THE CHINESE UNIVERSITY OF HONG KONG Department of Mathematics MATH2040A Solution to Homework 8

Compulsory Part

Sec. 5.4

- 4 Q: Let T be a linear operator on a vector space V, and let W be a T-invariant subspace of V. Prove that W is g(T)-invariant for any polynomial g(t).
 - Sol: Consider any polynomial g(t). \exists non-negative integer n and scalars $a_0, ..., a_n$ such that $g(t) = \sum_{i=1}^n a_i t^i$. Now fix $v \in W$. Note that $T^0(v) = \operatorname{Id}_V(v) = v \in W$. If k is a non-negative integer such that $T^k(v) \in W$, then $T^{k+1}(v) = T(T^k(v)) \in W$ because W is T-invariant. Hence, we have shown by mathematical induction that $T^i(v) \in W \forall$ non-negative integer i. Finally, $g(T)(v) = \sum_{i=1}^n a_i T^i(v) \in \operatorname{span}\{v, T(v), ..., Tn(v)\} \subset W$. To conclude, W is g(T)-invariant.
- 6 Q: For each linear operator T on the vector space V, find an ordered basis for the T-cyclic subspace generated by the vector z.

(b)
$$V = \mathsf{P}_3(\mathbb{R}), T(f(x)) = f''(x)$$
, and $z = x^3$.
(d) $V = \mathsf{M}_{2 \times 2}(\mathbb{R}), T(A) = \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix} A$, and $z = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.

Sol: (b)

$$T(z) = 6x;$$
 $T^{2}(z) = 0;$ $T^{3}(z) = 0.$

Then $(x^3, 6x)$ is an ordered basis for V.

(d)

$$T(z) = \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix}; \quad T^2(z) = \begin{pmatrix} 3 & 3 \\ 6 & 6 \end{pmatrix} = 3T(z).$$

Hence, the *T*-cyclic subspace *W* generated by *z* is $\text{span}\{z, T(z)\}$. If $a, b \in \mathbb{R}$ and az + bT(z) = O, then by comparing entries on both sides, we clearly see that a = b = 0. Hence, $\{z, T(z)\}$ is linearly independent. Therefore,

$$\{z, T(z)\} = \left\{ \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix} \right\}$$

is an ordered basis for W.

13 Q: Let T be a linear operator on a vector space V, let v be a nonzero vector in V, and let W be the T-cyclic subspace of V generated by v. For any $w \in V$, prove that $w \in W$ if and only if there exists a polynomial g(t) such that w = g(T)(v).

Sol: By the definition of *T*-cyclic subspace, if $w \in W$, then $\exists a_1, \dots, a_k$ and $i_1, \dots, i_k \in \{0, 1, 2, \dots\}$, s.t.

$$w = \sum_{j=1}^k a_j T^{i_j}(v).$$

Hence if we let $g(x) = \sum_{j=1}^{k} a_j x^{i_j}$, we have w = g(T)(v). If $w = g(T)(v) = \sum_{j=1}^{k} a_j x^j$, by the definition of *T*-cyclic subspace, we have $w \in W$.

17 Q: Let A be an $n \times n$ matrix. Prove that

$$\dim(\operatorname{span}(\{I_n, A, A^2 \cdots \})) \le n$$

Sol: Let $U = \text{span}(\{I, \ldots, A_{n-1}\})$. Then dim $U \leq n$. To show the proposition, we show that $\text{span}(\{I, A, \ldots\}) = U$. By definition, $\text{span}(\{I, A, \ldots\}) \supseteq U$. It then suffices to show that $A^k \in U$ for all $k \in \mathbb{N}$. The case where k < n is trivial from the definition of U.

Suppose there exists $l \ge n-1$ such that $I, A, \ldots, A^l \in U$. Let the characteristic polynomial of A be p(t). Then deg p = n. We may assume that $p(t) = \sum_{i=0}^{n} c_i t^i$ for some scalar c_0, \ldots, c_n with $c_n = (-1)^n$. By Cayley-Hamilton theorem, $p(A) = \sum_{i=0}^{n} c_i A^i = c_0 I + \ldots + c_n A^n = 0$. So $A^n = \sum_{i=0}^{n-1} -\frac{c_i}{c_n} A^i$, $A^{l+1} = A^{l-n+1}A^n = \sum_{i=0}^{n-1} -\frac{c_i}{c_n} A^{l-n+1+i} \in U$ as $A^{l-n+1}, \ldots, A^l \in U$. By induction, $A^k \in U$ for all $k \in \mathbb{N}$. So $\operatorname{span}(\{I, A, \ldots\}) = U$ and $\operatorname{dim}\operatorname{span}(\{I, A, \ldots\}) = \operatorname{dim} U \le n$.

- 23 Q: Let T be a linear operator on a finite-dimensional vector space V, and let W be a Tinvariant subspace of V. Suppose that $v_1, v_2, ..., v_k$ are eigenvectors of T corresponding to distinct eigenvalues. Prove that if $v_1 + v_2 + \cdots + v_k$ is in W, then $v_i \in W$ for all i.
 - Sol: We prove this statement by mathematical induction on k. The case for k = 1 is trivial. Assume that the statement is true for some positive integer k. Consider the statement for k+1. $\forall i \in \{1, ..., k+1\}$, let λ_i be the eigenvalue of T corresponding to the eigenvector v_i of T, i.e. $T(v_i) = \lambda_i v_i$. Let $w = v_1 + \cdots + v_k + v_{k+1} \in W$. As W is T-invariant, $T(w) = \lambda_1 v_1 + \cdots + \lambda_k v_k + \lambda_{k+1} v_{k+1} \in W$. We have

$$(\lambda_{k+1} - \lambda_1)v_1 + \dots + (\lambda_{k+1} - \lambda_k)v_k = \lambda_{k+1}w - T(w) \in W.$$

Since $\lambda_1, ..., \lambda_{k+1}$ are distinct, $\forall i \in \{1, ..., k\}$, $\lambda_{k+1} - \lambda_i \neq 0$ and thus $(\lambda_{k+1} - \lambda_i)v_i$ is an eigenvector of T corresponding to the eigenvalue λ_i . By induction hypothesis, $v_1, ..., v_k \in W$. Finally, $v_{k+1} = w - (v_1 + \cdots + v_k) \in W$. We are done.

Optional Part

Sec. 5.4

- 1 Q: Label the following statements as true or false.
 - (a) There exists a linear operator T with no T-invariant subspace.

- (b) If T is a linear operator on a finite-dimensional vector space V and W is a T-invariant subspace of V, then the characteristic polynomial of T_W divides the characteristic polynomial of T.
- (c) Let T be a linear operator on a finite-dimensional vector space V, and let v and w be in V. If W is the T-cyclic subspace generated by v, W' is the T-cyclic subspace generated by w, and W = W', then v = w.
- (d) If T is a linear operator on a finite-dimensional vector space V, then for any $v \in V$ the T-cyclic subspace generated by v is the same as the T-cyclic subspace generated by T(v).
- (e) Let T be a liner operator on an n-dimensional vector space. Then there exists a polynomial g(t) of degree n such that $g(T) = T_0$.
- (f) Any polynomial of degree n with leading coefficient $(-1)^n$ is the characteristic polynomial of some linear operator.
- (g) If T is a linear operator on a finite-dimensional vector space V, and V is the direct sum of k T-invariant subspaces, then there is an ordered basis β for V such that $[T]_{\beta}$ is a direct sum of k matrices.
- Sol: (a) False.
 - (b) True.
 - (c) False.
 - (d) False.
 - (e) True.
 - (f) True.
 - (g) True.
- 15 Q: Use the Cayley-Hamilton theorem (Theorem 5.23) to prove its corollary for matrices.

Sol: Let A be an $n \times n$ matrix, and let f(t) be the characteristic polynomial of A. We want to show that f(A) = O.

Consider the linear operator L_A . Then f(t) is the characteristic polynomial of L_A . By Cayley-Hamilton theorem (Theorem 5.23), $f(L_A)$ is the zero transformation. Then

$$f(L_A)(e_i) = f(A)e_i = \begin{pmatrix} 0\\ \vdots\\ 0 \end{pmatrix}.$$

for any $i \in \{1, ..., n\}$. So we have

$$f(A) = f(A) \begin{pmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{pmatrix} = O$$

as desired.

- 16 Q: Let T be a linear operator on a finite-dimensional vector space V.
 - (a) Prove that if the characteristic polynomial of T splits, then so does the characteristic polynomial of the restriction of T to any T-invariant subspace of V.

- (b) Deduce that if the characteristic polynomial of T splits, then any nontrivial T-invariant subspace of V contains an eigenvector of T.
- Sol: (a) Let W be a T-invariant subspace of V and let g(t) be the characteristic polynomial of T_W . By Theorem 5.21, g(t) divides the characteristic polynomial f(t) of T. As f(t) splits, so does g(t).
 - (b) We prove the statement by contradiction. Assume the contrary that there is a nontrivial *T*-invariant subspace *W* of *V* containing no eigenvectors of *T*. By (a) the characteristic polynomial g(t) of T_W splits, i.e. \exists scalars $a_1, ..., a_k$ such that

$$g(t) = (-1)^k (t - a_1) \cdots (t - a_k)$$

By Cayley-Hamilton theorem, $g(T_W)$ is the zero transformation. Fix $w \in W$. Then $(T_W - a_1 \operatorname{Id}_W) \cdots (T_W - a_k \operatorname{Id}_W)(w) = (-1)^k g(T_W)(w) = \vec{0}$. If $i \in \{1, ..., k\}$ and $(T_W - a_i \operatorname{Id}_W) \cdots (T_W - a_k \operatorname{Id}_W)(w) = \vec{0}$, then

$$(T_W - a_{i+1}\operatorname{Id}_W) \cdots (T_W - a_k\operatorname{Id}_W)(w) = \vec{0},$$

otherwise the left hand side of the above equality is an eigenvector of T_W corresponding to the eigenvalue a_i , where the expression $(T_W - a_{k+1} \operatorname{Id}_W) \cdots (T_W - a_k \operatorname{Id}_W)(w)$ denotes w by convention. By mathematical induction, $w = \vec{0}$. It leads to contradiction that W is a trivial subspace of V. We are done.

18 Q: Let A be an $n \times n$ matrix with characteristic polynomial

$$f(t) = (-1)^n t^n + a_{n-1} t^{n-1} + \dots + a_1 t + a_0.$$

- (a) Prove that A is invertible if and only if $a_0 \neq 0$.
- (b) Prove that if A is invertible, then

$$A^{-1} = (-1/a_0)[(-1)^n A^{n-1} + a_{n-1}A^{n-2} + \cdots + a_1I_n].$$

- Sol: (a) Note that $a_0 = f(0) \det(A 0I_n) = \det(A)$. Hence A is invertible if and only if $a_0 \neq 0$.
 - (b) By Cayley-Hamilton theorem, f(A) = O. By (a), $a_0 \neq 0$. Rearranging, we get

$$(-\frac{1}{a_0})[(-1)^n A^{n-1} + a_{n-1}A^{n-2} + \cdots + a_1I_n]A = I_n,$$

whence $A^{-1} = (-\frac{1}{a_0})[(-1)^n A^{n-1} + a_{n-1}A^{n-2} + \cdots + a_1I_n].$

19 Sol:

We show the proposition by induction on k.

For the case k = 1, the characteristic polynomial of the matrix $(-a_0)$ is $p(t) = -a_0 - t = (-1)^1(a_0 + t)$ for all scalar a_0 . So the proposition holds when k = 1.

Suppose for some $l \in \mathbb{Z}^+$ the proposition holds when k = l. Let a_0, \ldots, a_l be scalars,

and
$$A_{l+1} = \begin{pmatrix} 0 & 0 & \dots & 0 & -a_0 \\ 1 & 0 & \dots & 0 & -a_1 \\ 0 & 1 & \dots & 0 & -a_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & -a_l \end{pmatrix}$$
. Then the characteristic polynomial is

$$p(t) = \det(A_{l+1} - tI)$$

$$= \det\begin{pmatrix} -t & 0 & \dots & 0 & -a_0 \\ 1 & -t & \dots & 0 & -a_1 \\ 0 & 1 & \dots & 0 & -a_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & -a_l - t \end{pmatrix}$$

$$= -t \det\begin{pmatrix} -t & 0 & \dots & 0 & -a_1 \\ 1 & -t & \dots & 0 & -a_2 \\ 0 & 1 & \dots & 0 & -a_2 \\ 0 & 1 & \dots & 0 & -a_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & -a_l - t \end{pmatrix} + (-1)^l (-a_0) \det\begin{pmatrix} 1 & -t & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & -a_l - t \end{pmatrix}$$

$$= -t(-1)^l (a_1 + \dots + a_l t^{l-1} + t^l) - (-1)^l a_0$$

$$= (-1)^{l+1} (a_0 + a_1 t + \dots + a_l t^l + t^{l+1})$$

So the proposition holds when k = l + 1. By induction, the proposition holds for all k.

So det
$$\begin{pmatrix} 0 & 0 & \dots & 0 & -a_0 \\ 1 & 0 & \dots & 0 & -a_1 \\ 0 & 1 & \dots & 0 & -a_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & -a_{k-1} \end{pmatrix} = (-1)^k (a_0 + a_1 t + \dots + a_{k-1} t^{k-1} + t^k)$$
 for all k and all

scalars a_0, \ldots, a_{k-1} .

- 24 Q: Prove that the restriction of a diagonalizable linear operator T to any nontrivial T-invariant subspace is also diagonalizable.
 - Sol: Let W be a nontrivial T-invariant subspace of the domain V of T. Note that V is finite-dimensional.

Let $\lambda_1, ..., \lambda_k$ be all the distinct eigenvalues of T with respective eigenspaces $\mathsf{E}_{\lambda_1}, ..., \mathsf{E}_{\lambda_k}$. Since T is diagonalizable, we have by Theorem 5.11

$$V = \mathsf{E}_{\lambda_1} \oplus \cdots \oplus \mathsf{E}_{\lambda_k}.$$

Pick a finite subset $\{w_1, ..., w_n\}$ of W such that $W = \text{span}\{w_1, ..., w_n\}$ (say, a basis for W). $\forall i \in \{1, ..., n\}, \exists v_{i,1} \in \mathsf{E}_{\lambda_1}, ..., v_{i,k} \in \mathsf{E}_{\lambda_k}$ such that

$$w_i = v_{i,1} + \dots + v_{i,k} \in W_i$$

and therefore, by Q23, Sec. 5.4, $v_{i,1}, ..., v_{i,k} \in W$. We have

 $W = \operatorname{span}\{v_{1,1}, \dots, v_{1,k}, \dots, v_{n,1}, \dots, v_{n,k}\}.$

Then \exists ordered basis β for W such that every element in β is an eigenvector of T. Then $[T_W]_{\beta}$ is a diagonal matrix. T_W is therefore diagonalizable.

- 25 Q: (a) Prove the converse to Exercise 18(a) of Section 5.2: If T and U are diagonalizable linear operators on a finite-dimensional vector space V such that UT = TU, then T and U are simultaneously diagonalizable.
 - (b) State and prove a matrix version of (a).
 - Sol: (a) Let $\lambda_1, ..., \lambda_k$ be all the distinct eigenvalues of T with respective eigenspaces $\mathsf{E}_{\lambda_1}, ..., \mathsf{E}_{\lambda_k}$. Since T is diagonalizable, we have by Theorem 5.11

$$V = \mathsf{E}_{\lambda_1} \oplus \cdots \oplus \mathsf{E}_{\lambda_k}$$

Fix $i \in \{1, ..., k\}$. We claim that E_{λ_i} is U-invariant. $\forall v \in \mathsf{E}_{\lambda_i}$, as

$$T(U(v)) = U(T(v)) = U(\lambda_i v) = \lambda_i U(v),$$

 $U(v) \in \mathsf{E}_{\lambda_i}$. We have proved our claim. Now because U is diagonalizable, by (24) $U_{\mathsf{E}_{\lambda_i}}$ is also diagonalizable. Then \exists ordered basis β_i for E_{λ_i} such that $[U_{\mathsf{E}_{\lambda_i}}]_{\beta_i}$ is a diagonal matrix. In addition,

$$[T]_{\beta_i} = \begin{pmatrix} \lambda_i & \cdots & 0\\ \vdots & \ddots & \vdots\\ 0 & \cdots & \lambda_i \end{pmatrix}$$

is also a diagonal matrix. Let $\beta = \beta_1 \cup \cdots \cup \beta_k$. We now see that

$$[T]_{\beta} = [T_{\mathsf{E}_{\lambda_1}}]_{\beta_1} \oplus \cdots \oplus [T_{\mathsf{E}_{\lambda_k}}]_{\beta_k}, [U]_{\beta} \qquad = [U_{\mathsf{E}_{\lambda_1}}]_{\beta_1} \oplus \cdots \oplus [U_{\mathsf{E}_{\lambda_k}}]_{\beta_k}$$

are direct sum of diagonal matrices. Then obviously $[T]_{\beta}$ and $[U]_{\beta}$ are also diagonal matrices. Therefore, T, U are simultaneously diagonalizable.

(b) We shall prove that: If A and B are diagonalizable matrices such that AB = BA, then A and B are simultaneously diagonalizable.

Since A, B are diagonalizable, then linear operators L_A, L_B are diagonalizable. Also, as AB = BA, $L_A L_B = L_{AB} = L_{BA} = L_B L_A$. Then by (a), L_A, L_B are simultaneously diagonalizable, i.e. \exists ordered basis β for the common domain of L_A and L_B such that $[L_A]_{\beta}, [L_B]_{\beta}$ are diagonal matrices. Then \exists invertible matrix Q of the same size as A and B such that $Q^{-1}AQ = [L_A]_{\beta}$ and $Q^{-1}BA = [L_B]_{\beta}$. Hence, A, B are simultaneously diagonalizable.