Homework #6 Solutions

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

1 For submission to Thayer Anderson

Problem 1.1. Let V be a finite-dimensional real vector space, $f\colon V \to V$ a linear map, and $\lambda \in \mathbb{R}$ some real number. Show that there exists a second real number $\alpha \in \mathbb{R}$ with $|\lambda - \alpha| < \frac{1}{1000}$ such that $f - \alpha$ is invertible.

Solution. The linear map $f - \alpha$ will be invertible if it has no eigenvalues of 0. This follows because if there are no zero eigenvalues then the kernel is trivial and thus f must be an isomorphism as it is between two equal dimensional vector spaces. Suppose that λ is an eigenvalue of $f - \alpha$. I claim that $\lambda + \alpha$ is an eigenvalue of f. This follows from the definitions. Suppose that v is an eigenvector of $f - \alpha$ with eigenvalue λ, then:

$$(f - \alpha)(v) = \lambda v$$
$$\Rightarrow f(v) = (\lambda + \alpha)v$$

If α is not an eigenvalue of f, then 0 is not an eigenvalue of $f - \alpha$. The map f has finitely many eigenvalues and for each eigenvalue of f, k, there are infinitely many numbers α satisfying $|\alpha - k| < \frac{1}{1000}$ thus there exists an α with the desired property. (TA)

Problem 1.2. Let A be an $(n \times n)$-matrix presenting a linear function $\mathbb{R}^n \to \mathbb{R}^n$.

1. Suppose that the sum of the entries in each *row* of A equals 1. Show that 1 is an eigenvalue of A.
2. Suppose that the sum of the entries in each *column* of A equals 1. Show that 1 is an eigenvalue of A.

Solution. 1. The matrix A can be represented as:

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}$$

with the following relations on the entries:

$$\begin{aligned} a_{11} + a_{12} + \cdots + a_{1n} &= 1 \\ a_{21} + a_{22} + \cdots + a_{2n} &= 1 \\ &\vdots \\ a_{n1} + a_{n2} + \cdots + a_{nn} &= 1 \end{aligned}$$

Let us consider the action of A on an arbitrary vector $v = (x_1, \dots, x_n)$ such that $Av = v$. Applying A to v we obtain the following expression:

$$Av = \begin{pmatrix} \sum_{i=1}^n a_{1i}x_i \\ \vdots \\ \sum_{i=1}^n a_{ni}x_i \end{pmatrix}$$

We see that if $x_i = 1$ for all $1 \le i \le n$ then

$$Av = \begin{pmatrix} \sum_{i=1}^n a_{1i} \\ \vdots \\ \sum_{i=1}^n a_{ni} \end{pmatrix} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$$

And $v = (1, \ldots, 1)$ so v is an eigenvector of eigenvalue 1. This completes the proof.

2. Given the matrix A with the sum across the columns equal to 1 and encoding a map f, we take the dual map f^* which is encoded by the matrix A^T. We note that the sum of the entries in each row of A^T is equal to 1. It follows that f^* has an eigenvalue of 1. This means that $f^* - 1$ is not invertible. Applying Axler's theorem about the row rank and the column rank, we see that $(f^* - 1)^* = f - 1$ is not invertible if and only if $f^* - 1$ is not invertible. Therefore 1 is an eigenvalue of f.

$$\begin{aligned} \left[f^*(v^*)\right](v) &= v^*(v) \\ \Rightarrow v^* f(v) &= v^*(v \\ v^*(f(v)) &= 1 \end{aligned}$$

It follows that $f(v) = v$. Thus f has an eigenvalue of 1.

(TA)

Problem 1.3. Let V be finite dimensional and let $f \colon V \to V$ a linear function. Suppose that $v \in V$ is a non-zero vector, and suppose that p is a nonzero polynomial with $p(f)(v) = 0$, and suppose that there are no polynomials of degree less than that of p which have this property. Show that every zero of p is an eigenvalue of f.

Solution. Suppose that k is a zero of p. Then using our machinery from class, we can factor p as follows:

$$p(x) = (x - k)q(x)$$

where q is a polynomial of degree $\deg p - 1$. Then consider $p(f)(v)$:

$$p(f)(v) = (f - k)q(f)(v) = 0$$

Since q has degree less than that of p, it follows that $q(f)(v) := w \neq 0$. Then w is a vector such that $(f - k)(w) = 0$, then w is an eigenvector of f with eigenvalue k and the proof is complete. (TA)

2 For submission to Davis Lazowski

Problem 2.1. Suppose $f \colon V \to V$ is invertible. Show that λ is an eigenvalue of f if and only if λ^{-1} is an eigenvalue of f^{-1}, and show that v is an eigenvector of f if and only if it is also an eigenvector of f^{-1}.

Solution. Suppose $fv = \lambda v$.

Then $v = f^{-1}(fv) = f^{-1}(\lambda v) = \lambda f^{-1}(v)$, so that by diving by λ then $\lambda^{-1} v = f^{-1} v$.

This proves one direction, for both statements. But our choice of f was totally arbitrary, and we could do the same things for f^{-1}, so by symmetry done. (DL)

Problem 2.2. Suppose $f \colon V \to V$ is a linear transformation with $\dim \operatorname{im} f = k$. Show that f has at most $(k + 1)$ distinct eigenvalues.

Solution. If $fv = \lambda v$, and $fw = \lambda' w$, with $\lambda \neq \lambda'$, then v and w are linearly independent. There are at most $\dim \operatorname{im} \mathrm{f} = \mathrm{k}$ linearly independent vectors in the image, plus 0. So there are at most $k + 1$ distinct eigenvalues. (DL)

Problem 2.3. Suppose V is a *complex* vector space, $f\colon V \to V$ a linear function, and p a complex polynomial. Show that $\alpha \in \mathbb{C}$ is an eigenvalue of $p(f)$ if and only if $\alpha = p(\lambda)$ for some eigenvalue λ of f. Then, show that this result fails if V is merely assumed to be a real vector space and p a real polynomial.

Solution. Suppose $\lambda v = fv$.

Then

$$p(f)v = \sum_{j=1}^{n} f^j v = \sum_{j=1}^{n} \lambda^j v = p(\lambda)v = \alpha v$$

Finishing the first direction.

In the other direction, suppose $p(f)v = \alpha v$. In particular, $p - \alpha$ is a polynomial, and $[p-\alpha](f)v = 0$.

By problem 1.3, if $p - \alpha$ is the lowest degree polynomial such that $[p-\alpha](f)v = 0$, then we're done, because every zero of $[p-\alpha]$ is an eigenvalue of f, and if $p(\lambda) - \alpha = 0$, then $p(\lambda) = \alpha$.

Otherwise, suppose there exists $g(f)v = 0, \deg g < \deg[p-\alpha]$, so that $\deg g$ is minimal. Then by the division algorithm $[p-\alpha] = hg + r$, with $\deg r < \deg g$. But

$$r(f)v = [p-\alpha](f)v - hg(f)v = 0$$

So by our assumption of minimality, $r = 0$. Therefore, $[p-\alpha] = hg$.

Therefore, all the zeroes of g are also zeroes of $[p-\alpha]$, therefore done.

An example of how this fails for the reals.

Let $f : V \to V, f(w) = w - v$, for some $v \in V$.

Then let $p = x^2$. $p(f)(v) = -v$, so that -1 is an eigenvalue of $p(f)$. But there is no λ such that $\lambda^2 = -1$ over the reals.

Precisely, this problem fails over the reals because some polynomials might not have zeroes over the reals, for example $x^2 + 1$.

Side note about extending problem 1.3 to infinite dimensions.

Problem 1.3 assumes finite dimensionality but can easily be extended to the infinite dimensional case. Let $\langle v \rangle$ denote the subspace of V generated by v. Then let $n = \deg p$. The space $\tilde{V} = \langle v \rangle + \langle f(v) \rangle + \cdots + \langle f^n(v) \rangle$ is finite dimensional.

Apply problem 1.3 to $\tilde{f}(v) : \tilde{V} \to \tilde{V}$, with $\tilde{f}(v) = f(v)$. Then because $p(\tilde{f})(v) = p(f)(v)$, and $\tilde{f}w = \lambda w \iff fw = \lambda w$, this proves the infinite dimensional case. (DL)

3 For submission to Handong Park

Problem 3.1. Let $p\colon V \to V$ satisfy $p \circ p = p$. Show that $V = \ker p \oplus \operatorname{im} p$.

Solution. First, we express an arbitrary $v \in V$ as $v = pv + (v - pv)$, which is the sum of a vector $pv \in \operatorname{im} p$ and $(v - pv) \in \ker p$, since

$$p(v - pv) = pv - ppv = pv - pv = 0.$$

This shows $V = \ker p + \operatorname{im} p$. To show that the sum is direct, we show that $\ker p \cap \operatorname{im} p = 0$. So, suppose $v \in \ker p \cap \operatorname{im} p$ satisfies $pv = 0$ and also $pw = v$ for some $w \in V$. Then $ppw = pw$ gives $pv = v$ and hence we have calculated $v = 0$. (ECP)

Problem 3.2. Suppose that $f\colon V \to V$ is a linear operator with $f \circ f = \mathrm{id}$, and suppose that -1 is *not* an eigenvalue of f. Show that $f = \mathrm{id}$.

Solution. If $f \circ f = \mathrm{id}$, then $f \circ f - \mathrm{id} = 0$ factors as $(f - \mathrm{id})(f + \mathrm{id}) = 0$. Since -1 is not an eigenvalue of f, the operator $f + \mathrm{id}$ is invertible, hence we get $f - \mathrm{id} = (f - \mathrm{id})(f + \mathrm{id})(f + \mathrm{id})^{-1} = 0(f + \mathrm{id})^{-1} = 0$. It follows that $f = \mathrm{id}$. (ECP)

Problem 3.3.

1. Suppose that a subspace $U \le V$ is invariant under a linear function $f\colon V \to V$. Show that U is also invariant under $p(f)$, where p is any polynomial.

2. Now suppose V is a complex vector space with dimension $1 < \dim V < \infty$. Show that for any particular linear map $f\colon V \to V$, there is a *proper* subspace

$$\{p(f) \mid p \text{ a polynomial}\} < \mathcal{L}(V, V).$$

Solution. 1. We need only show that for $u \in U$, we have $p(f)(u) \in U$. This follows by direct calculation, beginning with setting $p(z) = a_0 + a_1 z + \cdots + a_n z^n$. Then, we have

$$p(f)(u) = a_0 \cdot u + a_1 f(u) + \cdots + a_n f^{\circ n}(u).$$

Each of the terms on the right is a member of U, since U is invariant and closed under scalar multiplication, and hence the whole sum is in U since U is closed under sums.

2. The map f is guaranteed to have an eigenvector v_1, hence an invariant subspace U of dimension 1. By the first part, every operator expressable as $p(f)$ also has U as an invariant subspace. To show properness, we thus only need to exhibit an operator $g\colon V \to V$ which does not have U invariant. Extending the eigenvector v of f to a basis $(v_1, v_2, \ldots, v_n)$ of V, we take g to be the operator that swaps v_1 and v_2 and leaves the other basis elements undisturbed. (ECP)

4 For submission to Rohil Prasad

Problem 4.1. Let $f : V \to V$ be a linear operator. Prove that $f/\ker f$ is injective if and only if

$$(\ker f) \cap (\operatorname{im} f) = 0.$$

Solution. Note that by definition, $(f/\ker f)(v + \ker f) = f(v) + \ker f$.

Furthermore, injectivity of $f/\ker f$ is equivalent to showing that $\ker(f/\ker f) = 0$.

First, assume $(\ker f) \cap (\operatorname{im} f) = 0$. This implies that for any $v \in V$, $f(v) \notin \ker f$. For any $v + U$, we have $(f/\ker f)(v + \ker f) = f(v) + \ker f$. Since $f(v) \notin \ker f$, this affine set is nonzero in $V/\ker f$, so $v + U \notin \ker(f/\ker f)$. Since this is true for any $v + U$, we find that the kernel of this map is 0.

For the other direction, we prove the contrapositive statement. If $v \in (\ker f) \cap (\operatorname{im} f)$ then let v' be such that $f(v') = v$. It follows that $(f/\ker f)(v' + \ker f) = v + \ker f$. Since $v \in \ker f$, this is the zero element of $V/\ker f$, so $v' + \ker f \in \ker(f/\ker f)$. (RP)

Problem 4.2. Suppose that $\lambda_1, \ldots, \lambda_n$ is a list of distinct real numbers. Show that $e^{\lambda_1 x}, \ldots, e^{\lambda_n x}$ is a list of linearly independent functions $\mathbb{R} \to \mathbb{R}$. (Hint: find a linear operator on the space of functions $\mathbb{R} \to \mathbb{R}$ for which these are eigenvectors of distinct eigenvalues.)

Solution. Let D be the subspace of everywhere-differentiable functions $\mathbb{R} \to \mathbb{R}$. Note that $e^{\lambda_i x} \in D$ for every i, and furthermore if the $e^{\lambda_i x}$ are linearly independent as elements of D, they are linearly independent as elements of the whole space.

Now consider the differentiation operator $\frac{d}{dx} : D \to D$. By definition, this operator is linear. Furthermore, $\frac{d}{dx}(e^{\lambda_i x}) = \lambda_i e^{\lambda_i x}$. Since the λ_i are distinct, the $e^{\lambda_i x}$ are eigenvectors of $\frac{d}{dx}$ with distinct eigenvalues.

We complete the proof by inducting on n. For $n = 1$, it is clear that the set $\{e^{\lambda_1 x}\}$ is linearly independent.

Now assume that the statement holds for $n - 1$. For the sake of contradiction, assume that the $e^{\lambda_i x}$ are linearly dependent. Therefore, there exists constants c_i not all zero such that $\sum_{i=1}^n c_i e^{\lambda_i x} = 0$.

Assume without loss of generality that $c_1 \neq 0$ and set $c'_j = -c_j/c_1$. Then by moving terms around and dividing by c_1, we have $e^{\lambda_1 x} = \sum_{j=2}^n c'_j e^{\lambda_j x}$.

Multiplying by λ_1, we find $\lambda_1 e^{\lambda_1 x} = \sum_{j=2}^n \lambda_1 c'_j e^{\lambda_j x}$.

If we instead apply $\frac{d}{dx}$ to the expression, we find $\lambda_1 e^{\lambda_1 x} = \sum_{j=2}^n \lambda_j c'_j e^{\lambda_j x}$.

Subtracting these two identities, we get $\sum_{j=2}^n (\lambda_j - \lambda_1) c'_j e^{\lambda_j x} = 0$. However, by our inductive hypothesis the set of functions $\{e^{\lambda_2 x}, \ldots, e^{\lambda_n x}\}$ is linearly independent. Since $\lambda_j \neq \lambda_1$ for any $j \neq 1$, this implies that

c_j' is equal to 0 for every j. This in turn implies $c_j = 0$ for every $j > 1$, so from our original identity we have that $c_1 e^{\lambda_1 x} = 0$. This is clearly false, so we arrive at a contradiction and $\{e^{\lambda_1 x}, \ldots, e^{\lambda_n x}\}$ is linearly independent. (RP)

Problem 4.3. Let V be an arbitrary vector space and let $f : V \to V$ be a linear function. Consider the following three situations:

1. Every nonzero vector is an eigenvector of f.
2. The vector space V is finite dimensional of dimension n, and every subspace $U \leq V$ with $\dim U = n-1$ is invariant under f.
3. The vector space V is finite dimensional of dimension $n \geq 3$, and every subspace $U \leq V$ with $\dim U = 2$ is invariant under f.

In each case, show that f is a scalar multiple of the identity operator.

Solution. 1. Note that if V is one-dimensional, then we are done since every vector is a scalar multiple of another, so they will all have the same eigenvalues.

Now assuming V has dimension > 2, we can pick v_1, v_2 such that v_2 is not a scalar multiple of v_1. For the sake of contradiction, assume $f(v_1) = \lambda_1 v_1$ and $f(v_2) = \lambda_2 v_2$ with $\lambda_1 \neq \lambda_2$. Let λ be the eigenvalue of $v_1 + v_2$. Now we have that $f(v_1 + v_2) = \lambda(v_1 + v_2)$.

By linearity, the left-hand side evaluates to $\lambda_1 v_1 + \lambda_2 v_2$. Rearranging, we find that

$$(\lambda_1 - \lambda)v_1 = (\lambda - \lambda_2)v_2$$

Therefore, v_2 is a scalar multiple of v_1 and so we arrive at a contradiction.

2. Let $\{v_1, v_2, \ldots, v_n\}$ be a basis of V. For $1 \leq i \leq n$, let U_i be the span of all the basis vectors except for v_i.

By our assumption, each of the U_i are invariant under f. Therefore, we have an induced linear map $f_i : V/U_i \to V/U_i$ defined by $f_i(v + U_i) = f(v) + U_i$ for every i. However, by definition V/U_i is one-dimensional, so f_i is multiplication by some scalar λ_i.

Therefore, it follows that $f_i(v_i + U_i) = \lambda_i v_i + U_i$, so for every i we have $f(v_i) = \lambda_i v_i + u_i$ for some $u_i \in U_i$.

Observe that $v_i \in U_j$ for every $j \neq i$. Therefore, by invariance we must have $f(v_i) \in U_j$ for every $j \neq i$. In order for $f(v_i)$ to be in U_j, we require its coefficient in the basis $\{v_1, \ldots, v_n\}$ at v_j to be 0. Therefore, if we write out $u_i = \sum_{j \neq i} c_j v_j$, we must have $c_j = 0$. Taking this over every $j \neq i$, it follows that $u_i = 0$.

Now it remains to show that all the λ_i are equal. We will show $\lambda_1 = \lambda_2$ and the proof is analogous for all others. Then observe that the span of $\{v_1 + v_2, v_3, \ldots, v_n\}$ is a subspace W of dimension $n-1$ and is therefore invariant under f. Therefore, $f(v_1 + v_2) = \lambda_1 v_1 + \lambda_2 v_2 \in W$.

Therefore, there exist constants $c, c_3, c_4, \ldots, c_n$ such that $\lambda_1 v_1 + \lambda_2 v_2 = c(v_1 + v_2) + \sum_{i \geq 3} c_i v_i$. It follows that $(c - \lambda_1)v_1 + (c - \lambda_2)v_2 + \sum_{i \geq 3} c_i v_i = 0$. By linear independence of the v_i, all of these coefficients are 0 and so $\lambda_1 = c = \lambda_2$.

3. Here we use the fact that an intersection of invariant subspaces is itself an invariant subspace.

Pick a basis $\{v_1, v_2, \ldots, v_n\}$ of V. Let U_i be the span of v_i, v_{i+1} for $1 \leq i \leq n-1$ and let U_n be the span of v_n, v_1.

By our assumption, each of the U_i are invariant. Furthermore, $U_i \cap U_{i-1}$ is the span of v_i for $i > 1$, and $U_1 \cap U_n$ is the span of v_1. These subspaces are also all invariant under f and one-dimensional, so the restriction of f to the span of v_i is multiplication by λ_i.

By a similar argument to Part 2, we show $\lambda_1 = \lambda_2$ and claim that the other equalities are analogous.

Note that the span of $v_1 + v_2, v_3$ is an invariant subspace W of dimension 2. Therefore, we have $f(v_1 + v_2 + v_3) \in W$, so there exist constants c, d such that $f(v_1 + v_2 + v_3) = cv_1 + cv_2 + dv_3$. By definition, $f(v_1 + v_2 + v_3) = \lambda_1 v_1 + \lambda_2 v_2 + \lambda_3 v_3$. By linear independence of v_1, v_2, v_3 it follows that $d = \lambda_3$, and $\lambda_1 = c = \lambda_2$ as desired. (RP)