

EXERCISES 7.2: SOLUTIONS

General notations: For a linear operator $T \in \text{End}(V)$, μ_T denotes its minimal polynomial and χ_T denotes its characteristic polynomial. For a T -invariant subspace $W \subset V$, the notation $S_T(\alpha; W)$ denotes the ideal $\{f \in F[x] : f(T)\alpha \in W\}$, which is called the conductor of α into W . In particular, if $W = 0$, $S_T(\alpha; 0) = \{f : f(T)\alpha = 0\}$. This is the T -annihilator of α , and it is also denoted by $M(\alpha; T)$ in §7.1. Let $I(T) = \cap_{\alpha \in V} S_T(\alpha; 0) = \{f \in F[x] : f(T)\alpha = 0, \forall \alpha \in V\}$. Note that μ_T is the monic generator of $I(T)$.

Exercise 2: Let $T : V \rightarrow V$ be a linear operator on a finite dimensional vector space V . Let R be the range of T and N be the null space of T . (a) Prove that R has a complementary T -invariant subspace if and only if R is independent of N . (b) If R and N are independent, prove that N is the unique T -invariant subspace complementary to R .

Proof. (a) The dimension theorem says that $\dim R + \dim N = \dim V$. If R and N are independent, we have $R \cap N = \{0\}$ and thus $\dim(R + N) = \dim R + \dim(N) = \dim(V)$, by dimension theorem. Thus $V = R + N$. Since $N \cap R = 0$, we get $V = R \oplus N$. Note that N is clearly T -invariant. Thus R has a T -invariant complementary subspace. Conversely, suppose that R has a T -invariant complementary subspace, and thus R is admissible. For any $\alpha \in V$, we have $T\alpha \in R$. The admissibility shows that there exists a $\beta \in R$ such that $T\alpha = T\beta$. Thus $\alpha - \beta \in N$. The equation $\alpha = \beta + \alpha - \beta$ implies that $V = R + N$. This means that $\dim(R \cap N) = \dim R + \dim N - \dim(R + N) = 0$. Thus $R \cap N = \{0\}$.

(b) Suppose that $V = R \oplus W$ for a T -invariant subspace $W \subset V$. We will show that $W = N$. Take $\alpha \in W$, we have $T\alpha \in W$ since W is T -invariant. On the other hand, $T\alpha \in R$ by definition. Thus $T\alpha \in R \cap W = \{0\}$, which implies that $T\alpha = 0$ and $\alpha \in N$. Thus $W \subset N$. On the other hand, we know that $\dim W = \dim V - \dim R = \dim N$. We must have $W = N$. \square

Exercise 8: Let $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ be the linear operator given by the matrix

$$\begin{bmatrix} 3 & -4 & -4 \\ -1 & 3 & 2 \\ 2 & -4 & -3 \end{bmatrix}.$$

Find nonzero vectors $\alpha_1, \dots, \alpha_r$ satisfying the conditions of Theorem 3.

Proof. We can compute that $\chi_T = (x-1)^3$ and $\mu_T = (x-1)^2$. Thus we have $V = Z(\alpha_1; T) \oplus Z(\alpha_2; T)$ and $p_1 = (x-1)^2, p_2 = (x-1)$. Note that α_2 is an eigenvector of 1, α_1 is in $\ker(p_1(T))$ but not an eigenvector of 1, but $Z(\alpha_1; T)$ contains an eigenvector of 1. Since $\dim Z(\alpha_1; T) = 2$, we have $T\alpha_1 \neq \alpha_1$, but $(T-I)^2\alpha_1 = 0$. We first compute the eigenspace of 1, namely, $E_T(1) = \ker(T-I)$. A simple calculation shows that

$$E_T(1) = \left\{ \begin{bmatrix} 2y+2z \\ y \\ z \end{bmatrix} : y, z \in \mathbb{R} \right\}.$$

Since $(T-I)^2 = 0$, α_1 can be taken as any vector with $\alpha_1 \notin E_T(1)$. For example, we can take $\alpha_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$. In this case $(T-I)\alpha_1 = \begin{bmatrix} 2 \\ -1 \\ 2 \end{bmatrix}$. The vector α_2 can be taken as any vector in $E_T(1)$

which is not proportional to $(T-I)\alpha_1$. For example, we can take $\alpha_2 = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}$. The choices of α_1, α_2 are not unique. \square

Exercise 9: Let A be the real matrix

$$A = \begin{bmatrix} 1 & 3 & 3 \\ 3 & 1 & 3 \\ -3 & -3 & -5 \end{bmatrix}.$$

Find an invertible real matrix $P \in \text{GL}_3(\mathbb{R})$ such that $P^{-1}AP$ is in rational form.

Proof. Let $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ be the linear operator defined by A . We can compute the characteristic polynomial of T is $\chi_T = (x+2)^2(x-1)$ and its minimal polynomial is $\mu_T = (x+2)(x-1)$. We have $V = Z(\alpha_1; T) \oplus Z(\alpha_2; T)$, with $p_1 = (x+2)(x-1)$ and $p_2 = x+2$. Similar as the last problem, we can take α_1 arbitrary other than eigenvectors of 1 or -2 , and α_2 an eigenvector of -2 . Take

$$\alpha_1 = [1, 0, 0]^T, T\alpha_1 = [1, 3, -3]^T; \alpha_2 = [1, -1, 0]^T,$$

and

$$P = [\alpha_1, T\alpha_1, \alpha_2] = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 3 & -1 \\ 0 & -3 & 0 \end{bmatrix}.$$

Then we have

$$AP = P \begin{bmatrix} 0 & 2 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & -2 \end{bmatrix}.$$

Again, the choice of P is not unique. □

Exercise 11: Prove that if A and B are 3×3 matrices over the field F , A is similar to B if and only if they have the same characteristic polynomial and the same minimal polynomial. Give an example which shows that this is false for 4×4 matrices.

Proof. If A and B are similar, then clearly they have the same characteristic and minimal polynomial (for the minimal polynomial part, it is easy to check $I(T_A) = I(T_B)$. A different argument is: A, B represent the same linear operators with different choice of basis). Now suppose that $A, B \in \text{Mat}_{3 \times 3}(F)$ such that $\chi_A = \chi_B$ and $\mu_A = \mu_B$. To show that A and B are similar, it suffices to show that A and B have the same invariant factors. We know that $\deg \chi_A = 3$ and we discuss degree of μ_A . If $\deg(\mu_A) = 3$, then $\mu_A = \chi_A$, and thus A has only a single invariant factor, which is μ_A . The same is true for B . The assumption shows that A, B have the same invariant factors. Next, we assume that $\deg(\mu_A) = 2$. In this case, $\chi_A = \mu_A q_A$ for a degree one factor q_A and the invariant factors of A are $\{\mu_A, q_A = \chi_A/\mu_A\}$. Again, the assumption shows that A and B have the same invariant factors. Finally, assume that $\deg(\mu_A) = 1$. Assume that $\mu_A = (x - a)$ for some $a \in F$. This implies that $A - aI_3 = 0$ and thus $A = aI_3$. Since $\mu_B = \mu_A$, we also have $B = aI_3$. Thus $A = B$ in this case.

In the 4×4 case, we can take A such that its invariant factors are x^2, x, x and take B such that its invariant factors are x^2, x^2 . Note that $\mu_A = \mu_B = x^2, \chi_A = \chi_B = x^4$. But A and B are not similar, because they have different invariant factors. Such matrices can be realized by

$$A = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

□

Exercise 12: Let F be a subfield of the field of complex numbers, and let $A, B \in \text{Mat}_{n \times n}(F)$. Prove that A and B are similar over the field of complex numbers, then they are similar over F .

We did not talk about how linear algebra behaves under field extension. Here we prove some simple useful facts regarding this problem. In the following, K is a field and F is a subfield of K , which means F is a subset of K and together with the addition and multiplication defined on K , F is also a field. You can think $K = \mathbb{C}$, F is either \mathbb{Q} or \mathbb{R} ; or $K = \{a + b\alpha + c\alpha^2 : \alpha = \sqrt[3]{2}, a, b, c \in \mathbb{Q}\}$, $F = \mathbb{Q}$.

Lemma 1. Let $A \in \text{Mat}_{m \times n}(F)$. If $Ax = 0$ has a nonzero solution $x \in K^n$, then $Ax = 0$ has a nonzero solution in F^n . Moreover, we have $\dim_K \{x \in K^n : Ax = 0\} = \dim_F \{x \in F^n : Ax = 0\}$.

Note that $F \subset K$, it is natural to view A as an element in $\text{Mat}_{m \times n}(K)$ and thus we can talk about solutions of $Ax = 0$ in K^n .

Proof. Let $R \in \text{Mat}_{m \times n}(F)$ be the row reduced echelon form of A . The key observation is when R is viewed as an element in $\text{Mat}_{m \times n}(K)$, it is still in row reduced echelon form. Note that $Ax = 0$ has a nonzero solution in K^n iff $Rx = 0$ has a nonzero solution in K^n iff the number of leading ones in R is less than n , or $\text{rank}(R) < n$. Thus $Ax = 0$ has a nonzero solution in F^n . Actually, the key observation shows that $\text{rank}_F(A) = \text{rank}_K(A)$, where $\text{rank}_K(A)$ denotes the rank of A when it is viewed as a matrix over K . The “moreover” part follows from

$$\dim_K \{x \in K^n : Ax = 0\} = n - \text{rank}(R) = \dim_F \{x \in F^n : Ax = 0\}.$$

□

Remark 2. The above proof used the fact that: after elementary row operations, every matrix A can be reduced to an elementary row echelon form R , and the linear system $Ax = 0$ is equivalent to $Rx = 0$. In particular, the elementary operation $R_i \rightarrow cR_i$ (replacing a row by c times this row) for $c \neq 0$ is invertible. This is a property of *field*. Think about the following example. Let $K = \mathbb{Z}/6\mathbb{Z}$, which consists of elements \bar{k} for $0 \leq k \leq 5$ and $k \in \mathbb{Z}$. Here $\bar{k} = k + 6\mathbb{Z}$ denotes the equivalence class. Consider its subset $F = \{\bar{0}, \bar{3}\} \subset K$. It should be easy to see that F is a field with the usual operations. In fact, $F = \mathbb{F}_2$, which is field consisting 2 elements. Note that K is not a field because $\bar{3}, \bar{2} \in K$ are nonzero, but $\bar{3} \cdot \bar{2} = \bar{0}$. Now consider the linear equation

$$x + x + x = 0.$$

Note that the above equation has a nontrivial solution $x = \bar{2}$ over K , but it does not have nontrivial solution over F . If you tried to go through the above proof, you will find that the main issue here is: while 3 is nonzero in K , it is not invertible in K .

Remark 3. In the terminology you will learn later, Lemma 1 can be restated as follows:

$$\ker(T_A) \otimes_F K = \ker(T_A \otimes_F K),$$

where $T_A : F^n \rightarrow F^m$ is the usual linear map defined by A and $T_A \otimes_F K$ is the linear map $F^n \otimes_F K = K^n \rightarrow K^m = F^m \otimes_F K$. In other words, the short sequence

$$0 \rightarrow \ker(T_A) \otimes_F K \rightarrow K^n \rightarrow K^m$$

is still exact. This reflects the fact that K is a *flat* F -module.

Lemma 4. Let $S = \{\alpha_1, \dots, \alpha_r\} \in F^n$. If S is linearly dependent over K , then it is also linearly dependent over F .

Since $F^n \subset K^n$, S can be viewed as a subset of K^n and thus we can consider linearly dependence of S over K .

Proof. Let A be the matrix $A = [\alpha_1, \dots, \alpha_r] \in \text{Mat}_{n \times r}(F) \subset \text{Mat}_{n \times r}(K)$. The assumption says that $Ax = 0$ has a nonzero solution in K^r . By Lemma 1, $Ax = 0$ has a nontrivial solution in F^r , which is equivalent to say that S is linearly dependent over F . □

First proof of Exercise 12. In this proof, we assume that the characteristic of K is zero, which is true if $K = \mathbb{C}$ as in the assumption of Ex 12. Later, we will see that this assumption is unnecessary. Let $V_K = \{X \in M_{n \times n}(K) : AX = XB\}$ and $V_F = \{X \in M_{n \times n}(F) : AX = XB\}$. The assumption says that V_K is not the zero space. Thus by Lemma 1, $\dim_F V_F = \dim_K V_K \geq 1$. Let $\mathcal{B} = \{\alpha_1, \dots, \alpha_k \in V_F\}$ be an F -basis of V_F . By Lemma 4, $\alpha_1, \dots, \alpha_k$ are also linearly independent over K . Let $W = \left\{ \sum_{i=1}^k c_i \alpha_i : c_i \in K \right\}$ be the K -span of \mathcal{B} . Then $W \subset V_K$ and $\dim_K W = k \geq 1$. Lemma 1 says that $\dim_F V_F = \dim_K V_K$, and thus we have $W = V_K$ by counting dimension. We need to show there exists a matrix $Q \in V_F$ such that $\det(Q) \neq 0$.

Consider the determinant function $\det : M_{n \times n}(K) \rightarrow K$ and restrict it to V_K . The assumption says that there exists a matrix $P \in V_K$ such that $\det(P) \neq 0$. For a general element $X = \sum_{i=1}^k x_i \alpha_i$ with $x_i \in F, \alpha_i \in \mathcal{B}$, a general fact says that $\det(X) = \det(\sum_{i=1}^k x_i \alpha_i)$ is a polynomial f on the variables x_1, \dots, x_k , whose coefficients in F . In other words, $f \in F[x_1, \dots, x_k]$. A very special case

is when $k = 1$ and in this case, $\det(x_1\alpha_1) = \det(\alpha_1)x_1^n$. The assumption says that there exists $x_1, \dots, x_k \in K$ such that $f(x_1, \dots, x_k) \neq 0$, and thus this polynomial f is nonzero. Since F has characteristic zero, there must be $y_1, \dots, y_k \in F$ such that $f(y_1, \dots, y_k) \neq 0$ (see Theorem 3, page 126 for this fact when there is only one variable). Note that $Q = \sum_{i=1}^k y_i \alpha_i \in V_F$ and $\det(Q) \neq 0$. We are done. \square

Remark 5. The above proof used some facts on determinant and polynomials of several variables. Moreover, it only works when characteristic of F is zero. See the following for a proof which works for more general situations.

Lemma 6. *Let $A \in \text{Mat}_{n \times n}(F)$, and let $\mu_{A,F}$ (resp. $\mu_{A,K}$) denote the minimal polynomial of A when viewed as a matrix over F (resp. over K). Then $\mu_{A,F} = \mu_{A,K}$.*

This fact is proved in page 192, but we did not cover the proof in class.

Proof. Denote $I(A, F) = \{f \in F[x] : f(A) = 0\}$ and $I(A, K) = \{f \in K[x] : f(A) = 0\}$. Then by definition $I(A, F) = \mu_{A,F}F[x]$, $I(A, K) = \mu_{A,K}K[x]$. Note that $\mu_{A,F} \in I(A, K)$ since $\mu_{A,F}(A) = 0$ and $\mu_{A,F} \in F[x] \subset K[x]$. This shows that $\mu_{A,K} | \mu_{A,F}$. Suppose that $\deg(\mu_{A,K}) = r$, then

$$S = \{I, A, \dots, A^r\}$$

is linearly dependent over K . Thus Lemma 4 shows that S is also linearly dependent over F . This shows that A satisfies a polynomial $f \in F[x]$ with $\deg(f) = r$. This shows $\deg(\mu_{A,F}) \leq r = \deg(\mu_{A,K})$. This condition plus $\mu_{A,K} | \mu_{A,F}$ imply that $\mu_{A,K} = \mu_{A,F}$. \square

Second proof of Exercise 12. Actually, the complex field \mathbb{C} can be replaced by any field K such that $F \subset K$. In the following argument, we just replace \mathbb{C} by K . We first show that the rational form for A is the same whether A is viewed as a matrix over F or over K . We consider the cyclic decomposition of $T : F^n \rightarrow F^n$, where $Tx = Ax$. We have

$$F^n = Z(\alpha_1; T; F) \oplus \dots \oplus Z(\alpha_r; T; F),$$

with invariant factors $p_1, p_2, \dots, p_r \in F[x]$, $p_i | p_{i-1}$, where $Z(\alpha_i; T; F) = \{f(T)\alpha_i : f \in F[x]\}$. Thus the canonical rational form of A (as a matrix in $M_{n \times n}(F)$) is

$$R = \begin{bmatrix} A_1 & & \\ & \ddots & \\ & & A_r \end{bmatrix},$$

where A_i is the companion matrix of p_i .

Let $T_i : Z(\alpha_i; T; F) \rightarrow Z(\alpha_i; T; F)$ be the restriction of T to $Z(\alpha_i; T; F)$. By Theorem 1 of page 228, p_i is the minimal polynomial of T_i , namely, $p_i = \mu_{T_i, F}$. Here we add an F in the subscript to emphasize that everything is viewed as an F -vector space. By Lemma 6, we also have $p_i = \mu_{T_i, K}$, namely p_i is the minimal polynomial of $T_i : Z(\alpha_i; T; K) \rightarrow Z(\alpha_i; T; K)$, when T_i is viewed as a linear operators of K -vector space. In particular, this shows that

$$\dim_K Z(\alpha_i; T; K) = \deg p_i = \dim_F Z(\alpha_i; T; F).$$

Assume that $\deg(p_i) = d_i$. Consider the basis $\mathcal{B}_i = \{\alpha_i, T\alpha_i, \dots, T^{d_i-1}\alpha_i\}$ of $Z(\alpha_i; T; F)$. Note that $\mathcal{B}_i \subset Z(\alpha_i; T; K)$, and by Lemma 4, \mathcal{B}_i is linearly independent over K . Since $\dim_K Z(\alpha_i; T; K) = d_i$, \mathcal{B}_i is also a K -basis of $Z(\alpha_i; T; K)$. Now consider $\mathcal{B} = \{\mathcal{B}_1, \dots, \mathcal{B}_r\}$, which is an F -basis of F^n by Lemma page 209. Since the set \mathcal{B} is linearly independent over F , it's linearly independent over K by Lemma 4 again. Since $|\mathcal{B}| = n$, $\mathcal{B} \subset F^n \subset K^n$ and \mathcal{B} is K -linearly independent, we get that

$$K^n = Z(\alpha_1; T; K) \oplus \dots \oplus Z(\alpha_r; T; K)$$

by Lemma page 209. Thus the above is indeed the cyclic decomposition of K^n by the uniqueness part of Theorem 3, page 233, and the invariant factors are still p_1, \dots, p_r . Thus the rational form of A (when viewed as a matrix in $\text{Mat}_{n \times n}(K)$) is still R .

Now suppose that $A, B \in \text{Mat}_{n \times n}(F)$ such that A and B are similar over K . This means that the rational form of A over K is the same as the rational form of B over K . By the above discussion, the rational forms of A, B over F are also the same. Thus A and B are similar over F . \square

The above proof is very complicate. Using Corollary of page 260, the proof can be greatly simplified. To do this, we prove the following

Lemma 7. *If $f, g \in F[x] \subset K[x]$. Write $\gcd_F(f, g)$ (resp. $\gcd_K(f, g)$) the gcd of f, g when they are viewed as elements of $F[x]$ (resp. of $K[x]$). Then*

$$\gcd_F(f, g) = \gcd_K(f, g).$$

This was a previous HW problem.

Proof. Suppose that $d_F = \gcd_F(f, g)$ and $d_K = \gcd_K(f, g)$. Recall that this means $d_F F[x] = fF[x] + gF[x]$ and $d_K K[x] = fK[x] + gK[x]$. Since there exists $f_1, g_1 \in F[x]$ with $d_F = ff_1 + gg_1$, and $ff_1 + gg_1 \in fK[x] + gK[x] = d_K K[x]$, we get $d_K | d_F$.

On the other hand, $d_F | f$ and $d_F | g$ in $F[x]$. Thus there exists $f', g' \in F[x]$ such that $f = d_F f', g = d_F g'$. By definition of d_K , there exists $f_2, g_2 \in K[x]$ such that $d_K = ff_2 + gg_2 = d_F(f'f_2 + g'g_2)$. Thus $d_F | d_K$. We are done. \square

Proof of Exercise 12 using Theorems in Section 7.4. Let $M = xI - A \in \text{Mat}_{n \times n}(F[x]) \subset \text{Mat}_{n \times n}(K[x])$ and let $\delta_k(M; F)$ (resp. $\delta_k(M; K)$) be the greatest common divisors of determinants of all $k \times k$ submatrices of M when viewed as a matrix over F (resp. over K). Let $p_1(F), \dots, p_r(F)$ be the invariant factors of A when viewed as a matrix over F . Similarly, we define $p_i(K)$. Section 7.4 told us that $p_i(F)$ can be computed using $\delta_k(M; F) / \delta_{k-1}(M; F)$ $1 \leq k \leq n$. Since gcd are independent of field extension by last lemma, we get $p_i(F) = p_i(K)$. This shows that the rational form of A is independent of the field we consider. \square

Comment: If you learn a little bit more algebra, you will find that the above proof can be greatly simplified further. In fact, since K is flat over F , for $p \in F[x]$ we have

$$(0.1) \quad (F[x]/pF[x]) \otimes_F K = K[x]/pK[x].$$

The cyclic decomposition of F^n is

$$\begin{aligned} F^n &= Z(\alpha_1; T; F) \oplus \cdots \oplus Z(\alpha_r; T; F) \\ &= F[x]/p_1 F[x] \times \cdots \times F[x]/p_r F[x]. \end{aligned}$$

After taking tensor product with $\otimes_F K$, we get

$$K^n = K[x]/p_1 K[x] \times \cdots \times K[x]/p_r K[x].$$

This shows that the invariant factors of a matrix is independent of field extension. The essential part of the above proof is just equation (0.1).

Exercise 13: Let $A \in \text{Mat}_{n \times n}(\mathbb{C})$ be a matrix such that every eigenvalue of A is real. Show that A is similar to a matrix with real entries.

Proof. Let $p_i, 1 \leq i \leq r$, be the invariant factors of A . Note that each p_i is a factor of f_A . By assumption, $f_A = \prod (x - c_i)^{e_i}$ with each $c_i \in \mathbb{R}$. Thus each factor of f_A has the form $\prod (x - c_i)^{s_i}$ with $0 \leq s_i \leq e_i$, which is in $\mathbb{R}[x]$. Thus $p_i \in \mathbb{R}[x]$ and its companion matrix has entries in \mathbb{R} . Thus the rational form of A has entries in \mathbb{R} . \square

Remark 8. Let us compare the terminologies used in Ex 12 and Ex 13. For $A, B \in \text{Mat}_{n \times n}(F)$, then “ A and B are similar **over** F ” means that there exists a matrix $P \in \text{GL}_n(F)$ such that $PAP^{-1} = B$. See Ex 12. For $A \in \text{Mat}_{n \times n}(\mathbb{C})$, then “ A is similar to a matrix with real entries” means that there exists a matrix $B \in \text{Mat}_{n \times n}(\mathbb{R})$ and there exists a matrix $P \in \text{GL}_n(\mathbb{C})$ such that $A = PBP^{-1}$. In Ex 13, we can say that A is similar to a matrix $B \in \text{Mat}_{n \times n}(\mathbb{R})$ over \mathbb{C} , not over \mathbb{R} .

Exercise 14: Let $T : V \rightarrow V$ with $\dim V < \infty$. Show that there is a vector $\alpha \in V$ with the property: if $f(T)\alpha = 0$ for $f \in F[x]$, then $f(T) = 0$. Such a vector is called a separating vector for the algebra $F[x]$. When T has a cyclic vector, give a direct proof that any cyclic vector is a separating vector.

Proof. We first assume that T has a cyclic vector, which means $V = Z(\alpha; T)$ for a cyclic vector α . We will show that the cyclic vector α is a separating vector. If $f(T)\alpha = 0$, then $f(T)h(T)\alpha = 0$ for any $h \in F[x]$ (because $f(T)$ commutes with $h(T)$). Since V is spanned by $h(T)\alpha$, we get that $f(T)v = 0$ for any $v \in V$. This shows that $f(T) = 0$ and thus α is a separating vector.

In general, consider the cyclic decomposition

$$V = Z(\alpha_1; T) \oplus \cdots \oplus Z(\alpha_r; T),$$

with invariant factors p_1, \dots, p_r , and $p_i | p_{i-1}$. Note that p_1 is the annihilator of α_1 and is also the minimal polynomial of T . We claim that α_1 is a separating vector. In fact, if $f \in F[x]$ and $f(T)\alpha_1 = 0$, we have $f \in S_T(\alpha_1; 0) = p_1 F[x]$. Thus $f = p_1 g$ for some $g \in F[x]$. We have $f(T) = p_1(T)g(T) = 0$ since $p_1(T) = 0$. (One can also show that $f(T)\alpha_i = 0$ for all $i \geq 1$ directly using $p_i | p_1$ and thus $p_i | f$. This also implies that $f(T)v = 0$ for any $v \in V$.) \square

Exercise 15: This is the above Lemma 6.

Exercise 16: Let A be an $n \times n$ matrix with real entries such that $A^2 + I = 0$. Prove that n is even, and if $n = 2k$, then A is similar over the field of real numbers to a matrix of the block form

$$\begin{bmatrix} 0 & -I \\ I & 0 \end{bmatrix},$$

where I is the $k \times k$ identity matrix.

Proof. Let $V = \mathbb{R}^n$ and $T : V \rightarrow V$ be the linear operator defined by $Tx = Ax$. Here an element in V is viewed as a column vector. Since $A^2 + I = 0$, we get $T^2 + I = 0$. Thus $f = x^2 + 1 \in I(T)$ and thus the minimal polynomial μ_T divides f . Since f is irreducible and $\mu_T \neq 1$, we get $\mu_T = f = x^2 + 1$. Let

$$V = Z(\alpha_1; T) \oplus Z(\alpha_2; T) \cdots \oplus Z(\alpha_k; T)$$

be the cyclic decomposition of V with $\alpha_1, \dots, \alpha_k \in V$. Let p_i be the T -annihilators of α_i , namely, p_1, \dots, p_k are the invariant factors of T . We have $p_1 = \mu_T = x^2 + 1$ and $p_i | p_{i-1}$ for $i \geq 2$. Since p_1 is irreducible, we have $p_i = x^2 + 1$ for each i . Since $\dim Z(\alpha_i; T) = \deg(p_i) = 2$, we get $\dim V = 2k$ is even. Let $\beta_i = T\alpha_i$. Then $\{\alpha_i, \beta_i\}$ is a basis of $Z(\alpha_i; T)$. Let $\mathcal{B} = \{\alpha_1, \dots, \alpha_k, \beta_1, \dots, \beta_k\}$, which is an ordered basis of V . Note that $T\alpha_i = \beta_i$, $T\beta_i = T^2\alpha_i = -\alpha_i$. We get

$$[T]_{\mathcal{B}} = \begin{bmatrix} 0 & -I \\ I & 0 \end{bmatrix}.$$

\square

Exercise 17: Let T be a linear operator on a finite-dimensional vector space V . Suppose that

- (a) the minimal polynomial for T is a power of an irreducible polynomial;
- (b) the minimal polynomial is equal to the characteristic polynomial.

Show that no non-trivial T -invariant subspace has a complementary T -invariant subspace.

Proof. We prove this by contradiction. Suppose that W_1 is a T -invariant nontrivial subspace ($W_1 \neq 0, W_1 \neq V$) and W_1 has a complementary T -invariant subspace W_2 . Let \mathcal{B}_i be an ordered basis of W_i . Then $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2$ is an ordered basis of V . Assume $A_i = [T]_{\mathcal{B}_i}$, we get

$$[T]_{\mathcal{B}} = \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix}.$$

This shows that $\chi_T = \chi_{T_1} \chi_{T_2}$, where $T_i = T|_{W_i}$ and χ_T denotes the characteristic polynomial of T . The assumption says that $\chi_T = p^r$ for an irreducible polynomial p and a positive integer r . Thus $\chi_{T_i} = p^{r_i}$ with $r_i > 0, r_1 + r_2 = r$. Let μ_{T_i} be the minimal polynomial of T_i . Then $\mu_{T_i} | \chi_{T_i}$. Thus $\mu_{T_i} = p^{s_i}$ for some integer s_i with $1 \leq s_i \leq r_i$. Let $s = \max\{s_1, s_2\}$ and $g = p^s \in F[x]$. By the choice of s , we have $g(A_1) = g(A_2) = 0$. Note that for any polynomial $h \in F[x]$, we have

$$h([T]_{\mathcal{B}}) = \begin{bmatrix} h(A_1) & \\ & h(A_2) \end{bmatrix}.$$

(Check this for monomials x^n first, which follows from a simple block matrix calculation.) In particular, since $g(A_1) = g(A_2) = 0$, we have $g([T]_{\mathcal{B}}) = 0$. This shows that the minimal polynomial of T divides $g = p^s$ (actually it is clear that the minimal polynomial is exactly $g = p^s$). Now since $s < s_1 + s_2 \leq r_1 + r_2$, we have $g \neq \chi_T = p^r$. This contradicts assumption (b). \square

Exercise 18: If T is a diagonalizable linear operator, then every T -invariant subspace has a complementary T -invariant subspace.

Proof. Let $W \subset V$ be a T -invariant subspace. We first show that $T|_W$ is diagonalizable. In fact $\mu_{T|_W}$ divides μ_T , which is a product of distinct linear factors. This shows that $T|_W$ is diagonalizable.

Let c_1, \dots, c_k be distinct eigenvalues of T and let $E_T(c_i) = \ker(T - c_i I)$. The condition T is diagonalizable means that

$$V = E_T(c_1) \oplus \cdots \oplus E_T(c_k).$$

Let $\mathcal{B}'_1 = \{\alpha_1, \dots, \alpha_{s_1}\}$ be a basis of W which consists of eigenvectors of T . We can assume this because $T|_W$ is diagonalizable. Since all distinct eigenvalues of T are c_1, \dots, c_k , we have $T\alpha_j = c_{i_j}\alpha_j$ for some index i_j with $1 \leq i_j \leq k$. After re-arrangement if necessary, we can assume that $\alpha_1, \dots, \alpha_{s_1} \in E_T(c_1), \alpha_{s_1+1}, \dots, \alpha_{s_1+s_2} \in E_T(c_2), \dots, \alpha_{s_1+s_2+\dots+s_{k-1}+1}, \dots, \alpha_{s_1+s_2+\dots+s_k} \in E_T(c_k)$. Here $s_k = s$. Assume that $\dim E_T(c_i) = r_i$, then $r_i \geq s_i$. Since α_i are linearly independent, we can extend $\alpha_{s_{i-1}+1}, \dots, \alpha_{s_i}$ to a basis

$$\alpha_{s_{i-1}+1}, \dots, \alpha_{s_i}, \beta_{s_i+1}, \dots, \beta_{r_i}$$

of $E_T(c_i)$. Let $W' = \text{Span}\{\beta_{s_i+1}, \dots, \beta_{r_i} : 1 \leq i \leq k\}$. Then clearly $V = W \oplus W'$ and W' is T -invariant. (Here W' is T -invariant because it has a basis which consists of eigenvectors of T). \square

A different proof. This exercise is a special case of Theorem 11 (page 264) of the textbook. The following is a proof based on the proof of Theorem 11.

Since T is diagonalizable, the minimal polynomial $\mu_T = (x - c_1) \cdots (x - c_k)$ for distinct c_1, \dots, c_k . Assume that $\chi_T = (x - c_1)^{r_1} \cdots (x - c_k)^{r_k}$ is the characteristic polynomial of T . Let $V = W_1 \oplus \cdots \oplus W_k$ be the primary decomposition of V , namely, $W_i = \ker(T - c_i I)^{r_i}$. Let W be a T -invariant subspace of V . We first claim that

$$W = (W \cap W_1) \oplus \cdots \oplus (W \cap W_k).$$

In fact, for any $\alpha \in W$, we can write $\alpha = \alpha_1 + \cdots + \alpha_k$ with each $\alpha_i \in W_i$. Let $E_i : V \rightarrow W_i$ be the projection map, which is known to have the form $h_i(T)$ for a polynomial h_i , see Corollary in page 221. We have $\alpha_i = E_i \alpha = h_i(T)\alpha \in W$ since W is T -invariant. This shows the above decomposition.

Next, we show that each $W \cap W_i$ has a T -invariant complement in W_i . For this, it suffices to show that $W \cap W_i$ is T -admissible subspace of W_i , namely, if $f \in F[x]$, $\alpha \in W_i$ with $f(T)\alpha \in W \cap W_i$, then there exists $\beta \in W \cap W_i$ such that $f(T)\alpha = f(T)\beta$. Note that, for $\alpha \in W_i$, we have $T\alpha = c_i\alpha$ and thus $f(T)\alpha = f(c_i)\alpha$. Suppose for some $\alpha \in W_i$ and $f \in F[x]$, we have $f(T)\alpha = f(c_i)\alpha \in W \cap W_i$. If $f(c_i) = 0$, we just take $\beta = 0$, which satisfies $f(T)\alpha = f(T)\beta = 0$. If $f(c_i) \neq 0$, the above condition means that $\alpha \in W \cap W_i$, and we just take $\beta = \alpha$, which satisfies $f(T)\alpha = f(T)\beta$.

Thus for each i , there is a T -invariant subspace $W'_i \subset W_i$ such that

$$W_i = (W \cap W_i) \oplus W'_i.$$

Take $W' = W'_1 \oplus \cdots \oplus W'_k$, which is still T -invariant. The above shows that

$$V = W_1 \oplus \cdots \oplus W_k = \bigoplus_i (W \cap W_i) \oplus W'_i = W \oplus W'.$$

This finishes the proof. \square

Remark 9. If T is diagonalizable, we actually have $W_i = \ker(T - c_i I)^{r_i} = \ker(T - c_i I)$. Thus the decompositions used in the above two different proofs are the same. Moreover, the first solution gives a direct proof that $W \cap W_i$ has a complement in W_i . Essentially, the above two proofs are the same. Apparently, the second approach works for more general case.

Exercise 19: Let T be a linear operator on the finite dimensional space V . Prove that T has a cyclic vector if and only if the following is true: Every linear operator U which commutes with T is a polynomial in T .

Proof. We assume that T has a cyclic vector α . Let $U : V \rightarrow V$ be a linear operator such that $TU = UT$. Note that, we have $UT^2 = UTT = TUT = T^2U$. Similarly, it is easy to check that $UT^i = T^iU$ for any $i \geq 0$. Since α is a cyclic vector, $V = \text{Span}\{\alpha, T\alpha, \dots, T^{n-1}\alpha\}$, where $n = \dim V$. Since $U(\alpha) \in V$, we can write

$$U(\alpha) = a_0\alpha + \dots + a_{n-1}T^{n-1}\alpha,$$

for some $a_0, a_1, \dots, a_{n-1} \in F$. (Here there is no requirement for a_i . If U is the zero operator, then all a_i are zero. If U is nonzero, there is at least one a_i is nonzero.)

Let $g = a_0 + a_1x + \dots + a_{n-1}x^{n-1} \in F[x]$. By choice, we have

$$U\alpha = g(T)\alpha.$$

We claim that $U = g(T)$, namely, $U\beta = g(T)\beta$ for all $\beta \in V$. Actually this follows easily from the above equation and the fact that $V = Z(\alpha; T)$. Here are some details. Since $V = \text{Span}\{\alpha, T\alpha, \dots, T^{n-1}\alpha\}$, it suffices to show that

$$U(T^i\alpha) = g(T)(T^i\alpha), i = 0, 1, \dots, n-1.$$

For $i = 0$, this follows from the definition of g . If $i = 1$, we have

$$U(T\alpha) = TU(\alpha) = T(g(T)\alpha) = g(T)(T\alpha).$$

Similarly, for any $i > 0$, we have

$$U(T^i\alpha) = T^i(U\alpha) = T^i(g(T)\alpha) = g(T)(T^i\alpha).$$

This shows that $U = g(T)$.

Conversely, suppose that T does not have a cyclic vector, we will construct a linear operator $U : V \rightarrow V$, which is not a polynomial of T . Consider the cyclic decomposition of V :

$$V = Z(\alpha_1; T) \oplus Z(\alpha_2; T) \oplus \dots \oplus Z(\alpha_r; T),$$

as in the cyclic decomposition theorem. The condition “ T does not have a cyclic vector” implies that $r \geq 2$. Let p_i be the annihilator of α_i , we have $p_2|p_1$.

Let $U = E_2$, the projection operator of V onto $Z(\alpha_2; T)$. Then $UT = TU$. This can be checked easily or it follows from Theorem 10, p214. We prove that U is not a polynomial of T by contradiction. Suppose that $U = g(T)$ for a polynomial $g \in F[x]$. Note that for any $\alpha \in Z(\alpha_1; T)$, we have $g(T)\alpha = U\alpha = 0$. Thus $p_1|g$ because p_1 is the annihilator of α_1 . On the other hand, $p_2|p_1$ and thus $p_2|g$. This means that g is a multiple of the annihilator of α_2 . Thus $g(T)\alpha_2 = 0$. This contradicts to $U\alpha_2 = \alpha_2$. We are done. \square

Exercise 20: Let V be a finite dimensional vector space over the field F and $T : V \rightarrow V$ be a linear operator. We ask when it is true that every non-zero vector in V is a cyclic vector for T . Prove that this is the case if and only if the characteristic polynomial for T is irreducible over F .

Proof. Assume that the characteristic polynomial χ_T of T is irreducible in $F[x]$. In particular, $\mu_T = \chi_T$. Given any $\alpha \in V, \alpha \neq 0$, we need to show that $Z(\alpha; T) = V$. Let p_α be the T -annihilator of α , we have $p_\alpha|\mu_T$. But μ_T is irreducible, and thus we have $p_\alpha = \mu_T$. Thus $\dim_F Z(\alpha; T) = \deg(p_\alpha) = \deg(\chi_T) = \dim V$. We have $Z(\alpha; T) = V$.

Conversely, suppose that every nonzero vector in V is a cyclic vector. Take $\alpha \neq 0$, we have $V = Z(\alpha; T)$. Suppose that μ_T is reducible, namely, $\mu_T = gh$ with $g, h \in F[x]$, $\deg(g) = k < n$, $\deg(h) = m < n$, where $n = \dim V$. Consider the vector $\beta = g(T)\alpha \neq 0$. Since $h(T)\beta = \mu_T(T)\alpha = 0$, the T -annihilator p_β of β divides $h(T)$. By Theorem 1 of page 228, we have $\dim Z(\beta; T) = \deg(p_\beta) \leq \deg h = m < n$. Thus $Z(\beta; T) \neq V$ and β is not a cyclic vector of V . \square

Exercise 21: Let $A \in \text{Mat}_{n \times n}(\mathbb{R})$. Let $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$ be the operator defined by A and $U : \mathbb{C}^n \rightarrow \mathbb{C}^n$ be the operator defined by A . If the only subspaces invariant under T are \mathbb{R}^n and the zero subspace, then U is diagonalizable.

Proof. Let $\alpha \in \mathbb{R}^n$ be any nonzero vector and consider $Z(\alpha; T)$. Since $Z(\alpha; T)$ is a nonzero T -invariant subspace of \mathbb{R}^n , the assumption says that $Z(\alpha; T) = \mathbb{R}^n$. This shows that every nonzero vector of \mathbb{R}^n is a cyclic vector. Exercise 20 says that $\mu_T = \chi_T$ is irreducible. We know that any irreducible polynomial over \mathbb{R} is either linear or quadratic $ax^2 + bx + c$ with $a, b, c \in \mathbb{R}, b^2 - 4ac < 0$.

Either case, $\mu_T = \chi_T$ has no repeated roots over \mathbb{C} . Thus U is diagonalizable. Note that $\mu_U = \mu_A = \mu_T$, namely no matter if you see A as a matrix over \mathbb{R} or over \mathbb{C} , its minimal polynomial is the same. See Exercise 12. \square

Remark 10. Exercise 21 seems too easy because in this case we can only have $n = 1$ or 2 . The following general case is true. Let F be a field of characteristic 0 and $A \in \text{Mat}_{n \times n}(F)$. Suppose that \overline{F} is an algebraically closed field such that $F \subset \overline{F}$. (Example: F is \mathbb{Q} or $\{a + b\alpha + c\alpha^2 : a, b, c \in \mathbb{Q}\}$ with $\alpha^3 = 2, \alpha \in \mathbb{R}$; and $\overline{F} = \mathbb{C}$.) Let $T : F^n \rightarrow F^n$ be the linear operator defined by A . If the only subspaces invariant under T are 0 and F^n itself, then A is diagonalizable over \overline{F} . In this general case, the dimension of V can be arbitrary. The proof is the same as the above once we know the following fact: if F has characteristic zero and $f \in F[x]$ is irreducible, then f has no repeated roots over \overline{F} . See Lemma of page 266 and Theorem 12 for its generalizations. If characteristic of F is finite, the above is false. In fact, if characteristic of F is finite, it is possible to find irreducible polynomial $f \in F[x]$, such that over an algebraic closure of F , $f = (x - c)^p$ for some positive integer p .