Matth 108B Selected Homework Solutions
Charles Martin
March 5, 2013

Homework 1

5.1.7  (a) If matrices \( A, B \) both represent \( T \) under different bases, then for some invertible matrix \( Q \) we have \( B = QAQ^{-1} \). Then
\[
\det(B) = \det(QAQ^{-1}) = \det(Q) \cdot \det(A) \cdot \det(Q^{-1}),
\]
since the determinant for matrices is multiplicative. Furthermore \( \det(Q^{-1}) = (\det(Q))^{-1} \), so
\[
\det(B) = \det(A),
\]
as desired.

(b) We know that \( T \) is invertible if and only if its matrix is invertible with respect to any basis; we also know that a matrix is invertible if and only if its determinant is zero. If \( A \) is a matrix representing \( T \) in some basis, then \( T \) is invertible if and only if \( A \) is invertible if and only if \( \det(T) = \det(A) \neq 0 \).

(c) If \( T \) has matrix \( A \) under some basis, then \( T^{-1} \) has matrix \( A^{-1} \) under the same basis. Thus
\[
\det(T^{-1}) = \det(A^{-1}) = \det(A)^{-1} = \det(T)^{-1}.
\]

(d) If operators \( T, U \) are represented by matrices \( A, B \) (respectively) in some basis, then the matrix for \( TU \) is \( AB \). From this we have
\[
\det(TU) = \det(AB) = \det(A) \det(B) = \det(T) \det(U),
\]
since the determinant for matrices is multiplicative.

(e) Given any ordered basis \( \beta = \{v_1, \ldots, v_n\} \), the identity map \( I_V \) maps each \( v_i \) to itself. That is, \( [I_V]_{\beta} = I \). Matrix subtraction is defined so that \( [T - \lambda I_V]_{\beta} = [T]_{\beta} - \lambda[I_V]_{\beta} = [T]_{\beta} - \lambda I \). Therefore
\[
\det(T - \lambda I_V) = \det([T - \lambda I_V]_{\beta}) = \det([T]_{\beta} - \lambda I).
\]

Homework 2

5.1.8 Let \( T : V \to V \) be a linear operator.

(a) \( \Rightarrow \) Assume that \( T \) is invertible and \( Tv = 0 \) for some vector \( v \). Multiplying by \( T^{-1} \) gives \( v = 0 \), so the equation \( Tv = 0v \) has only trivial solutions; that is, zero is not an eigenvalue.

\( \Leftarrow \) Suppose that the equation \( Tv = 0 \) has only the trivial solution \( v = 0 \). That is, \( \ker(T) = \{0\} \). Since \( V \) is finite-dimensional, this implies \( T \) is invertible.

(b) Suppose that \( T \) is invertible and that \( Tv = \lambda v \) for some \( \lambda \neq 0 \) and \( v \neq 0 \). Multiplying by \( T^{-1} \) and \( \lambda^{-1} \) gives \( T^{-1}v = \lambda^{-1}v \). Multiplying by \( T \) and \( \lambda \) reverses the argument.
5.1.14 We know that \( \det(A) = \det(A^t) \) for all matrices \( A \); furthermore \( (\lambda I)^t = \lambda I \). Since the transpose operation is additive we have
\[
\det(A - \lambda I) = \det((A - \lambda I)^t) = \det(A^t - \lambda I).
\]

5.1.15a If \( Tv = \lambda v \) for some nonzero \( v \), then \( T^2v = T(Tv) = \lambda(Tv) = \lambda^2v \). Similarly, \( T^n v = \lambda^n v \) for each \( n \); an inductive argument shows that \( T^n v = \lambda^n v \) for all integers \( n \geq 0 \).

5.2.20 Throughout the following, \( V \) is a finite-dimensional vector space with subspaces \( W_1, \ldots, W_k \) satisfying
\[
V = \sum_{i=1}^{k} W_i.
\]
Note that with no further assumptions we have
\[
\dim(V) \leq \sum_{i=1}^{k} \dim(W_i),
\]
which can be proved with induction and the readily-verified result
\[
\dim(W_1 + W_2) \leq \dim(W_1) + \dim(W_2).
\]
Let’s give a brief proof of the above inequality. Given bases \( \beta = \{v_1, \ldots, v_m\} \) and \( \gamma = \{w_1, \ldots, w_n\} \) for \( W_1 \) and \( W_2 \) respectively, we claim that \( \beta \cup \gamma \) spans \( W_1 + W_2 \). Indeed, given an arbitrary vector \( x + y \in W_1 + W_2 \) with \( x \in W_1 \) and \( y \in W_2 \), we find scalars so that \( x = \sum c_i v_i \) and \( y = \sum d_i w_i \). Then
\[
x + y = \left( c_1 v_1 + \cdots + c_m v_m \right) + \left( d_1 w_1 + \cdots + d_n w_n \right) \in \text{span}(\beta \cup \gamma).
\]
Thus the size of a basis for \( W_1 + W_2 \) must be at most the size of \( \beta \cup \gamma \), and we have
\[
\dim(W_1 + W_2) \leq |\beta \cup \gamma| \leq |\beta| + |\gamma| = \dim(W_1) + \dim(W_2).
\]
\( \Rightarrow \) Assume that \( V = W_1 \oplus \cdots \oplus W_k \). By theorem 5.10d, if we take an ordered basis \( \beta_i \) for each space \( W_i \), then \( \beta = \beta_1 \cup \cdots \cup \beta_k \) is a basis for \( V \). This gives
\[
|\beta| = \dim(V) \leq \sum_{i=1}^{k} \dim(W_i) = \sum_{i=1}^{k} |\beta_i|.
\]
Thus it suffices to prove that $|\beta| = \sum |\beta_i|$. Suppose that $v \in \beta_1 \cap \beta_2$ and note that we can write the zero vector nontrivially as $0 = v + (-v) + 0 + \cdots + 0 \in W_1 + W_2 + \cdots + W_k$.

This contradicts our assumption that the sum $\sum W_i$ is direct. Hence $\beta_1 \cap \beta_2 = \emptyset$; the same argument shows that $\beta_i \cap \beta_j = \emptyset$ for any $i \neq j$. Since the sets are pairwise disjoint, we have $|\beta| = \sum |\beta_i|$, as desired.

($\Leftarrow$) Now we assume that $\dim(V) = \sum_{i=1}^{k} \dim(W_i)$.

For each $W_i$ find an ordered basis $\beta_i$ and consider the set $\beta = \beta_1 \cup \cdots \cup \beta_k$; we do not know a priori that the bases $\beta_i$ are pairwise disjoint, so we only have an inequality:

$$|\beta| \leq \sum_{i=1}^{k} |\beta_i| = \sum_{i=1}^{k} \dim(W_i).$$

(1)

We claim that $\beta$ spans $V$. Indeed, given $x \in V = \sum W_i$ we can write

$$x = w_1 + \cdots + w_k$$

with $w_i \in W_i$ for each $i$. Each $w_i$ can be written as a linear combination of the elements of $\beta_i$, so $x$ can be written as a linear combination of the elements of $\beta$ — we could use explicit notation for this argument, but it would greatly clutter the simple idea at work. Since $\beta$ spans $V$, it must have size at least $\dim V$; combining this with equation (1) gives

$$\dim(V) \leq |\beta| \leq \sum_{i=1}^{k} \dim(W_i) = \dim(V).$$

We deduce that $|\beta| = \dim(V)$, and since $\beta$ spans $V$ we further conclude that $\beta$ is a basis of $V$. By theorem 5.10d, $V$ is the direct sum of the spaces $W_i$.

**Homework 3**

5.4.1 (a) False. Any operator $T : V \to V$ has $T(V) \subseteq V$.

(b) True. The roots of the characteristic polynomial are eigenvalues. An eigenvalue for $T_W$ is an eigenvalue for $T$, so the roots of the characteristic polynomial of $T_W$ are among those of the polynomial for $T$.

(c) False. We could have, for example, $v' = 2v$.

(d) False. The $T$–cyclic subspace generated by $v$ might not include $v$ in its span. For example, take $T = 0$ on $\mathbb{R}^2$ and any vector $v \neq 0$. The $T$–cyclic subspace generated by $v$ is the span of $v$, but the $T$–cyclic subspace of $Tv$ is trivial.

(e) True. This follows from the Cayley–Hamilton theorem.

(f) True. The matrix on displayed on page 316 has characteristic polynomial $(-1)^n(a_0 + a_1 t + \cdots + a_{k-1} t^{k-1} + t^k)$.

(g) True. Take a basis $\beta_i$ of each $T$–invariant subspace $W_i$ and consider the basis $\cup_i \beta_i$ for the whole space.

5.4.3 (a) Since $T : V \to V$ is linear, clearly $T(\{0\}) = \{0\}$ and $T(V) \subseteq V$. 
(b) Given \( x \in N(T) \) we have \( Tx = 0 \in N(T) \). Given \( y \in R(T) \) we have \( Ty \in R(T) \) by definition of the range.

(c) Given \( v \in E_\lambda \), we have \( Tv = \lambda v \in E_\lambda \).

5.4.18 (a) Given any \( x \in W \), we can find scalars \( c_0, \ldots, c_k \) so that \( x = c_0v + c_1Tv + \cdots + c_kT^kv \). Then

\[
Tx = c_0Tv + c_1T^2v + \cdots + c_kT^{k+1}v,
\]

which is also a linear combination of vectors of the form \( T^jv \). Hence \( Tx \in W \).

(b) Let \( S \) be a \( T \)-invariant subspace of \( V \) containing \( v \). An inductive argument shows that every iterate \( T^jv \) is in \( S \); indeed, \( T^0v = v \in S \) and whenever \( T^jv \in S \) we must also have \( T^{j+1}v = T(T^jv) \in S \) by virtue of \( T \)-invariance. Since \( S \) is a space, it must contain the span of the iterates \( T^jv \), which is precisely the space \( W \).

5.4.21 Let \( V \) be a two-dimensional space and \( T : V \rightarrow V \). There are two possibilities that can occur: either \( \{x, Tx\} \) is linearly independent for some vector \( x \in V \), or \( \{x, Tx\} \) is always linearly dependent. In the former situation we’d have that \( V \) is the \( T \)-cyclic space generated by \( x \), so we’d be done. Hence we assume that \( \{x, Tx\} \) is linearly dependent for every vector \( x \in V \). We will show in this case that \( T = cI \) for some scalar \( c \); that is, for some fixed scalar \( c \) every vector \( x \in V \) satisfies \( Tx = cx \). This is clearly true for \( x = 0 \), so from here on we only consider nonzero vectors \( x \).

Given a nonzero vector \( x \in V \), the linear dependence of the set \( \{x, Tx\} \) means we can find scalars \( c_1 \) and \( c_2 \) — not both zero — so that

\[
c_1x + c_2Tx = 0.
\]

If \( c_2 = 0 \) then \( c_1 \neq 0 \) and we must have \( x = 0 \). Since we wanted to exclude this case, we cannot have \( c_2 = 0 \). Dividing by \( c_2 \), we see that

\[
Tx = -(c_1/c_2)x.
\]

Rewriting \( -c_1/c_2 = c \), we find that every vector \( x \in V \) satisfies an eigenvalue equation \( Tx = cx \). We do not yet know, however, that the same scalar \( c \) works for all vectors.

Consider nonzero vectors \( x, y \in V \) and find scalars \( a, b \) so that \( Tx = ax \) and \( Ty = by \). If we show that \( a = b \), then we’re done. To this end, we can assume that \( x \) and \( y \) are not scalar multiples of each other; that is, assume that \( \{x, y\} \) is linearly independent. The vector \( x + y \) also satisfies an eigenvalue equation of the form \( T(x + y) = c(x + y) \), but linearity yields

\[
c(x + y) = T(x + y) = Tx + Ty = ax + by,
\]

so that \((c - a)x + (c - b)y = 0\). The linear independence of \( \{x, y\} \) implies that \( a = b \), as desired.

6.1.10 Assuming that \( x \) and \( y \) are orthogonal vectors in an inner product space \( V \), we have

\[
\|x + y\|^2 = \langle x + y, x + y \rangle = \langle x, x \rangle + \langle x, y \rangle + \langle y, x \rangle + \langle y, y \rangle = \langle x, x \rangle + \langle y, y \rangle = \|x\|^2 + \|y\|^2.
\]

The standard inner product in \( \mathbb{R}^2 \) yields the Pythagorean theorem.

6.1.17 Let \( T \) be a linear operator on an inner product space \( V \) satisfying \( \|Tx\| = \|x\| \) for all \( x \in V \) (such an operator is typically called an isometry). If \( x \in V \) is chosen so that \( Tx = 0 \), then \( \|x\| = \|Tx\| = 0 \), so \( x = 0 \) as well. That is, the null space of \( T \) is trivial, so \( T \) is injective.

Homework 4

6.2.6 Given a vector \( x \notin W \) we can use theorem 6.6 to write \( x = w + z \), with \( w \in W \) and \( z \in W^\perp \). If \( z \) were zero then we would have \( x = w \in W \), so \( z \neq 0 \). Notice that \( z \in W^\perp \) and

\[
\langle x, z \rangle = \langle w + z, z \rangle = \langle w, z \rangle + \langle z, z \rangle = 0 + \|z\|^2 \neq 0.
\]
6.2.11 Given \( i, j \in \{1, 2, \ldots, n\} \) the component formula for matrix multiplication gives

\[
(AA^*)_{ij} = \sum_{k=1}^{n} A_{ik}(A^*)_{kj} = \sum_{k=1}^{n} A_{ik}A_{jk}.
\]

This last expression is the inner product of rows \( i \) and \( j \); thus if the rows of \( AA^* \) are orthonormal, we have \((AA^*)_{ij} = 0\) when \( i \neq j \) and \((AA^*)_{ii} = 1\). That is, \( AA^* = I \). Conversely if \( AA^* = I \) then the rows are orthonormal by the same argument in reverse.

6.2.13 (a) Assume that \( S_0 \subseteq S \) and let \( y \in S^\perp \). Given an arbitrary \( x \in S_0 \) we have \( \langle x, y \rangle = 0 \). Thus \( y \in S_0^\perp \) and we conclude \( S^\perp \subseteq S_0^\perp \).

(b) Given \( x \in S \) note that \( x \) is orthogonal to every element of \( S^\perp \). But then \( x \in (S^\perp)^\perp \). Since the orthogonal complement of any set is a space, this implies \( \text{span}(S) \subseteq (S^\perp)^\perp \).

(c) It remains to show that \( (W^\perp)^\perp \subseteq W \). Suppose that \( x \notin W \); by problem 6.2.6 there exists \( y \in W^\perp \) with \( \langle x, y \rangle \neq 0 \). That is, \( x \) is not orthogonal to some element of \( W^\perp \) and \( x \notin (W^\perp)^\perp \).

(d) By theorem 6.6 any vector in \( V \) can be written as a sum of the form \( w + z \), with \( w \in W \) and \( z \in W^\perp \). Hence \( V = W + W^\perp \). Suppose that \( y \in W \cap W^\perp \). Since \( y \in W^\perp \), it must be orthogonal to every element of \( W \); in particular \( \langle y, y \rangle = 0 \). This implies \( y = 0 \), so \( W \cap W^\perp = \{0\} \). We conclude that \( V = W \oplus W^\perp \).

6.3.7 Consider \( \mathbb{R}^2 \) with the standard inner product and standard basis \( \{e_1, e_2\} \). Define the linear operator \( T : \mathbb{R}^2 \to \mathbb{R}^2 \) by \( T(e_1) = 0 \) and \( T(e_2) = e_1 \). Notice that \( e_2 \notin N(T) \); we'll show that \( e_2 \in N(T^*) \). We compute \( \langle e_1, T^*e_2 \rangle = \langle Te_1, e_2 \rangle = 0 \) and \( \langle e_2, T^*e_2 \rangle = \langle Te_2, e_2 \rangle = \langle e_1, e_2 \rangle = 0 \). Since \( T^*e_2 \) is orthogonal to both vectors in the basis, \( T^*e_2 = 0 \). Thus \( N(T) \neq N(T^*) \).

6.3.9 Let \( x \in V \) be arbitrary; we want to show that \( T^*x - Tx = 0 \). Since \( V = W \oplus W^\perp \) by assumption, we can write \( T^*x - Tx = w + z \) with \( w \in W \) and \( z \in W^\perp \). Note that

\[
\|T^*x - Tx\|^2 = \langle T^*x - Tx, T^*x - Tx \rangle = \langle T^*x - Tx, w \rangle + \langle T^*x - Tx, z \rangle,
\]

so we need only show that each inner product on the right vanishes. For the first product, we make two observations: \( w \in W \) implies \( Tw = w \) and \( x - Tx \in W^\perp \). With these in mind we have

\[
\langle T^*x - Tx, w \rangle = \langle T^*x, w \rangle - \langle Tx, w \rangle = \langle x, Tw \rangle - \langle Tx, w \rangle = \langle x, w \rangle - \langle Tx, w \rangle = \langle x - Tx, w \rangle = 0.
\]

For the second product, note that since \( z \in W^\perp \) we have both \( Tz = 0 \) and \( \langle Tx, z \rangle = 0 \). This gives

\[
\langle T^*x - Tx, z \rangle = \langle T^*x, z \rangle - \langle Tx, z \rangle = \langle x, Tz \rangle = 0.
\]

We conclude that \( T^*x - Tx = 0 \); as \( x \) was arbitrary, \( T = T^* \).

6.3.12 (a) Let \( x \in R(T^*)^\perp \). Then for any vector \( y \in V \) we have \( \langle x, T^*y \rangle = 0 \). Choosing \( y = Tx \) gives

\[
0 = \langle x, T^*Tx \rangle = \langle Tx, Tx \rangle = \|Tx\|^2,
\]

so \( Tx = 0 \). That is, \( x \in N(T) \) and \( R(T^*)^\perp \subseteq N(T) \).

Now suppose that \( x \in N(T) \) and \( T^*y \in R(T^*) \) is arbitrary. Then

\[
\langle x, T^*y \rangle = \langle Tx, y \rangle = 0,
\]

so \( x \in R(T^*)^\perp \). We conclude that \( N(T) \subseteq R(T^*)^\perp \), whence the spaces are equal.

(b) From exercise 6.2.13 we know that in a finite–dimensional space, \( (W^\perp)^\perp = W \) for any subspace \( W \). Using the previous part gives

\[
N(T)^\perp = (R(T^*)^\perp)^\perp = R(T^*),
\]

as desired.
Homework 5

6.4.1 (a) True.
(b) False. Consider the map $T : \mathbb{R}^2 \to \mathbb{R}^2$ with $T(e_1) = 0$ and $T(e_2) = e_1$. You can show that $T^*(e_1) = e_2$ and $T^*(e_2) = 0$. That is, $e_1$ is an eigenvector of $T$ with eigenvalue 0, but not for $T^*$.
(c) False. This only holds for orthonormal bases $\beta$.
(d) True.
(e) True.
(f) True.
(g) False; if the underlying field is real, self–adjointness is needed.
(h) True.

6.4.7 (a) Let $u,v \in W$. Then

$$\langle T_W u, v \rangle_W = \langle Tu, v \rangle = \langle u, Tv \rangle = \langle u, T_W v \rangle_W.$$ 

Hence $T_W$ is self–adjoint.
(b) Let $v \in W^\perp$. Given $u \in W$ we have

$$\langle T^* v, u \rangle = \langle v, Tu \rangle = 0,$$

since $Tu \in W$. Thus $T^* v \in W^\perp$ and $W^\perp$ is $T^*$–invariant.
(c) Let $u,v \in W$. Then

$$\langle u, (T_W)^* v \rangle_W = \langle T_W u, v \rangle_W = \langle Tu, v \rangle = \langle u, T^* v \rangle = \langle u, (T^*_W)v \rangle_W.$$ 

Thus $[(T_W)^* - (T^*_W)]v$ an element of $W \cap W^\perp = \{0\}$.
(d) Using the previous part, $T_W(T_W)^* = T_W(T^*_W) = (TT^*_W)$. Similarly $(T_W)^*T_W = (T^*T)_W$. The result follows from the assumption that $TT^* = T^*T$.

6.4.11 (a) Let $x \in V$ and note that

$$\langle Tx, x \rangle = \langle x, Tx \rangle = \overline{\langle Tx, x \rangle}.$$ 

Hence $\langle Tx, x \rangle$ equals its complex conjugate and is real.
(b) Let $x,y \in V$ and note that

$$0 = \langle T(x+y), x+y \rangle$$
$$= \langle Tx, x \rangle + \langle Tx, y \rangle + \langle Ty, x \rangle + \langle Ty, y \rangle$$
$$= \langle Tx, y \rangle + \langle y, Tx \rangle$$
$$= 2\text{Re}(Tx,y),$$

so $\langle Tx, x \rangle$ is purely imaginary. However we can also expand

$$0 = \langle T(x+iy), x+iy \rangle$$
$$= \langle Tx, x \rangle + \langle Tx, iy \rangle + \langle Ty, x \rangle + \langle Ty, iy \rangle$$
$$= -i\langle Tx, y \rangle + i\langle y, Tx \rangle$$
$$= 2\text{Im}(Tx,y),$$

whence $\langle Tx, y \rangle$ has zero imaginary part as well. Thus $\langle Tx, y \rangle = 0$ for any $x,y$; choosing $y = Tx$ shows $Tx = 0$. As $x$ was arbitrary, $T = 0$. 
(c) Let \( x \in V \) and compute
\[
\langle (T - T^*) x, x \rangle = \langle Tx, x \rangle - \langle T^* x, x \rangle = \langle Tx, x \rangle - \langle x, Tx \rangle.
\]
Since \( \langle Tx, x \rangle \in \mathbb{R} \), this last expression is zero. By the previous part, knowing that \( \langle (T - T^*) x, x \rangle = 0 \) for all \( x \) implies that \( T = T^* \).

6.5.6 Let \( f, g \in V \) and write
\[
\langle Tf, g \rangle = \int_0^1 (Tf)(t \overline{g(t)}) \, dt = \int_0^1 h(t)f(t)g(t) \, dt = \int_0^1 f(t)h(t)g(t) \, dt = \langle f, hg \rangle.
\]
From this we see that \( T^* g = hg \). Then \( TT^* f = T^* Tf = |h|^2 f \); this is the identity if and only if \( |h| = 1 \).

6.5.7 Choose an orthonormal ordered basis so that
\[
[T]_\beta = \begin{pmatrix}
\lambda_1 & 0 & \cdots & 0 \\
0 & \lambda_2 & \cdots & 0 \\
0 & 0 & \cdots & \lambda_n
\end{pmatrix}.
\]
Since \( T \) is unitary, each \( \lambda_k \) has modulus 1 and hence has a square root of modulus 1 (for those unfamiliar with complex arithmetic, a number has absolute value 1 if and only if it can be written as \( e^{it} \) with \( t \in \mathbb{R} \); then \( e^{it/2} \) has modulus one and squares to \( e^{it} \)). Defining \( U \) as
\[
[U]_\beta = \begin{pmatrix}
\sqrt{\lambda_1} & 0 & \cdots & 0 \\
0 & \sqrt{\lambda_2} & \cdots & 0 \\
0 & 0 & \cdots & \sqrt{\lambda_n}
\end{pmatrix}
\]
gives the required square root of \( T \).

6.5.10 We need the fact that similar matrices have equal traces. With this in hand, we can diagonalize \( A \); in diagonal form the nonzero entries are the eigenvalues and clearly \( \text{tr}(A) \) is their sum.

Furthermore, having diagonalized \( A \) we have that
\[
PAP^* = \begin{pmatrix}
\lambda_1 & 0 & \cdots & 0 \\
0 & \lambda_2 & \cdots & 0 \\
0 & 0 & \cdots & \lambda_n
\end{pmatrix} \quad \text{and} \quad PA^*P = \begin{pmatrix}
\lambda_1 & 0 & \cdots & 0 \\
0 & \lambda_2 & \cdots & 0 \\
0 & 0 & \cdots & \lambda_n
\end{pmatrix}.
\]
From this it follows that
\[
PA^*AP = \begin{pmatrix}
|\lambda_1|^2 & 0 & \cdots & 0 \\
0 & |\lambda_2|^2 & \cdots & 0 \\
0 & 0 & \cdots & |\lambda_n|^2
\end{pmatrix}.
\]
Clearly then \( \text{tr}(A^* A) = \sum |\lambda_k|^2 \).
Homework 6

6.6.1 (a) False. Only orthogonal projections are self-adjoint.
(b) True.
(c) True. (The Spectral Theorem)
(d) False. Only true for orthogonal projections.
(e) False. Most projections aren’t invertible, let alone unitary.

6.6.4 Since $W$ is finite dimensional, we have that $V = W \oplus W^\perp$. By definition of $T$, whenever $x \in W$ and $y \in W^\perp$, we have $T(x + y) = x$. But then $(I - T)(x + y) = y$. The finite dimensionality of $W$ implies that $(W^\perp)^\perp = W$, so $I - T$ is an orthogonal projection derived from the direct sum $V = W^\perp \oplus (W^\perp)^\perp$.

6.6.6 We’ve seen many times that for any projection $T$, the space decomposes as $V = R(T) \oplus N(T)$. Assume that $T$ is normal; to show that $T$ is an orthogonal projection, we need only show that $N(T) = R(T)^\perp$.

Let $x \in N(T)$ and consider arbitrary $y \in R(T)$. Then $y = Ty$ and

$$\langle x, y \rangle = \langle x, Ty \rangle = \langle T^*x, y \rangle = 0,$$

since $Tx = 0x \Leftrightarrow T^*x = 0x = 0$.

Next assume that $v \in R(T)^\perp$. Then

$$\langle Tv, Tv \rangle = \langle v, T^*Tv \rangle = \langle v, TT^*v \rangle = 0,$$

since we’ve previously shown that $R(T)^\perp = N(T^*)$. Hence $Tv = 0$ and $v \in N(T)$.

Homework 7

6.7.13 Assume that $A$ is positive semidefinite. The eigenvalues of $A^*A$ are $\sigma_i^2$, where $\sigma_i$ denotes the singular values of $A$. Part of the definition of positive definiteness requires $A$ to be self-adjoint, so $A^2$ has eigenvalues $\sigma_i^2$. If $\lambda_i$ denotes an eigenvalue of $A$, we know that $\lambda_i^2$ is an eigenvalue of $A^2$; after rearranging the lists, we have $\lambda_i^2 = \sigma_i^2$ for each $i$. Finally, the eigenvalues and singular values of a positive semidefinite matrix are nonnegative, so we conclude that $\sigma_i = \lambda_i$ for each $i$.

6.7.15 (a) $(\Rightarrow)$ Assume that $A$ is normal, and consider the polar decomposition $A = WP$ with $W$ unitary and $P$ positive semidefinite. Then


By definition we assume that positive operators are self-adjoint, so $P^* = P$. Furthermore, $WW^* = W^*W = I$, so we have

$$P^2 = WP^2W^*.$$

Multiplying on the right by $W$ gives the result.

$(\Leftarrow)$ Assume that $WP^2 = P^2W$ and note that

$$A^*AW = P^*W^*WPW = P^2W = WP^2 = WP^2W^*W = AA^*W.$$

Multiplying by $W^*$ on the right gives $AA^* = A^*A$.

(b) $(\Rightarrow)$ Assume that $A$ is normal. By the previous part, $WP^2 = P^2W$, which rearranges into

$$P^2 = W^*P^2W = W^*P(WW^*)PW = (W^*PW)^2.$$

Since $W^*PW$ is unitarily equivalent to $P$, it is also positive semidefinite. Positive semidefinite operators admit square roots, so the above equation becomes $P = W^*PW$, which is equivalent to $WP = PW$.

$(\Leftarrow)$ Assume that $WP = PW$. Then $WP^2 = PWP = P^2W$, so $A$ is normal by the previous part.
6.8.7  (a) Given $H \in B(W)$, the domain of $\hat{T}H$ is certainly $V \times V$. For bilinearity, we check the first argument:

\[
\hat{T}H(ax + y, z) = H(T(ax + y), Tz) = H(aTx + Ty, Tz) \\
= aH(Tx, Tz) + H(Ty, Tz) \\
= a\hat{T}H(x, z) + \hat{T}H(y, z)
\]

holds for any $x, y, z \in V$ and scalar $a$. Hence $\hat{T}H$ is linear in the first argument; the same reasoning shows $\hat{T}H$ to be linear in its second argument, so we conclude that $\hat{T}H \in B(V)$.

(b) This is similar to the previous part. Given bilinear forms $H, J \in B(W)$ and a scalar $a$, we have

\[
\hat{T}(aH + J)(x, y) = (aH + J)(Tx, Ty) = aH(Tx, Ty) + J(Tx, Ty) \\
= a\hat{T}H(x, y) + \hat{T}J(x, y)
\]

for any $x, y \in V$. Hence $\hat{T}$ is linear.

(c) We could show that $\hat{T}$ is bijective, but instead we can construct an obvious inverse map. Note that

\[
\hat{T}\hat{T}^{-1}H(x, y) = H(TT^{-1}x, TT^{-1}y) = H(x, y) \\
\hat{T}^{-1}\hat{T}H(x, y) = H(T^{-1}Tx, T^{-1}Ty) = H(x, y),
\]

so $\hat{T}^{-1}$ is an inverse map of $\hat{T}$, proving $\hat{T}$ to be invertible.