Amalgamated Worksheet # 2 Solutions

Various Artists

April 10, 2013

1 Peyam Tabrizian

Problem 1:

Find $T \in \mathcal{L}(\mathbb{R}^3)$ whose characteristic polynomial and minimal polynomial are not the same

Solution: Let $T = I$ (that is $T(x, y, z) = (x, y, z)$)

Then T has only one eigenvalue 1 with multiplicity 3, so the characteristic polynomial of T is $\boxed{p(z) = (z-1)^3}$

But since $T = I$, we know $T - I = 0$, so the minimal polynomial q of T divides $z - 1$.

Since q is monic (i.e. it's leading coefficient is 1) and of least degree, it follows that either $q(z) = z - 1$ (if q has degree 1) or $q(z) = 1$ (if q has degree 0).

However, if $q(z) = 1$, then we would have $0 = q(T) = I$ so $0 = I$, which is a contradiction.[1]

It follows that the minimal polynomial of T is $\boxed{q(z) = z - 1}$, which is different from $p(z) = (z-1)^3$ $\square$

[1] Remember this type of argument, it's very typical minimal polynomial argument

Problem 2:

Find all 2×2 matrices A such that $A^2 - 3A + 2I = 0$

Hint: You may use the following result: If $\mathbb{R}^2$ has a basis of eigenvectors of A, then there exists a matrix P such that $A = PDP^{-1}$, where D is the matrix of eigenvalues of A

Solution: Let $q(z)$ be the minimal polynomial of A.

Since $A^2 - 3A + 2I = 0$, we know that $q(z)$ divides $z^2 - 3z + 2 = (z-1)(z-2)$.

Since q is monic of least degree, we have three cases:
Case 1: $q(z) = z - 1$

Then by definition $q(T) = A - I = 0$, so in this case $\boxed{A = I}$

Case 2: $q(z) = z - 2$

Then by definition $q(T) = A - 2I = 0$, so $\boxed{A = 2I}$

Case 3: $q(z) = (z-1)(z-2)$

Then because the roots of q are the eigenvalues of A, it follows that A has two distinct eigenvalues $\lambda = 1$ and $\lambda = 2$, hence $\mathbb{R}^2$ has a basis of eigenvectors of A, namely (v_1, v_2), where v_1 is a nonzero eigenvector corresponding to $\lambda = 1$ and v_2 is a nonzero eigenvector corresponding to $\lambda = 2$

By the hint above, it follows that there exists a matrix P such that $\boxed{A = PDP^{-1}}$ where $D = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$, and you can also check that any matrix A of this form actually satisfies $A^2 - 3A + 2I = 0$.

Answer: $A = I, A = 2I$, or $A = PDP^{-1}$, where $D = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$

Problem 3:

If $dim(V) = n < \infty$, and $T \in \mathcal{L}(V)$ show that:

$$dim\left(Span\left\{I,T,T^2,\cdots\right\}\right) < n$$

Solution: Let $p(z) = a_0 + a_1 z + \cdots + a_n z^n$ be the characteristic polynomial of T.

Then by the Cayley-Hamilton theorem:

$$a_0 I + a_1 T + \cdots + a_n T^n = 0$$

Let k be the largest index such that $a_k \neq 0$, then we get:

$$\begin{aligned} a_0 I + \cdots + a_k T^k =& 0 \\ a_k T^k =& - a_0 I - \cdots - a_{k-1} T^{k-1} \\ T^k =& - \frac{a_0}{a_k} I - \frac{a_1}{a_k} T - \cdots - \frac{a_{k-1}}{a_k} T^{k-1} \end{aligned}$$

It follows that:

$$T^k \in Span\left\{I, T, \cdots, T^{k-1}\right\} \qquad (*)$$

Let's show by induction on i that $T^{k+i} \in Span\left\{I, T, \cdots, T^{k-1}\right\}$ (where $i = 0, 1, \cdots$)[2]

Base case: The case $i = 0$ follows from $(*)$

Induction step: Suppose $T^{k+i} \in Span\left\{I, T, \cdots, T^{k-1}\right\}$, that is:

$$T^{k+i} = b_0 I + b_1 T + \cdots + b_{k-1} T^{k-1}$$

for some constants $b_0, \cdots, b_{k-1}$.

Then applying T to the above equation, we get:

$$T^{k+i+1} = b_0 T + b_1 T^2 + \cdots + b_{k-2} T^{k-1} + b_{k-1} T^k$$

Now $b_0 T + \cdots + b_{k-2} T^{k-1} \in Span\left\{I, T, \cdots, T^{k-1}\right\}$. But by $(*)$, we also have $b_{k-1} T^k \in Span\left\{I, T, \cdots, T^{k-1}\right\}$, and therefore $T^{k+i+1} \in Span\left\{I, T, \cdots, T^{k-1}\right\}$ (being the sum of two vectors in that span) □

Hence, by what we've just shown, it follows that:

$$Span\left\{I, T, T^2, \cdots\right\} = Span\left\{I, T, \cdots, T^{k-1}\right\}$$

And hence:

[2] That is, higher powers of T still lie in the same span!

$$dim\left(Span\left\{I, T, T^2, \cdots\right\}\right) = dim\left(Span\left\{I, T, \cdots, T^{k-1}\right\}\right) \leq k \leq n = dim(V)$$

Problem 4:

Find a formula for T^{-1} in terms of the coefficients of the characteristic polynomial of T

Solution: If $p(z) = a_0 + a_1 z + \cdots + a_n z^n$ is the characteristic polynomial of T, then by the Cayley-Hamilton theorem, we get:

$$a_0 I + a_1 T \cdots + a_n T^n = 0$$

That is, $a_0 I = -a_1 T - \cdots - a_n T^n$.

Now applying T^{-1} to this equation, we get:

$$a_0 T^{-1} = -a_1 I - a_2 T - \cdots - a_n T^{n-1}$$

Now **if** $a_0 \neq 0$, we would get:

$$T^{-1} - \frac{a_1}{a_0} I - \frac{a_2}{a_0} T - \cdots - \frac{a_n}{a_0} T^{n-1}$$

and we would be done!

However, if $a_0 = 0$, then $p(0) = a_0 + a_1 0 + \cdots + a_n 0^n = a_0 = 0$, so $p(0) = 0$, hence 0 is an eigenvalue of T (by definition of the characteristic polymomial), hence T is not invertible, which is a contradiction with the fact that we assumed T is invertible!

Problem 5:

(if time permits) Given $v \in V$, a polynomial g is called the $T-$**annihilator of** v (or $T-$*killer* of v) if $g(t)$ is a monic polynomial of *least* degree such that $g(T)v = 0$.

Show that such a g divides the minimal polynomial q of T

Solution: Let p be the characteristic polynomial of T.

Then by the division algorithm for polynomials (Theorem 4.5 on page 66), we know that there exist polynomials s and r with $deg(r) < deg(g)$ such that:

$$q(z) = g(z)s(z) + r(z) = s(z)g(z) + r(z)$$

But then replacing z by T in the above equation and using the fact that $q(T) = 0$ (since q is the minimal polynomial of T), we get:

$$0 = s(T)g(T) + r(T)$$

Now applying this equation to v and using $g(T)v = 0$, we get:

$$0 = s(T)g(T)(v) + r(T)v = s(T)0 + r(T)v = r(T)v$$

Hence $r(T)v = 0$ (so r satisfies the same properties as the T-annihilator of V). However, since g is of *least* degree and $deg(r) < deg(g)$, it follows that $r(z) = 0$

But then from $q(z) = g(z)s(z) + r(z)$, we get that $q(z) = g(z)s(z)$, which means that $\boxed{g \text{ divides } q}$[3] □

Problem 6:

(if time permits) Find an (infinite-dimensional) vector space V and a linear operator $D \in \mathcal{L}(V)$ with no minimal polynomial.

Solution: Let $V = \mathcal{P}(\mathbb{R})$ (the space of polynomials), and define $D(p) = p'$.

Now suppose that there exists a minimal polynomial $q(z) = a_0 + a_1 z + \cdots + a_n z^n$ of D.

Then by definition of a minimal polynomial, we would have:

$$q(D) = a_0 I + a_1 D + \cdots + +a_{n-1} D^{n-1} + D^n = 0$$

That is, for every polynomial p, we would have:

$$0 = (a_0 I + a_1 D + \cdots + D^n)p = a_0 p + a_1 p' + \cdots + p^{(n)}$$

But choosing $p(z) = z^n \in V$, we get that:

$$0 = a_0 z^n + a_1 n z^{n+1} + \cdots + n!$$

[3]Remember this argument, it's a classical example of a division-algorithm question!

Plugging in $z = 0$, we get that:

$$0 = n!$$

which is a contradiction. $\square$

2 Daniel Sparks

Let T be a nilpotent operator, upper triangular with respect to the basis $\beta = \{v_1, \cdots, v_n\}$. Let $V_i = \text{Span}\{v_1 \cdots v_n\}$ and $V_0 = (0)$.

(a) Show that $T^i(V_i) = (0)$.

Solution # 1 (Yuval Gannot): V_i is T-invariant by Proposition 5.12. Then $T|_{V_i}$ is nilpotent on an i-dimensional vector space, so $T^i(V_i) = 0$.

In detail, notice that $T^m(v) = T^{m-1}(T|_{V_i}(v)) = T^{m-2}((T|_{V_i})^2(v)) = \cdots = (T|_{V_i})^m(v)$ for $v \in V_i$. Hence $(T|_{V_i})^n(v) = T^n(v) = 0$, so $T|_{V_i}$ is nilpotent. Hence by Corollary 8.8, $T^i(v) = (T|_{V_i})^i(v) = 0$ for $v \in V_i$, i.e. $T^i(V_i) = (0)$. $\square$

Since $M(T, \beta)$ (the matrix of T with respect to β) is upper triangular, the entries on the main diagonal are the eigenvalues by Proposition 5.18. Since T is nilpotent, it's only eigenvalue is 0. [$Tv = \lambda v \Rightarrow T^n v = \lambda^n v = 0 \Rightarrow \lambda = 0$.] Hence $M(T, \beta)$ is actually strictly upper triangular:

$$M(T, \beta) = \begin{pmatrix} 0 & a_{1,2} & a_{1,3} & \cdots & a_{1,n} \\ 0 & 0 & a_{2,3} & \cdots & a_{2,n} \\ 0 & 0 & 0 & \cdots & a_{3,n} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}$$

The following is a fundamental observation.

Lemma: $T(V_i) \subset V_{i-1}$.

Proof: Let $n \geq i > 0$. Let $v = a_1 v_1 + \cdots + a_i v_i \in V_i$. Write $u = a_1 v_1 + \cdots + a_{i-1} v_{i-1} \in V_{i-1}$, so that $v = u + a_i v_i$. Then $T(v) = T(u) + aT(v_i)$. Since V_{i-1} is T-invariant, $T(u) \in V_{i-1}$, and by the matrix displayed above, $T(v_i) = a_{1,i} v_1 + \cdots + a_{i-1,i} v_{i-1} \in V_{i-1}$. Therefore $T(v)$ is a linear combination of vectors in V_{i-1}, and hence is also. $\square$

Solution # 2: Special case of (b), below.

Solution # 3: First, $T^0(V_0) = \text{Id}((0)) = (0)$. [For a less pathological base case, notice $T(V_1) = (0)$ since v_1 is an eigenvector with eigenvalues 0.] Now, suppose as inductive hypothesis that $T^i(V_i) = (0)$ and that $i < n$. Then $T^{i+1}(V_{i+1}) = T^i(T(V_{i+1})) \subset T^i(V_i) = (0)$. The only subspace of (0) is (0), so we're done. $\square$

Solution # 4: $T^i(V_i) = T^{i-1}(T(V_i)) \subset T^{i-1}(V_{i-1}) = T^{i-2}T(V_{i-1}) \subset T^{i-2}(V_{i-2}) \subset \cdots \subset T(V_1) \subset V_0 = (0)$. [This is a less rigorous version of # 3.]

What this means in terms of the matrix of T^i is that the first i columns are all zero.
(b) Show that $T^i(V_j) \subset V_{j-i}$.

Solution: Let $1 \leq j \leq n$ be arbitrary and induct on i. The base case is $i = 1$ which is exactly the lemma above. Suppose $T^i(V_j) \subset V_{j-i}$, and that $i < n$. Then $T^{i+1}(V_j) = T(T^i(V_j)) \subset T(V_{j-i}) \subset V_{j-i-1} = V_{j-(i+1)}$. The first containment comes from the inductive hypothesis and the second from another use of the lemma.

What this means in terms of the matrix of T^i is that it has zeros on and below the i-th diagonal. (The k-th column is $T^i(v_k)$ which is a linear combination of $v_1 \cdots, v_{k-i}$. That means that the only nonzero entries $a_{j,k}$ occur when $j \leq k - i$, i.e. when the row is i entries higher than the column.)

(c) Let M be strictly upper-triangular and $n \times n$. Define an operator $\mathbf{F}^n \to \mathbf{F}^n$ to be the operator whose matrix (with respect to the standard basis) is M. We claim that M is nilpotent.

If $\mathbf{F} = \mathbf{C}$, this is automatic, since the only eigenvalue is 0, and all such operators on $\mathbf{C}^n$ are nilpotent. (However, one can prove it in general in various ways E.g., define $V_i = \text{Span}(e_1, \cdots, e_i)$ as above and show that $T(V_i) \subset V_{i-1}$, hence $T^n(V) = T^n(V_n) \subset (0)$ by reasoning identical to that in (a).)

Hence the analysis in (b) applies to our operator, and in fact the observation about the form of the matrix for T^i in (b) applies exactly to give us what we want. Composing operators corresponds to multiplying matrices, so M^i has zeros on and below the i-th diagonal.

Note in particular that $M^n = 0$, as expected.

(d) Let A, B be block diagonal matrices with blocks of matching size:

$$A = \begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & A_n \end{pmatrix}, B = \begin{pmatrix} B_1 & 0 & \cdots & 0 \\ 0 & B_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & B_n \end{pmatrix}$$

.

Suppose that there are only two blocks in each matrix, i.e. $n = 2$. Let T_A and T_B be operators whose matrices are A and B, respectively, with respect to the same basis

γ. Suppose A_1 is $m \times m$ and that A_2 is $k \times k$. Hence γ has $m+k$ vectors in it, which we name: $\gamma = \{v_1 \cdots v_m, w_1 \cdots, w_k\}$.

Consider the composition $(T_A \circ T_B)$. Since $V = \text{Span}(v_1, \cdots, v_m)$ is both T_A and T_B invariant, it is $(T_A \circ T_B)$ invariant as well. The same goes for the subspace $W = \text{Span}(w_1, \cdots, w_k)$. Hence AB, which is the matrix of $T_A \circ T_B$, is block diagonal with blocks of size m and k as well.

Now, if we restrict the operator T_A to V, we simply get T_{A_1}. Similarly, T_B restricts to V giving T_{B_1}. Since $T_{A_1}T_{B_1} = T_{A_1B_1}$, we see that T_AT_B restricts to V as $T_{A_1B_1}$. This means that the first block in T_AT_B is A_1B_1. Similarly, T_AT_B restricts to W as $T_{A_2B_2}$ implying that the second block is A_2B_2. Hence $\begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}\begin{pmatrix} B_1 & 0 \\ 0 & B_2 \end{pmatrix} = \begin{pmatrix} A_1B_1 & 0 \\ 0 & A_2B_2 \end{pmatrix}$.

One could expand the above proof to the case of many subspaces. Instead, use it as a base case for an induction. Now suppose that the result is true for n. Let A, B be in the above form but with $n+1$ blocks along the diagonal. Notice that we can group this in the following way:

$$\begin{pmatrix} \begin{pmatrix} A_1 & & \cdots & \\ & A_2 & \cdots & \\ \vdots & \vdots & & \vdots \\ & & \cdots & A_n \end{pmatrix} & \\ & A_{n+1} \end{pmatrix}\begin{pmatrix} \begin{pmatrix} B_1 & & \cdots & \\ & B_2 & \cdots & \\ \vdots & \vdots & & \vdots \\ & & \cdots & B_n \end{pmatrix} & \\ & B_{n+1} \end{pmatrix}$$

Therefore, the base case implies that this product is

$$\begin{pmatrix} \begin{pmatrix} A_1 & & \cdots & \\ & A_2 & \cdots & \\ \vdots & \vdots & & \vdots \\ & & \cdots & A_n \end{pmatrix}\begin{pmatrix} B_1 & & \cdots & \\ & B_2 & \cdots & \\ \vdots & \vdots & & \vdots \\ & & \cdots & B_n \end{pmatrix} & \\ & A_{n+1}B_{n+1} \end{pmatrix}$$

and now the inductive hypothesis, applied to the upper left block (which is a product of two block diagonal matrices with n blocks), gives the desired product

$$\begin{pmatrix} A_1B_1 & & \cdots & \\ & A_2B_2 & \cdots & \\ \vdots & \vdots & & \vdots \\ & & \cdots & A_{n+1}B_{n+1} \end{pmatrix}$$

(e) Let L be an operator on a complex space V of dimension n. Let $V = U_{\lambda_1} \oplus \cdots \oplus U_{\lambda_m}$ be the generalized eigenspace decomposition promised by the Jordan theorem. Write

$e_i = \dim U_{\lambda_i}$. If we concatenate bases from each U_λ to form a basis for V, we have the "weak" Jordan form consisting of a matrix M which is block diagonal with m blocks. Each block M_i is upper triangular with λ_i on the diagonal.

Now consider the matrix product

$$*(M - \lambda_1 I)^{e_1} \cdots (M - \lambda_n I)^{e_n}$$

By part (d) we can consider this product one block at a time. That is, the above product (*) is block diagonal with m blocks. The i-th block is

$$**\,(M_i - \lambda_1 I)^{e_1} \cdots (M_i - \lambda_i I)^{e_i} \cdots (M_i - \lambda_n I)^{e_n}$$

But notice that $M_i - \lambda_i I$ is strictly upper triangular of size $e_i \times e_i$. It follows from (c) that $(M_i - \lambda_i I)^{e_i} = 0$. Hence the entire product (**) is 0.

This is true for each of the blocks in (*), so the entire product is 0.