Math 413/513 Chapter 6 (from Friedberg, Insel, & Spence)

David Glickenstein

December 7, 2015

1 Inner product spaces

In this chapter, we will only consider the fields $\mathbb{R}$ and $\mathbb{C}$.

Definition 1 *Let* V *be a vector space over* $F = \mathbb{R}$ *or* $\mathbb{C}$*. An* inner product *on* V *is a function* $V \times V \to F$*, denoted* $(x, y) \to \langle x, y \rangle$ *, such that the following hold for all* $x, y, z \in V$ *and* $c \in F$*:*

1. $\langle x + z, y \rangle = \langle x, y \rangle + \langle z, y \rangle$.
2. $\langle cx, y \rangle = c \langle x, y \rangle$.
3. $\langle y, x \rangle = \overline{\langle x, y \rangle}$ *where the bar denotes complex conjugation.*
4. $\langle x, x \rangle > 0$ *if* $x \neq 0$.

Here are some observations:

- The first two conditions could be called being *linear in the first component.*
- The third condition is called being *conjugate symmetric* (or *symmetric* if $F = \mathbb{R}$).
- Linearity in the first component and conjugate symmetry imply linearity in the second component, and being linear in both components is called being *bilinear.*
- Notice that conjugate symmetry implies that $\langle x, x \rangle \in \mathbb{R}$ even if $F = \mathbb{C}$ since $\langle x, x \rangle = \overline{\langle x, x \rangle}$.
- The second condition also implies that $\left\langle \vec{0}, \vec{0} \right\rangle = 0$, which together with the fourth condition is called being *positive definite.*
- From all of this, we could have just specified that $\langle \cdot, \cdot \rangle$ is bilinear, conjugate symmetric, and positive definite.

- We often write $\|v\|^2$ to represent $\langle v, v\rangle$. Since $\|v\|^2 \geq 0$, it has a unique positive square root that we call $\|v\|$.

Example 2 *The first example is the dot product for* $\mathbb{R}^n$, *where if* $x = (x_1, \ldots, x_n)$ *and* $y = (y_1, \ldots, y_n)$ *then*

$$\langle x, y\rangle = x \cdot y = \sum_{i=1}^{n} x_i y_i.$$

Example 3 *The standard inner product on* $\mathbb{C}^n$ *is*

$$\langle x, y\rangle = \sum_{i=1}^{n} x_i \overline{y_i}.$$

Check the properties.

Example 4 *Given any inner product* $\langle \cdot, \cdot\rangle$, *we can multiply it by a positive real number* $r > 0$ *to get another inner product* $\langle x, y\rangle' = r\langle x, y\rangle$. *Note that it would not be an inner product if* $r \leq 0$ *or* r *is not a real number.*

Example 5 *For continuous, real-valued functions on* $[0, 1]$, *there is an inner product*

$$\langle f, g\rangle = \int_0^1 f(t)\, g(t)\, dt.$$

Note that it is important that the functions be continuous to ensure that $\langle f, f\rangle = \int_0^1 f(t)^2\, dt > 0$ *if* $f \neq \vec{0}$.

Example 6 *Consider the vector space* $\mathbb{C}^{n\times n}$. *We define the conjugate transpose, or adjoint,* A^* *of a matrix* A *by specifying the entries as:*

$$(A^*)_{ij} = \overline{A_{ji}}.$$

We can now define an inner product by

$$\langle A, B\rangle = tr\,(B^* A).$$

Let's check it is an inner product. First, $\langle A + A', B\rangle = tr\,(B^*(A + A')) = tr\,(B^*A + B^*A') = tr\,(B^*A) + tr\,(B^*A') = \langle A, B\rangle + \langle A', B\rangle$. *Also,*

$$\begin{aligned} \langle B, A\rangle &= tr\,(A^* B) \\ \overline{\langle A, B\rangle} &= tr\left(\overline{B^* A}\right) = tr\,(A^* B). \end{aligned}$$

Also,

$$\langle A, A\rangle = tr\,(A^* A) = \sum_i \sum_j A^*_{ij} A_{ji} = \sum_{j,i} \overline{A_{ji}} A_{ji} = \sum_{j,i} |A_{ji}|^2$$

and so this is nonnegative and equals zero if and only if $A = 0$.

Definition 7 *A vector space V together with an inner product is called an* inner product space. *If the field is $\mathbb{C}$ then we call it a* complex inner product space, *and if the field is $\mathbb{R}$ we call it a* real inner product space.

The following are basic properties of inner product spaces.

Theorem 8 *Let V be an inner product space. Then for $x, y, z \in V$ and $c \in F$ the following statements are true:*

1. $\langle x, y+z \rangle = \langle x, y \rangle + \langle x, z \rangle$.
2. $\langle x, cy \rangle = \bar{c} \langle x, y \rangle$.
3. $\left\langle x, \vec{0} \right\rangle = \left\langle \vec{0}, x \right\rangle = 0$.
4. $\langle x, x \rangle = 0$ *if and only if* $x = \vec{0}$.
5. *If* $\langle x, y \rangle = \langle x, z \rangle$ *for all* $x \in V$, *then* $y = z$.

Proof. Most of these follow pretty easily. The last one can be shown as follows. Suppose $\langle x, y \rangle = \langle x, z \rangle$ for all $x \in V$. Then $\langle x, y - z \rangle = 0$ for all x. In particular, taking $x = y - z$, we get that $\langle y - z, y - z \rangle = 0$. But that implies $y - z = \vec{0}$. ■

Note that the first two statements in the above theorem are called being *conjugate linear* in the second component.

Recall the definition of the length, or norm: $\|x\| = \sqrt{\langle x, x \rangle}$. This generalizes the Euclidean norm $\|(x_1, \ldots, x_n)\| = \sqrt{x_1^2 + \cdots + x_n^2}$. Many properties are still true for inner product spaces:

Theorem 9 *Let V be an inner product space over $F = \mathbb{R}$ or $\mathbb{C}$. The for $x, y \in V$ and $c \in F$, the following are true:*

1. $\|cx\| = |c| \cdot \|x\|$.
2. $\|x\| = 0$ *iff* $x = \vec{0}$. *In general,* $\|x\| \geq 0$.
3. *(Cauchy-Schwarz inequality)* $|\langle x, y \rangle| \leq \|x\| \|y\|$.
4. *(Triangle inequality)* $\|x + y\| \leq \|x\| + \|y\|$.

Proof. We will just prove the last two. Consider $x - cy$ and notice that

$$\begin{aligned} 0 &\leq \|x - cy\|^2 \\ &= \langle x - cy, x - cy \rangle \\ &= \|x\|^2 - \langle cy, x \rangle - \langle x, cy \rangle + \|cy\|^2 \\ &= \|x\|^2 - 2 \operatorname{Re} \bar{c} \langle x, y \rangle + |c|^2 \|y\|^2 . \end{aligned}$$

Notice that if we take $c = \frac{\langle x, y \rangle}{\|y\|^2}$ then

$$0 \leq \|x\|^2 - 2 \frac{|\langle x, y \rangle|^2}{\|y\|^2} + \frac{|\langle x, y \rangle|^2}{\|y\|^2} = \|x\|^2 - \frac{|\langle x, y \rangle|^2}{\|y\|^2},$$

which implies the Cauchy-Schwarz inequality.

We can now show

$$\begin{aligned} \|x+y\|^2 &= \langle x+y, x+y\rangle \\ &= \|x\|^2 + 2\operatorname{Re}\langle x,y\rangle + \|y\|^2 \\ &\le \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 \\ &= (\|x\|+\|y\|)^2 \end{aligned}$$

since $2a \le a^2+b^2 = |a+bi|^2$ and so $2\operatorname{Re}\langle x,y\rangle \le |\langle x,y\rangle| \le \|x\|\|y\|$. ■

Definition 10 *Let V be an inner product space. Vectors x and y in V are* orthogonal *(*perpendicular*) if $\langle x,y\rangle = 0$. As subset $S \subseteq V$ is* orthogonal *if any two distinct vectors in S are orthogonal. A vector x in V is a* unit vector *if $\|x\| = 1$. A subset $S \subseteq V$ is* orthonormal *if S is orthogonal and consists entirely of unit vectors.*

Note that $S = \{x_1, \dots, x_k\}$ is orthonormal iff $\langle x_i, x_j\rangle = \delta_{ij}$. Also note that we can make an orthonormal set from an orthogonal set by replacing each vector x by $\frac{1}{\|x\|}x$. This will not change the orthogonality since $\left\langle \frac{x}{\|x\|}, \frac{y}{\|y\|}\right\rangle = \frac{1}{\|x\|\|y\|}\langle x,y\rangle$ since $\|y\| \in \mathbb{R}$. We call this process *normalizing* the set.

Proposition 11 *If V is an inner product space and $S \subseteq V$ is orthogonal and nonzero, then S is linearly independent.*

Proof. We first note that if S is not the set consisting only of zero, then zero cannot be in S. Suppose $a_1x_1 + \cdots + a_kx_k = \vec{0}$ for scalars $a_1, \dots, a_k$ and vectors $x_1, \dots, x_k$ in S. Then we see that

$$0 = \langle a_1x_1 + \cdots + a_kx_k, x_i\rangle = a_i \|x_i\|^2$$

and since $\|x_i\|^2 \neq 0$, we must have $a_i = 0$. This can be done for all i. ■

2 Problems

- FIS Section 6.1 exercises 2, 3, 8, 10, 11, 16, 17, 19, 20, 22

3 Orthonormal bases

Definition 12 *Let V be an inner product space. A subset of V is an* orthonormal basis *for V if it is an ordered basis that is orthonormal.*

The standard basis is orthonormal for the usual inner product. So is any rotation of the standard basis!

Orthonormal bases make it easier to write the coefficients than in any old basis.

Theorem 13 *Let V be an inner product space and $S = \{v_1, v_2, \ldots, v_k\}$ be an orthogonal subset of V consisting of nonzero vectors. If $y \in \operatorname{span} S$, then*

$$y = \sum_{i=1}^{k} \frac{\langle y, v_i \rangle}{\langle v_i, v_i \rangle} v_i.$$

Proof. Since $y \in \operatorname{span} S$, we must have that there exist scalars $a_1, \ldots, a_k$ such that

$$y = \sum_{i=1}^{k} a_i v_i.$$

We can now take the inner product with v_j for $j = 1, \ldots, k$ and find that

$$\begin{aligned} \langle y, v_j \rangle &= \left\langle \sum_{i=1}^{k} a_i v_i, v_j \right\rangle \\ &= \sum_{i=1}^{k} a_i \langle v_i, v_j \rangle \\ &= a_j \langle v_j, v_j \rangle \end{aligned}$$

and so (since $\|v_j\| \neq 0$), $a_j = \frac{\langle y, v_j \rangle}{\langle v_j, v_j \rangle}$.

$$y = \sum_{i=1}^{k} \frac{\langle y, v_i \rangle}{\langle v_i, v_i \rangle} v_i.$$

■

Note that if S is orthonormal, then the denominators are all 1.

We can always take a linearly independent set and use it to find an orthogonal set with the same span. We do this by considering orthogonal projections of vectors onto a subspace.

Theorem 14 *Let W be a finite dimensional subspace of the inner product space V. Then for a vector $y \in V$, there is a unique vector $u \in W$ that minimizes $\|y - w\|^2$ for all $w \in W$.*

Proof. Suppose there is a $u \in W$ such that $\langle w, y - u \rangle = 0$ for any $w \in W$. Then if $w \in W$ (and hence so is $u - w$),

$$\begin{aligned} \|y - w\|^2 &= \|u + (y - u) - w\|^2 \\ &= \langle u - w + (y - u), u - w + (y - u) \rangle \\ &= \|u - w\|^2 + \langle u - w, y - u \rangle + \langle y - u, u - w \rangle + \|y - u\|^2 \\ &= \|u - w\|^2 + \|y - u\|^2 \\ &\geq \|y - u\|^2. \end{aligned}$$

We can do this if W is finite dimensional using the following theorem. ■

Definition 15 *We call the assignment of* u *the orthogonal projection of* y *onto* W*, denoted* $u = P_W(y)$*.*

It is an important fact that $\langle y - P_W(y), w\rangle = 0$ for all $w \in W$.

Definition 16 *The orthogonal complement of* W*, written* $W^\perp$ *(pronounced "W perp"), is the set of all vectors* $v \in V$ *such that* $\langle v, w\rangle$ *for all* $w \in W$*.*

Note that $W^\perp$ is a vector space.

Proposition 17 $W^\perp$ *is a vector space.*

Proof. It is straightforward to see that $\left\langle \vec{0}, w \right\rangle = 0$ for all $w \in W$, so $\vec{0} \in W^\perp$. if $v, u \in W^\perp$ and $c \in F$ then $\langle cv + u, w\rangle = c\langle v, w\rangle = \langle u, w\rangle = 0$ so $cv + u \in W^\perp$. ■

We can construct an orthogonal set from a linearly independent set by starting with the first vector and then projecting the next vector into

Theorem 18 *Let* V *be an inner product space and* $S = \{w_1, \ldots, w_n\}$ *be a linearly independent subset of* V*. Define* $S' = \{v_1, \ldots, v_n\}$ *by* $v_1 = w_1$ *and*

$$v_k = w_k - \sum_{j=1}^{k-1} \frac{\langle w_k, v_j\rangle}{\langle v_j, v_j\rangle} v_j$$

for $k = 2, \ldots, n$*. Then* S' *is an orthogonal set of nonzero vectors such that* $\operatorname{span} S' = \operatorname{span} S$.

Proof. We show inductively that v_{k+1} is orthogonal to $v_1, \ldots, v_k$. It is clear that

$$\begin{aligned}\langle v_2, v_1\rangle &= \left\langle w_2 - \frac{\langle w_2, v_1\rangle}{\langle v_1, v_1\rangle} v_1, v_1 \right\rangle \\ &= \langle w_2, v_1\rangle - \frac{\langle w_2, v_1\rangle}{\langle v_1, v_1\rangle}\langle v_1, v_1\rangle = 0.\end{aligned}$$

We then can use the inductive hypothesis to assume $\langle v_i, v_j\rangle = 0$ for $i, j \leq k$ and see that

$$\begin{aligned}\langle v_k, v_i\rangle &= \left\langle w_k - \sum_{j=1}^{k-1} \frac{\langle w_k, v_j\rangle}{\langle v_j, v_j\rangle} v_j, v_i \right\rangle \\ &= \langle w_k, v_i\rangle - \frac{\langle w_k, v_i\rangle}{\langle v_i, v_i\rangle}\langle v_i, v_i\rangle = 0.\end{aligned}$$

Thus S' is orthogonal. Hence S' is linearly independent and since each element of S' is in the span of S, $\operatorname{span} S' \subseteq \operatorname{span} S$, and hence $\operatorname{span} S' = \operatorname{span} S$ (since they have the same dimension). ■

Note: this process of producing an orthogonal set is called the *Gram-Schmidt process.*

Theorem 19 *Suppose that* $S = \{v_1, \ldots, v_k\}$ *is an orthonormal set in a* n*-dimensional inner product space* V*. Then*

1. S *can be extended to an orthonormal basis* $\{v_1, \ldots, v_k, v_{k+1}, \ldots, v_n\}$ *for* V*.*
2. *If* $W = \operatorname{span} S$*, then* $S_1 = \{v_{k+1}, \ldots, v_n\}$ *is an orthonormal basis for* $W^{\perp}$*.*
3. *If* W *is any subspace of* V*, then* $\dim V = \dim W + \dim W^{\perp}$*.*

Proof. By the replacement theorem, S can be extended into a basis, and then the Gram-Schmidt process can be used to turn this into an orthogonal set. Then normalizing gives an orthonormal set. S_1 is clearly a linearly independent subset of $W^{\perp}$. Since $\{v_1, \ldots, v_n\}$ is a basis, any vector in $W^{\perp}$ can be written as a linear combination of these vectors. However, since $w \in W^{\perp}$ satisfies $\langle w, v_i \rangle = 0$ for $i = 1, \ldots, k$, w is in the span of S_1, hence S_1 is a basis. The dimension statement is clear now that we know that S is a basis for S, S' is a basis for $W^{\perp}$, and $\{v_1, \ldots, v_n\}$ is a basis for V. ■

4 Problems

- FIS Section 6.2 exercises 4, 5, 7, 8, 10, 13, 22.

5 Adjoints and eigenvalues

Definition 20 *Suppose* T *is a linear operator on an inner product space* V*. If* T^* *is a linear operator on* V *such that*

$$\langle T(x), y \rangle = \langle x, T^*(y) \rangle$$

for all $x, y \in V$*, we say* T^* *is the* adjoint *of* T*. We read* T^* *as "*T *star."*

Theorem 21 *Let* V *be a finite-dimensional inner product space and let* T *be a linear operator on* V*. Then there exists a unique function* $T^* : V \to V$ *that for all* $x, y \in V$*,*

$$\langle T(x), y \rangle = \langle x, T^*(y) \rangle.$$

Furthermore, T^* *is linear and hence the adjoint of* T*.*

In order to prove this theorem, we need to show that every linear map on a finite-dimensional vector space can be represented in terms of the inner product. We will then use this idea to construct the adjoint.

Theorem 22 *Let* V *be a finite dimensional vector space over* F *and let* $g : V \to F$ *be a linear transformation. Then there exists a unique* $y \in V$ *such that* $g(x) = \langle x, y \rangle$ *for all* $x \in V$*.*

Proof. Let $\beta = \{v_1, \ldots, v_n\}$ be an orthonormal basis for V. For any $x \in V$, we must have $y = \sum b_i v_i$ for some b_i, so in order for $g(x) = \langle x, y \rangle$ we must have that

$$g(v_i) = \langle v_i, y \rangle = \bar{b}_i.$$

Hence any y that satisfies the equation must have $b_i = \overline{g(v_i)}$ (this proves uniqueness). We now confirm that $y = \sum \overline{g(v_i)} v_i$ satisfies the theorem: if $x = \sum a_i v_i$ then

$$\begin{aligned} \langle x, y \rangle &= \left\langle \sum_{i=1}^{n} a_i v_i, \sum_{j=1}^{n} \overline{g(v_j)} v_j \right\rangle \\ &= \sum_{i=1}^{n} \sum_{j=1}^{n} a_i g(v_j) \langle v_i, v_j \rangle \\ &= \sum_{i=1}^{n} a_i g(v_i) = g(x). \end{aligned}$$

■

Proof of Theorem 21. Let $\beta = \{v_1, \ldots, v_n\}$, be an orthonormal basis for V and define $T_i : V \to F$ by $T_i(x) = \langle T(x), v_i \rangle$. By Theorem 22, there exists a unique $y_i \in V$ such that $T_i(x) = \langle x, y_i \rangle$, and so

$$\langle T(x), v_i \rangle = \langle x, y_i \rangle.$$

We can thus define a linear tranformation T^* by defining it on the basis to be $T^*(v_i) = y_i$ and extending it linearly to a transformation on V. Now, for any vector $y \in V$, $y = \sum b_i v_i$ for some scalars $b_1, \ldots, b_n$ and we check that

$$\begin{aligned} \langle T(x), y \rangle &= \left\langle T(x), \sum b_i v_i \right\rangle \\ &= \sum \bar{b}_i \langle T(x), v_i \rangle \\ &= \sum \bar{b}_i \langle x, T^*(v_i) \rangle \\ &= \sum \langle x, b_i T^*(v_i) \rangle \\ &= \langle x, T^*(y) \rangle. \end{aligned}$$

T^* is unique since if there were another linear transformation U satisfying the properties, then for any $y \in V$,

$$\langle x, T^*(y) \rangle = \langle T(x), y \rangle = \langle x, U(y) \rangle$$

must be true for all $x \in V$, implying that $T^*(y) = U(y)$. (See 6.1, problem 9.)

■

Notice that the adjoint works the other way as well:

$$\langle x, T(y) \rangle = \overline{\langle T(y), x \rangle} = \overline{\langle y, T^*(x) \rangle} = \langle T^*(x), y \rangle.$$

Note our proof of the existence of an adjoint used a basis, and it turns out that for a linear operator on an infinite dimensional vector space, the existence of an adjoint is not guaranteed. However, most properties we will derive are true if the adjoint exists (sometimes it does).

Recall that for a matrix A, we denote its conjugate transpose by A^* (conjugate transpose means we take a transpose and replace each entry with its complex conjugate). This is related to the adjoint.

Theorem 23 *Let V be a finite dimensional inner product space and β be an orthonormal basis for V. If T is a linear operator on V, then*

$$[T^*]_\beta = [T]_\beta^* .$$

It might help to be clear about the content here. If one writes the matrix for the adjoint transformation T^*, this matrix is the same as the conjugate transpose of the matrix for the transformation T (provided the basis used in both cases is orthonormal).

Proof. If we write $A = [T]_\beta$ and $B = [T^*]_\beta$ then by orthonormality we have that

$$B_{ij} = \langle T^*(v_j), v_i\rangle = \langle v_j, T(v_i)\rangle = \overline{\langle T(v_i), v_j\rangle} = \bar{A}_{ji}.$$

It follows that $B = A^*$. ■

Corollary 24 *Left multiplication by an $n \times n$ matrix A satisfies*

$$L_{A^*} = L_A^*.$$

Theorem 25 *Let V be an inner product space and T, U be linear operators on V that have an adjoint (this is always true if V is finite dimensional). Then*

1. $(T+U)^* = T^* + U^*$
2. $(cT)^* = \bar{c}T^*$ *for any* $c \in F$
3. $(TU)^* = U^*T^*$
4. $T^{**} = T$
5. $I_V^* = I_V$.

Proof. The proofs are pretty straightforward. We prove the first and third: for any $x, y \in V$

$$\begin{aligned}\langle (T+U)x, y\rangle &= \langle T(x) + U(x), y\rangle \\ &= \langle T(x), y\rangle + \langle U(x), y\rangle \\ &= \langle x, T^*(y)\rangle + \langle x, U^*(y)\rangle \\ &= \langle x, T^*(y) + U^*(y)\rangle \\ &= \langle x, (T^* + U^*)(y)\rangle .\end{aligned}$$

Also,

$$\begin{aligned}\langle TU(x), y\rangle &= \langle T(U(x)), y\rangle \\ &= \langle U(x), T^*(y)\rangle \\ &= \langle x, U^*T^*(y)\rangle .\end{aligned}$$

■

6 Problems

- FIS Section 6.3 exercises 4, 6, 8, 10, 12, 13, 15

7 Schur's theorem and the (baby) spectral theorem

We would like to understand when there exists an orthonormal basis of eigenvectors. We first have Schur's theorem, that says that we can represent a linear transformation by an upper triangular matrx.

Theorem 26 (Schur) *Let T be a linear operator on a finite dimensional inner product space V. Suppose the characteristic polynomial of T splits. The there exists an orthonormal basis β of V such that $[T]_\beta$ is upper triangular.*

The main idea behind the proof is that we can restrict T to a subspace. If W is a subspace of V and $T(W) \subseteq W$, then we say that W is *T-invariant.* Hence there is a linear operator $T_W \in \mathcal{L}(W)$ that is just $T_W(x) = T(x)$. If W is T-invariant and β is a basis for V that is an extension of a basis for W, then $[T]_\beta$ has block form

$$\begin{bmatrix} [T_W] & * \\ 0 & * \end{bmatrix}$$

where the $*$'s are unknown but generally not zero.

We now see that if z is an eigenvector of T^* and if $z^\perp = \{x \in V : \langle x, z\rangle = 0\}$, then $z^\perp$ is T-invariant, i.e.,

$$\langle T(y), z\rangle = \langle y, T^*(z)\rangle = \langle y, \lambda z\rangle = \bar{\lambda}\langle y, z\rangle = 0.$$

Proof. We first notice that the characteristic polynomial of T^* satisfies

$$\det(T^* - \lambda I) = \det\left(T - \bar{\lambda} I\right)$$

since determinant is invariant under trace. Hence if the characteristic polynomial of T splits, then there is an eigenvector z of T^*. Then $z^\perp$ is T-invariant, and so we can use the inductive hypothesis to show that $[T]_\beta$ is upper triangular. This completes the proof. ■

Notice the following consequence:

Theorem 27 *Suppose $T \in \mathcal{L}(V)$ for a finite dimensional inner product space V. If T is self-adjoint, i.e., $T^* = T$, and its characteristic polynomial splits, then V has an orthonormal basis of eigenvectors of T.*

Proof. Since $T = T^*$, we have that $[T^*]_\beta = [T]_\beta^*$. If we choose the basis from the Schur theorem, $[T]_\beta$ is upper triangular and $[T]_\beta^*$ is lower triangular and since

$$[T]_\beta = [T^*]_\beta = [T]_\beta^*$$

we must have that $[T]_\beta$ is triangular. Note that we needed that the characteristic polynomials splits, which is not known for real inner product spaces, so we need the following lemma to complete the proof. ■

Lemma 28 *Let T be a self-adjoint operator on a finite dimensional inner product space V. Then*

1. *T has real eigenvalues*
2. *The characteristic polynomial of T splits.*

Proof. If λ is an eigenvalue with eigenvector v, then

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

and so $\lambda = \bar{\lambda}$ and λ is real. The second statement follows from the fundamental theorem of algebra if $F = \mathbb{C}$. If $F = \mathbb{R}$ then we can consider $A = [T]_\beta$ for some basis β, and then consider the linear transformation $L_A : \mathbb{C}^n \to \mathbb{C}^n$ given by $L_A(x) = Ax$. Since T is self-adjoint, $A = A^*$ and hence L_A is self-adjoint and hence has real eigenvalues. Since all polynomials split over the complex numbers, the characteristic polynomial for L_A splits, and since the eigenvalues are real, the polynomial has real coefficients. The characteristic polynomial for L_A is the same as that for T, and since the eigenvalues are real, that means that the characteristic polynomial for T splits. ■

Remark 29 *The theorem above is if and only if, actually, for real inner product spaces. For complex inner product spaces, self-adjoint can be relaxed to normal, which means that T and T^* commute, i.e. $TT^* = T^*T$. For a normal operator, any eigenvector x of T, with eigenvalue λ, is an eigenvector for T^* since*

$$\begin{aligned} 0 &= \langle Tx - \lambda x, Tx - \lambda x\rangle \\ &= \langle x, T^*Tx\rangle - \lambda \langle x, Tx\rangle - \bar{\lambda} \langle Tx, x\rangle + |\lambda|^2 \langle x, x\rangle \\ &= \langle x, TT^*x\rangle - \lambda \langle T^*x, x\rangle - \bar{\lambda} \langle x, T^*x\rangle + |\lambda|^2 \langle x, x\rangle \\ &= \left\langle T^*x - \bar{\lambda}x, T^*x - \bar{\lambda}x\right\rangle \end{aligned}$$

and so x is an eigenvector for T^* with eigenvalue $\bar{\lambda}$.

8 Problems