MAT247 Algebra II

Assignment 3

Solutions

1. Let T be a linear operator on a finite-dimensional vector space over a field F. Suppose that the characteristic polynomial of T is $f(t) = \pm t^2(t-1)^3(t^p+1)$, where p is a prime number.

(a) Suppose $F = \mathbb{R}$. Can T be diagonalizable? What are the eigenvalues of T? What are the possible values for the dimension of each eigenspace of T? Give an example for each possibility. (Suggestion: Construct your examples using block diagonal matrices. Exercise 19 of 5.4 gives a way of constructing a matrix with a given characteristic polynomial.)

(b) Suppose $F = \mathbb{C}$. Can T be diagonalizable? If yes, suppose T is diagonalizable. What is the dimension of each eigenspace of T?

(c) Suppose F has characteristic p. What are the eigenvalues of T? What are the possible values of the dimension of each eigenspace of T? (Suggestion: Expand $(t+1)^p$.)

(d) Back to a general F of arbitrary characteristic, show that T is not surjective.

Solution: (a) No, because the characteristic polynomial does not split over $\mathbb{R}$. The eigenvalues are $0, 1$ if $p = 2$ and $0, 1, -1$ if $p > 2$. The dimension of E_0 can be 1 or 2 (= multiplicity of eigenvalue 0), and the dimension of E_1 can be 1,2, or 3. When $p > 2$, the eigenspace E_{-1} is one-dimensional as -1 is an eigenvalue of multiplicity 1. We construct our examples for the case where p is odd. Consider the block-diagonal matrix

$$A = \begin{pmatrix} B_0 & & \\ & B_1 & \\ & & C \end{pmatrix},$$

where C is the companion matrix for the polynomial $t^p + 1$, i.e.

$$C = \begin{pmatrix} 0 & & & \cdots & & -1 \\ 1 & 0 & & \cdots & & 0 \\ 0 & 1 & 0 & \cdots & & 0 \\ \vdots & \ddots & & & & \vdots \\ & & & & 0 & 0 \\ 0 & & & & 1 & 0 \end{pmatrix} \in M_{p\times p}(\mathbb{R}),$$

and the diagonal blocks B_0 and B_1 are of the form

$$B_0 = \begin{pmatrix} 0 & a \\ 0 & 0 \end{pmatrix}$$

and

$$B_1 = \begin{pmatrix} 1 & b & 0 \\ 0 & 1 & c \\ 0 & 0 & 1 \end{pmatrix}.$$

Then the characteristic polynomial of A is exactly $f(t)$, and moreover by choosing a, b, c we can make sure E_0 and E_1 have dimensions as desired (up to the multiplicity of the eigenvalue, of course). If we take $a = 1$ (resp. $a = 0$), then $\dim(E_0) = 1$ (resp. $\dim(E_0) = 2$). If we take $b = 1$ and $c = 0$ (resp. $b = c = 1$ and $b = c = 0$), then the dimension of E_1 is 2 (resp. 1 and 3).

(b) It can. The complex roots of $f(t)$ are 0 (of mult. 2), 1 (of mult. 3), and the p-th roots $\alpha_1, \dots, \alpha_p$ of -1, each of multiplicity 1. The eigenspace E_{α_i} are each 1-dimensional, and T is diagonalizable if and only if $\dim(E_0) = 2$ and $\dim(E_1) = 3$.

(c) Consider

$$(t+1)^p = \sum_j \binom{p}{j} t^j.$$

For $1 < j < p$, the number $j!(p-j)!$ is not a multiple of p (the key being that p is a prime number), so that $\binom{p}{j} = \frac{p!}{j!(p-j)!}$ is a multiple of p. Thus if F is a field of characteristic p,

$$(t+1)^p = t^p + 1^p = t^p + 1.$$

If $p \neq 2$, then the distinct roots of $f(t)$ in F (or in other words, eigenvalues of T) are 0,1,-1, respectively of multiplicities 2, 3, and p. If $p = 2$, then $-1 = 1$ and the distinct roots of $f(t)$ in F are 0 and 1, respectively of multiplicities 2 and 5. In each case, the dimension of each eigenspace is bounded from below by 1 and from above by the multiplicity of the corresponding eigenvalue.

(d) Since T is a linear operator on a finite-dimensional vector space, T is surjective if and only if it is injective. Since 0 is an eigenvalue of T, $\ker(T)$ is not zero and hence T is not injective.

2. For $A \in M_{n\times n}(\mathbb{C})$, define

$$e^A = \sum_{k=0}^{\infty} \frac{1}{k!} A^k.$$

One can show that the sum converges for every A.

(a) Ignoring all convergence-related questions, show that if $AB = BA$, then $e^{A+B} = e^A e^B$.
(b) Show that if A is nilpotent, then the characteristic polynomial of e^A is $(-1)^n(t-1)^n$. (An element of $M_{n\times n}(F)$ with characteristic polynomial $(-1)^n(t-1)^n$ is called a *unipotent* matrix. Hint: How are the characteristic polynomials of B and $B + \lambda I$ related to one another?)
(c) Does it make sense to define e^A with the same formula as above for an $n \times n$ matrix A with entries in an arbitrary field of characteristic zero? What if A is nilpotent?
(d) Can we define e^A with the formula as above for a nilpotent matrix A with entries in a field of positive characteristic?

Solution: (a) Let us start with

$$e^{A+B} = \sum_{k=0}^{\infty} \frac{1}{k!} (A+B)^k.$$

Since A and B commute, this expands as

$$\sum_{k=0}^{\infty} \frac{1}{k!} \sum_{j=0}^{k} \binom{k}{j} A^j B^{k-j} = \sum_{k=0}^{\infty} \sum_{j=0}^{k} \frac{1}{k!} \binom{k}{j} A^j B^{k-j} = \sum_{k=0}^{\infty} \sum_{j=0}^{k} \frac{1}{j!(k-j)!} A^j B^{k-j}.$$

Setting $r = j$ and $s = k - j$, the sum can be rewritten as[†]

$$\sum_{r=0}^{\infty} \sum_{s=0}^{\infty} \frac{1}{r!s!} A^r B^s = \sum_{r=0}^{\infty} \left(\frac{1}{r!} A^r \sum_{s=0}^{\infty} \frac{1}{s!} B^s \right) = \left(\sum_{r=0}^{\infty} \frac{1}{r!} A^r \right) \left(\sum_{s=0}^{\infty} \frac{1}{s!} B^s \right) = e^A e^B.$$

[†]The reason we can do this is that each entry of the series above converges absolutely, so that we can rearrange the sum as we wish.

(b) Let A be nilpotent, say of nilpotency index m. Let $N = e^A - I$. We have

$$N = \sum_{k=1}^{\infty} \frac{1}{k!} A^k = \sum_{k=1}^{m} \frac{1}{k!} A^k = A \underbrace{\sum_{k=1}^{m} \frac{1}{k!} A^{k-1}}_{B}.$$

Note that A and B commute, and thus $N^m = A^m B^m = 0$. In particular, N is nilpotent and hence its characteristic polynomial is $(-1)^n t^n$. Then the characteristic polynomial of $e^A = N + I$ is $p_N(t-1) = (-1)^n (t-1)^n$. (Note that for any map or matrix M, $p_{M+\lambda I}(t) = \det(M + \lambda I - tI) = \det(M - (t-\lambda)I) = p_M(t-\lambda)$.)

(c) In general, for an arbitrary matrix A over an arbitrary field of characteristic zero, this definition does not make sense, because there is no notion of limit and hence infinite sums don't make sense. If A is nilpotent, then the definition does make sense because there are only finitely many nonzero terms in the sum anyways.

(d) Here even in the nilpotent case there might be a problem as $k!$ is zero in the field when k is greater than or equal to the characteristic, and hence does not have a multiplicative inverse.

3. Let T be a linear operator on a nonzero finite-dimensional vector space V over a field F. The *minimal polynomial* of T is the monic[†] polynomial $f(t)$ of smallest degree such that $f(T) = 0$. It is easy to show that the minimal polynomial exists and is unique, and by the Cayley-Hamilton theorem, its degree is $\leq \dim(V)$. You don't have to include the argument for these facts in your solution.

Suppose $f(t)$ is the minimal polynomial of T. Let $g(t) \in F[t]$ be any polynomial such that $g(T) = 0$. Show that $f(t) \mid g(t)$. (Suggestion: By the division algorithm, we can write $g(t) = q(t)f(t) + r(t)$ with $\deg(r(t)) < \deg(f(t))$. It might help to take a look at Exercise 1 of the extra practice problems below.)

Solution: By the division algorithm we can write $g(t) = f(t)q(t) + r(t)$ for (unique) $q(t), r(r) \in F[t]$ with $\deg(r(t)) < \deg(f(t))$. We then have $g(T) = f(T) \circ q(T) + r(T)$. Since $g(T) = f(T) = 0$, we get $r(T) = 0$. This implies $r(t) = 0$ (and hence $f(t) \mid g(t)$), as otherwise, if c is the leading coefficient of $r(t)$, then $r_1(t) = \frac{1}{c} r(t)$ is a monic polynomial of degree less that $\deg(f(t))$ and $r_1(T) = \frac{1}{c} r(T) = 0$, which contradicts the minimality property of the minimal polynomial of T.

4. Let F be a field. We say a polynomial $f(t) \in F[t]$ is *irreducible* (over F) if it satisfies (both of) the following conditions: (i) the degree of $f(t)$ is ≥ 1, and (ii) whenever $f(t) = g(t)h(t)$ for polynomials $g(t), h(t) \in F[t]$, then $g(t)$ or $h(t)$ has degree zero (i.e. we cannot write $f(t)$ as a product of two elements of $F[t]$ with positive degree). For instance, $t^2 + 1$ is an irreducible polynomial in $\mathbb{R}[t]$. The only irreducible polynomials in $\mathbb{C}[t]$ are the degree 1 polynomials.

Let T be a linear operator on a finite-dimensional vector space over F.

(a) Show that for any $f(t) \in F[t]$, the kernel of $f(T)$ is T-invariant.

(b) Suppose $f(t) \in F[t]$ is an irreducible factor of the characteristic polynomial $p_T(t)$ of T (by being a factor we mean $f(t)$ divides $p_T(t)$). Show that either $\ker(f(T)) = 0$ or

$$\dim \ker(f(T)) \geq \deg(f(t)).$$

(Suggestion: Question 3 can be useful. Note: In fact, $f(T)$ cannot be injective. You can try to prove this (but it is not mandatory).)

[†]We say a polynomial $f(t) \in F[t]$ is monic if its leading coefficient is 1.

Solution: (a) Given $v \in \ker(f(T))$, we have $f(T)((T(v)) = f(T) \circ T(v) = T \circ f(T)(v) = T(f(T)(v)) = T(0) = 0$. Thus $T(v) \in \ker(f(T))$ and $\ker(f(T))$ is T-invariant.

(b) Suppose $\ker(f(T)) \neq 0$. Set $W = \ker(f(T))$; then W is nonzero and T-invariant. Let $g(t)$ be the minimal polynomial of T_W. Since $W \neq 0$, $\deg(g(t)) > 0$. Now by definition of W, $f(T)$ is zero on W, so that $f(T_W) = 0$. Thus $g(t) \mid f(t)$. Since $f(t)$ is irreducible and $g(t)$ is not constant, it follows that $\deg(f(t)) = \deg(g(t))$. On the other hand, by Cayley-Hamilton, $p_{T_W}(T_W) = 0$, so that $\deg(g(t)) \leq \deg(p_{T_W}(t)) = \dim(W)$. Thus we have

$$\deg(f(t)) = \deg(g(t)) \leq \deg(p_{T_W}(t)) = \dim(W),$$

as desired. (Note: For the question as stated, we don't need to assume that $f(t)$ divides the characteristic polynomial of T. With that extra hypothesis, one can prove that $f(T)$ cannot be injective, so that we must have $\dim\ker(f(T)) \geq \deg(f(t))$ (rather than the weaker conclusion than either $\ker(f(T)) = 0$ or $\dim\ker(f(T)) \geq \deg(f(t))$). See the first problem on the extra practice list for Assignment 4.)

5. Let V be an n-dimensional vector space over a field F. Suppose $T : V \to V$ is a linear map such that the characteristic polynomial $p_T(t)$ is irreducible.

(a) Show that the only T-invariant subspaces of V are 0 and V.

(b) Let v be a nonzero element of V. Show that $\{v, T(v), \ldots, T^{n-1}(v)\}$ is a basis of V.

(c) Suppose $A \in M_{4\times4}(\mathbb{Q})$ has characteristic polynomial $p_A(t) = t^4 + t^3 + t^2 + t + 1$. Show that there exists a matrix $P \in M_{4\times4}(\mathbb{Q})$ such that

$$P^{-1}AP = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -1 \end{pmatrix}.$$

Note: For parts (b) and (c) you might want to wait until Tuesday. Otherwise, first read about T-cyclic subspaces on the bottom of page 313 and then read Theorem 5.22. For (c), you may take it for granted that $t^4 + t^3 + t^2 + t + 1$ is irreducible over $\mathbb{Q}$.)

Solution: (a) If W is a nonzero proper T-invariant subspace of V, then $0 < \deg(p_{T_W}(t)) = \dim(W) < \dim(V)$ and $p_T(t) = p_{T_W}(t)g(t)$ for some polynomial $g(t)$ of degree

$$\deg(g(t)) = \deg(p_T(t)) - \deg(p_{T_W}(t)) = \dim(V) - \dim(W) > 0,$$

contradicting the assumption that $p_T(t)$ is irreducible.

(b) Let W be the T-cyclic subspace generated by v. Then W is a nonzero T-invariant subspace of V. It follows from (a) that $W = V$, i.e $\dim(W) = n$. The assertion now follows from Theorem 5.22(a) of the textbook.

(c) Let v be a nonzero element of $\mathbb{Q}^4$. Since $p_A(t)$ is irreducible over $\mathbb{Q}$, by (c), $\beta = \{v, Av, A^2v, A^3v\}$ is a basis of $\mathbb{Q}^4$. By Theorem 5.22(b) of the textbook, we must have $A^4v + A^3v + A^2v + Av + v = 0$. The matrix of $L_A : \mathbb{Q}^4 \to \mathbb{Q}^4$ with respect to the basis β is then exactly the given matrix. Take $P = (v\ Av\ A^2v\ A^3v)$ (with $v, Av, \ldots$ as the columns). By the change of basis formula, $P^{-1}AP = [L_A]_\beta$.