MATH 110: LINEAR ALGEBRA
SPRING 2007/08
PROBLEM SET 7 SOLUTIONS

1. Let $A, B \in \mathbb{F}^{n\times n}$. Define the function $\mathcal{T} : \mathbb{F}^{n\times n} \to \mathbb{F}^{n\times n}$ by

$$\mathcal{T}(X) = AXB$$

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

(a) Show that $\mathcal{T} \in \operatorname{End}(\mathbb{F}^{n\times n})$.

Solution. Let $\alpha_1, \alpha_2 \in \mathbb{F}$ and $X_1, X_2 \in \mathbb{F}^{n\times n}$. Then by the distributive property of matrix multiplication,

$$\begin{aligned}\mathcal{T}(\alpha_1 X_1 + \alpha_2 X_2) &= A(\alpha_1 X_1 + \alpha_2 X_2)B \\ &= \alpha_1 A X_1 B + \alpha_2 A X_2 B \\ &= \alpha_1 \mathcal{T}(X_1) + \alpha_2 \mathcal{T}(X_2).\end{aligned}$$

Hence $\mathcal{T}$ is linear.

(b) Show that $\mathcal{T}$ is invertible if and only if A and B are nonsingular matrices.

Solution. Suppose A and B are nonsingular. Define the map $\mathcal{S} : \mathbb{F}^{n\times n} \to \mathbb{F}^{n\times n}$ by

$$\mathcal{S}(X) = A^{-1}XB^{-1}$$

for all $X \in \mathbb{F}^{n\times n}$. By the associativity of matrix multiplication, we have

$$\begin{aligned}\mathcal{S}(\mathcal{T}(X)) &= A^{-1}(AXB)B^{-1} = (A^{-1}A)X(BB^{-1}) = IXI = X, \\ \mathcal{T}(\mathcal{S}(X)) &= A(A^{-1}XB^{-1})B = (AA^{-1})X(B^{-1}B) = IXI = X\end{aligned}$$

for all $X \in \mathbb{F}^{n\times n}$. So $\mathcal{S} \circ \mathcal{T} = \mathcal{I} = \mathcal{T} \circ \mathcal{S}$ and so $\mathcal{T}^{-1} = \mathcal{S}$. Suppose $\mathcal{T}$ is invertible, then $\ker(\mathcal{T}) = \{O\}$. We will prove by contradiction. Suppose A or B is singular. Without loss of generality, we may assume that A is singular (the argument for singular B is similarly). Then there exists a non-zero $\mathbf{x} \in \operatorname{nullsp}(A)$, ie. $A\mathbf{x} = \mathbf{0}$ but $\mathbf{x} \neq \mathbf{0}$. Define the matrix $X \in \mathbb{F}^{n\times n}$ all of whose columns are $\mathbf{x}$, ie.

$$X = [\mathbf{x}, \dots, \mathbf{x}].$$

Then $X \neq O$ but

$$\mathcal{T}(X) = AXB = A[\mathbf{x}, \dots, \mathbf{x}]B = [A\mathbf{x}, \dots, A\mathbf{x}]B = [\mathbf{0}, \dots, \mathbf{0}]B = OB = O.$$

So $\ker(\mathcal{T}) \neq \{O\}$ and so $\mathcal{T}$ is not invertible.

2. Let V be a finite dimensional vector space and $\mathcal{T} \in \operatorname{End}(V)$. Let $\dim(V) = n$. Let $\mathbf{v} \in V$ be such that

$$\mathcal{T}^{n-1}(\mathbf{v}) \neq \mathbf{0} \quad \text{and} \quad \mathcal{T}^{n}(\mathbf{v}) = \mathbf{0}.$$

(a) Show that the vectors

$$\mathbf{v}, \mathcal{T}(\mathbf{v}), \mathcal{T}^2(\mathbf{v}), \dots, \mathcal{T}^{n-1}(\mathbf{v})$$

form a basis for V.

Solution. Let $\alpha_0, \dots, \alpha_{n-1} \in \mathbb{F}$ be such that

$$\alpha_0 \mathbf{v} + \alpha_1 \mathcal{T}(\mathbf{v}) + \alpha_2 \mathcal{T}^2(\mathbf{v}) + \cdots + \alpha_{n-1}\mathcal{T}^{n-1}(\mathbf{v}) = \mathbf{0}. \tag{2.1}$$

Date: May 7, 2008 (Version 1.0).

Applying $\mathcal{T}^{n-1}$ to both sides of (2.1), we obtain

$$\mathcal{T}^{n-1}(\alpha_0\mathbf{v} + \alpha_1\mathcal{T}(\mathbf{v}) + \alpha_2\mathcal{T}^2(\mathbf{v}) + \cdots + \alpha_{n-1}\mathcal{T}^{n-1}(\mathbf{v})) = \mathcal{T}^{n-1}(\mathbf{0}),$$
$$\alpha_0\mathcal{T}^{n-1}(\mathbf{v}) + \alpha_1\mathcal{T}^{n}(\mathbf{v}) + \alpha_2\mathcal{T}^{n+1}(\mathbf{v}) + \cdots + \alpha_{n-1}\mathcal{T}^{2n-2}(\mathbf{v})) = \mathbf{0}.$$

Note that for all $m \geq n$, $\mathcal{T}^m(\mathbf{v}) = \mathcal{T}^{m-n}(\mathcal{T}^n(\mathbf{v})) = \mathcal{T}^{m-n}(\mathbf{0}) = \mathbf{0}$. So the last equation becomes

$$\alpha_0\mathcal{T}^{n-1}(\mathbf{v}) = \mathbf{0}.$$

Since $\mathcal{T}^{n-1}(\mathbf{v}) \neq \mathbf{0}$, we must have $\alpha_0 = 0$. Now, (2.1) becomes

$$\alpha_1\mathcal{T}(\mathbf{v}) + \alpha_2\mathcal{T}^2(\mathbf{v}) + \cdots + \alpha_{n-1}\mathcal{T}^{n-1}(\mathbf{v}) = \mathbf{0}$$

and we apply $\mathcal{T}^{n-2}$ to both sides and use the same argument above to conclude that $\alpha_1 = 0$. Repeating this argument n times gives

$$\alpha_0 = \alpha_1 = \cdots = \alpha_{n-1} = 0.$$

Hence $\{\mathbf{v}, \mathcal{T}(\mathbf{v}), \mathcal{T}^2(\mathbf{v}), \ldots, \mathcal{T}^{n-1}(\mathbf{v})\}$ is linearly independent and since $\dim(V) = n$, it forms a basis of V.

(b) Let $\mathscr{B}$ be the basis in (a). What is the matrix representation $[\mathcal{T}]_{\mathscr{B},\mathscr{B}}$?

Solution. We apply $\mathcal{T}$ to each vector of $\mathscr{B}$ in turn to get

$$\begin{aligned}
\mathcal{T}(\mathbf{v}) &= 0\mathbf{v} + 1\mathcal{T}(\mathbf{v}) + 0\mathcal{T}^2(\mathbf{v}) + \cdots + 0\mathcal{T}^{n-1}(\mathbf{v}),\\
\mathcal{T}(\mathcal{T}(\mathbf{v})) &= 0\mathbf{v} + 0\mathcal{T}(\mathbf{v}) + 1\mathcal{T}^2(\mathbf{v}) + \cdots + 0\mathcal{T}^{n-1}(\mathbf{v}),\\
&\;\vdots\\
\mathcal{T}(\mathcal{T}^{n-2}(\mathbf{v})) &= 0\mathbf{v} + 0\mathcal{T}(\mathbf{v}) + 0\mathcal{T}^2(\mathbf{v}) + \cdots + 1\mathcal{T}^{n-1}(\mathbf{v}),\\
\mathcal{T}(\mathcal{T}^{n-1}(\mathbf{v})) &= 0\mathbf{v} + 0\mathcal{T}(\mathbf{v}) + 0\mathcal{T}^2(\mathbf{v}) + \cdots + 0\mathcal{T}^{n-1}(\mathbf{v}).
\end{aligned}$$

Hence writing the coefficients as *columns* yield the required matrix representation

$$[\mathcal{T}]_{\mathscr{B},\mathscr{B}} = \begin{bmatrix}
0 & 1 & 0 & 0 & \cdots & \cdots & 0\\
0 & 0 & 1 & 0 & \cdots & \cdots & 0\\
\vdots & & \ddots & \ddots & \ddots & & \vdots\\
\vdots & & & \ddots & \ddots & \ddots & \vdots\\
0 & 0 & \cdots & \cdots & 0 & 1 & 0\\
0 & 0 & \cdots & \cdots & 0 & 0 & 1\\
0 & 0 & \cdots & \cdots & 0 & 0 & 0
\end{bmatrix},$$

a matrix with 1's on the super-diagonal the 0's everywhere else.

3. Let V and W be finite-dimensional vector spaces over $\mathbb{F}$.

(a) Let $\mathcal{T} \in \mathrm{Hom}(V, W)$. Prove the following.

(i) If $\mathcal{T}$ is injective, then $\dim(V) \leq \dim(W)$.

(ii) If $\mathcal{T}$ is surjective, then $\dim(V) \geq \dim(W)$.

(iii) If $\mathcal{T}$ is bijective, then $\dim(V) = \dim(W)$.

Solution. Since $\mathrm{im}(\mathcal{T})$ is a subspace of W, we must have

$$\dim(W) \geq \dim(\mathrm{im}(\mathcal{T})) = \mathrm{rank}(\mathcal{T}) = \dim(V) - \mathrm{nullity}(\mathcal{T}). \tag{3.2}$$

If $\mathcal{T}$ is injective, then $\mathrm{nullity}(\mathcal{T}) = 0$ and so

$$\dim(W) \geq \dim(V).$$

If $\mathcal{T}$ is surjective, then $W = \mathrm{im}(\mathcal{T})$ and equality holds in (3.2), ie.

$$\dim(W) = \dim(V) - \mathrm{nullity}(\mathcal{T}).$$

Hence

$$\dim(V) = \dim(W) + \text{nullity}(\mathcal{T}) \geq \dim(W).$$

If $\mathcal{T}$ is biective, then being both injective and surjective, we have $\dim(V) \leq \dim(W)$ and $\dim(V) \geq \dim(W)$ and so

$$\dim(V) = \dim(W).$$

(b) Show that if $\dim(V) = \dim(W)$, then there exists a bijective $\mathcal{T} \in \text{Hom}(V, W)$. [Together with (iii), this shows that 'V and W are *isomorphic* if and only if $\dim(V) = \dim(W)$'.]
SOLUTION. Let $n = \dim(V) = \dim(W)$. Let $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ and $\{\mathbf{w}_1, \dots, \mathbf{w}_n\}$ be bases of V and W respectively. We define $\mathcal{T} : V \to W$ by

$$\mathcal{T}(\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \cdots + \alpha_n\mathbf{v}_n) = \alpha_1\mathbf{w}_1 + \alpha_2\mathbf{w}_2 + \cdots + \alpha_n\mathbf{w}_n$$

for all $\alpha_1, \alpha_2, \dots, \alpha_n \in \mathbb{F}$ (note that every $\mathbf{v} \in V$ may be written in this form). $\mathcal{T}$ is linear since for $\lambda, \mu \in \mathbb{F}$ and $\alpha_1\mathbf{v}_1 + \cdots + \alpha_n\mathbf{v}_n, \beta_1\mathbf{v}_1 + \cdots + \beta_n\mathbf{v}_n \in V$,

$$\begin{aligned}
\mathcal{T}(\lambda(\alpha_1\mathbf{v}_1 + \cdots + \alpha_n\mathbf{v}_n) + \mu(\beta_1\mathbf{v}_1 + \cdots + \beta_n\mathbf{v}_n)) &= \mathcal{T}((\lambda\alpha_1 + \mu\beta_1)\mathbf{v}_1 + \cdots + (\lambda\alpha_n + \mu\beta_n)\mathbf{v}_n) \\
&= (\lambda\alpha_1 + \mu\beta_1)\mathbf{w}_1 + \cdots + (\lambda\alpha_n + \mu\beta_n)\mathbf{w}_n \\
&= \lambda(\alpha_1\mathbf{w}_1 + \cdots + \alpha_n\mathbf{w}_n) + \mu(\beta_1\mathbf{w}_1 + \cdots + \beta_n\mathbf{w}_n) \\
&= \lambda\mathcal{T}(\alpha_1\mathbf{v}_1 + \cdots + \alpha_n\mathbf{v}_n) + \mu\mathcal{T}(\beta_1\mathbf{v}_1 + \cdots + \beta_n\mathbf{v}_n).
\end{aligned}$$

Let $\alpha_1\mathbf{v}_1 + \cdots + \alpha_n\mathbf{v}_n \in \ker(\mathcal{T})$. Then

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

Since $\mathbf{w}_1, \dots, \mathbf{w}_n$ are linearly independent, we get $\alpha_1 = \cdots = \alpha_n = 0$ and so

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

Hence $\ker(\mathcal{T}) = \{\mathbf{0}_V\}$ and so $\mathcal{T}$ is injective.

(c) Let $\dim(V) = \dim(W)$. Let $\mathcal{T} \in \text{Hom}(V, W)$ and $\mathcal{S} \in \text{Hom}(W, V)$. Show that

$$\mathcal{S} \circ \mathcal{T} = \mathcal{I}_V \tag{3.3}$$

if and only if

$$\mathcal{T} \circ \mathcal{S} = \mathcal{I}_W.$$

SOLUTION. Let $\mathcal{S} \circ \mathcal{T} = \mathcal{I}_V$. Then $\mathcal{T}$ is injective since if $\mathcal{T}(\mathbf{v}) = \mathbf{0}_W$, then $\mathbf{v} = \mathcal{I}_V(\mathbf{v}) = \mathcal{S}(\mathcal{T}(\mathbf{v})) = \mathcal{S}(\mathbf{0}_W) = \mathbf{0}_V$ (ie. $\ker(\mathcal{T}) = \{\mathbf{0}_V\}$). Since V is finite-dimensional, we may apply Theorem **4.12** to conclude that $\mathcal{T}$ is invertible. Let $\mathcal{T}^{-1}$ be the inverse of $\mathcal{T}$. Then composing $\mathcal{T}^{-1}$ on the right of (3.3), we get

$$\begin{aligned}
(\mathcal{S} \circ \mathcal{T}) \circ \mathcal{T}^{-1} &= \mathcal{I}_V \circ \mathcal{T}^{-1}, \\
\mathcal{S} \circ (\mathcal{T} \circ \mathcal{T}^{-1}) &= \mathcal{T}^{-1}, \\
\mathcal{S} \circ \mathcal{I}_V &= \mathcal{T}^{-1}, \\
\mathcal{S} &= \mathcal{T}^{-1}.
\end{aligned}$$

Hence

$$\mathcal{T} \circ \mathcal{S} = \mathcal{T} \circ \mathcal{T}^{-1} = \mathcal{I}_W$$

as required. For the converse, just swap the roles of $\mathcal{T}$ and $\mathcal{S}$.

(d) Show that (b) and (c) are false if $\dim(V) \neq \dim(W)$.
SOLUTION. If $\dim(V) \neq \dim(W)$ and there exists a bijective $\mathcal{T}$, then this would contradict (iii) in (a).

4. Let $\mathbb{P}$ be the vector space of all polynomials over $\mathbb{R}$. Define the functions $\mathcal{D} : \mathbb{P} \to \mathbb{P}$ and $\mathcal{M} : \mathbb{P} \to \mathbb{P}$ by

$$\mathcal{D}(p)(x) = p'(x) \quad \text{and} \quad \mathcal{M}(p)(x) = xp(x)$$

for all $p \in \mathbb{P}$, ie. the 'differentiation with respect to x' and 'multiplication by x' functions. Explicitly, if $p(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_d x^d$ with $a_0, a_1, a_2, \ldots, a_d \in \mathbb{R}$, then

$$\mathcal{D}(p)(x) = a_1 + 2a_2 x + \cdots + d a_d x^{d-1},$$
$$\mathcal{M}(p)(x) = a_0 x + a_1 x^2 + \cdots + a_d x^{d+1}.$$

(a) Show that $\mathcal{D} \in \text{End}(\mathbb{P})$ and $\mathcal{M} \in \text{End}(\mathbb{P})$.

Solution. Let $\lambda, \mu \in \mathbb{F}$ and $p(x) = a_0 + a_1 x + \cdots + a_d x^d, q(x) = b_0 + b_1 x + \cdots + b_d x^d \in \mathbb{P}$ where $d = \max\{\deg(p(x)), \deg(q(x))\}$. Then

$$\begin{aligned}\mathcal{D}(\lambda p + \mu q)(x) &= (\lambda a_1 + \mu b_1) + 2(\lambda a_2 + \mu b_2)x + \cdots + d(\lambda a_d + \mu b_d)x^{d-1} \\ &= \lambda(a_1 + 2a_2 x + \cdots + d a_d x^{d-1}) + \mu(b_1 + 2b_2 x + \cdots + d b_d x^{d-1}) \\ &= \lambda \mathcal{D}(p)(x) + \mu \mathcal{D}(q)(x)\end{aligned}$$

and

$$\begin{aligned}\mathcal{M}(\lambda p + \mu q)(x) &= (\lambda a_0 + \mu b_0)x + (\lambda a_2 + \mu b_2)x^2 + \cdots + (\lambda a_d + \mu b_d)x^{d+1} \\ &= \lambda(a_0 x + a_1 x^2 + \cdots + a_d x^{d+1}) + \mu(b_0 x + b_1 x^2 + \cdots + b_d x^{d+1}) \\ &= \lambda \mathcal{M}(p)(x) + \mu \mathcal{M}(q)(x).\end{aligned}$$

So $\mathcal{D}$ and $\mathcal{M}$ are linear.

(b) Show that

$$\text{im}(\mathcal{D}) = \mathbb{P}, \quad \ker(\mathcal{D}) \neq \{0(x)\}, \quad \text{im}(\mathcal{M}) \neq \mathbb{P}, \quad \ker(\mathcal{M}) = \{0(x)\},$$

where $0(x)$ denotes the zero polynomial.

Solution. Let $p(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_d x^d \in \mathbb{P}$. If we let

$$q(x) = a_0 x + \frac{a_1}{2}x^2 + \frac{a_2}{3}x^3 + \cdots + \frac{a_d}{d+1}x^{d+1},$$

we see that

$$\mathcal{D}(q)(x) = p(x).$$

So $\mathcal{D}$ is surjective and $\text{im}(\mathcal{D}) = \mathbb{P}$. Let $c(x) = 1$. Then $\mathcal{D}(c)(x) = 0(x)$ and so $c(x) \in \ker(\mathcal{D})$. Since $c(x) \neq 0(x)$, $\ker(\mathcal{D}) \neq \{0(x)\}$. Let $p(x) \in \ker(\mathcal{M})$, then $\mathcal{M}(p)(x) = 0(x)$, ie.

$$a_0 x + a_1 x^2 + \cdots + a_d x^{d+1} = 0x + 0x^2 + \cdots + 0x^{d+1}.$$

So $a_0 = a_1 = \cdots = a_d = 0$ and so $p(x) = 0(x)$. Note that $c(x) \notin \text{im}(\mathcal{M})$ since if

$$\mathcal{M}(p)(x) = c(x),$$

then $\deg(\mathcal{M}(p)(x)) = \deg(c(x)) = 1$, which is only possible of $p(x) = 0(x)$ but clearly $\mathcal{M}(0)(x) = 0(x) \neq c(x)$.

(c) Are $\mathcal{D}$ and $\mathcal{M}$ surjective, injective, or bijective? Why would these observations *not* contradict Theorem **4.12** from the lectures?

Solution. Theorem **4.12** applies only to finite-dimensional vector spaces whereas $\mathbb{P}$ is infinite-dimensional.

(d) Show that

$$\mathcal{D} \circ \mathcal{M} - \mathcal{M} \circ \mathcal{D} = \mathcal{I}_{\mathbb{P}}$$

and more generally

$$\mathcal{D}^n \circ \mathcal{M} - \mathcal{M} \circ \mathcal{D}^n = n\mathcal{D}^{n-1}$$

for all $n \in \mathbb{N}$.

SOLUTION. Let $p(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_d x^d \in \mathbb{P}$. Note that

$$\mathcal{M}(\mathcal{D}(p))(x) = a_1 x + 2a_2 x^2 + \cdots + d a_d x^d,$$
$$\mathcal{D}(\mathcal{M}(p))(x) = a_0 + 2a_1 x + \cdots + (d+1) a_d x^d,$$

and so

$$\begin{aligned}(\mathcal{D} \circ \mathcal{M} - \mathcal{M} \circ \mathcal{D})(p)(x) &= \mathcal{D}(\mathcal{M}(p))(x) - \mathcal{M}(\mathcal{D}(p))(x) \\ &= a_0 + a_1 x + a_2 x^2 + \cdots + a_d x^d \\ &= p(x).\end{aligned}$$

Hence $\mathcal{D} \circ \mathcal{M} - \mathcal{M} \circ \mathcal{D} = \mathcal{I}_{\mathbb{P}}$. For the general case, we will use induction. We have already shown that it is true for $n = 1$. Suppose it is true for all $n = 1, \ldots, k$, we will use this to deduce that it is also true for $n = k + 1$,

$$\begin{aligned}\mathcal{D}^{k+1} \circ \mathcal{M} - \mathcal{M} \circ \mathcal{D}^{k+1} &= \mathcal{D} \circ (\mathcal{D}^k \circ \mathcal{M}) - \mathcal{M} \circ \mathcal{D}^{k+1} \\ &= \mathcal{D} \circ (k\mathcal{D}^{k-1} + \mathcal{M} \circ \mathcal{D}^k) - \mathcal{M} \circ \mathcal{D}^{k+1} \\ &= k\mathcal{D}^k + \mathcal{D} \circ \mathcal{M} \circ \mathcal{D}^k - \mathcal{M} \circ \mathcal{D}^{k+1} \\ &= k\mathcal{D}^k + (\mathcal{D} \circ \mathcal{M} - \mathcal{M} \circ \mathcal{D}) \circ \mathcal{D}^k \\ &= k\mathcal{D}^k + \mathcal{I}_{\mathbb{P}} \circ \mathcal{D}^k \\ &= k\mathcal{D}^k + \mathcal{D}^k \\ &= (k+1)\mathcal{D}^k\end{aligned}$$

as required.