Homework #1 Solutions

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

1 For submission to Thayer Anderson

Problem 1.1. Consider the matrix

$$M = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 & 0 \\ 0 & 3 & 1 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 4 & 1 \end{pmatrix}$$

Compuate a square root of M.

Solution. We see that we can represent the matrix M as follows:

$$M = I + \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 4 & 0 \end{pmatrix}$$

Let N be the matrix above such that N is nilpotent and $M = N + I$. Then we do a Taylor series expansion on $\sqrt{1+x}$ and obtain a square root as described in Axler:

$$A = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ -3/4 & 3/2 & 1 & 0 & 0 \\ 3/8 & 3/8 & -1/2 & 1 & 0 \\ 15/16 & -3/4 & 1/2 & 2 & 1 \end{pmatrix}$$

with $A^2 = M$. (TA)

Problem 1.2. Prove or give a counterexample: for V a complex vector space, $f : V \to V$ a linear operator, and $n = \dim V < \infty$, the operator $f^{\circ n}$ is diagonalizable.

Solution. Let $V = \mathbb{R}^2$. Consider the matrix

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

We see that

$$A^2 = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$$

But this matrix is not diagonalizable. This can be proved as follows. From its upper triangular presentation, we see that its only eigenvalue is 1. Then suppose the vector $\begin{pmatrix} a \\ b \end{pmatrix}$ is an eigenvector of eigenvalue 1. Then we have

$$\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} a+b \\ b \end{pmatrix} = \begin{pmatrix} a \\ b \end{pmatrix}$$

Thus we obtain the relation that $b = 0$. This provides only one degree of freedom and so there can be at most one linearly independent eigenvector. Therefore A^2 is not diagonalizable. (TA)

Problem 1.3. Suppose $f : V \to V$ is a linear operator on a finite-dimensional complex vector space, and let $v \in V$ be a vector. Prove that there exists a unique monic polynomial p of smallest degree such that $p(f)(v) = 0$¿ Then, prove that this $p(z)$ divide sthe minimal polynomial of f.

Solution. Consider the minimal polynomial, m. We have $m(f) = 0$ and therefore $m(f)(v) = 0$. Thus there exists such a polynomial p' satisfying $p'(f)(v) = 0$. Take p to be the minimal degree monic polynomial (we can always clear the leading coefficient and the relation will remain true). We wis hto show uniqueness. Suppose that s is another monic polynomial of minimal degree that satisfies $s(f)(v) = 0$. Then consider:

$$0 = p(f)(v) - s(f)(v) = (p - s)(f)(v)$$

Thus the polynomial $p - s$ satisfies the condition that $(p - s)(f)(v) = 0$. Since p and s were monic, it follows that $p - s$ has degree strictly less than the degree of p or s. Thus $(p - s) = 0$ or we have a contradiction. It follows that p is unique.

Now to prove that p divides the minimal polynomial, we decompose as follows:

$$m(z) = p(z)m(z) + r(z)$$

for some polynomial q and some polynomial r with degree strictly less than the degree of p. We wish to show that r is equal to zero. Consider

$$0 = m(f)(v) = qp(f)(v) + r(f)(v) = q \circ 0(v) + r(f)(v) = r(f)(v)$$

Thus r satisfies $r(f)(v) = 0$. But if r is non-zero it has degree strictly less than that of p and this is a contradiction. (Note that we can always take r to be monic). (TA)

2 For submission to Davis Lazowski

Problem 2.1. Let $f\colon V \to V$ be a linear operator on a finite-dimensional vector space. Prove that $\ker f^{\circ m} = \ker f^{\circ(m+1)}$ if and only if $\operatorname{im} f^{\circ m} = \operatorname{im} f^{\circ(m+1)}$. Conclude that $\operatorname{im} f^{\circ N}$ is stable for $N \gg 0$.

Solution. If $\ker f^{\circ m} = \ker f^{\circ(m+1)}$, then $\dim\ker f^{\circ m} = \dim\ker f^{\circ(m+1)}$. Therefore, by rank-nullity,

$$\dim \mathrm{im}\, \mathrm{f}^{\circ \mathrm{m}} = \dim \mathrm{V} - \dim\ker \mathrm{f}^{\circ \mathrm{m}} = \dim \mathrm{V} - \dim\ker \mathrm{f}^{\circ(\mathrm{m}+1)} = \dim \mathrm{im}\, \mathrm{f}^{\circ(\mathrm{m}+1)}$$

Also,

$$\mathrm{im}\, \mathrm{f}^{\circ(\mathrm{m}+1)} = \mathrm{f}^{\circ(\mathrm{m}+1)}(\mathrm{V}) = \mathrm{f}^{\circ\mathrm{m}}(\mathrm{f}(\mathrm{V})) \subset \mathrm{f}^{\circ\mathrm{m}}(\mathrm{V}) = \mathrm{im}\, \mathrm{f}^{\circ\mathrm{m}}$$

Therefore, in this case we have two subspaces of equal dimension, one including the other, therefore they are equal. So $\mathrm{im}\, \mathrm{f}^{\circ(\mathrm{m}+1)} = \mathrm{im}\, \mathrm{f}^{\circ\mathrm{m}}$, as required.

If $\mathrm{im}\, \mathrm{f}^{\circ\mathrm{m}} = \mathrm{im}\, \mathrm{f}^{\circ(\mathrm{m}+1)}$, then by the symmetric argument $\dim\ker f^{\circ m} = \dim\ker f^{\circ(m+1)}$. Also, clearly $\ker f^{\circ m} \subset \ker f^{\circ(m+1)}$. So again $\ker f^{\circ m} = \ker f^{\circ(m+1)}$, therefore done. (DL)

Problem 2.2. Give an example of an operator $f\colon V \to V$ on a finite-dimensional *real* vector space such that 0 is the only eigenvalue of f but f is not nilpotent.

Solution. To construct such a function, I will first find a surjective linear operator which admits no eigenvalues. I will then tack on a nilpotent part.

First, let g be the two-dimensional rotation operator by $\frac{\pi}{2}$:

$$g\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} -b \\ a \end{pmatrix}$$

Over the reals, g has no eigenvalues. Specifically: the solutions to the equation $\begin{pmatrix} a \\ b \end{pmatrix} = \lambda \begin{pmatrix} -b \\ a \end{pmatrix}$ are precisely the solutions to $\lambda^2 = -1$, which has no solutions over the reals.

g is also surjective, because $g^4 = id$.

Now, let's tack on the nilpotent part. Define

$$f\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} -b \\ a \\ 0 \end{pmatrix}$$

Clearly, $\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ is an eigenvector with eigenvalue 0.

But $f^4 \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} a \\ b \\ 0 \end{pmatrix}$ so is not nilpotent. (DL)

Problem 2.3. Let $f\colon V \to V$ be an operator on a finite-dimensional complex vector space. Prove that there exist operators $D, N\colon V \to V$ such that D is diagonalizable, N is nilpotent, $f = D + N$, and they satisfy the commutation relation $DN = ND$.

Solution. Induct on the number of generalised eigenspaces of f. Recall that

$$f_{|G(\lambda_j,f)} - \lambda_j I_{|G(\lambda_j,f)}$$

is nilpotent.

Therefore, if there is exactly one generalised eigenspace, $f_{|G(\lambda_j,f)} - \lambda_j I_{|G(\lambda_j,f)} = f - \lambda_j I$. Then $(f - \lambda_j I)$ is nilpotent and $\lambda_j I$ is diagonal, so $(f - \lambda_j I) + \lambda_j I = f$, as required. Finally, I commutes with every matrix, so $DN = ND$.

Let it be true that for all g with $n-1$ generalised eigenspaces, then there are N nilpotent, D diagonlisable, such that $g = N + D$, and such that $DN = ND$.

Consider some f with n generalised eigenspaces. Choose the first. Then $f_{|G(\lambda_j,f)^\perp}$ has $n-1$ generalised eigenspaces. Therefore, applying the hypothesis, $f_{|G(\lambda_j,f)^\perp} = \tilde{N} + \tilde{D}$. Also by the hypothesis, $f_{|G(\lambda_j,f)} = \overline{N} + \overline{D}$.

Define the inclusion map for a subspace U of V to be $i_U : U \to V$, the map which sends $u \in U$ to $u \in V$. Define $p_U : V \to U$ to be the projection map onto U. For brevity, let $G(\lambda_j, f) = U$.

Then

$$\begin{aligned} f &= i_{U^\perp} \circ f_{|U^\perp} \circ p_{U^\perp} + i_U \circ f_{|U} \circ p_U \\ &= i_{U^\perp} \circ \tilde{D} \circ p_{U^\perp} + i_{U^\perp} \circ \tilde{N} \circ p_{U^\perp} + i_U \circ \overline{D} \circ p_U + i_U \circ \overline{N} \circ p_U \end{aligned}$$

Now,

$$D := i_{U^\perp} \circ \tilde{D} \circ p_{U^\perp} + i_U \circ \overline{D} \circ p_U$$

Is diagonalisable because we can choose an eigenbasis for $U^\perp$ and an eigenbasis for U. It is easy to verify that any of these vectors is still an eigenvector of D.

Also,

$$N := i_{U^\perp} \circ \tilde{N} \circ p_{U^\perp} + i_U \circ \overline{N} \circ p_U$$

is nilpotent because $N^j = i_{U^\perp} \circ \tilde{N}^j \circ p_{U^\perp} + i_U \circ \overline{N}^j \circ p + U$ and each of $\overline{N}, \tilde{N}$ are nilpotent.

Finally,

$$\begin{aligned} &DN \\ &= i_{U^\perp} \circ (\tilde{D}\tilde{N}) \circ p_{U^\perp} + i_U \circ (\overline{DN}) \circ p_U \\ &= i_{U^\perp} \circ (\tilde{N}\tilde{D}) \circ p_{U^\perp} + i_U \circ (\overline{ND}) \circ p_U \\ &= ND \end{aligned}$$

as required. Therefore done. (DL)

Problem 2.4. Suppose that $f\colon V \to V$ has characteristic polynomial

$$p(z) = 4 + 5z - 6z^2 - 7z^3 + 2z^4 + z^5.$$

Calculate the minimal polynomial of f. Prove that f^{-1} must exist and calculate its minimal polynomial as well.

Solution. This polynomial has five distinct zeroes. Therefore the characteristic polynomial is the minimal polynomial of f. Therefore V is of dimension five and f has five eigenvalues. Therefore f is diagonalisable, with nonzero eigenvalues, therefore invertible.

In general, let $(z-\alpha_1)(z-\alpha_2)...$ be a minimal polynomial. Then the minimal polynomial of the inverse is $(z-\alpha_1^{-1})(z-\alpha_2^{-1})....$

A linear multiple of this polynomial is $(\alpha_1\alpha_2....\alpha_n)(z-\alpha_1^{-1})(z-\alpha_2^{-1})...$, that is $(\alpha_1 z - 1)(\alpha_2 z - 1).....$

Observe this is the formula of the minimal polynomial, except with the coefficients swapped from 1 to z. This implies that, if the minimal polynomial of f has $\deg n$, then the coefficient of z^j in the minimal polynomial, up to a global multiple, is the coefficient of z^{n-j} in the minimal polynomial of f^{-1}.

So the minimal polynomial of the inverse is just some multiple of

$$4z^5 + 5z^4 - 6z^3 - 7z^2 + 2z + 1$$

And so in particular is

$$z^5 + \frac{5}{4}z^4 - \frac{3}{2}z^3 - \frac{7}{4}z^2 + \frac{1}{2}z + \frac{1}{4}$$ (DL)

3 For submission to Handong Park

Problem 3.1. Suppose A and B are block diagonal matrices of the form

$$A = \begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_m \end{pmatrix}, \qquad B = \begin{pmatrix} B_1 & & 0 \\ & \ddots & \\ 0 & & B_m \end{pmatrix},$$

where A_j and B_j are of size $n_j \times n_j$ for $j = 1, \ldots, m$. Show that AB is a block diagonal matrix of the form

$$AB = \begin{pmatrix} A_1B_1 & & 0 \\ & \ddots & \\ 0 & & A_mB_m \end{pmatrix}.$$

Solution. This is just an exercise in matrix multiplication - but we can prove it by looking at indices of the rows and columns in which the blocks inhabit.
To do so, suppose we have the corresponding blocks A_j and B_j. Let's say that they are located at rows $l, ..., l+n_j$ and columns $l, ..., l+n_j$, where $l \leq l+n_j$ and both are between 1 and n, where n is the total number of rows in each matrix.
By the definition of matrix multiplication, we have that for the entry $(AB)_{ij}$ in the product matrix, if A_{ij} is the entry in the i'th row and j'th column of A, and B_{ij} is the similar entry for B, then

$$(AB)_{ij} = \sum_{k=1}^{n} A_{ik}B_{kj}$$

First, consider an entry in the product, $(AB)_{ij}$, such that $i < l$ but $l \leq j \leq l+n_j$. In that case, we have

$$(AB)_{ij} = \sum_{k=1}^{n} A_{ik}B_{kj} = 0$$

From a geometric standpoint, looking at the matrices, we see that the $A_{ik} = 0$ precisely when the B_{kj} are in one of the blocks and the $B_{kj} = 0$ precisely when the A_{ik} are in one of the blocks. The same exact argument holds for $(AB)_{ij}$ such that $i > l+n_j$. It also holds for the cases where we have columns that are not in the block's columns (even though the rows are in the block's rows), i.e. when $j < l$ or $j > l+n_j$. Details are left to the reader on this end.
Meanwhile, now that we see "block"-ness, we can look at the entries in the actual block. For entries in the block, we just get

$$(AB)_{ij} = \sum_{k=1}^{n} A_{ik}B_{kj} = \sum_{k=1}^{l-1} 0 + \sum_{k=l}^{l+n_j} A_{ik}B_{kj} + \sum_{k=l+n_j+1}^{n} 0 = \sum_{k=l}^{l+n_j} A_{ik}B_{kj}$$

which is just the just gives us the same entries as if we just multiplied the block matrices by themselves.
(In general, the key thing here is to just convince yourself via entry-wise calculation and the formula for matrix multiplication that this indeed works.) (HP)

Problem 3.2. Suppose V is a finite-dimensional complex vector space and $f: V \to V$ is a linear operator.

1. Prove that V has a basis consisting of eigenvectors of f if and only if every generalized eigenvector of f is a classical eigenvector of f.
2. Prove that V has a basis consisting of eigenvectors of f if and only if the minimal polynomial of f has no repeated zeroes.

Solution. 1. The first part of this proof is quick: if every generalized eigenvector of f is a classical eigenvector, then take a basis for V of generalized eigenvectors of f - which we know to be possible from class. Now, each generalized eigenvector is already a classical eigenvector, so we actually just have a basis of classical eigenvectors to begin with, and we are done.
For the other direction, suppose we have a basis of eigenvectors of f. We want to prove that every generalized eigenvector of f is actually a classical eigenvector. Well, if we have a basis, let's take it - call it $v_1, ..., v_n$. Suppose that we have a generalized eigenvector w, which then has a corresponding eigenvalue which we will call λ_w. Since $w \in V$, we can write w in terms of the basis vectors $v_1, ..., v_n$, so that

$$w = \sum_{i=1}^{n} c_i v_i$$

for scalars $c_1, ..., c_n \in \mathbb{C}$. Now, since w is a generalized eigenvector, we know that if we raise $(f - \lambda_w(id_V))$ to the n'th power (where id_V is the identity map on V), we get 0 - in other words,

$$(f - \lambda_w(id_V))^n(w) = 0$$

We then get that

$$(f-\lambda_w(id_V))^n(w)=\sum_{i=1}^{n}c_i(f-\lambda_w)^n(v_i)$$

Since each v_i has an associated eigenvalue λ_i, we get

$$(f-\lambda_w(id_V))^n(w)=\sum_{i=1}^{n}c_i(\lambda_i-\lambda_w)^n(v_i)=0$$

For this last expression to equal 0, we must have that $c_i(\lambda_i-\lambda_w)^n=0$ for each i, since the v_i's are linearly independent. But then, we must have that $c_i=0$ for all i where $\lambda_w\neq\lambda_i$, which allows us to write w as a linear combination of classical eigenvectors corresponding to λ_w. Thus, w is just a classical eigenvector, and we are done.

2. First, suppose that the minimal polynomial for f has distinct roots. We prove that V has a basis of eigenvectors of f so that f is diagonalizable.
If we have distinct roots in the polynomial (call it m), then

$$m_f(z)=(z-\lambda_1)(z-\lambda_2)...(z-\lambda_k)$$

is what our minimal polynomial should look like, where $\lambda_i\neq\lambda_j$ for any of the i and j, and these λ_i are the eigenvalues of f.
Consider the eigenspace E_i for each λ_i - i.e., all vectors $v\in V$ such that $f(v)=\lambda_i v$. To show that f has a basis of eigenvectors, we show that

$$V=\bigoplus_{1\leq i\leq k}E_i$$

To do this, given any $v\in V$, we want to write it as a linear combination of eigenvectors: i.e.:

$$v=w_1+w_2+...+w_k$$

where each w_i is in E_i.
Take a linearly independent and spanning set of vectors for each E_i, an eigenbasis for each E_i. Since eigenvectors for different eigenvalues are linearly independent to each other already, we then get that

$$E_1+...+E_k$$

is a direct sum. Now all we need to show is that

$$V=E_1+...+E_k$$

Suppose we had the following scenario: we had polynomials $h_1,...h_k$ such that

$$v=[h_1(f)](v)+...+[h_k(f)](v)$$

so that each $[h_i(f)](v)$ gives us an E_i-eigenvector component of v, the w_i component we seek above. For the above to hold, we would need h_i's such that

$$id_V=h_1(f)+...+h_k(f)$$

(so that when we apply v to both sides, we get that $v=\sum[h_i(f)](v)$). We would also need h_i's such that if $m(f)$ is our minimal polynomial, and λ_i is the eigenvalue for E_i,

$$h_i(f)\text{is divisible by}\frac{m(f)}{(f-\lambda_i)}$$

which implies that

$$(f - \lambda_i)[h_i(f)](v) = 0$$

so that each h_i actually gives us an eigenvector from E_i.
Now all we need to do is find h_i's that actually work. Take

$$p_i = \prod_{j \neq i} (f - \lambda_j)$$

These p_i's each do not divide each other (since we have distinct λ_i's due to no repeated roots in the minimal polynomial). We then have that there exists polynomials q_i such that

$$\sum_{i=1}^{k} p_i q_i = 1$$

Now, if we take $p_i q_i = h_i$, we are done.
For the other direction, suppose that we have a basis of eigenvectors of f, and let their distinct eigenvalues be $\lambda_1, ..., \lambda_k$. We then have that

$$V = \bigoplus_{1 \leq i \leq k} E_i$$

for the eigenspaces E_i related to each λ_i eigenvalue.
We want to prove that the minimal polynomial is just

$$(z - \lambda_1)...(z - \lambda_k)$$

Since there exists a basis of eigenvectors, we know that these eigenvectors form a spanning set for V. Given any eigenvector $w \in V$, by definition, it has an eigenvalue λ such that $f(w) = \lambda w$, so that $(f - \lambda(id_V))(w) = 0$. We know from the proof of Jordan Normal Form that

$$(f - \lambda_i)(f - \lambda_j)(v) = (f - \lambda_j)(f - \lambda_i)(v)$$

i.e., that the $f - \lambda_i$'s commute with each other as functions (and when represented as matrices). So consider the product

$$(f - \lambda_1)...(f - \lambda_k)$$

Since w is an eigenvector, it must be killed by one of these factors, and these factors all commute, so w is killed by this product. But then, this product can successfully send any eigenvector w to 0 - and since the eigenvectors span the space, we know that this product sends all of V to 0. So at the very least, we have that this product is at least a multiple of the minimal polynomial, since it satisfies the conditions for the minimal polynomial besides being the polynomial of smallest degree that kills all eigenvectors.
However, we also have that each root of this product MUST be a root of the minimal polynomial (otherwise, the minimal polynomial will miss an eigenvalue, which is not allowed). In other words, this product is the simplest polynomial that kills each eigenvalue in question, since we need all of these factors to be in the minimal polynomial to capture all of the possible eigenvalues.
Since this polynomial is the polynomial of smallest degree that captures all of our eigenvalues, we get that this must indeed be the minimal polynomial - and as we wanted to prove, it has distinct roots.

(HP)

Problem 3.3. Suppose V is a finite-dimensional complex vector space and $f\colon V \to V$ is a linear operator. Prove that there does not exist a direct sum decomposition of V into two proper invariant subspaces if and only if the minimal polynomial of f is of the form $(z - \lambda)^{\dim V}$ for some $\lambda \in \mathbb{C}$.

Solution. To prove this statement, let's start off by simplifying our problem - we will prove first that there does not exist a direct sum decomposition of V into two (or more) proper invariant subspaces if and only if the Jordan Normal Form of our operator f is a single Jordan block.

First, suppose that we have two or more Jordan blocks in the Jordan normal form of f. Immediately, by the block upper triangular nature of the matrix of f, we can find two proper f-invariant subspaces of V by just taking the subspace spanned by the basis vectors that correspond to the columns/rows of the first block, and then for our other subspace, just take the subspace spanned by the basis vectors that correspond to the columns/rows of the second block. Recall that these are literally just generalized eigenspaces, which we know to be f-invariant already - so we have at least two distinct proper invariant subspaces in V, and the sum of these generalized eigenspaces is direct and adds to V, so we're done with this direction.

Now, suppose that we have a direct sum of two or more proper f-invariant subspaces of V. For now, let's just use two, but we can extend this method to more subspaces. Take S_1 to be our first subspace and S_2 to be our second, so we have $V = S_1 \oplus S_2$. Now, take $f|_{S_1} : S_1 \to S_1$ and $f|_{S_2} : S_2 \to S_2$. Find the Jordan Normal Form for each of these two, and then we can write, just as when we prove Jordan Normal Form by restricting our given f to generalized eigenspaces and splicing the matrices together as upper triangular blocks:

$$[f] = \begin{bmatrix} [f|_{S_1}] & 0 \\ 0 & [f|_{S_2}] \end{bmatrix}$$

where $[f|_{S_1}]$ and $[f|_{S_2}]$ are in JNF. But then, we just have a JNF for f with at least two blocks, so we are done.

Now, we must a matrix has only one Jordan block if and only if the minimal polynomial of f is of the form $(z - \lambda)^{\dim V}$ for some $\lambda \in \mathbb{C}$.

Suppose that the minimal polynomial of f is $(z - \lambda)^{\dim V}$ for some $\lambda \in \mathbb{C}$. Then there is only one eigenvalue, so our Jordan Normal Form for f has only one block!

For the opposite direction, suppose we have only one Jordan block. Then, there is only one eigenvalue, so the minimal polynomial must be of the form

$$(z - \lambda_1)^k$$

for some power k. Moreover, since the single block takes up the whole matrix for f in Jordan Normal Form, which has dimensions $\dim V \times \dim V$, to kill the whole block would require $k = \dim V$ at the very least here in the minimal polynomial. So we must have that the minimal polynomial is indeed $(z - \lambda_1)^{\dim V}$ and we are done. (HP)

Problem 3.4. For coefficients $a_0, \ldots, a_{n-1} \in \mathbb{C}$, consider the matrix

$$\begin{pmatrix} 0 & & & & -a_0 \\ 1 & 0 & & & -a_1 \\ & 1 & \ddots & & -a_2 \\ & & \ddots & & \vdots \\ & & & 0 & -a_{n-2} \\ & & & 1 & -a_{n-1} \end{pmatrix}.$$

Calculate the characteristic and minimal polynomials of this matrix.

Solution. In class, we did an example with "marching ones" on an off-diagonal, and we showed that taking powers of such a matrix moves the ones further and further into the lower-left corner. In particular, this shows that the set of vectors $(M^0, M^1, M^2, \ldots, M^{n-1})$ is linearly independent. However, we know that the characteristic polynomial of M is of degree exactly n, so it must agree with the minimal polynomial.

What remains is to actually calculate a monic polynomial of degree n which kills M. We have actually seen this matrix before: for a fixed polynomial u, you have previously considered the vector space quotient

$$V = \{\text{polynomials of all degrees}\}/\{f \mid f = g \cdot u \text{ for some other polynomial } g\},$$

which you showed to be of dimension $\deg u$, spanned by the set $(1, x, x^2, \ldots, x^{n-1})$. Writing $u(x) = a_0 + a_1 x + \cdots + x^n$, the matrix M encodes multiplication by x in this basis. Since $u \equiv 0$ in the quotient space V, it follows that $u(M) = 0$. (ECP)

Solution. Here is a solution using determinants. To calculate the characteristic polynomial of this matrix, we want to find

$$\det(zI - [f])$$

In other words, we want to find the determinant of the matrix given by

$$\begin{pmatrix} z & & & & a_0 \\ -1 & z & & & a_1 \\ & -1 & \ddots & & a_2 \\ & & \ddots & & \vdots \\ & & & z & a_{n-2} \\ & & & -1 & z + a_{n-1} \end{pmatrix}.$$

To find the determinant of this matrix, we can expand by minors along the last column. For any given a_i, when we remove the last column and the i'th row, the resulting associated minor is an upper triangular matrix. Now, the determinant of each minor is just the multiplication of the resulting minor's diagonal. Summing over the diagonals for each associated minor, and accounting for the sign change with each minor, with some bashing, we eventually get

$$a_0 + a_1 z + a_2 z^2 + \ldots + a_{n-1} z^{n-1} + z^n$$

as our characteristic polynomial. (HP)

4 For submission to Rohil Prasad

Problem 4.1. Let $f, g : V \to V$ be two linear operators related by an isomorphism $h : V \to V$ via the equation $f = hgh^{-1}$. Relate the generalized eigenspaces for f and g through h. Conclude that f and g have the same eigenvalues with the same algebraic multiplicities (i.e., these are *invariants* of f and g.

Solution. Let $G(\lambda, f)$ and $G(\lambda, g)$ be the generalized eigenspaces of f and g with eigenvalue λ.

We will show that h maps $G(\lambda, g)$ surjectively onto $G(\lambda, f)$. Since h is injective, this will show that both of these generalized eigenspaces have the same dimension, and therefore f, g will have the same algebraic multiplicity for any $\lambda \in K$. By definition of algebraic multiplicity, it follows that f, g have the same eigenvalues as well.

Pick $v \in G(\lambda, g)$. Then, there exists some n such that $(g - \lambda)^n(v) = 0$.

Note that for any k, $f^k = (hgh^{-1})^k = hg^k h^{-1}$. Therefore, $f^k(h(v)) = hg^k(v)$.

We also have $\lambda(h(v)) = h(\lambda(v))$. As a result, we can calculate $(f - \lambda)^n(h(v)) = h(g - \lambda)^n(v) = 0$.

We must also show for any $v \in G(\lambda, f)$ that $h^{-1}(v) \in G(\lambda, g)$. This follows from the exact same reasoning as above using the fact that $g = h^{-1} f h$. (RP)

Problem 4.2. Suppose $f : \mathbb{C}^4 \to \mathbb{C}^4$ is such that the eigenvalues of f are 3, 5, and 8. Prove that

$$(f - 3)^2 (f - 5)^2 (f - 8)^2 = 0$$

Solution. Recall that the sum of the algebraic multiplicities of an operator must add up to the dimension of the space. Since f has three distinct eigenvalues and $\mathbb{C}^4$ has dimension 4, we see that we must have two eigenvalues with an algebraic multiplicity of 1 and one with an algebraic multiplicity of 2.

Without loss of generality, suppose the generalized eigenspace $G(3, f)$ has dimension 2. Therefore, we have $\ker(f - 3)^2$ has dimension 1 or 2. If it has dimension 1, then $\ker(f - 3)$ has dimension 1 as well and by stabilization of the kernel $G(3, f)$ would have dimension 1. Therefore, we must have it has dimension 2, and therefore $G(3, f) = \ker(f - 3)^2$.

Thus, any vector $v \in \mathbb{C}^4$ by the generalized eigenspace decomposition can be written as a sum of an eigenvector of eigenvalue 5, an eigenvector of eigenvalue 8, and an element of $G(3, f)$. We have the operators $(f - 5)^2$, $(f - 8)^2$, and $(f - 3)^2$ kill each of these parts respectively, so the problem's assertion follows immediately. (RP)

Problem 4.3. Suppose $f : V \to V$ is invertible. Prove that there exists a polynomial $p(z)$ (dependent upon f) such that $f^{-1} = p(f)$.

Solution. Let $P(z) = \sum_{i=0}^{n} a_i z^i$ be the characteristic polynomial of f. Since f is invertible, it cannot have an eigenvalue of 0. Therefore, its constant term a_0 must be nonzero.

Since $P(f) = 0$, we find $\sum_{i=1}^{n} a_i f^i + a_0 = 0$. Therefore, $\sum_{i=1}^{n} a_i f^i = -a_0$, which implies $\sum_{i=1}^{n} -\frac{a_i}{a_0} f^i = 1$.

Factoring out f, we find that $p(z) = \sum_{i=1}^{n} -\frac{a_i}{a_0} z^{i-1}$ is our desired polynomial. (RP)