Linear Transformations $T : \mathbb{C}^n \to \mathbb{C}^n$

Algebra of Linear Transformations

Definition 1. Let $F = \mathbb{R}$ or $\mathbb{C}$. We define the algebra of linear transformations on F^n to be the set $\mathcal{A}(F^n) = \{F^n \to^T F^n : T$ is a linear transformation.

Theorem 2. For the algebra $\mathcal{A}(F^n)$, the following operations are well-defined.
(A) $\boxplus : \mathcal{A}(F^n) \times \mathcal{A}(F^n) \to \mathcal{A}(F^n)$ given by $(T \boxplus S)(\mathbf{v}) = T(\mathbf{v}) + S(\mathbf{v})$.
(M) $\circ : \mathcal{A}(F^n) \times \mathcal{A}(F^n) \to \mathcal{A}(F^n)$ given by $(T \circ S)(\mathbf{v}) = T(S(\mathbf{v}))$.
(S) $\cdot : F \times \mathcal{A}(F^n) \to \mathcal{A}(F^n)$ given by $(cT)(\mathbf{v}) = cT(\mathbf{v})$.

Proof. We prove (A) and leave the other two as an exercise. (WD1) We must show that if $S, T \in \mathcal{A}(F^n)$, then $S \boxplus T \in \mathcal{A}(F^n)$. That is, we must show that $(S \boxplus T) : F^n \to F^n$ is a linear transformation. If $\mathbf{v}, \mathbf{w} \in F^n$, then

$$\begin{aligned}
& (S \boxplus T)(\mathbf{v} + \mathbf{w}) \\
= \;& S(\mathbf{v} + \mathbf{w}) + T(\mathbf{v} + \mathbf{w}) \quad (\text{Defn of } \boxplus \text{ in } \mathcal{A}(F^n)) \\
= \;& (S(\mathbf{v}) + S(\mathbf{w})) + (T(\mathbf{v}) + T(\mathbf{w})) \quad (\text{Since } S, T \text{ are linear maps}) \\
= \;& (S(\mathbf{v}) + T(\mathbf{v})) + (S(\mathbf{w}) + T(\mathbf{w})) \quad (\text{Usual properties of } + \text{ in } F^n) \\
= \;& (S \boxplus T)(\mathbf{v}) + (S \boxplus T)(\mathbf{w})
\end{aligned}$$

(WD2) We must show that if $S_1 = S_2$ and $T_1 = T_2$, then $S_1 \boxplus T_1 = S_2 \boxplus T_2$. To prove that these two maps are equal, we must show that they agree at every point in F^n. So choose any $\mathbf{v} \in F^n$. We have

$$\begin{aligned}
& S_1 = S_2 \text{ and } T_1 = T_2 \\
\Rightarrow \;& S_1(\mathbf{v}) = S_2(\mathbf{v}) \text{ and } T_1(\mathbf{v}) = T_2(\mathbf{v}) \\
\Rightarrow \;& S_1(\mathbf{v}) + T_1(\mathbf{v}) = S_2(\mathbf{v}) + T_2(\mathbf{v}) \quad (\text{Since } + \text{ is well-defined on } F^n) \\
\Rightarrow \;& (S_1 \boxplus T_1)(\mathbf{v}) = (S_2 \boxplus T_2)(\mathbf{v}) \quad (\text{Defn of } \boxplus \text{ in } \mathcal{A}(F^n)) \\
\Rightarrow \;& S_1 \boxplus T_1 = S_2 \boxplus T_2
\end{aligned}$$

Theorem 3. For the algebra $\mathcal{A}(F^n)$, the following properties hold.
(A1) $\boxplus$ is associative.
(A2) $\boxplus$ is commutative.
(A3) $\boxplus$ has an identity $0 : F^n \to F^n$ given by $0(\mathbf{v}) = \mathbf{0}$.
(A4) $\boxplus$ has an inverses $(-T) : F^n \to F^n$ given by $(-T)(\mathbf{v}) = -T(\mathbf{v})$.
(M1) $\circ$ is associative.
(M2) $\circ$ has the identity $1_{F^n} : F^n \to F^n$ given by $1_{F^n}(\mathbf{v}) = \mathbf{v}$.
(D) $R \circ (S \boxplus T) = (R \circ S) \boxplus (R \circ T)$ for all $R, S, T \in \mathcal{A}(F^n)$.
(S1) $1_F T = T$ for all $T \in \mathcal{A}(F^n)$.
(S2) $(a + b)T = (aT) \boxplus (bT)$ for all $a, b \in F$ and $T \in \mathcal{A}(F^n)$.
(S3) $a(S \boxplus T) = (aS) \boxplus (aT)$ for all $a \in F$ and $S, T \in \mathcal{A}(F^n)$.
(S4) $a(bT) = (ab)T$ for all $a, b \in F$ and $T \in \mathcal{A}(F^n)$.

Proof. We prove (S3) and leave the remaining parts as an exercise. Choose any $\mathbf{v} \in F^n$ and check that

$$\begin{aligned} & (a(S \boxplus T))(\mathbf{v}) \\ = \;& a(S \boxplus T)(\mathbf{v}) \quad \text{(Theorem 2(S))} \\ = \;& a(S(\mathbf{v}) + T(\mathbf{v})) \quad \text{(Theorem 2(A))} \\ = \;& aS(\mathbf{v}) + aT(\mathbf{v}) \quad \text{(Usual scalar mult in } F^n\text{)} \\ = \;& (aS)(\mathbf{v}) + (aT)(\mathbf{v}) \quad \text{(Theorem 2(S))} \\ = \;& ((aS) \boxplus (aT))(\mathbf{v}) \quad \text{(Theorem 2(A)).} \end{aligned}$$

Remarks 4. From now on, we will write $S+T$ instead of $S \boxplus T$ and ST instead of $S \circ T$. Therefore, $T^k = T \circ T \circ ... \circ T$ composed k times. If $p(x) \in F[x]$ and $p(x) = a_0 + a_1 x + ... + a_n x^n$, then

$$p(T)(v) = (a_0 1_{F^n} + a_1 T + ... + a_n T^n)(\mathbf{v}) = a_0 \mathbf{v} + a_1 T(\mathbf{v}) + ... + a_n T^n(\mathbf{v}).$$

Complex Eigenvalues

Lemma 5. Let $S, T \in \mathcal{A}(F^n)$. If $\ker ST \neq \{\mathbf{0}\}$, then $\ker S \neq \{\mathbf{0}\}$ or $\ker T \neq \{\mathbf{0}\}$. More generally, if $T_1, T_2, ..., T_m \in \mathcal{A}(F^n)$ and $\ker T_1 T_2 \cdots T_m \neq \{\mathbf{0}\}$, then there exists a $j \leq m$ such that $\ker T_j \neq \{\mathbf{0}\}$.

Proof. We prove the first part using the contrapositive. The second part follows by induction. If $\ker S = \{\mathbf{0}\}$ and $\ker T = \{\mathbf{0}\}$, then both S, T are 1-1. It follows that the composition ST is 1-1. This is equivalent to $\ker ST = \{\mathbf{0}\}$.

Theorem 6. Every $T \in \mathcal{A}(\mathbb{C}^n)$ has a complex eigenvalue $\lambda \in \mathbb{C}$.

Proof. Choose any non-zero $\mathbf{v} \in \mathbb{C}^n$. Since $\dim(\mathbb{C}^n) = n$, the set

$$S = \{\mathbf{v}, T(\mathbf{v}), T^2(\mathbf{v}) ... T^n(\mathbf{v})\}$$

of $n+1$ many vectors is linearly dependent in $\mathbb{C}^n$. We can find $a_0, a_1, a_2, ..., a_n \in \mathbb{C}$ not all zero such that

$$a_0 \mathbf{v} + a_1 T(\mathbf{v}) + a_2 T^2(\mathbf{v}) + ... + a_n T^n(\mathbf{v}) = \mathbf{0}.$$

After removing any leading zeros, we can write

$$a_0 \mathbf{v} + a_1 T(\mathbf{v}) + a_2 T^2(\mathbf{v}) + ... + a_m T^m(\mathbf{v}) = \mathbf{0} \text{ where } m \leq n \text{ and } a_m \neq 0.$$

Using the algebra operations, we find that

$$(a_0 1_{\mathbb{C}^n} + a_1 T + a_2 T^2 + ... + a_m T^m)(\mathbf{v}) = \mathbf{0}.$$

Let

$$p(x) = a_0 + a_1 x + a_2 x^2 + ... + a_m x^m \in \mathbb{C}[x].$$

By The Fundamental Theorem of Algebra, we can factor p as

$$p(x) = a_m(x - \lambda_1)(x - \lambda_2) \cdots (x - \lambda_m).$$

It now follows that

$$p(T)(\mathbf{v}) = (a_m(T - \lambda_1)(T - \lambda_2) \cdots (T - \lambda_m))\,(\mathbf{v}) = \mathbf{0}.$$

Since $\mathbf{v}$ is non-zero,

$$\ker(T - \lambda_1)(T - \lambda_2) \cdots (T - \lambda_m) \neq \{\mathbf{0}\}.$$

By Lemma 5,

$$\ker(T - \lambda_1) \neq \{\mathbf{0}\} \text{ for some } j \leq m.$$

Therefore, there exists a non-zero $\mathbf{w} \in \mathbb{C}^n$ such that

$$(T - \lambda_j)(\mathbf{w}) = \mathbf{0} \Longleftrightarrow T(\mathbf{w}) = \lambda_j \mathbf{w}.$$

Definition 7. A matrix $A \in M_n$ is called upper triangular if $\mathrm{ent}_{ij}(A) = 0$ whenever $i > j$. That is, if

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ 0 & a_{22} & a_{23} & \cdots & a_{2n} \\ 0 & 0 & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & a_{nn} \end{bmatrix}. \qquad (*)$$

A matrix $B \in M_n$ is called triangularizable if there exists an invertible matrix P such that $B = PAP^{-1}$ (similar to a triangular matrix). A linear transformation $T \in \mathcal{A}(\mathbb{C}^n)$ is called triangularizable if there exists a basis B of $\mathbb{C}^n$ such that $M(T, B, B)$ is triangular.

Theorem 8. Let $B = \{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n\}$ be a basis for F^n. The following statements are equivalent for a linear transformation $T \in \mathcal{A}(F^n)$.
(1) $M(T, B, B)$ is upper triangular.
(2) $T(\mathbf{v}_k) \in \mathrm{Span}(\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_k)$ for each $k \leq n$.
(3) $\mathrm{Span}(\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_k)$ is invariant under T for each $k \leq n$.

Proof.
(1)$\Rightarrow$(2) If $M(T, B, B)$ has the form $(*)$, then

$$\begin{aligned} T(\mathbf{v}_1) &= a_{11}\mathbf{v}_1 \in \mathrm{Span}(\mathbf{v}_1) \\ T(\mathbf{v}_2) &= a_{12}\mathbf{v}_1 + a_{22}\mathbf{v}_2 \in \mathrm{Span}(\mathbf{v}_1, \mathbf{v}_2) \\ &\vdots \\ T(\mathbf{v}_n) &= a_{1n}\mathbf{v}_1 + a_{2n}\mathbf{v}_2 + ... + a_{nn}\mathbf{v}_n \in \mathrm{Span}(\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n). \end{aligned}$$

(2)$\Rightarrow$(3) Choose any $\mathbf{v} \in \text{Span}(\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_k)$ and write $\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + ... + c_n\mathbf{v}_n$. We have

$$\begin{aligned} & T(\mathbf{v}) \\ = \; & T(c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + ... + c_n\mathbf{v}_n) \\ = \; & c_1T(\mathbf{v}_1) + c_2T(\mathbf{v}_2) + ... + c_nT(\mathbf{v}_n) \\ \in \; & \text{Span}(\mathbf{v}_1) + \text{Span}(\mathbf{v}_1, \mathbf{v}_2) + ... + \text{Span}(\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n) \;\; \text{(since (2) holds)} \\ = \; & \text{Span}(\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n) \;\; \text{(since } \text{Span}(\mathbf{v}_1) \subseteq \text{Span}(\mathbf{v}_1, \mathbf{v}_2) \subseteq ... \subseteq \text{Span}(\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n)\text{)}. \end{aligned}$$

(3)$\Rightarrow$(1) We have

$$\begin{aligned} \mathbf{v}_1 \;\; &\in \;\; \text{Span}(\mathbf{v}_1) \Longrightarrow T(\mathbf{v}_1) \in \text{Span}(\mathbf{v}_1) \text{ (since (2) holds)} \Longrightarrow T(\mathbf{v}_1) = a_{11}\mathbf{v}_1 \\ \mathbf{v}_2 \;\; &\in \;\; \text{Span}(\mathbf{v}_1, \mathbf{v}_2) \Longrightarrow T(\mathbf{v}_2) \in \text{Span}(\mathbf{v}_1, \mathbf{v}_2) \Longrightarrow T(\mathbf{v}_2) = a_{12}\mathbf{v}_1 + a_{22}\mathbf{v}_2 \\ & \;\; \vdots \\ \mathbf{v}_n \;\; &\in \;\; \text{Span}(\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n) \Longrightarrow T(\mathbf{v}_n) = a_{1n}\mathbf{v}_1 + a_{2n}\mathbf{v}_2 + ... + a_{nn}\mathbf{v}_n. \end{aligned}$$

Since

$$M(T, B, B) = \begin{bmatrix} T(\mathbf{v}_1) & T(\mathbf{v}_2) & \cdots & T(\mathbf{v}_n) \end{bmatrix}$$

we conclude that $M(T, B, B)$ has the form $(*)$ and is upper triangular.

Lemma 9. If $T \in \mathcal{A}(F^n)$ and $\lambda \in F$, then $\text{Im}(T - \lambda 1_{F^n})$ is invariant under T.

Proof. Let $V = \text{Im}(T - \lambda 1_{F^n})$ and choose any $\mathbf{v} \in V$. Then

$$\begin{aligned} & T(\mathbf{v}) \\ = \; & T(\mathbf{v}) - \lambda\mathbf{v} + \lambda\mathbf{v} \\ = \; & (T - \lambda 1_{F^n})(\mathbf{v}) + \lambda\mathbf{v} \\ \in \; & \text{Im}(T - \lambda 1_{F^n}) + V \\ = \; & V \end{aligned}$$

Definition 10. Let $V \leq F^n$ and define $\mathcal{A}(V) = \{V \to^T V : T \text{ is a linear transformation}\}$.

Lemma 11. If $V \leq F^n$ is invariant under $T \in \mathcal{A}(F^n)$, then the restriction map $T_V : V \to V$ given by $T_V(\mathbf{v}) = T(\mathbf{v})$ belongs to $\mathcal{A}(V)$.

Proof. Since V is invariant under T, we have that $T(V) \subseteq V$. It follows that $T_V : V \to V$ is a well-defined function (in particular, WD1 is satisfied). Since $T_V(\mathbf{v}) = T(\mathbf{v})$ for all $\mathbf{v} \in V$, it is certainly true that T_V is a linear transformation. That is, $T_V \in \mathcal{A}(V)$.

Theorem 12. Every $T \in \mathcal{A}(\mathbb{C}^n)$ is triangularizable.

Proof. The proof goes by induction on $n = \dim(\mathbb{C}^n)$ and we dispense with the trivial case when $n = 1$. Suppose that $n > 1$ the theorem holds for all linear transformations $S \in \mathcal{A}(W)$ with $\dim(W) < n$. By Theorem 6, T has an eigenvalue $\lambda \in \mathbb{C}$. It follows that $\ker(T - \lambda 1_{\mathbb{C}^n})$ is non trivial and so $T - \lambda 1_{\mathbb{C}^n}$ is not 1-1. But then, $T - \lambda 1_{\mathbb{C}^n}$ is not onto (square) and so the subspace $\mathrm{Im}(T - \lambda 1_{\mathbb{C}^n})$ is *properly* contained in $\mathbb{C}^n$. Setting $V = \mathrm{Im}(T - \lambda 1_{\mathbb{C}^n})$ we have that $\dim(V) < n$, say $\dim(V) = m$. By Lemma 11, V is invriant under T. By Lemma 12, the restriction $T_V : V \to V$ belongs to $\mathcal{A}(V)$. The induction assumption says that we can find a basis $B_V = \{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_m\}$ such that $M(T_V, B, B)$ is triangular. It follows that for each $j \le m$, we have $T(\mathbf{v}_j) = T_V(\mathbf{v}_j) \in \mathrm{Span}(\mathbf{v}_1, ..., \mathbf{v}_j)$. Now extend B_V to a basis $B = \{\mathbf{v}_1, ..., \mathbf{v}_m, \mathbf{x}_{m+1}, ..., \mathbf{x}_n\}$ of all of $\mathbb{C}^n$. We now have

$$
\begin{aligned}
& T(\mathbf{x}_j) \\
= \;& T(\mathbf{x}_j) - \lambda \mathbf{x}_j + \lambda \mathbf{x}_j \\
= \;& (T - \lambda 1_{F^n})(\mathbf{x}_j) + \lambda \mathbf{x}_j \\
\in \;& \mathrm{Im}(T - \lambda 1_{F^n}) + \mathrm{Span}(\mathbf{x}_j) \\
= \;& V + \mathrm{Span}(\mathbf{x}_j) \\
= \;& \mathrm{Span}(\mathbf{v}_1, ..., \mathbf{v}_m) + \mathrm{Span}(\mathbf{x}_j) \\
\subseteq \;& \mathrm{Span}(\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_m, \mathbf{x}_1, ..., \mathbf{x}_j)
\end{aligned}
$$

and so T is triangularizable in the basis B.

Theorem 13. The upper triangular matrix in $(*)$ above is invertible if and only if $a_{jj} \neq 0$ for each $j \in \{1, 2, ..., n\}$.

Proof. Let A be the triangular matrix in $(*)$ and let $E_n = \{\mathbf{e}_1, ..., \mathbf{e}_n\}$ be the standard basis. We let $T : F^n \to F^n$ be the usual linear transformation such that $M(T, E_n, E_n) = A$.

($\Rightarrow$) We will prove that if $a_{jj} = 0$ for some $j \in \{1, 2, ..., n\}$, then A is not invertible. If $a_{11} = 0$, then

$$T(\mathbf{e}_1) = a_{11}\mathbf{e}_1 = 0.$$

which means that $\ker T \neq \{\mathbf{0}\}$. It follows that T is not 1-1 and hence, not invertible. Suppose now that $a_{jj} = 0$ for some $j \in \{2, 3, ..., n\}$. It follows that

$$T(\mathbf{e}_j) = a_{1j}\mathbf{e}_1 + a_{2j}\mathbf{e}_2 + ... + a_{j-1,j}\mathbf{e}_{j-1} + 0\mathbf{e}_j \in \mathrm{Span}(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_{j-1}).$$

Let T_j be the restriction of the map T to the domain $\mathrm{Span}(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_{j-1})$. It follows from the previous equation and Theorem 8 that $T(\mathrm{Span}(\mathbf{e}_1, ...\mathbf{e}_{j-1}, \mathbf{e}_j)) \subseteq \mathrm{Span}(\mathbf{e}_1, ...\mathbf{e}_{j-1})$. Hence,

$$T_j : \mathrm{Span}(\mathbf{e}_1, ..., \mathbf{e}_{j-1}, \mathbf{e}_j) \to \mathrm{Span}(\mathbf{e}_1, ..., \mathbf{e}_{j-1}).$$

Since $\{\mathbf{e}_1, ...\mathbf{e}_{j-1}, \mathbf{e}_j\}$ is a linearly independent set, we have that

$$\dim(\mathrm{Span}(\mathbf{e}_1, ..., \mathbf{e}_{j-1}, \mathbf{e}_j)) = j > j - 1 = \dim(\mathrm{Span}(\mathbf{e}_1, ..., \mathbf{e}_{j-1})).$$

It is therefore impossible that T_j is 1-1. Since T_j is a restriction of T, it is impossible that T is 1-1 and hence T is not invertible.

($\Leftarrow$) If A is not invertible, then it is certainly true that $A_T = A$ is not invertible. It follows that T is not 1-1 (since T is square) and there exists a *nonzero* $\mathbf{v} \in F^n$ such that $T(\mathbf{v}) = \mathbf{0}$. Let us write $\mathbf{v} = c_1\mathbf{e}_1 + c_2\mathbf{e}_2 + ... + c_m\mathbf{e}_m$ where $m \leq n$ and m is the largest index of a nonzero scalar coefficient. That is, $c_m \neq 0$ and $c_k = 0$ for all $k \in \{m+1, m+2, ..., n\}$. Then

$$\begin{aligned} & \mathbf{0} \\ = \ & T(\mathbf{v}) \\ = \ & T(c_1\mathbf{e}_1 + ... + +c_{m-1}\mathbf{e}_{m-1} + c_m\mathbf{e}_m) \\ = \ & T(c_1\mathbf{e}_1 + ... + +c_{m-1}\mathbf{e}_{m-1}) + c_mT(\mathbf{e}_m). \end{aligned}$$

Since $c_m \neq 0$, we can write

$$T(\mathbf{e}_m) = -\frac{1}{c_m}T(c_1\mathbf{e}_1 + ... + +c_{m-1}\mathbf{e}_{m-1}).$$

But T is triangular making $\text{Span}(\mathbf{e}_1, ...\mathbf{e}_{m-1})$ invriant under T and we obtain

$$T(\mathbf{e}_m) \in \text{Span}(\mathbf{e}_1, ...\mathbf{e}_{m-1}).$$

On the other hand, the equation

$$T(\mathbf{e}_m) = a_{1m}\mathbf{e}_1 + ... + a_{m-1,m}\mathbf{e}_{m-1} + a_{mm}\mathbf{e}_m$$

forces $a_{mm} = 0$ as needed.

Theorem 14. If $A \in \mathcal{M}_n$ is an upper triangular matrix, then the elemenets on the main diagonal are precisely the eigenvalues of the linear transformation $T \in \mathcal{A}(F^n)$ such that $A_T = A$.

Proof. We have that

$$\begin{aligned} & \lambda \text{ is an eigenvalue of } T \\ \Leftrightarrow \ & \ker(\theta - \lambda 1_V) \neq \{0\} \\ \Leftrightarrow \ & T - \lambda 1_V \text{ is not invertible} \\ \Leftrightarrow \ & A_{T-\lambda 1_V} \text{ is not invertible} \\ \Leftrightarrow \ & a_{j,j} - \lambda = 0 \text{ for some } j \in \{1, 2, ..., n\} \\ \Leftrightarrow \ & \lambda = a_{j,j} \text{ for some } j \in \{1, 2, ..., n\}. \end{aligned}$$

Definition 15. Let $T \in \mathcal{A}(\mathbb{C}^n)$ and let B be a basis for $\mathbb{C}^n$ such that $M = M(T, B, B)$ is triangular with diagonal entries $\text{ent}_{ii}(M) = a_{ii}$. We define the characteristic polynomial $\chi_T(x) \in \mathbb{C}[x]$ to be the polynomial

$$\chi_T(x) = (x - a_{11})(x - a_{22}) \cdots (x - a_{nn})$$

Theorem 16. (Cayley-Hamilton Theorem) If $T \in \mathcal{A}(\mathbb{C}^n)$, then $\chi_T(T) = 0$. Equivalently, $\chi_T(A_T) = [0]$.

Proof. By Theorem 12, T is a triangular linear transformation with respect to some basis $B = \{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n\}$. Using the matrix $A = M(T, B, B)$ in equation $(*)$ of Definition 7 and rearranging some terms, we have

$$\begin{aligned}
(T - a_{11})(\mathbf{v}_1) &= \mathbf{0} \\
(T - a_{22})(\mathbf{v}_2) &= T(\mathbf{v}_2) - a_{22}\mathbf{v}_2 = a_{12}\mathbf{v}_1 \\
(T - a_{33})(\mathbf{v}_3) &= T(\mathbf{v}_3) - a_{33}\mathbf{v}_3 = a_{13}\mathbf{v}_1 + a_{23}\mathbf{v}_2 \\
&\vdots \\
(T - a_{nn})(\mathbf{v}_n) &= T(\mathbf{v}_n) - a_{nn}\mathbf{v}_n = a_{1n}\mathbf{v}_1 + a_{2n}\mathbf{v}_2 + ... + a_{n-1,n}\mathbf{v}_{n-1}.
\end{aligned}$$

We will show that $\chi_T(T) = (T - a_{11})(T - a_{22}) \cdots (T - a_{nn})$ annihilates every basis element $\mathbf{v}_j$, $1 \leq j \leq n$. We will use the fact that the linear transformations $(T - a_{ii})$ and c commute for all $i, j \in \{1, 2, ..., n\}$. That is,

$$[(T - a_{jj})(T - a_{ii})](\mathbf{v}) = [(T - a_{ii})(T - a_{jj})](\mathbf{v})$$

for all $\mathbf{v} \in \mathbb{C}^n$ (verify!). We have

$$(T - a_{11})(\mathbf{v}_1) = \mathbf{0}$$

and so by commutativity,

$$\begin{aligned}
& \chi_T(T)(\mathbf{v}_1) \\
= \; & \left[(T - a_{11})(T - a_{22}) \cdots (T - a_{nn})\right](\mathbf{v}_1) \\
= \; & \left[(T - a_{22}) \cdots (T - a_{nn})(T - a_{11})\right](\mathbf{v}_1) \\
= \; & \left[(T - a_{22}) \cdots (T - a_{nn})\right](\mathbf{0}) \\
= \; & \mathbf{0}.
\end{aligned}$$

At the next step,

$$\begin{aligned}
& [(T - a_{11})(T - a_{22})](\mathbf{v}_2) \\
= \; & (T - a_{11})(a_{12}\mathbf{v}_1) \\
= \; & a_{12}(T - a_{11})(\mathbf{v}_1) \\
= \; & a_{12}\mathbf{0} \\
= \; & \mathbf{0}
\end{aligned}$$

and similarly,

$$\chi_T(T)(\mathbf{v}_2) = \mathbf{0}.$$

One more for good measure.

$$\begin{aligned}
& [(T-a_{11})(T-a_{22})(T-a_{33})](\mathbf{v}_3) \\
= \; & [(T-a_{11})(T-a_{22})](a_{13}\mathbf{v}_1 + a_{23}\mathbf{v}_2) \\
= \; & a_{13}[(T-a_{11})(T-a_{22})](\mathbf{v}_1) + a_{23}[(T-a_{11})(T-a_{22})](\mathbf{v}_2) \\
= \; & a_{13}[(T-a_{22})(T-a_{11})](\mathbf{v}_1) + a_{23}[(T-a_{11})(T-a_{22})](\mathbf{v}_2) \\
= \; & \mathbf{0}.
\end{aligned}$$

Continuing on in this fashion, we find that

$$[\chi_T(T)](\mathbf{v}_j) = [(T-a_{11})(T-a_{22})\cdots(T-a_{nn})](\mathbf{v}_j) = \mathbf{0}$$

for all basis elements $\mathbf{v}_j$. It follows that $\chi_T(T)$ kills every vector in $\mathbb{C}^n$ so that $\chi_T(T) = 0$.

Exercises

1. Let

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 2 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 3 \end{bmatrix}$$

Justify all responses
(a) Is A invertible?
(b) Write down the characteristicpolynomial $\chi_A(x)$.
(c) Write down all eigenvalues of A.
(d) Write down the basis vectors for each eigenspace.
(e) Is A diagonalizable? If so, find an invertible matrix P and a diagonal matrix D such that $D = PAP^{-1}$.

2. Let

$$A = \begin{bmatrix} 2 & 1 & 0 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 & 1 \\ 0 & 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 3 \end{bmatrix}$$

Justify all responses
(a) Is A invertible?
(b) Write down the characteristicpolynomial $\chi_A(x)$.
(c) Write down all eigenvalues of A.
(d) Write down the basis vectors for each eigenspace.
(e) Is A diagonalizable? If so, find an invertible matrix P and a diagonal matrix D such that $D = PAP^{-1}$.

3. Let $f(x) \in F[x]$ and let $\lambda \in F$ be an eigenvalue of $T \in \mathcal{A}(F^n)$. Prove the following statements.
(a) $f(\lambda)$ is an eigenvalue of $f(T)$.
(b) $f(T) = 0$ implies $f(\lambda) = 0$.
(c) $f(T) = 0$ and $f(\mu) = 0$ does not imply μ is an eigenvalue of T. (Give a counterexample and a careful explanation.)

4. Upper triangularity is really needed in Theorem 13! Give an example of a square matrix $A \in \mathcal{M}_n$ with the following properties.
(a) $\text{ent}_{ii}(A) \neq 0$ for each $i \leq n$ but A is not invertible.
(b) $\text{ent}_{ii}(A) = 0$ for each $i \leq n$ but A is invertible.

5. Let $S, T \in \mathcal{A}(F^n)$ with S invertible. Given any polynomial $p(x) \in F[x]$, prove that $p(STS^{-1}) = Sp(T)S^{-1}$.

6. Let $T \in \mathcal{A}(\mathbb{C}^n)$ and $p(x) \in \mathbb{C}[x]$. Prove that $\lambda \in \mathbb{C}$ is an eigenvalue of $p(T)$ if and only if T has an eigenvalue $\mu \in \mathbb{C}$ such that $p(\mu) = \lambda$. Does the result hold if $\mathbb{C}$ is replaced by $\mathbb{R}$?

7. Let $T \in \mathcal{A}(F^n)$. Prove that, for each $k \in \{1, 2, ..., n\}$, there is a T-invariant subspace $U_k \leq F^n$ such that $\dim(U_k) = k$.

Exercises

1. Prove that a polynomial $f(x) \in F[x]$ has a root $\alpha \in F$ if and only if there exists a polynomial $g(x) \in F[x]$ such that $f(x) = (x - \alpha)g(x)$. Hint: Use Theorem 5 of the polynomial notes.

2. Prove that if $f(x) \in F[x]$ with $\deg(f) = n \in \mathbb{N}$, then $f(x)$ has at most n roots in F. Hint: Use Exercise 1 and induction.

3. For every polynomial $f(x) \in \mathbb{C}[x]$ with $\deg(f) = n$, there exists a unique set $\{c, z_1, z_2, ..., z_n\} \in \mathbb{C}$ such that $f(x) = c(x - z_1)(x - z_2) \cdots (x - z_n)$. Hint: Use Theorem 6 of the polynomial notes and Exercise 1 for existence.

4. A complex number $z \in \mathbb{C}$ is a root of the polynomial $f(x) \in \mathbb{R}[x]$ if and only if the conjugate $\overline{z}$ is a root of $f(x)$. Hint: Theorem 3 of the polynomial notes.

5. Let $f(x) = x^2 + bx + c$ be a polynomial in $\mathbb{R}[x]$. Prove that there is a polynomial factorization of the form $f(x) = (x - r_1)(x - r_2)$ for some $r_1, r_2 \in \mathbb{R}$ if and only if $b^2 - 4c \geq 0$.