Chapter 10

Spectral Theorems in Euclidean and Hermitian Spaces

10.1 Introduction: What's with Lie Groups and Lie Algebras?

The purpose of the next three chapters is to give a concrete introduction to Lie groups and Lie algebras. Our ulterior motive is to present some beautiful mathematical concepts that can also be used as tools for solving practical problems arising in computer science, more specifically in robotics, motion planning, computer vision, and computer graphics.

10.2 Normal Linear Maps

Recall that every linear map $f: E \to E$ has an *adjoint* f^* which is a linear map $f^*: E \to E$ 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 10.2.1 Given a Euclidean space E, a linear map $f: E \to E$ is *normal* iff

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

A linear map $f: E \to E$ 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 $f: E \to E$, 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.

Lemma 10.2.2 *Given a Euclidean space E, if $f: E \to E$ is a normal linear map, then*
$\operatorname{Ker} f = \operatorname{Ker} f^*$.

The next step is to show that for every linear map $f: E \to E$, 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 10.2.3 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

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

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 $\overline{w} = u - iv$.

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

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

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:

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

Then, given any linear map $f: E \to E$, 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 Lemma 10.2.2 also holds.

Lemma 10.2.4 *Given a Hermitian space E, for any normal linear map $f: E \to E$, 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 lemma shows a very important property of normal linear maps: eigenvectors corresponding to distinct eigenvalues are orthogonal.

Lemma 10.2.5 *Given a Hermitian space E, for any normal linear map $f: E \to E$, 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.

Lemma 10.2.6 *Given a Hermitian space E, the eigenvalues of any self-adjoint linear map $f: E \to E$ 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 $E = W \oplus W^\perp$ (see Lemma 9.2.5, or construct an orthonormal basis of E using the Gram–Schmidt orthonormalization procedure). The same result also holds for Euclidean spaces, see Lemma 4.2.8.

The following lemma provides the key to the induction that will allow us to show that a normal linear map can diagonalized. It actually holds for any linear map.

Lemma 10.2.7 *Given a Hermitian space E, for any linear map $f: E \to E$, if W is any subspace of E such that $f(W) \subseteq W$ and $f^*(W) \subseteq W$, then $f(W^\perp) \subseteq W^\perp$ and $f^*(W^\perp) \subseteq W^\perp$.*

The above Lemma also holds for Euclidean spaces. Although we are ready to prove that for every normal linar 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 lemma.

Lemma 10.2.8 *Given a Euclidean space E, for any normal linear map $f: E \to E$, 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 $f^*(W) \subseteq W$.

If f is a normal linear map, the proof of Lemma 10.2.8 shows that λ, μ, u, and v, satisfy the equations

$$\begin{aligned} f(u) &= \lambda u - \mu v, \\ f(v) &= \mu u + \lambda v, \\ f^*(u) &= \lambda u + \mu v, \\ f^*(v) &= -\mu u + \lambda v, \end{aligned}$$

from which we get

$$\begin{aligned} \frac{1}{2}\left(f + f^*\right)(u) &= \lambda u, \\ \frac{1}{2}\left(f + f^*\right)(v) &= \lambda v, \\ \frac{1}{2}\left(f^* - f\right)(u) &= \mu v, \\ \frac{1}{2}\left(f - f^*\right)(v) &= \mu u. \end{aligned}$$

Using the above equations, we also get

$$\left(\frac{1}{2}\left(f - f^*\right)\right)^2(u) = -\mu^2 u,$$

$$\left(\frac{1}{2}\left(f - f^*\right)\right)^2 (v) = -\mu^2 v.$$

Thus, we observe that λ is an eigenvalue of $1/2\,(f + f^*)$, that $-\mu^2$ is an eigenvalue of $(1/2\,(f - f^*))^2$, and that u and v are both eigenvectors of $1/2\,(f + f^*)$ for λ and of $(1/2\,(f - f^*))^2$ for $-\mu^2$.

It is easily verified that $1/2\,(f + f^*)$ and $(1/2\,(f - f^*))^2$ are self-adjoint, and we proved earlier that self-adjoint maps have real eigenvalues (Lemma 10.2.6).

Furthermore, there are good numerical methods for finding the eigenvalues of symmetric matrices.

Thus, it should be possible to compute λ, μ, u, and v, from the eigenvalues and the eigenvectors of the self-adjoint maps $1/2\,(f + f^*)$ and $(1/2\,(f - f^*))^2$.

Note that if we have λ and u, then we get μ from

$$\left(\frac{1}{2}\,(f - f^*)\right)^2 (u) = -\mu^2 u,$$

and if $\mu \neq 0$, we get v from

$$\frac{1}{2}\,(f^* - f)\,(u) = \mu v.$$

I am not aware of a good method (i.e., numerically stable) to compute the block diagonal form of a normal matrix, but this seems an interesting problem.

We can finally prove our first main theorem.

Theorem 10.2.9 *Given a Euclidean space E of dimension n, for every normal 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_i is either a one-dimensional matrix (i.e., a real scalar) or a two-dimensional matrix of the form

$$A_i = \begin{pmatrix} \lambda_i & -\mu_i \\ \mu_i & \lambda_i \end{pmatrix}$$

where $\lambda_i, \mu_i \in \mathbb{R}$, with $\mu_i > 0$.

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 (and since we have all the tools to so do), we go back to the case of a Hermitian space and show that normal linear maps can be diagonalized with respect to an orthonormal basis.

Theorem 10.2.10 *Given a Hermitian space E of dimension n, for every normal linear map $f: E \to E$, 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*

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

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

Remark: There is a converse to Theorem 10.2.10, namely, if there is an orthonormal basis $(e_1, \ldots, e_n)$ of eigenvectors of f, then f is normal. We leave the easy proof as an exercise.

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

We begin with self-adjoint maps.

Theorem 10.3.1 *Given a Euclidean space E of dimension n, for every self-adjoint linear map $f: E \to E$, 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*

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

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

Theorem 10.3.1 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 $i \neq j$.

Next, we consider skew self-adjoint maps.

Theorem 10.3.2 *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_i is either 0 or a two-dimensional matrix of the form

$$A_i = \begin{pmatrix} 0 & -\mu_i \\ \mu_i & 0 \end{pmatrix}$$

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

Remark: One will note that if f is skew self-adjoint, then $if_{\mathbb{C}}$ is self-adjoint w.r.t. $\langle -, - \rangle_{\mathbb{C}}$.

By Lemma 10.2.6, the map $if_{\mathbb{C}}$ has real eigenvalues, which implies that the eigenvalues of $f_{\mathbb{C}}$ are pure imaginary or 0.

Finally, we consider orthogonal linear maps.

Theorem 10.3.3 *Given a Euclidean space E of dimension n, for every orthogonal 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_i is either 1, -1, or a two-dimensional matrix of the form

$$A_i = \begin{pmatrix} \cos\theta_i & -\sin\theta_i \\ \sin\theta_i & \cos\theta_i \end{pmatrix}$$

where $0 < \theta_i < \pi$.

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

It is obvious that we can reorder the orthonormal basis of eigenvectors given by Theorem 10.3.3, so that the matrix of f w.r.t. this basis is a block diagonal matrix of the form

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

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

$$A_i = \begin{pmatrix} \cos\theta_i & -\sin\theta_i \\ \sin\theta_i & \cos\theta_i \end{pmatrix}$$

with $0 < \theta_i < \pi$.

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.

Theorem 10.3.3 can be used to prove a sharper version of the Cartan-Dieudonné Theorem, as claimed in remark (3) after Theorem 5.2.1.

Theorem 10.3.4 *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.

10.4 Normal, Symmetric, Skew Symmetric, Orthogonal, Hermitian, Skew Hermitian, and Unitary Matrices

First, we consider real matrices.

Definition 10.4.1 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
$$A\,A^\top = A^\top A = I_n.$$

Theorems 10.2.9 and 10.3.1–10.3.3 can be restated as follows.

Theorem 10.4.2 *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*

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

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

$$D_i = \begin{pmatrix} \lambda_i & -\mu_i \\ \mu_i & \lambda_i \end{pmatrix}$$

where $\lambda_i, \mu_i \in \mathbb{R}$, with $\mu_i > 0$.

Theorem 10.4.3 *For every symmetric matrix A, there is an orthogonal matrix P and a diagonal matrix D such that $A = PDP^\top$, where D is of the form*

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

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

Theorem 10.4.4 *For every skew symmetric matrix A, there is an orthogonal matrix P and a block diagonal matrix D such that $A = PDP^\top$, where D is of the form*

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

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

$$D_i = \begin{pmatrix} 0 & -\mu_i \\ \mu_i & 0 \end{pmatrix}$$

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

Theorem 10.4.5 *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*

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

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

$$D_i = \begin{pmatrix} \cos\theta_i & -\sin\theta_i \\ \sin\theta_i & \cos\theta_i \end{pmatrix}$$

where $0 < \theta_i < \pi$.

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

We now consider complex matrices.

Definition 10.4.6 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 10.2.10 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 10.4.7 *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.