MATH 110: LINEAR ALGEBRA
SPRING 2007/08
PROBLEM SET 5 SOLUTIONS

1. Let U, W, V be finite-dimensional vector spaces over $\mathbb{F}$. Let $\alpha, \beta \in \mathbb{F}$.

(a) Let $\mathcal{T} : V \to W$ be a linear transformation. Show that

$$\operatorname{rank}(\mathcal{T}) \leq \dim(V).$$

Solution. Let $\dim(V) = n$. Suppose $\dim(\operatorname{im}(\mathcal{T})) = \operatorname{rank}(\mathcal{T}) > n$. Then there must be at least $n+1$ vectors $\mathbf{w}_1, \dots, \mathbf{w}_{n+1} \in \operatorname{im}(\mathcal{T})$ that are linearly independent (eg. choose the first $n+1$ vectors in a basis of $\operatorname{im}(\mathcal{T})$). Since $\mathbf{w}_1, \dots, \mathbf{w}_{n+1}$ are in $\operatorname{im}(\mathcal{T})$, there exist $\mathbf{v}_1, \dots, \mathbf{v}_{n+1} \in V$ such that $\mathcal{T}(\mathbf{v}_i) = \mathbf{w}_i$, $i = 1, \dots, n+1$. Since $\dim(V) = n$, $\mathbf{v}_1, \dots, \mathbf{v}_{n+1}$ must be linearly dependent, so there exist $\alpha_1, \dots, \alpha_{n+1} \in \mathbb{F}$, not all zero, such that

$$\alpha_1 \mathbf{v}_1 + \cdots + \alpha_{n+1} \mathbf{v}_{n+1} = \mathbf{0}_V.$$

Hence we have found $\alpha_1, \dots, \alpha_{n+1} \in \mathbb{F}$, not all zero, such that

$$\mathcal{T}(\alpha_1 \mathbf{v}_1 + \cdots + \alpha_{n+1} \mathbf{v}_{n+1}) = \mathcal{T}(\mathbf{0}_V),$$

ie.

$$\alpha_1 \mathcal{T}(\mathbf{v}_1) + \cdots + \alpha_{n+1} \mathcal{T}(\mathbf{v}_{n+1}) = \mathbf{0}_W,$$

ie.

$$\alpha_1 \mathbf{w}_1 + \cdots + \alpha_{n+1} \mathbf{w}_{n+1} = \mathbf{0}_W,$$

which implies that $\mathbf{w}_1, \dots, \mathbf{w}_{n+1}$ are linearly dependent — a contradiction. Hence our original assumption must have been false, ie. we must have $\dim(\operatorname{im}(\mathcal{T})) \leq n$.

(b) Let $\mathcal{S} : U \to V$ and $\mathcal{T} : V \to W$ be linear transformations. Show that

$$\operatorname{rank}(\mathcal{T} \circ \mathcal{S}) \leq \operatorname{rank}(\mathcal{T})$$

and

$$\operatorname{rank}(\mathcal{T} \circ \mathcal{S}) \leq \operatorname{rank}(\mathcal{S}).$$

Solution. Recall the following elementary fact from set theory: $A \subseteq B$ implies $f(A) \subseteq f(B)$ for any function f. In particular, since $\mathcal{S}(U) \subseteq V$, we must have $\mathcal{T}(\mathcal{S}(U)) \subseteq \mathcal{T}(V)$, ie. $\mathcal{T} \circ \mathcal{S}(U) \subseteq \mathcal{T}(V)$, ie.

$$\operatorname{im}(\mathcal{T} \circ \mathcal{S}) \subseteq \operatorname{im}(\mathcal{T})$$

(recall that $\operatorname{im}(\varphi)$ is just another way of writing $\varphi(V)$ — both denote the range of φ). Now both of these are subspaces of W and so

$$\dim(\operatorname{im}(\mathcal{T} \circ \mathcal{S})) \leq \dim(\operatorname{im}(\mathcal{T})),$$

ie.

$$\operatorname{rank}(\mathcal{T} \circ \mathcal{S}) \leq \operatorname{rank}(\mathcal{T}).$$

For the second part, we define the function $\varphi : \operatorname{im}(\mathcal{S}) \to W$ by $\varphi(\mathbf{v}) = \mathcal{T}(\mathbf{v})$ for all $\mathbf{v} \in \operatorname{im}(S)$. Note that φ is a linear transformation since for all $\mathbf{v}_1, \mathbf{v}_2 \in \operatorname{im}(\mathcal{S})$ and $\alpha, \beta \in \mathbb{F}$, we have $\varphi(\alpha\mathbf{v}_1 + \beta\mathbf{v}_2) = \mathcal{T}(\alpha\mathbf{v}_1 + \beta\mathbf{v}_2) = \alpha\mathcal{T}(\mathbf{v}_1) + \beta\mathcal{T}(\mathbf{v}_2) = \alpha\varphi(\mathbf{v}_1) + \beta\varphi(\mathbf{v}_2)$. So applying part (a) to φ, we get

$$\operatorname{rank}(\varphi) \leq \dim(\operatorname{im}(\mathcal{S})) = \operatorname{rank}(\mathcal{S}). \tag{1.1}$$

Date: May 3, 2008 (Version 1.0).

But note that $\operatorname{im}(\mathcal{T} \circ \mathcal{S}) \subseteq \operatorname{im}(\varphi)$ — since if $\mathbf{w} \in \operatorname{im}(\mathcal{T} \circ \mathcal{S})$, then $\mathbf{w} = \mathcal{T}(\mathcal{S}(\mathbf{u}))$ for some $\mathbf{u} \in U$ and therefore $\mathbf{w} = \mathcal{T}(\mathbf{v}) = \varphi(\mathbf{v})$ where $\mathbf{v} = \mathcal{S}(\mathbf{u}) \in \operatorname{im}(\mathcal{S})$; hence $\mathbf{w} \in \operatorname{im}(\varphi)$. So

$$\dim(\operatorname{im}(\mathcal{T} \circ \mathcal{S})) \leq \dim(\operatorname{im}(\varphi)),$$

ie.

$$\operatorname{rank}(\mathcal{T} \circ \mathcal{S}) \leq \operatorname{rank}(\varphi). \tag{1.2}$$

Combining (1.1) and (1.2) then yields

$$\operatorname{rank}(\mathcal{T} \circ \mathcal{S}) \leq \operatorname{rank}(\mathcal{S}).$$

(c) Let $\mathcal{S}_1 : U \to V$, $\mathcal{S}_2 : U \to V$, and $\mathcal{T} : V \to W$ be linear transformations. Show that

$$\mathcal{T} \circ (\alpha\mathcal{S}_1 + \beta\mathcal{S}_2) = \alpha\mathcal{T} \circ \mathcal{S}_1 + \beta\mathcal{T} \circ \mathcal{S}_2.$$

Solution. Let $\mathbf{u} \in U$. Then

$$\begin{aligned}
\mathcal{T} \circ (\alpha\mathcal{S}_1 + \beta\mathcal{S}_2)(\mathbf{u}) &= \mathcal{T}((\alpha\mathcal{S}_1 + \beta\mathcal{S}_2)(\mathbf{u})) \\
&= \mathcal{T}(\alpha\mathcal{S}_1(\mathbf{u}) + \beta\mathcal{S}_2(\mathbf{u})) \\
&= \alpha\mathcal{T}(\mathcal{S}_1(\mathbf{u})) + \beta\mathcal{T}(\mathcal{S}_2(\mathbf{u})) \\
&= \alpha\mathcal{T} \circ \mathcal{S}_1(\mathbf{u}) + \beta\mathcal{T} \circ \mathcal{S}_2(\mathbf{u}) \\
&= (\alpha\mathcal{T} \circ \mathcal{S}_1 + \beta\mathcal{T} \circ \mathcal{S}_2)(\mathbf{u}).
\end{aligned}$$

Since this is true for any $\mathbf{u} \in U$, we have that

$$\mathcal{T} \circ (\alpha\mathcal{S}_1 + \beta\mathcal{S}_2) = \alpha\mathcal{T} \circ \mathcal{S}_1 + \beta\mathcal{T} \circ \mathcal{S}_2.$$

(d) Let $\mathcal{S} : U \to V$, $\mathcal{T}_1 : V \to W$, and $\mathcal{T}_2 : V \to W$ be linear transformations. Show that

$$(\alpha\mathcal{T}_1 + \beta\mathcal{T}_2) \circ \mathcal{S} = \alpha\mathcal{T}_1 \circ \mathcal{S} + \beta\mathcal{T}_2 \circ \mathcal{S}.$$

Solution. Let $\mathbf{u} \in U$. Then

$$\begin{aligned}
(\alpha\mathcal{T}_1 + \beta\mathcal{T}_2) \circ \mathcal{S}(\mathbf{u}) &= (\alpha\mathcal{T}_1 + \beta\mathcal{T}_2)(\mathcal{S}(\mathbf{u})) \\
&= \alpha\mathcal{T}_1(\mathcal{S}(\mathbf{u})) + \beta\mathcal{T}_2(\mathcal{S}(\mathbf{u})) \\
&= \alpha\mathcal{T}_1 \circ \mathcal{S}(\mathbf{u}) + \beta\mathcal{T}_2 \circ \mathcal{S}(\mathbf{u}) \\
&= (\alpha\mathcal{T}_1 \circ \mathcal{S} + \beta\mathcal{T}_2 \circ \mathcal{S})(\mathbf{u}).
\end{aligned}$$

Since this is true for any $\mathbf{u} \in U$, we have that

$$(\alpha\mathcal{T}_1 + \beta\mathcal{T}_2) \circ \mathcal{S} = \alpha\mathcal{T}_1 \circ \mathcal{S} + \beta\mathcal{T}_2 \circ \mathcal{S}.$$

(e) Let $\mathcal{S}_1 : \mathbb{R}^3 \to \mathbb{R}^2$, $\mathcal{S}_2 : \mathbb{R}^3 \to \mathbb{R}^2$, $\mathcal{T} : \mathbb{R}^2 \to \mathbb{R}^2$ be defined by

$$\mathcal{S}_1(x, y, z) = (y, x + z), \quad \mathcal{S}_2(x, y, z) = (2z, x - y), \quad \mathcal{T}(x, y) = (y, 2x)$$

respectively. Find the formula that defines $\mathcal{T}^2 \circ (3\mathcal{S}_1 - 5\mathcal{S}_2)$ and state its domain and codomain. Is this a linear transformation?

Solution. $3\mathcal{S}_1 - 5\mathcal{S}_2 : \mathbb{R}^3 \to \mathbb{R}^2$ and its formula is given by

$$\begin{aligned}
(3\mathcal{S}_1 - 5\mathcal{S}_2)(x, y, z) &= 3\mathcal{S}_1(x, y, z) - 5\mathcal{S}_2(x, y, z) \\
&= 3(y, x + z) - 5(2z, x - y) \\
&= (3y - 10z, -2x + 3z + 5y).
\end{aligned}$$

$\mathcal{T}^2 \circ (3\mathcal{S}_1 - 5\mathcal{S}_2) : \mathbb{R}^3 \to \mathbb{R}^2$ and its formula is given by

$$\begin{aligned}
\mathcal{T}^2 \circ (3\mathcal{S}_1 - 5\mathcal{S}_2)(x, y, z) &= \mathcal{T}^2(3\mathcal{S}_1 - 5\mathcal{S}_2)(x, y, z)) \\
&= \mathcal{T}^2(3y - 10z, -2x + 3z + 5y) \\
&= \mathcal{T}(-2x + 3z + 5y, 6y - 20z) \\
&= (6y - 20z, -4x + 6z + 10y).
\end{aligned}$$

It is clearly a linear transformation.

2. Let V and W be finite-dimensional vector spaces over $\mathbb{F}$. Let $\mathscr{B}_V = \{\mathbf{v}_1, \dots, \mathbf{v}_m\}$ and $\mathscr{B}_W = \{\mathbf{w}_1, \dots, \mathbf{w}_n\}$ be bases of V and W respectively, where $m = \dim(V)$ and $n = \dim(W)$.

(a) Let $i \in \{1, \dots, m\}$ and $j \in \{1, \dots, n\}$ be fixed. Show that we can define a linear transformation $\mathcal{F}_{ij} : V \to W$ such that

$$\mathcal{F}_{ij}(\mathbf{v}_i) = \mathbf{w}_j,$$

and for all $k \neq i$,

$$\mathcal{F}_{ij}(\mathbf{v}_k) = \mathbf{0}.$$

Solution. Let $\mathbf{v} \in V$. Then since $\mathscr{B}_V$ is a basis, $\mathbf{v}$ may be uniquely expressed as

$$\mathbf{v} = \alpha_1 \mathbf{v}_1 + \cdots + \alpha_m \mathbf{v}_m$$

for some $\alpha_1, \dots, \alpha_m \in \mathbb{F}$. We will define $\mathcal{F}_{ij}(\mathbf{v})$ to be

$$\mathcal{F}_{ij}(\mathbf{v}) = \alpha_i \mathbf{w}_j.$$

Clearly $\mathcal{F}_{ij}$ is well-defined by the uniqueness of the coefficient α_i for each given $\mathbf{v}$. Also $\mathcal{F}_{ij}$ has the required properties since

$$\mathcal{F}_{ij}(\mathbf{v}_k) = \begin{cases} 1 \cdot \mathbf{w}_j & k = i, \\ 0 \cdot \mathbf{w}_j & k \neq i, \end{cases} = \begin{cases} \mathbf{w}_j & k = i, \\ \mathbf{0} & k \neq i. \end{cases}$$

It remains to show that $\mathcal{F}_{ij}$ is a linear transformation. Let $\mathbf{v}, \mathbf{u} \in V$ and $\lambda, \mu \in \mathbb{F}$. Suppose also that

$$\mathbf{u} = \beta_1 \mathbf{v}_1 + \cdots + \beta_m \mathbf{v}_m.$$

Then

$$\begin{aligned} \mathcal{F}_{ij}(\lambda \mathbf{v} + \mu \mathbf{u}) &= \mathcal{F}_{ij}(\lambda(\alpha_1 \mathbf{v}_1 + \cdots + \alpha_m \mathbf{v}_m) + \mu(\beta_1 \mathbf{v}_1 + \cdots + \beta_m \mathbf{v}_m)) \\ &= \mathcal{F}_{ij}((\lambda\alpha_1 + \mu\beta_1)\mathbf{v}_1 + \cdots + (\lambda\alpha_m + \mu\beta_m)\mathbf{v}_m) \\ &= (\lambda\alpha_i + \mu\beta_i)\mathbf{w}_j \\ &= \lambda(\alpha_i \mathbf{w}_j) + \mu(\beta_i \mathbf{w}_j) \\ &= \lambda\mathcal{F}_{ij}(\alpha_1 \mathbf{v}_1 + \cdots + \alpha_m \mathbf{v}_m) + \mu\mathcal{F}_{ij}(\beta_1 \mathbf{v}_1 + \cdots + \beta_m \mathbf{v}_m) \\ &= \lambda\mathcal{F}_{ij}(\mathbf{v}) + \mu\mathcal{F}_{ij}(\mathbf{u}) \end{aligned}$$

as required.

(b) Show that the set $\{\mathcal{F}_{ij} \mid i = 1, \dots, m;\ j = 1, \dots, n\}$ is a basis for $\mathrm{Hom}_{\mathbb{F}}(V, W)$.

Solution. Let $\alpha_{ij} \in \mathbb{F}$ for $i = 1, \dots, m$ and $j = 1, \dots, n$. Suppose the linear combination

$$\sum\nolimits_{i,j=1}^{m,n} \alpha_{ij} \mathcal{F}_{ij} = \mathcal{O}.$$

Then

$$\left(\sum\nolimits_{i,j=1}^{m,n} \alpha_{ij} \mathcal{F}_{ij}\right)(\mathbf{v}) = \mathcal{O}(\mathbf{v}) \tag{2.3}$$

for all $\mathbf{v} \in V$. By the definition of a linear combination of linear transformation, the LHS is just

$$\left(\sum\nolimits_{i,j=1}^{m,n} \alpha_{ij} \mathcal{F}_{ij}\right)(\mathbf{v}) = \sum\nolimits_{i,j=1}^{m,n} \alpha_{ij} \mathcal{F}_{ij}(\mathbf{v})$$

and the RHS is of course just $\mathbf{0}_W$. Choosing $\mathbf{v} = \mathbf{v}_k$ for a $k \in \{1, \dots, n\}$, we get

$$\sum\nolimits_{i,j=1}^{m,n} \alpha_{ij} \mathcal{F}_{ij}(\mathbf{v}_k) = \sum\nolimits_{j=1}^{n} \alpha_{kj} \mathcal{F}_{ij}(\mathbf{v}_k) = \sum\nolimits_{j=1}^{n} \alpha_{kj} \mathbf{w}_j.$$

So (2.3) becomes

$$\sum\nolimits_{j=1}^{n} \alpha_{kj} \mathbf{w}_j = \mathbf{0}_W$$

and since $\mathscr{B}_W$ is a basis (thus linearly independent),

$$\alpha_{k1} = \cdots = \alpha_{kn} = 0.$$

Since $k \in \{1, \ldots, n\}$ is arbitrary, we conclude that $\alpha_{ij} = 0$ for $i = 1, \ldots, m$ and $j = 1, \ldots, n$. Hence $\{\mathcal{F}_{ij} \mid i = 1, \ldots, m;\ j = 1, \ldots, n\}$ is linearly independent. We will now show that it also spans $\mathrm{Hom}_{\mathbb{F}}(V, W)$. Let $\mathcal{T} \in \mathrm{Hom}_{\mathbb{F}}(V, W)$. For each $\mathbf{v}_i \in \mathscr{B}_V$, $i \in \{1, \ldots, n\}$, we express $\mathcal{T}(\mathbf{v}_i)$ in terms of basis vectors in $\mathscr{B}_W$,

$$\mathcal{T}(\mathbf{v}_i) = \gamma_{i1}\mathbf{w}_1 + \cdots + \gamma_{in}\mathbf{w}_n.$$

For a general $\mathbf{v} \in V$ with $\mathbf{v} = \delta_1\mathbf{v}_1 + \cdots + \delta_m\mathbf{v}_m$,

$$\begin{aligned}
\mathcal{T}(\mathbf{v}) &= \delta_1\mathcal{T}(\mathbf{v}_1) + \cdots + \delta_m\mathcal{T}(\mathbf{v}_m) \\
&= \sum\nolimits_{i=1}^{m} \delta_i(\gamma_{i1}\mathbf{w}_1 + \cdots + \gamma_{in}\mathbf{w}_n) \\
&= \sum\nolimits_{i,j=1}^{m,n} \delta_i\gamma_{ij}\mathbf{w}_j \\
&= \sum\nolimits_{i,j=1}^{m,n} \delta_i\gamma_{ij}\mathcal{F}_{ij}(\mathbf{v}_i) \\
&= \sum\nolimits_{i,j=1}^{m,n} \gamma_{ij}\mathcal{F}_{ij}(\delta_1\mathbf{v}_1 + \cdots + \delta_m\mathbf{v}_m) \\
&= \sum\nolimits_{i,j=1}^{m,n} \gamma_{ij}\mathcal{F}_{ij}(\mathbf{v}).
\end{aligned}$$

Since this holds for all $\mathbf{v} \in V$, we have that

$$\mathcal{T} = \sum\nolimits_{i,j=1}^{m,n} \gamma_{ij}\mathcal{F}_{ij},$$

ie. $\mathcal{T} \in \mathrm{span}\{\mathcal{F}_{ij} \mid i = 1, \ldots, m;\ j = 1, \ldots, n\}$, as required.

(c) What are the values of

$$\dim(\mathrm{Hom}_{\mathbb{F}}(V, W)), \quad \dim(\mathrm{End}_{\mathbb{F}}(V)), \quad \text{and} \quad \dim(\mathrm{End}_{\mathbb{F}}(W))?$$

Solution. By (b),

$$\begin{aligned}
\dim(\mathrm{Hom}_{\mathbb{F}}(V, W)) &= mn, \\
\dim(\mathrm{End}_{\mathbb{F}}(V)) &= m^2, \\
\dim(\mathrm{End}_{\mathbb{F}}(W)) &= n^2.
\end{aligned}$$

3. Let V and W be finite-dimensional vector spaces over $\mathbb{F}$. Let $\mathcal{T} : V \to W$ be a linear transformation. Prove or disprove the following statements.

(a) If $\mathbf{v}_1, \ldots, \mathbf{v}_n \in V$ are linearly independent, then $\mathcal{T}(\mathbf{v}_1), \ldots, \mathcal{T}(\mathbf{v}_n) \in W$ are also linearly independent.

Solution. False. For a counter example, let $\mathcal{T} = \mathcal{O}$, any set that contains $\mathbf{0}_W = \mathcal{T}(\mathbf{v}_i)$ cannot be linearly independent.

(b) If $\mathcal{T}(\mathbf{v}_1), \ldots, \mathcal{T}(\mathbf{v}_n) \in W$ are linearly independent, then $\mathbf{v}_1, \ldots, \mathbf{v}_n \in V$ are also linearly independent.

Solution. True. Let $\alpha_1, \ldots, \alpha_n \in \mathbb{F}$ be such that

$$\alpha_1\mathbf{v}_1 + \cdots + \alpha_n\mathbf{v}_n = \mathbf{0}_V.$$

Then

$$\mathcal{T}(\alpha_1\mathbf{v}_1 + \cdots + \alpha_n\mathbf{v}_n) = \mathcal{T}(\mathbf{0}_V),$$

and so

$$\alpha_1\mathcal{T}(\mathbf{v}_1) + \cdots + \alpha_n\mathcal{T}(\mathbf{v}_n) = \mathbf{0}_W.$$

Since $\mathcal{T}(\mathbf{v}_1), \ldots, \mathcal{T}(\mathbf{v}_n)$ are linearly independent, it follows that

$$\alpha_1 = \cdots = \alpha_n = 0.$$

Hence $\mathbf{v}_1, \dots, \mathbf{v}_n$ are linearly independent.

(c) If $\mathbf{v}_1, \dots, \mathbf{v}_n$ span V, then $\mathcal{T}(\mathbf{v}_1), \dots, \mathcal{T}(\mathbf{v}_n)$ span W.
SOLUTION. False. For a counter example, let W be any vector space that is not the zero vector space; let $\mathcal{T} = \mathcal{O}$. Then $\operatorname{span}\{(\mathbf{v}_1), \dots, \mathcal{T}(\mathbf{v}_n)\} = \{\mathbf{0}_W\} \neq W$.

(d) If $\mathcal{T}(\mathbf{v}_1), \dots, \mathcal{T}(\mathbf{v}_n)$ span W, then $\mathbf{v}_1, \dots, \mathbf{v}_n$ span V.
SOLUTION. False. Let $V = \mathbb{R}^3$ and $W = \mathbb{R}^2$. Let $\mathbf{v}_1 = (1,0,0)^\top$ and $\mathbf{v}_2 = (0,1,0)^\top$. Let $\mathcal{T} : \mathbb{R}^3 \to \mathbb{R}$ be given by $\mathcal{T}(x,y,z) = (x,y)$. Then $\mathcal{T}(\mathbf{v}_1) = (1,0)^\top, \mathcal{T}(\mathbf{v}_2) = (0,1)^\top$ span $\mathbb{R}^2$ but then $\mathbf{v}_1, \mathbf{v}_2$ do not span $\mathbb{R}^3$.

4. Let $A \in \mathbb{F}^{n\times n}$ be a fixed matrix and $A \neq O$, the zero matrix. Define the function $\mathcal{T}_A : \mathbb{F}^{n\times n} \to \mathbb{F}^{n\times n}$ by

$$\mathcal{T}_A(X) = AX - XA \tag{4.4}$$

for all $X \in \mathbb{F}^{n\times n}$.

(a) Show that $\mathcal{T}_A$ is a linear transformation.
SOLUTION. Let $X_1, X_2 \in \mathbb{F}^{n\times n}$ and $\alpha_1, \alpha_2 \in \mathbb{F}$. Then

$$\begin{aligned}\mathcal{T}_A(\alpha_1 X_1 + \alpha_2 X_2) &= A(\alpha_1 X_1 + \alpha_2 X_2) - (\alpha_1 X_1 + \alpha_2 X_2)A \\ &= \alpha_1(AX_1 - X_1 A) + \alpha_2(AX_2 - X_2 A) \\ &= \alpha_1 \mathcal{T}_A(X_1) + \alpha_2 \mathcal{T}_A(X_2)\end{aligned}$$

by the properties of matrix arithmetic. So $\mathcal{T}_A \in \operatorname{End}(\mathbb{F}^{n\times n})$.

(b) Show that if $A^m = O$, then $\mathcal{T}_A^{2m} = \mathcal{O}$.
SOLUTION. Note that if $A^m = O$, then

$$A^{m+1} = A^{m+2} = \cdots = O.$$

We will try to find a formula for $\mathcal{T}_A^k$. For $k = 2,3,4$, we get

$$\begin{aligned}\mathcal{T}_A^2(X) &= A(AX - XA) - (AX - XA)A \\ &= A^2X - 2AXA + XA^2, \\ \mathcal{T}_A^3(X) &= A^2(AX - XA) - 2A(AX - XA)A + (AX - XA)A^2 \\ &= A^3X - 3A^2XA + 3AXA^2 - XA^3, \\ \mathcal{T}_A^4(X) &= A^2(A^2X - 2AXA + XA^2) - 2A(A^2X - 2AXA + XA^2)A \\ &\qquad + (A^2X - 2AXA + XA^2)A^2 \\ &= A^4X - 4A^3XA + 6A^2XA^2 - 4AXA^3 + XA^4.\end{aligned}$$

Observe that the coefficients are binomial coefficients and so the pattern suggests a general formula

$$\mathcal{T}_A^k(X) = \sum_{i=0}^{k} (-1)^i \binom{k}{i} A^{k-i} X A^i$$

which we can then prove by induction. When $k = 2m$, then

$$A^{k-i} X A^i = O$$

for all $i = 0, 1, \dots, 2m$, since either $k - i \geq m$ or $i \geq m$ (and so either A^{k-i} or A^i must be O). Hence

$$\mathcal{T}_A^{2m}(X) = \mathcal{O}.$$

(c) For any $X, Y \in \mathbb{R}^{n\times n}$, show that

$$\mathcal{T}_A(XY) = X\mathcal{T}_A(Y) + \mathcal{T}_A(X)Y.$$

Solution. This follows from a direct verification:

$$\begin{aligned} X\mathcal{T}_A(Y) + \mathcal{T}_A(X)Y &= X(AY - YA) + (AX - XA)Y \\ &= AXY - XYA \\ &= \mathcal{T}_A(XY). \end{aligned}$$

(d) Let $B \in \mathbb{F}^{n\times n}$ be a fixed non-zero matrix such that

$$AB = BA.$$

Let $\mathcal{T}_B$ be defined as in (4.4). Show that

$$\mathcal{T}_A \circ \mathcal{T}_B = \mathcal{T}_B \circ \mathcal{T}_A.$$

Solution. Note that $\mathcal{T}_A \circ \mathcal{T}_B = \mathcal{T}_B \circ \mathcal{T}_A$ iff for all $X \in \mathbb{F}^{n\times n}$,

$$ABX + XBA = BAX + XAB,$$

iff for all $X \in \mathbb{F}^{n\times n}$,

$$(AB - BA)X = X(AB - BA). \tag{4.5}$$

Now note that (4.5) clearly holds if $AB = BA$.

(e) Prove that $\mathcal{T}_A = \mathcal{O}$ if and only if $A = \lambda I$ for some $\lambda \in \mathbb{F}$.

Solution. Let $A = [a_{ij}]_{i,j=1}^{n}$. Let $E_{ij} \in \mathbb{F}^{n\times n}$ be the matrix with 1 in the (i, j)-entry and 0 elsewhere. Since

$$\mathcal{T}_A(X) = O$$

for all $X \in \mathbb{F}^{n\times n}$,

$$\mathcal{T}_A(E_{ij}) = O$$

and so

$$AE_{ij} = E_{ij}A. \tag{4.6}$$

Now observe that AE_{ij} is the matrix whose jth column is the ith column of A and 0 elsewhere while $E_{ij}A$ is the matrix whose ith row is the jth row of A and 0 elsewhere. So for $i \neq j$, (4.6) implies that $a_{ii} = a_{jj}$ and $a_{kl} = 0$ for all $(k, l) \neq (i, i)$ or (j, j). Since i and j are arbitrary, we conclude that A is a diagonal matrix.

(f) Is the converse of (d) true, ie. is it true that if

$$\mathcal{T}_A \circ \mathcal{T}_B = \mathcal{T}_B \circ \mathcal{T}_A,$$

then

$$AB = BA.$$

Solution. The answer depends on the choice of $\mathbb{F}$. Note that (4.5) holds for all $X \in \mathbb{F}^{n\times n}$ iff

$$\mathcal{T}_{AB-BA} = \mathcal{O}.$$

By (e), the previous equation holds iff

$$AB - BA = \lambda I.$$

Taking trace, we see that

$$\begin{aligned} \operatorname{tr}(AB - BA) &= \operatorname{tr}(\lambda I), \\ \operatorname{tr}(AB) - \operatorname{tr}(BA) &= \lambda \operatorname{tr}(I), \\ 0 &= \lambda n. \end{aligned}$$

If $\mathbb{F} = \mathbb{R}$, then this implies that $\lambda = 0$ and so $AB = BA$. Hence the converse of (d) is true when $\mathbb{F} = \mathbb{R}$. But if $\mathbb{F} = \mathbb{Z}/2\mathbb{Z}$, then letting

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

we see that

$$AB = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}, \quad BA = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix},$$

and so

$$AB - BA = I.$$

Hence $\mathcal{T}_A \circ \mathcal{T}_B = \mathcal{T}_B \circ \mathcal{T}_A$ but $AB \neq BA$, ie. the converse of (d) is false when $\mathbb{F} = \mathbb{Z}/2\mathbb{Z}$.