Homework #9 Solutions

Thayer Anderson, Davis Lazowski, Handong Park, Rohil Prasad
Eric Peterson

1 For submission to Thayer Anderson

Problem 1.1. Suppose for $f\colon V \to V$ that $v \in V$ and $m \in \mathbb{N}$ satisfy $f^m(v) = 0$ but $f^{m-1}(v) \neq 0$. Prove that the following list is linearly independent:

$$v, fv, f^2v, \dots, f^{m-1}v$$

Solution. Suppose for the sake of contradiction that the given list is linearly dependent. Then there exist some $a_0, \dots, a_{m-1}$ not all 0 such that

$$a_0 v + a_1 f(v) + \cdots + a_{m-1} f^{m-1}(v) = 0$$

Let k be minimal such that the $a_k \neq 0$. Then consider

$$\begin{aligned} 0 = f^{m-k-1}(0) &= f^{m-k-1}\left(a_k f^k(v) + \cdots + a_{m-1} f^{m-1}(v)\right) \\ &= a_k f^{m-1}(v) + a_{k+1} f^m(v) + \cdots + a_{m-1} f^{m+k-1}(v) \\ &= a_k f^{m-1}(v) \end{aligned}$$

Since $f^{m-1}(v) \neq 0$ it follows that $a_k = 0$. This is a contradiction and therefore the list must be linearly independent. (TA)

Problem 1.2. Suppose $f\colon V \to W$ is a linear function between inner product spaces. Show that f^*f is a positive operator on V and ff^* is a positive operator on W.

Solution. To prove that f^*f is a positive operator we will prove that f^*f is self-adjoint and satisfies $\langle fv, v\rangle \geq 0$ for all $v \in V$. We prove the the second criteria first. Consider

$$\langle f^*fv, v\rangle = \langle fv, fv\rangle$$

This equality comes from Problem 2.2 on Homework 8. It follows that $\langle fv, fv\rangle \geq 0$ by positive-definiteness. Moreover, $\langle fv, fv\rangle \in \mathbb{R}$ - we will need this fact in a moment. To prove that f^*f is self-adjoint, let g be its adjoint and consider

$$\langle (f^*f - g)v, v\rangle$$

for arbitary $v \in V$. We have

$$\langle (f^*f - g)v, v\rangle = \langle f^*f, v\rangle - \langle gv, v\rangle$$

Then applying the definition of the adjoint:

$$\langle f^*f, v\rangle - \langle gv, v\rangle = \langle f^*f, v\rangle - \langle v, f^*fv\rangle$$

Since $\langle f^*fv, v\rangle \in \mathbb{R}$ it follows that $\langle f^*fv, v\rangle = \langle v, f^*fv\rangle$ which gives:

$$\langle (f^*f - g)v, v\rangle = 0$$

for all $v \in V$. This means that $f^*f = g$. This completes the proof. The proof of the other direction is virtually identical. Let me know if you want to see it. (TA)

Problem 1.3. Suppose $f\colon V \to V$ is a linear function on a finite-dimensional inner product space. Define a new pairing by

$$\langle u, v\rangle_f = \langle fu, v\rangle.$$

Show that $\langle -, -\rangle_f$ is an inner product on V if and only if f is an invertible positive operator.

Solution. First we prove the forwards direction. Suppose that f is an invertible positive operator. Then we need to prove that $\langle -, -\rangle_f$ is a conjugate-symmetric, positive-definite, sesqui-linear form. First we prove conjugate-symmetry. Consider

$$\langle u, v\rangle_f = \langle fu, v\rangle = \overline{\langle v, fu\rangle}$$

Since f is positive, it is self adjoint, so we have

$$\overline{\langle v, fu\rangle} = \overline{\langle fv, u\rangle} = \overline{\langle v, u\rangle_f}$$

This proves the desired result. Next we prove positive-definiteness. Suppose $v \in V$. Then

$$\langle v, v\rangle_f = \langle fv, v\rangle \geq 0$$

with inequality by the definition of positivity. Furthermore, f has a positive (and hence self-adjoint) square root. Call that square root g. Then we have

$$\langle fv, v\rangle = \langle g \circ gv, v\rangle = \langle gv, gv\rangle$$

Thus suppose $\langle v, v\rangle_f = 0$. It follows that $\langle gv, gv\rangle = 0$ and so $gv = 0$. But g is invertible so $v = 0$. This proves positive-definiteness. Now we prove linearity in the first argument. Suppose that $c \in K$ (the field of scalars - be it $\mathbb{C}$ or $\mathbb{R}$.) THen

$$\langle cu, v\rangle_f = \langle f(cu), v\rangle = \langle cf(u), v\rangle = c\langle fu, v\rangle = c\langle u, v\rangle_f$$

We see that $\langle -, -\rangle_f$ inherits linearity from f. This will remain the case for addition. This completes the first direction of the proof.

Suppose instead that f induces an inner product as given. First we consider

$$0 \leq \langle v, v\rangle_f = \langle fv, v\rangle$$

this gives us that $\langle fv, v\rangle \geq 0$ for all v, as required in the definition of positivity. Note also that $\langle fv, v\rangle$ is real and so

$$\langle fv, v\rangle = \overline{\langle fv, v\rangle}$$

Then consider the following, for arbitrary $v \in V$:

$$\begin{aligned}\langle (f - f^*)v, v\rangle &= \langle fv, v\rangle - \langle f^*v, v\rangle \\ &= \langle fv, v\rangle - \langle v, fv\rangle \\ &= \langle fv, v\rangle - \overline{\langle fv, v\rangle} = 0\end{aligned}$$

with the final inequality coming from our last lemma. Thus $f - f^* = 0$ and therefore f is self-adjoint. Combined with our other proof, we see that f is positive. This completes the proof. (TA)

2 For submission to Davis Lazowski

Problem 2.1. Suppose $f, g\colon V \to V$ are linear operators and suppose that fg is nilpotent. Prove that gf is also nilpotent.

Solution. If fg is nilpotent, then $(fg)^n = 0$ for some n. In this case,

$$0 = g0f = g(fg)^n f = (gf)^{n+1} \tag{DL}$$

Problem 2.2. Suppose that $f\colon V \to V$ is a linear function on an inner product space, and suppose that there exists an orthonormal basis $e_1, \ldots, e_n$ of V such that $\|fe_j\| = 1$ for each j. Either show that f must be an isometry or give a counterexample.

Solution. Fix i. Let $f(e_j) = e_i$ for all i.

Then

$$||f(e_j)|| = ||e_i|| = 1$$

But $\langle e_j, e_{j+1}\rangle = 0$, yet $\langle fe_j, fe_{j+1}\rangle = \langle e_i, e_i\rangle = 1$. (DL)

Problem 2.3. Fix vectors $u, x \in V$ in a finite-dimensional vector space V with $u \neq 0$. Consider an operator $f\colon V \to V$ defined by

$$f(v) = \langle v, u\rangle \cdot x$$

(as in one of the summands in singular value decomposition). Prove the following:

$$\sqrt{f^*f}(v) = \frac{\|x\|}{\|u\|}\langle v, u\rangle \cdot u.$$

Solution. We can write $f = g \circ \sqrt{f^*f}$ by SVD. Therefore,

$$g^{-1}(f(v)) = g^{-1}(\langle v, u\rangle\, x) = \langle v, u\rangle\, g^{-1}(x) = \sqrt{f^*f}(v)$$

Now, $g^{-1}(x) = \frac{||x||}{||u||}\tilde{u}$, for some $\tilde{u}$, because it is an isometry. We need to show that $\tilde{u} = u$. It's enough to show that $g^{-1}(x) \in \mathrm{span(u)}$, or, equivalently, that $g(u) \in \mathrm{span(x)}$. This is true by $f = g \circ \sqrt{f^*f}$ if $u \in \mathrm{im}\sqrt{\mathrm{f^*f}}$. Equivalently, we need to show that $u \in \mathrm{im\, f^*f}$. But by definition

$$\begin{aligned} \langle fv, w\rangle &= \langle v, f^*w\rangle \\ \langle fv, w\rangle &= \langle v, u\rangle\, \langle x, w\rangle \\ \implies f^* &= \langle w, x\rangle\, u \end{aligned}$$

So that $u \in \mathrm{im\, f^*f}$, therefore done. (DL)

Problem 2.4. Suppose $f\colon V \to V$ has singular value decomposition given by

$$f(v) = s_1\langle v, e_1\rangle f_1 + \cdots + s_n\langle v, e_n\rangle f_n$$

for $s_1, \ldots, s_n$ the singular values of f and $e_1, \ldots, e_n$ and $f_1, \ldots, f_n$ orthonormal bases of V. Prove the following effects:

1. $f^*(v) = s_1\langle v, f_1\rangle e_1 + \cdots + s_n\langle v, f_n\rangle e_n$.
2. $f^*f(v) = s_1^2\langle v, e_1\rangle e_1 + \cdots + s_n^2\langle v, e_n\rangle e_n$.
3. $\sqrt{f^*f}(v) = s_1\langle v, e_1\rangle e_1 + \cdots + s_n\langle v, e_n\rangle e_n$.
4. Lastly, suppose f is invertible. Show also $f^{-1}(v) = s_1^{-1}\langle v, f_1\rangle e_1 + \cdots + s_n^{-1}\langle v, f_n\rangle e_n$.

Solution. 1. In polar decomposition, $f = g \circ \sqrt{f^* f}$. Therefore, $f^* = \sqrt{f^* f}^* \circ g^* = \sqrt{f^* f} \circ g^{-1}$.

We can decompose v as

$$v = \sum_{j=1}^{n} \langle v, f_j \rangle f_j$$

Because $g(e_j) = f_j$, therefore $g^{-1}(f_j) = e_j$. Therefore

$$g^{-1} v = \sum_{j=1}^{n} \langle v, f_j \rangle e_j$$

Applying the singular values:

$$\sqrt{f^* f} \circ g^{-1} v = \sum_{j=1}^{n} s_j \langle v, f_j \rangle f_j$$

2. We have that

$$f^* f(v) = f^* (\sum_{j=1}^{n} s_j \langle v, e_j \rangle f_j)$$

So it's enough to show that $f^*(f_j) = s_j e_j$. But $f^*(f_j) = \sqrt{f^* f} \circ g^{-1}(f_j) = \sqrt{f^* f} e_j = s_j e_j$, as required.

3. $g^{-1} \circ f = \sqrt{f^* f}$. $g^{-1}(f_j) = e_j$. The rest comes out linearly, so

$$g^{-1}(f(v)) = g^{-1}(\sum_{j=1}^{n} s_j \langle v, e_j \rangle f_j) = \sum_{j=1}^{n} s_j \langle v, e_j \rangle e_j$$

As required.

4. We have that

$$f^{-1} = \sqrt{f^* f}^{-1} \circ g^{-1}$$

Then

$$\begin{aligned} f^{-1}(v) &= f^{-1}(\sum_{j=1}^{n} \langle v, f_j \rangle f_j) \\ &= \sqrt{f^* f}^{-1} (\sum_{j=1}^{n} \langle v, f_j \rangle e_j) \\ &= \sum_{j=1}^{n} s_j^{-1} \langle v, f_j \rangle e_j \end{aligned}$$

As required. (DL)

3 For submission to Handong Park

Problem 3.1. Prove or give a counterexample: the set of nilpotent operators on V is a vector subspace of $\mathcal{L}(V, V)$.

Solution. This statement is false, and here's an interesting counterexample that demonstrates why. Consider the following field, which we'll call $\mathbb{Z}/2\mathbb{Z}$: in other words, the integers with modular arithmetic mod 2. This is a field consisting of exactly 0 and 1, and is defined as follows:

$$0+0=0, 0+1=1, 1+1=0$$

and

$$0 \cdot 0=0, 0 \cdot 1=0, 1 \cdot 1=1$$

We can easily check that all the rules for a field to satisfy hold here. For instance, each non-zero element has a multiplicative inverse (only 1 is non-zero, and $1 \cdot 1 = 1$), and each element has an additive inverse ($0 + 0 = 0$, and $1 + 1 = 0$). The other rules are left as an exercise.
Now, taking K, the field of scalars, to be $\mathbb{Z}/2\mathbb{Z}$, consider the vector space $V = K^2$, and consider $\mathcal{L}(V, V)$ and the set $N \subset \mathcal{L}(V, V)$ of nilpotent operators on V. We have that

$$M_1 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \in N$$

and that

$$M_2 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \in N$$

since both of these matrices, when squared on (given our $\mathbb{Z}/2\mathbb{Z}$ field), give us 0.
However, consider

$$M_1 + M_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

We have that, by the modular arithmetic of the scalars in our field,[1]

$$(M_1 + M_2)^2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

But then, this means that $(M_1 + M_2)^c$, for any $c \in \mathbb{N}$, is not 0 - it's either $M_1 + M_2$ or the identity matrix I (depending on whether we have an even or odd power).
Thus, we find that N is not closed under addition in this case, so that $N \subset \mathcal{L}(V, V)$ is not necessarily always a vector subspace. (HP)

Problem 3.2. For $f\colon V \to V$ a linear operator on a finite-dimensional inner product space, write $s_{\min}$ for its smallest singular value and $s_{\max}$ for its largest singular value.

1. Prove the inequalities $s_{\min}\|v\| \le \|fv\| \le s_{\max}\|v\|$.
2. For any eigenvalue λ of f, show $s_{\min} \le |\lambda| \le s_{\max}$.
3. Let $g\colon V \to V$ be another linear operator with minimum and maximum singular values $t_{\min}$ and $t_{\max}$ respectively. Show that the maximum singular value of the composite gf is bounded above by $s_{\max} \cdot t_{\max}$ and that the maximum singular value of the sum $g + f$ is bounded above by $s_{\max} + t_{\max}$.

Solution. 1. We have that f has a singular value decomposition as follows: if $s_1, ..., s_n$ are the singular values for f,

$$f(v) = s_1\langle v, e_1\rangle f_1 + ... + s_n\langle v, e_n\rangle f_n$$

for any $v \in V$, $e_1, ..., e_n$ an orthonormal basis of V and $f_1, ..., f_n$ an orthonormal basis of V as well. Since the f_i's form an orthonormal basis, we have that

$$||f(v)||^2 = (s_1\langle v, e_1\rangle f_1)^2 + ... + (s_n\langle v, e_n\rangle f_n)^2$$

[1] ECP: Actually, this is true even over $\mathbb{R}$ or $\mathbb{C}$. For instance, since this matrix is symmetric, $(M_1+M_2)^2 = (M_1+M_2)^*(M_1+M_2)$ computes the matrix of inner products, which is the identity matrix. The mod–2 thing is also interesting, though: this is an important case of $M_1M_2 \ne M_2M_1$ (since otherwise $(M_1+M_2)^2$ "=" $M_1^2 + 2M_1M_2 + M_2^2$ would give 0).

But then, we know that if $s_{\max}$ is the largest of these singular values, we must have

$$||f(v)||^2 \leq s_{\max}^2(\langle v, e_1\rangle f_1 + ... + \langle v, e_n\rangle f_n)^2$$

so that

$$||f(v)||^2 \leq s_{\max}^2||v||^2$$

which proves that

$$||f(v)|| \leq s_{\max}||v||$$

Now that we have one inequality, we can prove the other inequality by almost the same process. We have

$$f(v) = s_1\langle v, e_1\rangle f_1 + ... + s_n\langle v, e_n\rangle f_n$$

for any $v \in V$, $e_1, ..., e_n$ an orthonormal basis of V and $f_1, ..., f_n$ an orthonormal basis of V as well. Since the f_i's form an orthonormal basis, we have that

$$||f(v)||^2 = (s_1\langle v, e_1\rangle f_1)^2 + ... + (s_n\langle v, e_n\rangle f_n)^2$$

This time, if $s_{\min}$ is the smallest singular value for f, we know that

$$||f(v)||^2 \geq s_{\min}^2(\langle v, e_1\rangle f_1 + ... + \langle v, e_n\rangle f_n)^2$$

so that

$$||f(v)||^2 \geq s_{\min}^2||v||^2$$

which proves that

$$||f(v)|| \geq s_{\min}||v||$$

and we are done.

2. We know that for any eigenvalue λ of f, we have

$$|\lambda| = ||\lambda v|| = ||\lambda||||v_1|| + ... + ||\lambda||||v_n||$$

where $v_1, ..., v_n$ are the coordinates of v in V, if V is an n-dimensional inner product space. Similarly, we just have that for a singular value s, we have

$$|s| = ||sv|| = s||v_1|| + ...s||v_n||$$

But then, if we have $s_{\min}$, we just have that each $s_{\min}||v_i|| \leq \lambda||v_i||$ each time, giving us that $s_{\min} \leq |\lambda$. And if we have $s_{\max}$, we just have that each $s_{\max}||v_i|| \geq \lambda||v_i||$ each time, giving us that $s_{\max} \geq \lambda$, as hoped.

3. Suppose we have an SVD of gf, then we have

$$gf(v) = r_1\langle v, e_1\rangle h_1 + ... + r_n\langle v, e_n\rangle h_n$$

for some orthonormal bases $e_1, ..., e_n$ and $h_1, ..., h_n$. Suppose without loss of generality that r_1 is $r_{\max}$, our maximum singular value for gf.
Then consider that

$$||gf(v)|| = ||g(f(v))|| = t_{\max} \cdot ||f(v)|| = t_{\max} \cdot s_{\max} \cdot ||v||$$

by what we proved above. If we plug in $v = e_1$, we just have

$$||r_{\max}|| = r_1||h_1|| = r_1 \leq t_{\max} \cdot s_{\max} \cdot 1$$

which proves the statement as desired. Similarly, if we have $g+f$, we have an SVD

$$(g+f)(v) = r_1\langle v, e_1\rangle h_1 + ... + r_n\langle v, e_n\rangle h_n$$

for some orthonormal bases $e_1, ..., e_n$ and $h_1, ..., h_n$. Suppose without loss of generality that r_1 is $r_{\max}$, our maximum singular value for $g + ff$.
Then we have by the triangle inequality that

$$||(g+f)(v)|| \leq ||g(v)|| + ||f(v)|| \leq t_{\max}||v|| + s_{\max}||v||$$

Then substituting e_1 as before gives us

$$r_{\max} \leq t_{\max} + s_{\max}$$

as desired.

(HP)

Problem 3.3. Suppose that V is a finite-dimensional inner product space, $f\colon V \to V$ is a linear operator, $g\colon V \to V$ is an isometry, and $h\colon V \to V$ a positive operator satisfying $f = g \circ h$. Show that $h = \sqrt{f^*f}$.

Solution. To begin, we consider that

$$f = g \circ h$$

However, suppose we take the adjoint of both sides. Then we have

$$f^* = (g \circ h)^* = h^* \circ g^*$$

Knowing that h is a positive operator, we know that h is also self-adjoint, meaning that

$$f^* = h \circ g^*$$

Now multiply both sides by f on the right to get

$$f^* \circ f = h \circ g^* \circ g \circ h = h \circ id_V \circ h = h^2$$

since g is an isometry. But then, we have

$$\sqrt{f^* \circ f} = h$$

as we hoped to prove. (HP)

4 For submission to Rohil Prasad

Problem 4.1. Suppose $f : V \to V$ is a linear operator on a finite-dimensional inner product space. Show that $\dim \mathrm{im} f$ equals the number of nonzero singular values of f.

Solution. Recall by the polar decomposition there exists an isometry s such that $f = s\sqrt{f^*f}$.

Since isometries are invertible, we find that s is injective. It follows that the kernel of $s\sqrt{f^*f}$ is equal to the kernel of $\sqrt{f^*f}$, so by rank nullity the image of f and of $\sqrt{f^*f}$ have the same dimension.

By the Spectral Theorem, we find that $\sqrt{f^*f}$ is diagonalizable. Therefore, it is immediate that the dimension of its image is equal to the total number of nonzero eigenvalues. Therefore, by the above reasoning the dimension of the image of f is equal to the total number of nonzero singular values. (RP)

Problem 4.2. Last week in Problem 4.2, you considered the inner product space of continuous functions on $[-\pi, \pi]$ as well as the subspace

$$U_n = \mathrm{span}\{1, \cos x, \cos 2x, \ldots, \cos nx, \sin x, \sin 2x, \ldots, \sin nx\}$$

and the double-derivative operator $D^2 : U_n \to U_n$. Show that $-D^2$ is a positive operator.

Solution. Recall last week we showed $D^* = -D$. Therefore, we find $(-D^2)^* = -(D^*)^2 = -(-D)^2 = -D^2$, so $-D^2$ is self-adjoint.

By Axler 7.35, to show $-D^2$ is positive it suffices to show there exists an operator R such that $-D^2 = R^*R$. Picking $R = D$ works, since $D^*D = (-D)D = -D^2$. (RP)

Problem 4.3. Define $f : \mathbb{C}^3 \to \mathbb{C}^3$ by $f(z_1, z_2, z_3) = (z_2, z_3, 0)$. Prove that f has no square-root.

Solution. We will first show for a linear operator g that if $\ker g^i = \ker g^{i+1}$ for some $i > 0$, then $\ker g^i = \ker g^j$ for all $j \geq i$.

It suffices to show by induction that given $\ker g^i = \ker g^{i+1}$, then $\ker g^{i+2}$ is equal to both as well. Since $\ker g^{i+1} \subset \ker g^{i+2}$, we need only show the reverse inclusion. Pick $v \in V$ such that $g^{i+2}(v) = 0$. Then we have $g^{i+1}(g(v)) = 0$, which by our above assumption implies $g^i(g(v)) = 0$, which implies $g^{i+1}(v) = 0$ as desired.

Now assume for the sake of contradiction that there exists g such that $g^2 = f$. Since $f(1, 0, 0) = 0$, we have $g(1, 0, 0) \in \ker(g)$. Therefore, we have that the dimension of the kernel of g is ≥ 1. Since the dimension of the kernel of $f = g^2$ is clearly 1, we must have the dimension of the kernel of g is 1 as well. However, this implies from the above reasoning that the dimension of the kernel of g^i is 1 for all i.

Since $g^6 = f^3 = 0$, we arrive at a contradiction and thus f cannot have a square root. (RP)