MATH 23A SOLUTION SET #10 (PART B)

ISIDORA MILIN

Problem (3). *Let $A : V \to V$ be a linear transformation on a finite dim. vector space, and by a slight abuse of notation, let A also be the matrix for this transformation with respect to a fixed basis. Using the following method, we determine the eigenvalues of A:*

$$\begin{aligned} \lambda \in spec(A) \quad &\Leftrightarrow \quad V_\lambda = Ker(A - \lambda I) \neq \{0\} \\ &\Leftrightarrow \quad A - \lambda I \text{ is not invertible} \\ &\Leftrightarrow \quad det(A - \lambda I) = 0 \end{aligned}$$

$p_A(\lambda) = det(A - \lambda I)$ is called the characteristic polynomial of A.

(a). *Prove that no scalar $\lambda_0 \in F$ is an eigenvalue for A unless it is a root of $p_A(\lambda)$*

Solution. The argument for this claim was basically provided in the statement of the problem. Suppose λ_0 is an eigenvalue for A, and let v be a λ_0-eigenvector. Then, $(A - \lambda_0 I)v = Av - \lambda_0 v = 0$, so that $ker(A - \lambda_0 I)$ is nontrivial, and hence $A - \lambda_0 I$ is not invertible. Thus, $p_A(\lambda_0) = det(A - \lambda_0 I) = 0$ and λ_0 is a root of p_A.

In fact, the other direction holds as well - if $\lambda_0 \in F$ is such that $p_A(\lambda_0) = 0$, then λ_0 is an eigenvalue for A. For, if $p_A(\lambda_0) = 0$, it follows that $A - \lambda_0 I$ is noninvertible, so that λ_0 is an eigenvalue for A. □

(b). *If $p_A(\lambda) = (\lambda - \lambda_0)^k q(\lambda)$ with $q(\lambda) \neq 0$, then we say that the eigenvalue λ_0 has algebraic multiplicity equal to k. (That is, λ_0 is a root of $p_A(\lambda)$ of order k.) Show that the geometric multiplicity (which, by definition, is the dimension of the corresponding eigenspace) of an eigenvalue is less than or equal to its algebraic multiplicity.*

Solution. Let $dim(V) = n$ and let m be the geometric multiplicity of λ_0 as an eigenvalue for A. Since $V_{\lambda_0} \subseteq V$, we have $m = dim(V_{\lambda_0}) = m \leq n$. Let $\{v_1, \ldots, v_m\}$ be a basis for V_{λ_0}. Extend it to a basis $\mathfrak{B} = \{v_1, \ldots, v_m, v_{m+1}, \ldots, v_n\}$ for V. We consider the matrix of the linear transformation A with respect to $\mathfrak{B}$. Since $v_i \in V_{\lambda_0}$ for all $1 \leq i \leq m$, we have complete information about the first m *columns* of that matrix. (Recall that a matrix of a linear transformation is written down by writing down the images of basis vectors as *columns, not rows.*) Thus, the matrix of A with respect to $\mathfrak{B}$ will look something like this:

$$[A]_{\mathfrak{B}} = \begin{pmatrix} \lambda_0 & 0 & 0 & \ldots & 0 & * & \ldots & * \\ 0 & \lambda_0 & 0 & \ldots & 0 & * & \ldots & * \\ 0 & 0 & \lambda_0 & \ldots & 0 & * & \ldots & * \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & \lambda_0 & * & \ldots & * \\ 0 & 0 & 0 & \ldots & 0 & * & \ldots & * \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & 0 & * & \ldots & * \end{pmatrix} = \begin{pmatrix} \lambda_0 I_m & B \\ 0 & C \end{pmatrix}$$

where $*$ represents any element of F, B is some $m \times (n-m)$ matrix, and C is some $(n-m) \times (n-m)$ matrix. It follows that

$$[A - \lambda I]_{\mathfrak{B}} = \begin{pmatrix} \lambda_0 - \lambda & 0 & 0 & \dots & 0 & * & \dots & * \\ 0 & \lambda_0 - \lambda & 0 & \dots & 0 & * & \dots & * \\ 0 & 0 & \lambda_0 - \lambda & \dots & 0 & * & \dots & * \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \lambda_0 - \lambda & * & \dots & * \\ 0 & 0 & 0 & \dots & 0 & * & \dots & * \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 0 & * & \dots & * \end{pmatrix} = \begin{pmatrix} (\lambda_0 - \lambda)I_m & B' \\ 0 & C' \end{pmatrix}$$

Using the expansion into minors method to calculate $det(A - \lambda I)$, we get that

$$det(A - \lambda I) = (\lambda_0 - \lambda)^m det(C') = (\lambda - \lambda_0)^m q(\lambda)$$

Now, $q(\lambda)$, as a polynomial in λ, can be written as $(\lambda - \lambda_0)^l r(\lambda)$, where $r(\lambda_0) \neq 0$ and $l \geq 0$. It follows that

$$p_A(\lambda) = det(A - \lambda I) = (\lambda - \lambda_0)^m q(\lambda) = (\lambda - \lambda_0)^{m+l} r(\lambda)$$

so that the algebraic multiplicity of λ_0 is $k = m + l \geq m$. □

(**c**). *Use this method to find all eigenvalues of the real matrix*

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

Is the matrix diagonalizable? Explain.

Solution. Recall from class that a matrix is diagonalizable iff:

$$\sum_{\lambda \in spec(A)} dim(V_\lambda) = dim(V)$$

Thus, in order to check whether A is diagonalizable, we need to find its eigenvalues and the dimensions of their corresponding eigenspaces (that is, their geometric multiplicities). By part (a), all eigenvalues of A are roots of $p_A(\lambda) = det(A - \lambda I)$.

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

so that the expansion into minors yields:

$$\begin{aligned} p_A(\lambda) = det(A - \lambda I) &= (-\lambda) det \begin{pmatrix} -\lambda & 0 & 0 \\ 0 & -\lambda & 1 \\ 0 & -1 & -\lambda \end{pmatrix} + (-1) det \begin{pmatrix} 1 & 0 & 0 \\ 0 & -\lambda & 1 \\ 0 & -1 & -\lambda \end{pmatrix} \\ &= (\lambda^2 + 1) det \begin{pmatrix} -\lambda & 1 \\ -1 & -\lambda \end{pmatrix} \\ &= (\lambda^2 + 1)(\lambda - 1)(\lambda + 1) \end{aligned}$$

Thus, $spec(A) \subseteq \{1, -1, i, -i\}$. If we consider A as a real matrix (as the problem suggests we should), we see that A has two real eigenvalues, each with algebraic multiplicity one. From part (b) we deduce that

$$dim(V_1) + dim(V_{-1}) \leq 1 + 1 = 2 < 4 = dim(V)$$

so that A is not diagonalizable over $\mathbb{R}$. □

Note 1. Note that the fact that A just has two real eigenvalues does not in itself imply A is not diagonalizable over $\mathbb{R}$, as many of you stated in your solutions. Just think of I_n - the $n \times n$ identity matrix. Its only eigenvalue is 1, but with (both algebraic and geometric) multiplicity n, so that I_n is diagonalizable even though it does not have n distinct eigenvalues.

Note 2. Many people stated that A is diagonalizable over $\mathbb{C}$ with $1, -1, i, -i$ along the diagonal. That is technically correct, and was (if you did everything else correctly) enough to get you full credit, but it was "cheating", as it was clear the problem asked you to explain geometric and algebraic multiplicity concepts from part (b) in this particular example. So if you recognize yourself here, please make sure you understand the above reasoning!

Problem (6). *Given real vector space V with convergent sequence $\{v_n\}$ such that $\lim_{n\to\infty} v_n = v$, and a sequence of real numbers $\{c_n\}$ such that $\lim_{n\to\infty} c_n = c$, prove that $\lim_{n\to\infty} c_n v_n = cv$.*

Solution. The proof comes down to using the triangle inequality in a clever way, and, what many people did not realize, the fact that a convergent sequence in $\mathbb{R}$ is necessarilly bounded. (The proof of this last fact is not hard. Do it as an exercise or look it up in Fitzpatrick!)

To begin, we fix $\epsilon > 0$ and we want to show $\exists N \in \mathbb{N}$ such that $\forall n \geq N$, $||c_n v_n - cv|| \leq \epsilon$. We first use the triangle inequality to put $||c_n v_n - cv||$ into a more convinient form:

$$\begin{aligned} ||c_n v_n - cv|| &= ||c_n v_n - c_n v + c_n v - cv|| \\ &= ||c_n(v_n - v) + (c_n - c)v|| \\ &\leq ||c_n(v_n - v)|| + ||(c_n - c)v|| \\ &= |c_n|\, ||v_n - v|| + |c_n - c|\, ||v|| \end{aligned}$$

where $|\cdot|$ is the usual absolute value on $\mathbb{R}$, and $||\cdot||$ the norm on V. It is a good habit to use different notation for different norms - it makes your proofs easier to read! Now, since $\{c_n\}$ is a convergent sequence, it is bounded. Let $M \in \mathbb{R}$ be such that $|c_n| \leq M$ for all $n \in \mathbb{N}$.

Since $\lim_{n\to\infty} v_n = v$, we can find an N_1 such that:

$$\forall n \geq N_1, \ \ ||v_n - v|| \leq \frac{\epsilon}{2M}$$

Since $\lim_{n\to\infty} c_n = c$, we can find an N_2 such that:

$$\forall n \geq N_2, \ \ ||c_n - c|| \leq \frac{\epsilon}{2||v||}$$

Note that $||v|| = 0$ or $M = 0$ poses no problem above - it just makes N_1 and $N - 2$ easier to find! Now let $N = max\{N_1, N_2\}$. For $n \geq N$ we have that:

$$||c_n v_n - cv|| \leq |c_n|\, ||v_n - v|| + |c_n - c|\, ||v|| \leq M\frac{\epsilon}{2M} + \frac{\epsilon}{2||v||}||v|| = \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$$

which shows that $lim_{n\to\infty} c_n v_n = cv$. □

Note. Another way to prove this was to treat it coordinatewise - in each coordinate, the problem reduces to convergence of a product sequence in $\mathbb{R}$ - product of the limits is the limit of the product, calc BC! Now, since the i-th coordinate of $c_n v_n$ converges to the i-th coordinate of cv, for all i, it follows that $c_n v_n$ converges to cv.