Solution to Homework 5

Sec. 5.4

2. (e) No. Note that $A = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$ is in W, but

$$T(A) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 0 & 2 \\ 1 & 0 \end{pmatrix}$$

is not in W. Hence, W is not a T-invariant subspace of V.

3. Note that $T : V \to V$ is a linear operator on V. To check that W is a T-invariant subspace of V, we need to know if $T(w) \in W$ for any $w \in W$.

   (a) Since we have

$$T(0) = 0 \in \{0\} \quad \text{and} \quad T(v) \in V,$$

   so both of $\{0\}$ and V to be T-invariant subspaces of V.

   (b) Note that $0 \in N(T)$. For any $u \in N(T)$, we have

$$T(u) = 0 \in N(T).$$

   Hence, $N(T)$ is a T-invariant subspace of V.
   For any $v \in R(T)$, as $R(T) \subset V$, we have $v \in V$. So, by definition,

$$T(v) \in R(T).$$

   Hence, $R(T)$ is also a T-invariant subspace of V.

   (c) Note that for any $v \in E_\lambda$, λv is a scalar multiple of v, so $\lambda v \in E_\lambda$ as E_λ is a subspace. So we have

$$T(v) = \lambda v \in E_\lambda.$$

   Hence, E_λ is a T-invariant subspace of V.

4. For any w in W, we know that $T(w)$ is in W as W is a T-invariant subspace of V. Then, by induction, we know that $T^k(w)$ is also in W for any k. Suppose $g(T) = a_k T^k + \cdots + a_1 T + a_0$, we have

$$g(T)(w) = a_k T^k(w) + \cdots + a_1 T(w) + a_0(w) \in W$$

   because it is just a linear combination of elements in W.

   Hence, W is a $g(T)$-invariant subspace of V.

6. (d) To find an ordered basis for the T-cyclic subspace generated by the vector z, we check if $T^k(z)$ is spanned by $\{z, T(z), \ldots, T^{k-1}(z)\}$. By direct calculation, we have the following.

$$\begin{aligned} z &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\ T(z) &= \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix} \\ T^2(z) = T(T(z)) &= \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix} = \begin{pmatrix} 3 & 3 \\ 6 & 6 \end{pmatrix} \end{aligned}$$

We see that $T^2(z) = 3T(z)$, so the dimension of the T-cyclic subspace generated by z is just 2 and $\{z, T(z)\}$, which is $\left\{ \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix} \right\}$, is a basis for the subspace.

18. (a) Note that $f(t) = \det(A - tI)$ and

$$a_0 = f(0) = \det(A).$$

Hence, A is invertible if and only if $a_0 \neq 0$.

(b) Suppose A is invertible. By (a), we have $a_0 \neq 0$. By Cayley-Hamilton Theorem, we have $f(A) = O$, where O is the $n \times n$ zero matrix.

$$\begin{aligned} f(A) = (-1)^n A^n + a_{n-1}A^{n-1} + \cdots + a_1 A + a_0 I &= O \\ (-1)^n A^n + a_{n-1}A^{n-1} + \cdots + a_1 A &= -a_0 I \\ A\left((-1)^n A^{n-1} + a_{n-1}A^{n-2} + \cdots + a_1 I\right) &= -a_0 I \\ -\frac{1}{a_0}\left((-1)^n A^{n-1} + a_{n-1}A^{n-2} + \cdots + a_1 I\right) &= A^{-1} \end{aligned}$$

(c) To use the result of (b), we need to find the characteristic polynomial of A and compute A^2, A^1 and I (as n is now 3).

$$\det(A - tI) = (1-t)(2-t)(-1-t) = -t^3 + 2t^2 + t - 2$$

$$I = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad A = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 2 & 3 \\ 0 & 0 & -1 \end{pmatrix}, \quad A^2 = \begin{pmatrix} 1 & 6 & 6 \\ 0 & 4 & 3 \\ 0 & 0 & 1 \end{pmatrix},$$

By the formula, we can compute A^{-1} directly.

$$A^{-1} = -\frac{1}{(-2)}\left[-A^2 + 2A + I\right] = \frac{1}{2}\begin{pmatrix} 2 & -2 & -4 \\ 0 & 1 & 3 \\ 0 & 0 & -2 \end{pmatrix}$$

19. For 1×1 matrices, the statement obviously holds.

$$\det(A - tI) = -a_0 - t = (-1)^1(a_0 + t^1)$$

Suppose the statement holds for $(m-1)\times(m-1)$ matrices, we want to show that the statement is also true for $m \times m$ matrices. For an $m \times m$ matrix A, we find the characteristic polynomial by expanding the determinant along the first row.

$$\begin{aligned}
\det(A - tI) &= \begin{pmatrix} -t & 0 & \cdots & 0 & 0 & -a_0 \\ 1 & -t & \cdots & 0 & 0 & -a_1 \\ 0 & 1 & \cdots & 0 & 0 & -a_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -t & -a_{m-2} \\ 0 & 0 & \cdots & 0 & 1 & -a_{m-1} - t \end{pmatrix} \\
&= (-t)\underbrace{\det\begin{pmatrix} -t & \cdots & 0 & 0 & -a_1 \\ 1 & \cdots & 0 & 0 & -a_2 \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & \cdots & 1 & -t & -a_{m-2} \\ 0 & \cdots & 0 & 1 & -a_{m-1} - t \end{pmatrix}}_{(-1)^{m-1}(a_1 + a_2 t + \cdots + a_{m-1}t^{m-2} + t^{m-1})} \\
&\quad (-a_0)(-1)^{k+1}\underbrace{\det\begin{pmatrix} 1 & -t & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -t \\ 0 & 0 & \cdots & 0 & 1 \end{pmatrix}}_{1} \\
&= (-t)\left((-1)^{m-1}\left(a_1 + a_2 t + \cdots + a_{m-1}t^{m-2} + t^{m-1}\right)\right) \\
&\quad + (-1)^k a_0 \\
&= (-1)^m(a_0 + a_1 t + \cdots + a_{m-1}t^{m-1} + t^m)
\end{aligned}$$

Hence, by induction, the statement holds for any $k \times k$ matrices.

23. Let's prove the statement by induction on k. When $k = 1$, it is obviously true.

Assume the statement hold for $k = m-1$. When $k = m$, if $v_1 + v_2 + \cdots + v_m$ is in W, we want to show that $v_i \in W$ for all i.

Note that W is a T-invariant subspace of V, write $u = v_1 + v_2 + \cdots + v_m$, we have $T(u) \in W$, that is

$$\lambda_1 v_1 + \lambda_2 v_2 + \cdot + \lambda_m v_m \in W.$$

For a particular v_i, we know that $\lambda_i u \in W$ as u is in W. Consider $T(u) - \lambda_i u$, note that $T(u) - \lambda_i u \in W$ as $T(u)$ and $\lambda_i u$ are in W, that is

$$(\lambda_1 - \lambda_i)v_1 + (\lambda_2 - \lambda_i)v_2 + \cdots + (\lambda_k - \lambda_i)v_m \in W$$

As eigenvalues are distinct, $\lambda_j - \lambda_i \neq 0$ for each $j \neq i$. So each term is a non-zero multiple of an eigenvector of T, which is also an eigenvector. Now that their sum is in W, using the induction hypothesis, each $(\lambda_j - \lambda_i)v_j$ is in W, so is each v_j. Lastly, this particular v_i is also in W.

$$v_i = u - \sum_{j \neq i} v_j \in W$$

Hence, by induction, the statement is true.

24. Suppose T is a diagonalizable linear operator on V. Then V is a direct sum of eigenspaces corresponding to distinct eigenvalues.

$$V = E_{\lambda_1} \oplus E_{\lambda_2} \oplus \cdots \oplus E_{\lambda_k}$$

Let T_W be the restriction of T on W. Consider $W_{\lambda_i} = W \cap E_{\lambda_i}$. For every $w \in W$, note that w is also in V, so w can be expressed as a linear combination of eigenvectors of distinct eigenvalues.

$$w = a_1 u_1 + a_2 u_2 + \cdots + a_k u_k, \text{ where } u_i \in E_{\lambda_i}$$

So u_i is in both W and E_{λ_i}, that is, $u_i \in W \cap E_{\lambda_i} = W_{\lambda_i}$. Hence, W is a direct sum of W_{λ_i}.

$$W = W_{\lambda_1} \oplus W_{\lambda_2} \oplus \cdots \oplus W_{\lambda_k}$$

But for each W_{λ_i}, we can find a basis β_{λ_i} for W_{λ_i}. Then $\bigcup_i \beta_{\lambda_i}$ will be a basis consisting of eigenvectors T_W. In other words, T_W is diagonalizable.

25. (a) For any eigenvalue λ_i of T and $v \in E_{\lambda_i}$, we have the following.

$$TU(v) = UT(v) = \lambda_i U(v)$$

So we see that $U(v)$ is an eigenvector of T, that is, $U(v) \in E_{\lambda_i}$. That means E_{λ_i} is an U-invariant subspace of V. Then, by the above exercise, the restriction of U on E_{λ_i} is digonalizable. In other words, we may find a basis β_{λ_i} consisting of eigenvectors of restriction of U on E_{λ_i}. Now consider the union β of all these β_{λ_i}.

$$\beta = \bigcup_i \beta_{\lambda_i}$$

Note that it consists of eigenvectors of both T and U as β are eigenvectors from E_{λ_i} and β_{λ_i} are eigenvectors of U by construction. Hence, T and U are simultaneously diagonalizable.

(b) *"If A and B are diagonalizable matrices of dimension $n \times n$ such that $BA = AB$, then A and B are simultaneously diagonalizable."*

To prove the statement, we consider the linear transformations L_A and L_B. As A and B are diagonalizable, L_A and L_B are also diagonalizable. Note that $L_B L_A = L_A L_B$, by (a), we know that L_A and L_B are simultaneously diagonalizable. By previous exercise (Sec. 5.2 Exercise 17), we know that A and B are simultaneously diagonalizable.