

Chapter 16

Adjoint Linear Transformations

The presence of an inner product has significant consequences. In particular, it enables us to map a vector space into its dual space, and gives rise to the notion of the *adjoint* of a linear transformation, which is a cornerstone of several applications of linear algebra, such as to quantum mechanics.

To discuss the adjoint we first investigate the relation between a vector space and its dual in the presence of an inner product.

Recall that if V is a vector space over the field \mathbb{F} , then its dual space, V^* , is $\text{Hom}_{\mathbb{F}}(V, \mathbb{F})$, the \mathbb{F} vector space of all \mathbb{F} -linear transformations $V \rightarrow \mathbb{F}$. Such a linear transformation is often called a *1-form* or a *linear form*.

Lemma 16.1. *Let $(V, \langle \cdot, \cdot \rangle)$ be an inner product space over \mathbb{F} . For each $\mathbf{v} \in V$*

$$\langle \cdot, \mathbf{v} \rangle: V \longrightarrow \mathbb{F}, \quad \mathbf{x} \longmapsto \langle \mathbf{x}, \mathbf{v} \rangle$$

is a linear transformation.

Proof. Take $\mathbf{x}, \mathbf{y} \in V$ and $\lambda, \mu \in \mathbb{F}$. By the definition of inner product,

$$\langle \lambda \mathbf{x} + \mu \mathbf{y}, \mathbf{v} \rangle = \lambda \langle \mathbf{x}, \mathbf{v} \rangle + \mu \langle \mathbf{y}, \mathbf{v} \rangle$$

□

We use Lemma 16.1 to embed V in V^* .

Lemma 16.2. *Let $(V, \langle \cdot, \cdot \rangle)$ be an inner product space over \mathbb{F} . Then*

$$R: V \longrightarrow V^*, \quad \mathbf{v} \longmapsto \langle \cdot, \mathbf{v} \rangle$$

is injective.

Proof. Take $\mathbf{u}, \mathbf{v} \in V$. Then $R(\mathbf{u}) = R(\mathbf{v})$ if and only if for all $\mathbf{x} \in V$, $(R(\mathbf{u}))(\mathbf{x}) = (R(\mathbf{v}))(\mathbf{x})$, or, equivalently, $\langle \mathbf{x}, \mathbf{u} \rangle = \langle \mathbf{x}, \mathbf{v} \rangle$.

By the definition of inner product, this is equivalent to $\langle \mathbf{x}, \mathbf{u} - \mathbf{v} \rangle = 0$ for all $\mathbf{x} \in V$.

In particular $\langle \mathbf{u} - \mathbf{v}, \mathbf{u} - \mathbf{v} \rangle = 0$, whence $\mathbf{u} = \mathbf{v}$ since $\langle \cdot, \cdot \rangle$ is an inner product.

□

16.1 Riesz Representation Theorem

For general vector spaces, there are no “natural” linear forms, that is linear transformations from the given space to the field of scalars.

The preceding discussion shows that for inner product spaces, there is a rich supply of them, at least one for each vector in V . The Riesz Representation Theorem states that under some additional conditions, these are the only linear forms. We prove the Riesz Representation Theorem for finitely generated inner product spaces.

Theorem 16.3 (Riesz Representation Theorem). *Let $(V, \langle \cdot, \cdot \rangle)$ be a finitely generated inner product space over \mathbb{F} . Given a linear transformation $\varphi: V \rightarrow \mathbb{F}$, there is a unique vector, $\mathbf{v}_\varphi \in V$, such that for all $\mathbf{x} \in V$*

$$\varphi(\mathbf{x}) = \langle \mathbf{x}, \mathbf{v}_\varphi \rangle$$

Proof. We first establish the uniqueness of such a vector, \mathbf{v}_φ .

Suppose that $\langle \mathbf{x}, \mathbf{u} \rangle = \langle \mathbf{x}, \mathbf{v} \rangle$ for all $\mathbf{x} \in V$. Since this is equivalent to $\langle \mathbf{x}, \mathbf{v} - \mathbf{u} \rangle = 0$ for all $\mathbf{x} \in V$, it follows that $\langle \mathbf{v} - \mathbf{u}, \mathbf{v} - \mathbf{u} \rangle = 0$, whence $\mathbf{v} = \mathbf{u}$.

Next we show the existence of \mathbf{v}_φ .

Since $\text{im}(\varphi)$ is a vector subspace of \mathbb{F} , either $\text{im}(\varphi) = \{0\}$ or $\text{im}(\varphi) = \mathbb{F}$.

In the former case, $\varphi(\mathbf{x}) = 0$ for all $\mathbf{x} \in V$, and so we may choose $\mathbf{v}_\varphi := \mathbf{0}_V$.

In the latter case, choose an orthonormal basis for $\ker(\varphi)$, say $\{\mathbf{e}_2, \dots, \mathbf{e}_n\}$, and extend to an orthonormal basis, $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ of V .

Observe that since $\mathbf{e}_1 \notin \ker(\varphi)$, $\varphi(\mathbf{e}_1) \neq 0$.

Take $\mathbf{x} \in V$. By Theorem 14.7,

$$\mathbf{x} = \sum_{j=1}^n \langle \mathbf{x}, \mathbf{e}_j \rangle \mathbf{e}_j,$$

so that

$$\begin{aligned} \varphi(\mathbf{x}) &= \varphi\left(\sum_{j=1}^n \langle \mathbf{x}, \mathbf{e}_j \rangle \mathbf{e}_j\right) \\ &= \sum_{j=1}^n \langle \mathbf{x}, \mathbf{e}_j \rangle \varphi(\mathbf{e}_j) && \text{as } \varphi \text{ is linear} \\ &= \langle \mathbf{x}, \mathbf{e}_1 \rangle \varphi(\mathbf{e}_1) && \text{as } \varphi(\mathbf{e}_j) = 0 \text{ for } j > 1 \\ &= \langle \mathbf{x}, \overline{\varphi(\mathbf{e}_1)} \mathbf{e}_1 \rangle && \text{as } \varphi(\mathbf{e}_1) \in \mathbb{F}. \end{aligned}$$

So $\mathbf{v}_\varphi := \overline{\varphi(\mathbf{e}_1)} \mathbf{e}_1$ clearly has the required property. □

Corollary 16.4. *Let $(V, \langle \cdot, \cdot \rangle)$ be a finite dimensional inner product space over \mathbb{F} . Then the function*

$$R: V \longrightarrow \text{Hom}(V, \mathbb{F}), \quad \mathbf{v} \longmapsto \langle \cdot, \mathbf{v} \rangle,$$

is an additive bijection, which is an isomorphism of vector spaces whenever $\mathbb{F} \subseteq \mathbb{R}$.

Proof. That R is well defined follows from the fact that for all $\mathbf{x}, \mathbf{y}, \mathbf{v} \in V$, $\alpha, \beta \in \mathbb{F}$, $\langle\langle \alpha\mathbf{x} + \beta\mathbf{y}, \mathbf{v} \rangle\rangle = \alpha\langle\langle \mathbf{x}, \mathbf{y} \rangle\rangle + \beta\langle\langle \mathbf{y}, \mathbf{v} \rangle\rangle$.

That R is bijective is a restatement of the Riesz Representation Theorem.

Finally, take $\alpha, \beta \in \mathbb{F}$ and $\mathbf{u}, \mathbf{v} \in V$. Then, for any $\mathbf{x} \in V$,

$$\begin{aligned} (R(\alpha\mathbf{u} + \beta\mathbf{v}))(\mathbf{x}) &= \langle\langle \mathbf{x}, \alpha\mathbf{u} + \beta\mathbf{v} \rangle\rangle \\ &= \bar{\alpha}\langle\langle \mathbf{x}, \mathbf{u} \rangle\rangle + \bar{\beta}\langle\langle \mathbf{x}, \mathbf{v} \rangle\rangle. \\ &= \bar{\alpha}(R(\mathbf{u}))(\mathbf{x}) + \bar{\beta}(R(\mathbf{v}))(\mathbf{x}) \\ &= (\bar{\alpha}R(\mathbf{u}) + \bar{\beta}R(\mathbf{v}))(\mathbf{x}) \end{aligned}$$

Thus $R(\alpha\mathbf{u} + \beta\mathbf{v}) = \bar{\alpha}R(\mathbf{u}) + \bar{\beta}R(\mathbf{v})$.

If, in fact, $\mathbb{F} \subseteq \mathbb{R}$, then $R(\alpha\mathbf{u} + \beta\mathbf{v}) = \alpha R(\mathbf{u}) + \beta R(\mathbf{v})$ □

Example 16.5. Our version of the Riesz Representation Theorem is not the original one. While the restriction to finitely generated inner product spaces is not necessary, some restriction (on either the inner product space, V , or on the class of linear forms, $\varphi : V \rightarrow \mathbb{F}$) is required, as we now show, recalling Example 14.14.

Take $V := \mathbb{R}[t]$ with the inner product

$$\langle\langle \cdot, \cdot \rangle\rangle : V \times V \longrightarrow \mathbb{F}, \quad (f, g) \longrightarrow \int_0^1 f(x)g(x) dx$$

Take the linear transformation

$$\varphi : V \longrightarrow \mathbb{R}, \quad f \longmapsto f(0).$$

For any $h \in V$, $f := t^2h \in \ker(\varphi)$ and

$$\langle\langle f, h \rangle\rangle = \langle\langle t^2h, h \rangle\rangle = \int_0^1 x^2(h(x))^2 dx = \langle\langle th, th \rangle\rangle.$$

Thus $\varphi(f) = \langle\langle f, h \rangle\rangle$ if and only if $\varphi(f) = \|th\|^2$.

Since $\varphi(f) = 0$, this is the case if and only if $th = 0$.

As t is not the zero polynomial, this implies that h is the zero polynomial. Then $\langle\langle p, h \rangle\rangle = 0$ for all $p \in V$.

But φ is not the zero transformation, since $\varphi(1) = 1$.

Hence there is no $\mathbf{v}_\varphi \in V$ with $\varphi(\mathbf{x}) = \langle\langle \mathbf{x}, \mathbf{v}_\varphi \rangle\rangle$ for all $\mathbf{x} \in V$.

We consider the effect of an endomorphism on the above.

Let $T : V \longrightarrow V$ be an endomorphism of the finitely generated inner product space $(V, \langle\langle \cdot, \cdot \rangle\rangle)$ and take $\mathbf{v} \in V$. Then

$$L_T^\mathbf{v} : V \longrightarrow \mathbb{F}, \quad \mathbf{x} \longmapsto \langle\langle T(\mathbf{x}), \mathbf{v} \rangle\rangle$$

is a linear form on V .

By the Riesz Representation Theorem, there is a unique $\mathbf{v}_{L_T^\mathbf{v}} \in V$ with $L_T^\mathbf{v}(\mathbf{x}) = \langle\langle \mathbf{x}, \mathbf{v}_{L_T^\mathbf{v}} \rangle\rangle$ for all $\mathbf{x} \in V$. In other words,

$$\langle\langle T(\mathbf{x}), \mathbf{v} \rangle\rangle = \langle\langle \mathbf{x}, \mathbf{v}_{L_T^\mathbf{v}} \rangle\rangle$$

for all $\mathbf{x} \in V$. Given a fixed endomorphism $T : V \longrightarrow V$ we obtain a function

$$T^* : V \longrightarrow V, \quad \mathbf{v} \longmapsto \mathbf{v}_{L_T^*},$$

characterised by

$$\langle\langle T(\mathbf{x}), \mathbf{y} \rangle\rangle = \langle\langle \mathbf{x}, T^*(\mathbf{y}) \rangle\rangle \quad \text{for all } \mathbf{x}, \mathbf{y} \in V.$$

Lemma 16.6. *For each endomorphism $T : V \longrightarrow V$,*

$$T^* : V \longrightarrow V, \quad \mathbf{v} \longmapsto \mathbf{v}_T^*$$

is a linear transformation

Proof. Take $\mathbf{u}, \mathbf{v} \in V$ and $\alpha, \beta \in \mathbb{F}$. Then, for each $\mathbf{x} \in V$,

$$\begin{aligned} \langle\langle \mathbf{x}, T^*(\alpha\mathbf{u} + \beta\mathbf{v}) \rangle\rangle &= \langle\langle T(\mathbf{x}), \alpha\mathbf{u} + \beta\mathbf{v} \rangle\rangle \\ &= \alpha\langle\langle T(\mathbf{x}), \mathbf{u} \rangle\rangle + \beta\langle\langle T(\mathbf{x}), \mathbf{v} \rangle\rangle \\ &= \alpha\langle\langle \mathbf{x}, T^*(\mathbf{u}) \rangle\rangle + \beta\langle\langle \mathbf{x}, T^*(\mathbf{v}) \rangle\rangle \\ &= \langle\langle \mathbf{x}, \alpha T^*(\mathbf{u}) + \beta T^*(\mathbf{v}) \rangle\rangle. \end{aligned}$$

By the uniqueness of $\mathbf{v}_{L_T^*}$, $T^*(\alpha\mathbf{u} + \beta\mathbf{v}) = \alpha T^*(\mathbf{u}) + \beta T^*(\mathbf{v})$. □

Definition 16.7. Given a linear transformation $T : V \longrightarrow V$, the linear transformation, $T^* : V \longrightarrow V$, is the *adjoint* of T .

Lemma 16.8.

$$id_V^* = id_V$$

and for $S, T : V \longrightarrow V$,

$$(S \circ T)^* = T^* \circ S^*.$$

Proof. The former statement is obvious.

As for the latter, take $\mathbf{v}, \mathbf{x} \in V$. Then

$$\begin{aligned} \langle\langle \mathbf{x}, (S \circ T)^*(\mathbf{v}) \rangle\rangle &= \langle\langle (S \circ T)(\mathbf{x}), \mathbf{v} \rangle\rangle \\ &= \langle\langle S(T(\mathbf{x})), \mathbf{v} \rangle\rangle \\ &= \langle\langle T(\mathbf{x}), S^*(\mathbf{v}) \rangle\rangle \\ &= \langle\langle \mathbf{x}, T^*(S^*(\mathbf{v})) \rangle\rangle \\ &= \langle\langle \mathbf{x}, (T^* \circ S^*)(\mathbf{v}) \rangle\rangle. \end{aligned}$$

By uniqueness, $(S \circ T)^*(\mathbf{v}) = (T^* \circ S^*)(\mathbf{v})$ for all $\mathbf{v} \in V$. Hence $(S \circ T)^* = T^* \circ S^*$. □

There is a useful and important relationship between subspaces invariant under an endomorphism and the adjoint of the endomorphism.

Theorem 16.9. *Let $T : V \rightarrow V$ be an endomorphism of the inner product space $(V, \langle\langle \cdot, \cdot \rangle\rangle)$. If the subspace W of V is invariant under T , then W^\perp is invariant under T^* .*

In other words, if $T(\mathbf{w}) \in W$ for all $\mathbf{w} \in W$, then $T^(\mathbf{x}) \in W^\perp$ for all $\mathbf{x} \in W^\perp$.*

Proof. Take $\mathbf{w} \in W$ and $\mathbf{x} \in W^\perp$. Then

$$\begin{aligned}\langle\langle \mathbf{w}, T^*(\mathbf{x}) \rangle\rangle &= \langle\langle T(\mathbf{w}), \mathbf{x} \rangle\rangle \\ &= 0 \quad \text{since } T(\mathbf{w}) \in W \text{ and } \mathbf{x} \in W^\perp\end{aligned}$$

Thus $T^*(\mathbf{x}) \in W^\perp$. □

An important class of endomorphisms are those which agree with their adjoints.

Definition 16.10. The endomorphism $T : V \rightarrow V$ is *self-adjoint* if and only if $T^* = T$.

Being self-adjoint has significant consequences for the eigenvalues of an endomorphism.

Theorem 16.11. *The eigenvalues of a self-adjoint endomorphism are all real.*

Proof. Let $\mathbf{v} \neq \mathbf{0}_V$ be an eigenvector of the self-adjoint endomorphism T for the eigenvalue λ . Then

$$\lambda \langle\langle \mathbf{v}, \mathbf{v} \rangle\rangle = \langle\langle \lambda \mathbf{v}, \mathbf{v} \rangle\rangle = \langle\langle T(\mathbf{v}), \mathbf{v} \rangle\rangle = \langle\langle \mathbf{v}, T^*(\mathbf{v}) \rangle\rangle = \langle\langle \mathbf{v}, T(\mathbf{v}) \rangle\rangle = \langle\langle \mathbf{v}, \lambda \mathbf{v} \rangle\rangle = \bar{\lambda} \langle\langle \mathbf{v}, \mathbf{v} \rangle\rangle.$$

Since $\mathbf{v} \neq \mathbf{0}_V$, it follows that $\lambda = \bar{\lambda}$, which is the case if and only if λ is real. □

Corollary 16.12. *Every self-adjoint endomorphism has eigenvalues, whenever the ground field contains all real numbers.*

Proof. By the Fundamental Theorem, the characteristic polynomial factors into linear factors over the complex numbers, so that

$$\chi_T(t) = \prod_{j=1}^n (t - \lambda_j).$$

The λ_j 's are precisely the eigenvalues of T . Since these are all real, this is, in fact, a factorisation over the reals. Hence T has n real eigenvalues (with multiplicities). □

Corollary 16.13. *Eigenvectors to distinct eigenvalues of a self-adjoint endomorphism are mutually orthogonal.*

Proof. Let $T : V \rightarrow V$ be self-adjoint. Take $\mathbf{u}, \mathbf{v} \in V \setminus \{\mathbf{0}_V\}$ with $T(\mathbf{u}) = \lambda \mathbf{u}$ and $T(\mathbf{v}) = \mu \mathbf{v}$ with $\lambda \neq \mu$. Then

$$\begin{aligned}\lambda \langle\langle \mathbf{u}, \mathbf{v} \rangle\rangle &= \langle\langle \lambda \mathbf{u}, \mathbf{v} \rangle\rangle \\ &= \langle\langle T(\mathbf{u}), \mathbf{v} \rangle\rangle \\ &= \langle\langle \mathbf{u}, T(\mathbf{v}) \rangle\rangle && \text{as } T \text{ is self-adjoint} \\ &= \langle\langle \mathbf{u}, \mu \mathbf{v} \rangle\rangle \\ &= \mu \langle\langle \mathbf{u}, \mathbf{v} \rangle\rangle && \text{as } \mu \in \mathbb{R}.\end{aligned}$$

Since $\lambda \neq \mu$, $\langle\langle \mathbf{u}, \mathbf{v} \rangle\rangle = 0$. □

We come to our main result on self-adjoint endomorphisms.

Theorem 16.14. *If $T : V \rightarrow V$ is a self-adjoint endomorphism of the finitely generated inner product space $(V, \langle\langle \cdot, \cdot \rangle\rangle)$, then V has an orthonormal basis of eigenvectors of T .*

Proof. We use induction on $\dim(V)$.

If $\dim(V) = 1$, then there is nothing to prove.

Suppose that the result holds for self-adjoint endomorphisms of inner product spaces of dimension less than n .

Let $\dim(V) = n$ and take a self-adjoint endomorphism $T : V \rightarrow V$.

By Corollary 16.12, T has an eigenvalue, say λ . Let $\mathbf{v} \neq \mathbf{0}_V$ be an eigenvector. Then $W := \mathbb{F}\mathbf{v}$ is a T -invariant subspace of V .

Since V is finitely generated, $V \cong W \oplus W^\perp = \mathbb{F}\mathbf{v} \oplus (\mathbb{F}\mathbf{v})^\perp$, whence $\dim W = n - 1$.

By Theorem 16.9, W^\perp is a T^* -invariant subspace of V .

Since T is self-adjoint, W^\perp is an $(n - 1)$ -dimensional T -invariant subspace of V .

By the inductive hypothesis, W^\perp has an orthonormal basis, $\{\mathbf{e}_2, \dots, \mathbf{e}_n\}$, consisting of eigenvectors of T .

Putting $\mathbf{e}_1 := \frac{\mathbf{v}}{\|\mathbf{v}\|}$, $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ is an orthonormal basis for V comprising eigenvectors of T . \square

Corollary 16.15. *If $\underline{\mathbf{A}}$ is a Hermitian $n \times n$ complex matrix, then there is a unitary matrix, $\underline{\mathbf{B}}$, such that $\underline{\mathbf{B}}^t \underline{\mathbf{A}} \underline{\mathbf{B}}$ is a (real) diagonal matrix.*

Proof. Recall that if take $\mathbb{C}_{(n)}$ with the standard (Euclidean) inner product and regard the $n \times n$ complex matrix $\underline{\mathbf{A}}$ as the linear transformation

$$\mathbb{C}_{(n)} \longrightarrow \mathbb{C}_{(n)}, \quad \mathbf{x} \longmapsto \underline{\mathbf{A}}\mathbf{x},$$

then $\underline{\mathbf{A}}$ is self-adjoint if and only if it is Hermitian.

In that case $\mathbb{C}_{(n)}$ has an orthonormal $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ basis comprising eigenvectors of $\underline{\mathbf{A}}$.

Let $\underline{\mathbf{B}}$ be the $n \times n$ complex matrix whose j^{th} column is \mathbf{e}_j .

Since $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ is orthonormal, $\underline{\mathbf{B}}^t \underline{\mathbf{B}} = \underline{\mathbf{1}}_n$.

Thus $\underline{\mathbf{B}}^{-1} = \underline{\mathbf{B}}^t$, that is to say, $\underline{\mathbf{B}}$ is unitary.

Moreover, since for each j there is a $\lambda_j \in \mathbb{R}$ with $\underline{\mathbf{A}}\mathbf{e}_j = \lambda_j \mathbf{e}_j$, we have

$$\underline{\mathbf{A}} \underline{\mathbf{B}} = \underline{\mathbf{B}} \underline{\text{diag}}(\lambda_1, \dots, \lambda_n),$$

where $\underline{\text{diag}}(\lambda_1, \dots, \lambda_n)$ is the $n \times n$ matrix $[x_{ij}]_{n \times n}$ defined by

$$x_{ij} = \begin{cases} \lambda_j & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}$$

Thus $\underline{\mathbf{B}}^t \underline{\mathbf{A}} \underline{\mathbf{B}} = \underline{\mathbf{B}}^{-1} \underline{\mathbf{A}} \underline{\mathbf{B}} = \underline{\text{diag}}(\lambda_1, \dots, \lambda_n)$ is a (real) diagonal matrix. \square

Corollary 16.16. *If $\underline{\mathbf{A}}$ is a symmetric $n \times n$ real matrix, then there is an orthogonal matrix, $\underline{\mathbf{B}}$, such that $\underline{\mathbf{B}}^t \underline{\mathbf{A}} \underline{\mathbf{B}}$ is a diagonal matrix.*

16.2 Exercises

Exercise 16.1. Take $V := \mathcal{P}_3$, the set of all polynomials of degree at most 2 in the indeterminate t with real coefficients, with inner product $\langle\langle \ , \ \rangle\rangle$ given by

$$\langle\langle p, q \rangle\rangle := \int_0^1 p(x)q(x)dx$$

Consider the linear form

$$\varphi : V \longrightarrow \mathbb{R}, \quad p \longmapsto p(0)$$

Find the element of \mathcal{P}_3 which represents φ .

Exercise 16.2. Find an orthogonal matrix which diagonalises the real matrix

$$\begin{bmatrix} 1 & 2 \\ 2 & 5 \end{bmatrix}$$

Exercise 16.3. Find a unitary matrix which diagonalises the complex matrix

$$\begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}$$