u, u\rangle}$ and the map $u \mapsto \|u\|$ is a *norm*.

Recall that every linear map, $f\colon E \to E$, has an *adjoint* f^* which is a linear map, $f^*\colon E \to E$, such that

$$\langle f(u), v\rangle = \langle u, f^*(v)\rangle,$$

for all $u, v \in E$.

Since $\langle -, -\rangle$ is symmetric, it is obvious that $f^{**} = f$.

Definition 10.1. Given a Euclidean (or Hermitian) space, E, a linear map $f\colon E \to E$ is *normal* iff

$$f \circ f^* = f^* \circ f.$$

A linear map $f\colon E \to E$ is *self-adjoint* if $f = f^*$, *skew-self-adjoint* if $f = -f^*$, and *orthogonal* if $f \circ f^* = f^* \circ f = \mathrm{id}$.

Our first goal is to show that for every *normal* linear map $f: E \to E$ (where E is a Euclidean space), there is an *orthonormal basis* (w.r.t. $\langle -, - \rangle$) such that the matrix of f over this basis has an especially nice form:

It is a *block diagonal matrix* in which the blocks are either one-dimensional matrices (i.e., single entries) or two-dimensional matrices of the form

$$\begin{pmatrix} \lambda & \mu \\ -\mu & \lambda \end{pmatrix}$$

This normal form can be further refined if f is self-adjoint, skew-self-adjoint, or orthogonal.

As a first step, we show that f and f^* have the same kernel when f is normal.

Proposition 10.1. *Given a Euclidean space E, if $f: E \to E$ is a normal linear map, then* $\mathrm{Ker}\, f = \mathrm{Ker}\, f^*$.

The next step is to show that for *every linear map* $f\colon E \to E$, there is some subspace W of dimension 1 or 2 such that $f(W) \subseteq W$.

When $\dim(W) = 1$, W is actually an eigenspace for some real eigenvalue of f.

Furthermore, when f is normal, there is a subspace W of dimension 1 or 2 such that $f(W) \subseteq W$ and $f^*(W) \subseteq W$.

The difficulty is that the eigenvalues of f are not necessarily real. One way to get around this problem is to *complexify* both the vector space E and the inner product $\langle -, - \rangle$.

First, we need to embed a real vector space E into a complex vector space $E_{\mathbb{C}}$.

Definition 10.2. Given a real vector space E, let $E_{\mathbb{C}}$ be the structure $E \times E$ under the addition operation

$$(u_1, u_2) + (v_1, v_2) = (u_1 + v_1, u_2 + v_2),$$

and multiplication by a complex scalar $z = x + iy$ defined such that

$$(x + iy) \cdot (u, v) = (xu - yv, yu + xv).$$

The space $E_{\mathbb{C}}$ is called the *complexification* of E.

It is easily shown that the structure $E_{\mathbb{C}}$ is a complex vector space.

It is also immediate that

$$(0, v) = i(v, 0),$$

and thus, identifying E with the subspace of $E_{\mathbb{C}}$ consisting of all vectors of the form $(u, 0)$, we can write

$$(u, v) = u + iv.$$

Given a vector $w = u + iv$, its *conjugate* $\overline{w}$ is the vector $\overline{w} = u - iv$.

Observe that if $(e_1, \ldots, e_n)$ is a basis of E (a real vector space), then $(e_1, \ldots, e_n)$ is also a basis of $E_{\mathbb{C}}$ (recall that e_i is an abreviation for $(e_i, 0)$).

Given a linear map $f\colon E \to E$, the map f can be extended to a linear map $f_{\mathbb{C}}\colon E_{\mathbb{C}} \to E_{\mathbb{C}}$ defined such that

$$f_{\mathbb{C}}(u + iv) = f(u) + if(v).$$

For any basis $(e_1, \ldots, e_n)$ of E, the matrix $M(f)$ representing f over $(e_1, \ldots, e_n)$ is identical to the matrix $M(f_{\mathbb{C}})$ representing $f_{\mathbb{C}}$ over $(e_1, \ldots, e_n)$, where we view $(e_1, \ldots, e_n)$ as a basis of $E_{\mathbb{C}}$.

As a consequence, $\det(zI - M(f)) = \det(zI - M(f_{\mathbb{C}}))$, which means that *$f$ and $f_{\mathbb{C}}$ have the same characteristic polynomial* (which has real coefficients).

We know that every polynomial of degree n with real (or complex) coefficients always has n complex roots (counted with their multiplicity), and the roots of $\det(zI - M(f_{\mathbb{C}}))$ that are real (if any) are the eigenvalues of f.

Next, we need to extend the inner product on E to an inner product on $E_{\mathbb{C}}$.

The inner product $\langle -, - \rangle$ on a Euclidean space E is extended to the Hermitian positive definite form $\langle -, - \rangle_{\mathbb{C}}$ on $E_{\mathbb{C}}$ as follows:

$$\begin{aligned}&\langle u_1 + iv_1, u_2 + iv_2 \rangle_{\mathbb{C}} \\ &\qquad\qquad = \langle u_1, u_2 \rangle + \langle v_1, v_2 \rangle + i(\langle u_2, v_1 \rangle - \langle u_1, v_2 \rangle).\end{aligned}$$

Then, given any linear map $f\colon E \to E$, it is easily verified that the map $f_{\mathbb{C}}^*$ defined such that

$$f_{\mathbb{C}}^*(u + iv) = f^*(u) + if^*(v)$$

for all $u, v \in E$, is the *adjoint* of $f_{\mathbb{C}}$ w.r.t. $\langle -, - \rangle_{\mathbb{C}}$.

Assuming again that E is a Hermitian space, observe that Proposition 10.1 also holds.

Proposition 10.2. *Given a Hermitian space E, for any normal linear map $f\colon E \to E$, a vector u is an eigenvector of f for the eigenvalue λ (in $\mathbb{C}$) iff u is an eigenvector of f^* for the eigenvalue $\overline{\lambda}$.*

The next proposition shows a very important property of normal linear maps: eigenvectors corresponding to distinct eigenvalues are orthogonal.

Proposition 10.3. *Given a Hermitian space E, for any normal linear map $f\colon E \to E$, if u and v are eigenvectors of f associated with the eigenvalues λ and μ (in $\mathbb{C}$) where $\lambda \neq \mu$, then $\langle u, v\rangle = 0$.*

We can also show easily that the eigenvalues of a self-adjoint linear map are real.

Proposition 10.4. *Given a Hermitian space E, the eigenvalues of any self-adjoint linear map $f\colon E \to E$ are real.*

There is also a version of Proposition 10.4 for a (real) Euclidean space E and a self-adjoint map $f\colon E \to E$.

Proposition 10.5. *Given a Euclidean space E, if $f\colon E \to E$ is any self-adjoint linear map, then every eigenvalue λ of $f_{\mathbb{C}}$ is real and is actually an eigenvalue of f (which means that there is some real eigenvector $u \in E$ such that $f(u) = \lambda u$). Therefore, all the eigenvalues of f are real.*

Given any subspace W of a Hermitian space E, recall that the *orthogonal* $W^{\perp}$ of W is the subspace defined such that

$$W^{\perp} = \{u \in E \mid \langle u, w\rangle = 0, \text{ for all } w \in W\}.$$

Recall that $E = W \oplus W^{\perp}$ (construct an orthonormal basis of E using the Gram–Schmidt orthonormalization procedure). The same result also holds for Euclidean spaces.

As a warm up for the proof of Theorem 10.9, let us prove that every self-adjoint map on a Euclidean space can be diagonalized with respect to an orthonormal basis of eigenvectors.

Theorem 10.6. *Given a Euclidean space E of dimension n, for every self-adjoint linear map $f\colon E \to E$, there is an orthonormal basis $(e_1, \ldots, e_n)$ of eigenvectors of f such that the matrix of f w.r.t. this basis is a diagonal matrix*

$$\begin{pmatrix} \lambda_1 & & \ldots & \\ & \lambda_2 & \ldots & \\ \vdots & \vdots & \ddots & \vdots \\ & & \ldots & \lambda_n \end{pmatrix},$$

with $\lambda_i \in \mathbb{R}$.

One of the key points in the proof of Theorem 10.6 is that we found a subspace W with the property that $f(W) \subseteq W$ implies that $f(W^{\perp}) \subseteq W^{\perp}$.

In general, this does not happen, but normal maps satisfy a stronger property which ensures that such a subspace exists.

The following proposition provides a condition that will allow us to show that a normal linear map can be diagonalized. It actually holds for any linear map.

Proposition 10.7. *Given a Hermitian space E, for any linear map $f\colon E \to E$ and any subspace W of E, if $f(W) \subseteq W$, then $f^*\big(W^{\perp}\big) \subseteq W^{\perp}$.*

Consequently, if $f(W) \subseteq W$ and $f^*(W) \subseteq W$, then $f\big(W^{\perp}\big) \subseteq W^{\perp}$ and $f^*\big(W^{\perp}\big) \subseteq W^{\perp}$.

The above Proposition *also holds for Euclidean spaces*. Although we are ready to prove that for every normal linear map f (over a Hermitian space) there is an orthonormal basis of eigenvectors, we now return to real Euclidean spaces.

If $f\colon E \to E$ is a linear map and $w = u + iv$ is an eigenvector of $f_{\mathbb{C}}\colon E_{\mathbb{C}} \to E_{\mathbb{C}}$ for the eigenvalue $z = \lambda + i\mu$, where $u, v \in E$ and $\lambda, \mu \in \mathbb{R}$, since

$$f_{\mathbb{C}}(u + iv) = f(u) + if(v)$$

and

$$\begin{aligned} f_{\mathbb{C}}(u + iv) &= (\lambda + i\mu)(u + iv) \\ &= \lambda u - \mu v + i(\mu u + \lambda v), \end{aligned}$$

we have

$$f(u) = \lambda u - \mu v \quad \text{and} \quad f(v) = \mu u + \lambda v,$$

from which we immediately obtain

$$f_{\mathbb{C}}(u - iv) = (\lambda - i\mu)(u - iv),$$

which shows that $\overline{w} = u - iv$ is an eigenvector of $f_{\mathbb{C}}$ for $\overline{z} = \lambda - i\mu$. Using this fact, we can prove the following proposition:

Proposition 10.8. *Given a Euclidean space E, for any normal linear map $f\colon E \to E$, if $w = u + iv$ is an eigenvector of $f_{\mathbb{C}}$ associated with the eigenvalue $z = \lambda + i\mu$ (where $u, v \in E$ and $\lambda, \mu \in \mathbb{R}$), if $\mu \neq 0$ (i.e., z is not real) then $\langle u, v\rangle = 0$ and $\langle u, u\rangle = \langle v, v\rangle$, which implies that u and v are linearly independent, and if W is the subspace spanned by u and v, then $f(W) = W$ and $f^*(W) = W$. Furthermore, with respect to the (orthogonal) basis (u, v), the restriction of f to W has the matrix*

$$\begin{pmatrix} \lambda & \mu \\ -\mu & \lambda \end{pmatrix}.$$

If $\mu = 0$, then λ is a real eigenvalue of f and either u or v is an eigenvector of f for λ. If W is the subspace spanned by u if $u \neq 0$, or spanned by $v \neq 0$ if $u = 0$, then $f(W) \subseteq W$ and $f^*(W) \subseteq W$.

Theorem 10.9. *(Main Spectral Theorem) Given a Euclidean space E of dimension n, for every normal linear map $f\colon E \to E$, there is an orthonormal basis $(e_1, \ldots, e_n)$ such that the matrix of f w.r.t. this basis is a block diagonal matrix of the form*

$$\begin{pmatrix} A_1 & & \ldots & \\ & A_2 & \ldots & \\ \vdots & \vdots & \ddots & \vdots \\ & & \ldots & A_p \end{pmatrix}$$

such that each block A_j is either a one-dimensional matrix (i.e., a real scalar) or a two-dimensional matrix of the form

$$A_j = \begin{pmatrix} \lambda_j & -\mu_j \\ \mu_j & \lambda_j \end{pmatrix}$$

where $\lambda_j, \mu_j \in \mathbb{R}$, with $\mu_j > 0$.

After this relatively hard work, we can easily obtain some nice normal forms for the matrices of self-adjoint, skew-self-adjoint, and orthogonal, linear maps.

However, for the sake of completeness, we state the following theorem.

Theorem 10.10. *Given a Hermitian space E of dimension n, for every normal linear map $f\colon E \to E$, there is an orthonormal basis $(e_1, \ldots, e_n)$ of eigenvectors of f such that the matrix of f w.r.t. this basis is a diagonal matrix*

$$\begin{pmatrix} \lambda_1 & & \ldots & \\ & \lambda_2 & \ldots & \\ \vdots & \vdots & \ddots & \vdots \\ & & \ldots & \lambda_n \end{pmatrix}$$

where $\lambda_j \in \mathbb{C}$.

Remark: There is a *converse* to Theorem 10.10, namely, if there is an orthonormal basis $(e_1, \ldots, e_n)$ of eigenvectors of f, then f is normal.

10.2 Self-Adjoint, Skew-Self-Adjoint, and Orthogonal Linear Maps

Theorem 10.11. *Given a Euclidean space E of dimension n, for every self-adjoint linear map $f\colon E \to E$, there is an orthonormal basis $(e_1, \ldots, e_n)$ of eigenvectors of f such that the matrix of f w.r.t. this basis is a diagonal matrix*

$$\begin{pmatrix} \lambda_1 & & \ldots & \\ & \lambda_2 & \ldots & \\ \vdots & \vdots & \ddots & \vdots \\ & & \ldots & \lambda_n \end{pmatrix}$$

where $\lambda_i \in \mathbb{R}$.

Theorem 10.11 implies that if $\lambda_1, \ldots, \lambda_p$ are the distinct real eigenvalues of f and E_i is the eigenspace associated with λ_i, then

$$E = E_1 \oplus \cdots \oplus E_p,$$

where E_i and E_j are othogonal for all $i \neq j$.

Theorem 10.12. *Given a Euclidean space E of dimension n, for every skew-self-adjoint linear map $f : E \to E$, there is an orthonormal basis $(e_1, \ldots, e_n)$ such that the matrix of f w.r.t. this basis is a block diagonal matrix of the form*

$$\begin{pmatrix} A_1 & & \ldots & \\ & A_2 & \ldots & \\ \vdots & \vdots & \ddots & \vdots \\ & & \ldots & A_p \end{pmatrix}$$

such that each block A_j is either 0 or a two-dimensional matrix of the form

$$A_j = \begin{pmatrix} 0 & -\mu_j \\ \mu_j & 0 \end{pmatrix}$$

where $\mu_j \in \mathbb{R}$, with $\mu_j > 0$. In particular, the eigenvalues of $f_{\mathbb{C}}$ are pure imaginary of the form $\pm i\mu_j$, or 0.

Theorem 10.13. *Given a Euclidean space E of dimension n, for every orthogonal linear map $f\colon E \to E$, there is an orthonormal basis $(e_1, \ldots, e_n)$ such that the matrix of f w.r.t. this basis is a block diagonal matrix of the form*

$$\begin{pmatrix} A_1 & & \ldots & \\ & A_2 & \ldots & \\ \vdots & \vdots & \ddots & \vdots \\ & & \ldots & A_p \end{pmatrix}$$

such that each block A_j is either 1, -1, or a two-dimensional matrix of the form

$$A_j = \begin{pmatrix} \cos\theta_j & -\sin\theta_j \\ \sin\theta_j & \cos\theta_j \end{pmatrix}$$

where $0 < \theta_j < \pi$.

In particular, the eigenvalues of $f_{\mathbb{C}}$ are of the form $\cos\theta_j \pm i\sin\theta_j$, or 1, or -1.

It is obvious that we can reorder the orthonormal basis of eigenvectors given by Theorem 10.13, so that the matrix of f w.r.t. this basis is a block diagonal matrix of the form

$$\begin{pmatrix} A_1 & \dots & & & \\ \vdots & \ddots & \vdots & & \vdots \\ & \dots & A_r & & \\ & & & -I_q & \\ \dots & & & & I_p \end{pmatrix}$$

where each block A_j is a two-dimensional rotation matrix $A_j \neq \pm I_2$ of the form

$$A_j = \begin{pmatrix} \cos\theta_j & -\sin\theta_j \\ \sin\theta_j & \cos\theta_j \end{pmatrix}$$

with $0 < \theta_j < \pi$.

The linear map f has an eigenspace $E(1, f) = \mathrm{Ker}\,(f - \mathrm{id})$ of dimension p for the eigenvalue 1, and an eigenspace $E(-1, f) = \mathrm{Ker}\,(f + \mathrm{id})$ of dimension q for the eigenvalue -1.

If $\det(f) = +1$ (f is a rotation), the dimension q of $E(-1, f)$ must be even, and the entries in $-I_q$ can be paired to form two-dimensional blocks, if we wish.

Remark: Theorem 10.13 can be used to prove a sharper version of the Cartan-Dieudonné Theorem.

Theorem 10.14. *Let E be a Euclidean space of dimension $n \geq 2$. For every isometry $f \in \mathbf{O}(E)$, if $p = \dim(E(1, f)) = \dim(\mathrm{Ker}\,(f - \mathrm{id}))$, then f is the composition of $n - p$ reflections and $n - p$ is minimal.*

The theorems of this section and of the previous section can be immediately applied to matrices.

10.3 Normal, Symmetric, Skew-Symmetric, Orthogonal, Hermitian, Skew-Hermitian, and Unitary Matrices

First, we consider real matrices.

Definition 10.3. Given a real $m \times n$ matrix A, the *transpose $A^\top$ of A* is the $n \times m$ matrix $A^\top = (a^\top_{i\,j})$ defined such that

$$a^\top_{i\,j} = a_{j\,i}$$

for all i, j, $1 \leq i \leq m$, $1 \leq j \leq n$. A real $n \times n$ matrix A is

1. *normal* iff
$$A\,A^\top = A^\top A,$$

2. *symmetric* iff
$$A^\top = A,$$

3. *skew-symmetric* iff
$$A^\top = -A,$$

4. *orthogonal* iff
$$A\,A^\top = A^\top A = I_n.$$

Theorem 10.15. *For every normal matrix A, there is an orthogonal matrix P and a block diagonal matrix D such that $A = PD\,P^\top$, where D is of the form*

$$D = \begin{pmatrix} D_1 & & \ldots & \\ & D_2 & \ldots & \\ \vdots & \vdots & \ddots & \vdots \\ & & \ldots & D_p \end{pmatrix}$$

such that each block D_j is either a one-dimensional matrix (i.e., a real scalar) or a two-dimensional matrix of the form

$$D_j = \begin{pmatrix} \lambda_j & -\mu_j \\ \mu_j & \lambda_j \end{pmatrix}$$

where $\lambda_j, \mu_j \in \mathbb{R}$, with $\mu_j > 0$.

Theorem 10.16. *For every symmetric matrix A, there is an orthogonal matrix P and a diagonal matrix D such that $A = PD\,P^\top$, where D is of the form*

$$D = \begin{pmatrix} \lambda_1 & & \ldots & \\ & \lambda_2 & \ldots & \\ \vdots & \vdots & \ddots & \vdots \\ & & \ldots & \lambda_n \end{pmatrix}$$

where $\lambda_i \in \mathbb{R}$.

Theorem 10.17. *For every skew-symmetric matrix A, there is an orthogonal matrix P and a block diagonal matrix D such that $A = PD\,P^{\top}$, where D is of the form*

$$D = \begin{pmatrix} D_1 & & \dots & \\ & D_2 & \dots & \\ \vdots & \vdots & \ddots & \vdots \\ & & \dots & D_p \end{pmatrix}$$

such that each block D_j is either 0 or a two-dimensional matrix of the form

$$D_j = \begin{pmatrix} 0 & -\mu_j \\ \mu_j & 0 \end{pmatrix}$$

where $\mu_j \in \mathbb{R}$, with $\mu_j > 0$. In particular, the eigenvalues of A are pure imaginary of the form $\pm i\mu_j$, or 0.

Theorem 10.18. *For every orthogonal matrix A, there is an orthogonal matrix P and a block diagonal matrix D such that $A = PD\,P^\top$, where D is of the form*

$$D = \begin{pmatrix} D_1 & & \dots & \\ & D_2 & \dots & \\ \vdots & \vdots & \ddots & \vdots \\ & & \dots & D_p \end{pmatrix}$$

such that each block D_j is either 1, -1, or a two-dimensional matrix of the form

$$D_j = \begin{pmatrix} \cos\theta_j & -\sin\theta_j \\ \sin\theta_j & \cos\theta_j \end{pmatrix}$$

where $0 < \theta_j < \pi$.

In particular, the eigenvalues of A are of the form $\cos\theta_j \pm i\sin\theta_j$, or 1, or -1.

We now consider complex matrices.

Definition 10.4. Given a complex $m \times n$ matrix A, the *transpose $A^\top$ of A* is the $n \times m$ matrix $A^\top = (a^\top_{i\,j})$ defined such that

$$a^\top_{i\,j} = a_{j\,i}$$

for all i, j, $1 \leq i \leq m$, $1 \leq j \leq n$. The *conjugate $\overline{A}$ of A* is the $m \times n$ matrix $\overline{A} = (b_{i\,j})$ defined such that

$$b_{i\,j} = \overline{a}_{i\,j}$$

for all i, j, $1 \leq i \leq m$, $1 \leq j \leq n$. Given an $n \times n$ complex matrix A, the *adjoint A^* of A* is the matrix defined such that

$$A^* = \overline{(A^\top)} = (\overline{A})^\top.$$

A complex $n \times n$ matrix A is

1. *normal* iff
$$AA^* = A^*A,$$
2. *Hermitian* iff
$$A^* = A,$$
3. *skew-Hermitian* iff
$$A^* = -A,$$
4. *unitary* iff
$$AA^* = A^*A = I_n.$$

Theorem 10.10 can be restated in terms of matrices as follows. We can also say a little more about eigenvalues (easy exercise left to the reader).

Theorem 10.19. *For every complex normal matrix A, there is a unitary matrix U and a diagonal matrix D such that $A = UDU^*$. Furthermore, if A is Hermitian, D is a real matrix, if A is skew-Hermitian, then the entries in D are pure imaginary or null, and if A is unitary, then the entries in D have absolute value 1.*