Diagonalizability, Triangularization, Invariant and Nilpotent Subspaces

Question 1:

Let F be a field and $A \in M_{n\times n}(F)$. Show that L_A (left multiplication by A) is diagonalizable over F if and only if there exist an invertble matrix $Q \in M_{n\times n}(F)$ such that QAQ^{-1} is diagonal.

Proof:

First, we prove the forward implication. So suppose A is diagonalizable. Thus, we can pick a basis of A such that A is a diagonal matrix of eigenvectors in the basis.
Let $\beta = \{v_1, ..., v_n\}$ with corresponding eigenvalues $\{\lambda_1, ..., \lambda_n\}$
We can create a matrix Q made up of the n eigenvectors viewed as column vectors. Since, the eigenvectors are in a basis, they are linearly independent and thus Q^{-1} exists. Let D be a diagonal matrix made up of the eigenvalues of A. We will prove that $QD = AQ$.
Consider each column of the equation. The i^{th} column of QD is $\lambda_i v_i$. The same holds for AQ since the i^{th} column of Q is an eigenvector, and by definition, $Av_i = \lambda_i v_i$.
Since $QD = AQ$, multiply both sides by Q^{-1} which we know exists, and we get $D = Q^{-1}AQ$ as wanted.

To prove the backward implication, suppose the matrix $Q^{-1}AQ = D$ where D is diagonal. Let $\beta = \{e_1, ..., e_n\}$ denote the standard basis. Note, $QQ^{-1}AQ = IAQ = QD$.

$$\begin{aligned}
QDe_i &= QD_{ii}e_i = D_{ii}Qe_i = D_{ii}q_i && \text{(where } q_i \text{ is the } i^{th} \text{ column vector of Q.)} \\
&\Rightarrow AQe_i = D_{ii}q_i && \text{(since AQ = QD)} \\
&\Rightarrow Aq_i = D_{ii}q_i
\end{aligned}$$

Thus, q_i is an eigenvector, since $D_{ii} \in \mathcal{F}$. Since $i \in [1, n]$ is arbitrary, we can see that q_i is an eigenvector for all i. But since Q is invertible, q_i are all linearly independent. Thus, we can form a basis made up of the column vectors of Q such that $Aq_i = \lambda_i q_i$. Thus, A is diagonalizable. ■

Question 2:

Let V be a vector space over $\mathbb{C}$. Then, V can also be considered as a vector space over $\mathbb{R}$. Show that if $\{v_1, ..., v_n\}$ is a basis of V over $\mathbb{C}$ then $\{v_1, ..., v_n, iv_1, ..., iv_n\}$ is a basis of V over $\mathbb{R}$.

Proof:

We want to show that $\{v_1, iv_1, ..., v_n, iv_n\}$ is a basis for V over $\mathbb{R}$. First, we show linear independence:

$$\begin{aligned} c_1v_1 + d_1iv_1 + ... + c_nv_n + d_niv_n &= 0 && \text{(for some } c, d \in \mathbb{R}) \\ (c_1 + d_1i)v_1 + ... + (c_n + d_ni)v_n &= 0 \end{aligned}$$

But $c + di \in \mathbb{C}$ and $\{v_1, ..., v_n\}$ forms a linearly independent basis for V over $\mathbb{C}$. Thus, $c_j + d_ji = 0 \Rightarrow c_j = d_j = 0 \ \forall j \in [1, n]$ which is the definition of linear independence.

Now we want to prove that $\{v_1, iv_1, ..., v_n, iv_n\}$ spans V. Well, let $v \in V$. Since V is a vector space over $\mathbb{C}$ there exists $(c + di) \in \mathbb{C}$ such that:

$$\begin{aligned} (c_1 + d_1i)v_1 + ... + (c_n + d_ni)v_n &= v \\ c_1v_1 + d_1iv_1 + ... + c_nv_n + d_niv_n &= v \end{aligned}$$

But $c_j, \ d_j \in \mathbb{R}, \forall j \in [1, n]$.

Thus, since any $v \in V$ can be expressed as a linear combination of $\{v_1, iv_1, ..., v_n, iv_n\}$ with scalars in $\mathbb{R}$, $\text{span}(\{v_1, iv_1, ..., v_n, iv_n\}) = V$.

Therefore, $\{v_1, iv_1, ..., v_n, iv_n\}$ forms a basis for V over $\mathbb{R}$ $\blacksquare$

Question 3:

Let V be a finite-dimensional subspace and W a non-zero proper T-invariant subspace. Let $f(t),\ g(t),\ h(t)$ be the characteristic polynomials of $T,\ T_W,\ \bar{T}$ respectively. Show that $f(t) = g(t)h(t)$

Proof:

First we will show that $\bar{T} : V \setminus W \to V \setminus W$ where $\bar{T}(v+W) = T(v) + W$ is well defined and linear. Suppose $v + W = v' + W$

$$\begin{aligned}
\Rightarrow \quad & v + W - (v' - W) = \vec{0} + W \\
& (v - v') + W = \vec{0} + W \\
& v - v \sim \vec{0} \\
\Rightarrow \quad & (v - v') \in W \\
\Rightarrow \quad & T(v - v') \in W && \text{(since W is T-invariant)} \\
\Rightarrow \quad & T(v) - T(v') \in W \\
\Rightarrow \quad & T(v) + W - (T(v') + W) = \vec{0} + W \\
& T(v) + W = T(v') + W \\
& \bar{T}(v + W) = \bar{T}(v' + W) && \text{(thus, } \bar{T} \text{ is well-defined.)}
\end{aligned}$$

Now, let $\beta = \{v_1, ..., v_k\}$ be a basis for W. Extend it to get $\gamma = \{v_1, ..., v_k, ...v_n\}$ a basis for V (guaranteed to not be equal to β since W is a proper subset of V). From exercise 35, we know $\delta = \{v_{k+1} + W, ..., v_n + W\}$ forms a basis for V/W.

Now, let $w \in W,\ u \in V/W$. $[w]\gamma = \begin{pmatrix} a_1 \\ \vdots \\ a_k \\ 0_{k+1} \\ \vdots \\ 0_n \end{pmatrix}$.

Notice $[T]_\gamma[w]_\gamma = \begin{pmatrix} [T_W]_\beta[w]_\beta & 0_{k+1} & ... & 0_n \end{pmatrix}$ since W is T-invariant. Also, $[T]_\gamma([u]_\gamma + W) = \begin{pmatrix} 0_1 & ... & 0_k & [\bar{T}]_\delta[u]_\delta \end{pmatrix}$ since the first k vectors are mapped to $\vec{0}$ in the quotient space.

Thus, $[T]_\gamma = \begin{pmatrix} [T_W]_\beta & C \\ 0 & [\bar{T}]_\delta \end{pmatrix}$ where each component are a square matrices.

Characteristic polynomial of $[T]_\gamma$:

$$f(t) = \det([T]_\gamma - tI) = \det([T_W]_\beta - tI)\det([\bar{T}]_\delta - tI) = g(t)h(t)$$

as wanted. ■

Question 4:

Let V be an n-dimensional vector space over field F and $T : V \to V$ be a linear map whose characteristic polynomial splits over F. Show that there exist a basis β of V such that $[T]_\beta$ is upper-triangular.

Proof:

We will prove the statement "If $f(t)$ splits, $\exists\beta$ of V such that $[T]_\beta$ is upper triangular" with induction on $\dim(V)$.

Base Case: If $\dim(V) = 1$, pick any basis β and will $[T]_\beta$ will be upper-triangular trivially.

Now suppose the statement is true for $\dim(V) = n$. We will prove it for $\dim(V) = n + 1$

Suppose $f(t)$ splits into $c(t - \lambda_1)...(t - \lambda_n)(t - \lambda_{n+1})$. λ_1 is an eigenvalue with at least one eigenvector w. Let $\text{span}(\{w\}) = W$. Define $T_W : W \to W$, $T_W(w) = T(w)$. We know this is possible since $T(w) = \lambda_1 w \in W$. Thus, W is T-invariant.

Then, define $\bar{T} : V/W : V/W$ such that $\bar{T}(v + W) = T(v) + W$. The characteristic polynomial of $\bar{T}$ also splits ($h(t) = c(t - \lambda_2)...(t - \lambda_{n+1})$) by the result in Question 3.

By Induction Hypothesis, $\exists$ β such that $[\bar{T}]_\beta$ is upper triangular. Also, we know that $[\bar{T}]_\beta \in M_{n\times n}$ since $\dim(V/W) = \dim(V) - \dim(W) = n + 1 - 1 = n$. Also, w is linearly independent of all vectors in β.

By the result in $Q3(b)$, $[T]_{\{w\}\cup\beta} = \begin{pmatrix} [T_W]_{\{w\}} & C \\ 0 & [\bar{T}]_\beta \end{pmatrix}$ for some β basis of V/W. Since $[\bar{T}]_\beta$ is upper triangular, $\{w\} \cup \beta$ is a basis at which $[T]$ is upper-triangular. $\blacksquare$

Question 5:

(a) Let V be a finite-dimensional vector space and $T : V \to V$ be a linear map such that for every $v \in V$ there exist an integer $k \geq 1$ (possibly depending on v) such that $T^(k)(v) = 0$. Show T is nilpotent.

(b) Let $\dim(V) = n$ and $T : V \to V$ be a nilpotent linear map. Show that if λ is an eigenvalue of T, then $\lambda = 0$

Proof:

(a) Let $\mathbb{K} = \{k_1, ..., k_n\}$ be the set such that $T^{k_i}(v_i) = 0$. Well, choose $k' = \max \mathbb{K}$.

Since T is linear, for any $v_i \in V$, $T^{k'}(v_i) = T^{(k'-k_i)}T^{k_i}(v_i) = T^{(k'-k_i)}(\vec{0}) = \vec{0}$.

Thus, T is nilpotent. ■

(b) Consider $T(v) = \lambda v$ where $v \neq \vec{0}$, which we know to exist since λ is an eigenvalue of T.

Since $T(v)$ is nilpotent, $\exists\ k \in \mathbb{N}$. With the fact that $v \neq \vec{0}$, we have that

$$\vec{0} = T^k(v) = \lambda^k(v) \Rightarrow \lambda = 0$$

Thus all eigenvalues of T is 0.

Since $0 \in F$ for any F, T contains all it's eigenvalues. Thus, the characteristic polynomial splits and $f(t) = (-1)^n(t - \lambda)^n = (-1)^n(t)^n$ ■

Question 6:

Let T and S be linear operators on a vector space V such that $TS = ST$. Show that the kernel and image of S are T-invariant.

Proof:

We want to show that $\ker(S)$ is T-invariant. That is, $v \in \ker(S) \Rightarrow T(v) \in \ker(S)$.
So, suppose $v \in \ker(S)$. Thus, $S(v) = 0$. Applying T to both sides of the equation:

$$T(S(v)) = 0 \Rightarrow S(T(v)) = 0 \Rightarrow T(v) \in \ker(S)$$

as wanted.

Now, we want to show that $\mathrm{Im}(S)$ is T-invariant. That is, $w \in \mathrm{Im}(S) \Rightarrow T(w) \in \mathrm{Im}(S)$.
So suppose $w \in \mathrm{Im}(S)$. Thus, $\exists\ u \in V$ s.t. $S(u) = w$. Well,

$$T(w) = T(S(u)) = S(T(u)) \in \mathrm{Im}(S)$$

completing the proof.

Question 7:

Let T be a nilponent linear operator on a (possibly infinite-dimensional) vector space V. Suppose the nilpotency index of T is k. Show that if $0 \leq i < k$, then $Im(T^i + 1) \subsetneq Im(T^i)$ and $\ker(T^i) \subsetneq \ker(T^i + 1)$.

Proof:

Let's prove $\operatorname{Im}(T^{i+1}) \subsetneq \operatorname{Im}(T^i)$.

Clearly, T commutes with T^i. Thus, by Q1, $\operatorname{Im}(T^i)$ is T-invariant. In other words, $T(\operatorname{Im}(T^i)) = \operatorname{Im}(T^{i+1}) \subset \operatorname{Im}(T^i)$

To show $\operatorname{Im}(T^{i+1}) \neq \operatorname{Im}(T^i)$, suppose for sake of contradiction otherwise. Notice that

$$\operatorname{Im}(T^{i+1}) = \operatorname{Im}(T^i) \Rightarrow \operatorname{Im}(T^{i+2}) = \operatorname{Im}(T^{i+1}) = \operatorname{Im}(T^i)$$

is an immediate result of $T(\operatorname{Im}(T^i)) = \operatorname{Im}(T^{i+1})$.

Apply T k-times in the equation, and we get the result $\operatorname{Im}(T^i) = \operatorname{Im}(T^{k+i}) = 0$, but by the minimality of k, $i < k \Rightarrow \operatorname{Im}(T^i) \neq 0$ which is a contradiction.

Therefore $\operatorname{Im}(T^{i+1}) \subsetneq \operatorname{Im}(T^i)$.

Now let's prove $\ker(T^i) \subsetneq \ker(T^{i+1})$.

Let $v \in \ker(T^i)$. Thus, $T^{i+1}(v) = T(T^i(v)) = T(0) = 0$. Thus, $v \in \ker(T^{i+1})$ and it follows that $\ker(T^i) \subset \ker(T^{i+1})$.

Now we only need to prove $\ker(T^i) \neq \ker(T^{i+1})$. Well, by dimension theorem,

$$N(T^i) + R(T^i) = \dim V = N(T^{i+1}) + R(T^{i+1})$$

Since $\operatorname{Im}(T^{i+1}) \subsetneq \operatorname{Im}(T^i)$, $R(T^i) - R(T^{i+1}) \neq 0$ (Rank is defined as dimension of it's image). We then get $N(T^{i+1}) - N(T^i) \neq 0$. Thus, the dimension of their kernels are different and therefore $\ker(T^i) \neq \ker(T^{i+1})$ completing the proof.

Question 8:

Let V be a nonzero finite-dimensional vector space over $\mathbb{C}$. Denote the identity map on V by I. Let T be a linear operator on V such that $T^k = I$ for some positive integer k. Show that T is diagonalizable.

Proof:

Let the number of distinct eigenvectors of T be n. We know that the characteristic polynomial of T splits over $\mathbb{C}$ and so the minimal polynomial of T,

$$m(t) = \prod_{i=1}^{n} (t - \lambda_i)^{d_i}$$

since the minimal polynomial have the same zeros as the characteristic. All that's left to show is $d_i = 1$ for all $i \in [1, n] \cap \mathbb{N}$.

Consider the polynomial $p(t) = t^k - 1$. Over $\mathbb{C}$, $p(t)$ splits. The roots of $(t^k - 1)$ are $e^{i2n\pi/k}$ $\forall n \in \mathbb{Z}$. Considering that $|\mathbb{Z}_k| = k$ (where $\mathbb{Z}_k$ is the set of integers modulo k), the set $\{\frac{2\pi n}{k} \mod 2\pi \mid n \in \mathbb{Z}\}$ has cardinality k. So, $(t^k - 1)$ has k unique roots - the maximal number by the fundamental theorem of algebra, and thus no root is repeated.

Since $\deg p(t) = k$, $p(t) = \prod_{i=1}^{k} (t - r_i)$ where r_i is a unique root of $p(t)$.

We can see that $p(T) = T^k - I = 0$ and thus $m(t) | (t^k - 1)$. But the multiplicities of the linear products of $p(t)$ are all 1. So, the same must go for $m(t)$ (or else, $m(t)$ won't divide $p(t)$), completing our proof that $d_i = 1$ for all $i \in [1, n] \cap \mathbb{N}$.

Question 9:

Let V be a nonzero finite-dimensional vector space and T a diagonalizable linear operator on T. Let W be a T-invariant subspace of V. Show that T_W (the restriction of T to W) is diagonalizable.

Proof:

Since T is diagonalisable, its characteristic polynomial splits. Let $m(t)$ be the the minimal polynomial of $T|_W$. By the same reasoning in Q2, we need only prove that $m(t)$ splits and has no repeated factors.

Well, we know that $P_{T|_W}(t)|P_T(t)$. So $P_{T|_W}(t)$ must also be made up solely of linear factors. Thus the minimal polynomial $m(t)$ must also split.

Now let $\mu(t)$ be the minimal polynomial of T. Since $\mu(T) = 0$, it follows that $0 = \mu(T)|_W = \mu(T|_W)$. So, by the property of minimal polynomials, $m(t)$ must divide $\mu(t)$. But the linear factors of $\mu(t)$ all have multiplicities of 1 since T is diagonalisable. So, $m(t)$ must also have no repeated factors, completing our proof.

Question 10:

Let V be a nonzero finite-dimensional vector space. Let S be a collection of diagonalizable linear operators on V such that any two maps in S commute with each other. Show that the maps in S can be simultaneously diagonalized. That is, show that there exists a basis β of V such that for every $T \in S$, the matrix $[T]_\beta$ is diagonal.

Proof:

We will prove the statement via strong induction on $\dim V$.

Base case: If $\dim V = 1$, then every $T \in \mathcal{S}$ is diagonalisable regardless of basis.

Inductive Hypothesis: Suppose that you can simultaneously diagonalise any commuting set of diagonalisable matrices on vector spaces of degree $< n$. We will prove the result holds for $\dim V = n$.

Case 1: Suppose each $T \in \mathcal{S}$ only has 1 eigenvector.

Since any T is diagonalisable, then every T is in the form λI regardless of basis. Done.

Case 2: Suppose there exists $T \in \mathcal{S}$ with $k > 1$ eigenvalues.

Decompose $V = \bigoplus_{i=1}^{k} E_{\lambda_i}$. We can do this since T is diagonalisable. Pick an arbitrary E_{λ_i} and let $T|_{\lambda_i}$ be the restriction of T to subspace E_{λ_i}.

First, we will show that E_{λ_i} is S-invariant for any arbitrary $S \in \mathcal{S}$. So let $v \in E_{\lambda_i}$. We need to show $S(v) \in E_{\lambda_i}$. Indeed, $T(Sv) = STv = S\lambda_i v = \lambda_i(Sv) \Rightarrow S(v) \in E_{\lambda_i}$.

Thus, E_{λ_i} is S-invariant for any $S \in \mathcal{S}$. We can then restrict any S to E_{λ_i}. From Q3, we know that $S|_{\lambda_i}$ is diagonalisable (for S is diagonalisable). And since $\dim S|_{\lambda_i} < n$ ($E_{\lambda_i} \subsetneq V$), there exists a basis where all $S|_{\lambda_i}$ are simultaneously diagonalisable by the induction hypothesis.

We can repeat this for all the eigenspaces of T whose sum is directly V. We can then combine each basis that simultaneously diagonalise $\mathcal{S}$. That set is indeed a basis for V because it is a basis that diagonalises T. Each S would look like

$$\begin{pmatrix} [S|_{\lambda_1}] & 0 & & 0 \\ 0 & [S|_{\lambda_2}] & & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & & [S|_{\lambda_k}] \end{pmatrix}$$

where each $[S|_{\lambda_i}]$ is diagonal. And hence we're done.