Non-book problems for Assignment 10

Some of the facts you're asked to prove in these problems were proven in class—perhaps recently, perhaps several weeks ago. Do any such proofs over again (for practice and reinforcement); don't just say, "This is true because we proved it in class."

NB 10.1. Let V and W be finite-dimensional vector spaces, and let $\mathsf{T} : V \to W$ be a linear map.

(a) Show that

$$\text{rank}(\mathsf{T}) \leq \min\{\dim(V), \dim(W)\}. \tag{1}$$

(b) Show that T is one-to-one if and only if $\text{rank}(\mathsf{T}) = \dim(V)$.

(c) Show that T is onto if and only if $\text{rank}(\mathsf{T}) = \dim(W)$.

Before starting on part (a), prove the following trivial lemma for yourself: for any real numbers x, y, z, the inequality "$x \leq \min\{y, z\}$" is equivalent to "$x \leq y$ and $x \leq z$." After proving this lemma for yourself, don't bother citing it when you use it. Thus, for example, in the exercise above, to show $\text{rank}(\mathsf{T}) \leq \min\{\dim(V), \dim(W)\}$, you should show that $\text{rank}(\mathsf{T}) \leq \dim(V)$ and that $\text{rank}(\mathsf{T}) \leq \dim(W)$, and then say something like "Therefore $\text{rank}(\mathsf{T}) \leq \min\{\dim(V), \dim(W)\}$;" you should not break the argument down into cases according to which of $\dim(V)$ and $\dim(W)$ is the larger, and you don't need to cite the trivial lemma above.

****

Below, you should find problems NB 10.2 and 10.3(a) *very quick* and *very easy* to do. Everything that you're asked to check is an immediate consequence of *definitions*. If you're on top of the relevant definitions, you should be able to do most or all of the problem-parts in your head. The main purpose of these problems is to make you conscious of various facts that otherwise might not occur to you—however "obvious" or "trivial" they may seem once they're pointed out to you—and that can be very useful in proving less trivial statements.

NB 10.2. Let V and W be vector spaces and let $\mathsf{T} : V \to W$ be a linear transformation.

(a) Define $\overline{\mathsf{T}} : V \to \mathsf{R}(\mathsf{T})$ by $\overline{\mathsf{T}}(v) = \mathsf{T}(v)$ (for each $v \in V$). (Thus, the only difference between T and $\overline{\mathsf{T}}$ is in their *codomains*.[1]) Check that the following are true:

(i) $\overline{\mathsf{T}} : V \to \mathsf{R}(\mathsf{T})$ is linear.

[1]Recall that a function f has three "parts": a specified domain, say A; a specified codomain, say B; and an assignment of an element $f(x) \in B$ to each element $x \in A$. Changing any of these "parts" creates a different function.

(ii) $\mathsf{R}(\overline{\mathsf{T}}) = \mathsf{R}(\mathsf{T}) = \mathsf{T}(V)$.

(Recall that for any function $f : A \to B$, and any subset $S \subseteq A$, the notation $f(S)$ means the set $\{f(x) : x \in A\}$.)

(iii) $\overline{\mathsf{T}}$ is onto. (This highlights the fact that T and $\overline{\mathsf{T}}$ are not the same function, even though $\overline{\mathsf{T}}(v) = \mathsf{T}(v)$ for every $v \in V$.)

(iv) $\mathsf{N}(\overline{\mathsf{T}}) = \mathsf{N}(\mathsf{T})$.

(b) Let $V_0 \subseteq V$ be a subspace. Recall that the *restriction of* T *to* V_0, denoted $\mathsf{T}|_{V_0}$, is the function from V_0 to W defined by $\mathsf{T}|_{V_0}(v) = \mathsf{T}(v)$ (for each $v \in V_0$). (Thus, the only difference between T and $\mathsf{T}|_{V_0}$ is in their *domains*.) Check that the following are true:

(i) $\mathsf{T}|_{V_0} : V_0 \to W$ is linear.

(ii) $\mathsf{R}(\mathsf{T}|_{V_0}) = \mathsf{T}(V_0)$.

(iii) $\mathsf{N}(\mathsf{T}|_{V_0}) = \mathsf{N}(\mathsf{T}) \cap V_0$.

NB 10.3. Let V, W, and Z be vector spaces, let $\mathsf{T} : V \to W$ and $\mathsf{U} : W \to Z$ be linear transformations, and define $\overline{\mathsf{T}}$ as in problem NB 10.2.

Below, if you find that the notation "$\mathsf{U}|_{\mathsf{R}(\mathsf{T})}$" makes any of the equations below hard to follow, then temporarily introduce simpler notation for this transformation. E.g., if you temporarily write $\overline{\mathsf{U}}$ for $\mathsf{U}|_{\mathsf{R}(\mathsf{T})}$, then the equations in (a)(i) read $\mathsf{U} \circ \mathsf{T} = \overline{\mathsf{U}} \circ \mathsf{T} = \overline{\mathsf{U}} \circ \overline{\mathsf{T}}$, while the equations in (a)(ii) read $\mathsf{R}(\mathsf{U} \circ \mathsf{T}) = \mathsf{U}(\mathsf{R}(\mathsf{T})) = \mathsf{R}(\overline{\mathsf{U}})$.

(a) Check that the following are true:

(i) $\mathsf{U} \circ \mathsf{T} = \mathsf{U}|_{\mathsf{R}(\mathsf{T})} \circ \mathsf{T} = \mathsf{U}|_{\mathsf{R}(\mathsf{T})} \circ \overline{\mathsf{T}}$.

(ii) $\mathsf{R}(\mathsf{U} \circ \mathsf{T}) = \mathsf{U}(\mathsf{R}(\mathsf{T})) = \mathsf{R}(\mathsf{U}|_{\mathsf{R}(\mathsf{T})})$

(Note: for general functions f and g, the composition "$g \circ f$" is defined as long as the *range* of f is contained in the domain of g; it is not necessary to have even if codomain(f) = domain(g), In particular, the composition $\mathsf{U}|_{\mathsf{R}(\mathsf{T})} \circ \mathsf{T}$ is defined even though $\mathsf{R}(\mathsf{T})$ [the domain of $\mathsf{U}|_{\mathsf{R}(\mathsf{T})}$] may be not be all of W [the codomain of T].)

(b) Assume that V and W are finite-dimensional. Show that

$$\text{rank}(\mathsf{U} \circ \mathsf{T}) \leq \min\{\text{rank}(\mathsf{U}), \text{rank}(\mathsf{T})\}.$$

Note: There is more than one way to do this, but the purpose of NB 10.3(a), together with certain parts of problems NB 10.1 and NB 10.2, was to help lead you to a nice, easy way of doing NB 10.3 (b).

NB 10.4. Let V and W be finite-dimensional vector spaces and assume that $\dim(V) < \dim(W)$.

(a) Show that there exists a one-to-one linear transformation from V to W. (*Hint*: Theorem 2.6 in FIS Chapter 2.)

Note that the word "unique" does not appear in this statement, or in part (b) or (c)!

(b) Show that if a linear transformation $\mathsf{T} : V \to W$ is one-to-one, then T has a linear *left inverse*, i.e. a linear transformation $\mathsf{U} : W \to V$ such that $\mathsf{U} \circ \mathsf{T} = I_V$.

(c) Show that if a linear transformation $\mathsf{U} : W \to V$ is onto, then U has a linear *right inverse*, i.e. a linear transformation $\mathsf{T} : V \to W$ such that $\mathsf{U} \circ \mathsf{T} = I_V$. Show also that any such T is one-to-one. (*Note*: The U and T in this problem-part are *new*; they're not carried over from part (b). In fact, **the wording of part (b) tells you this. How?**)

(d) Give two proofs, as indicated below, that there do not exist any linear transformations $\mathsf{T} : V \to W$ and $\mathsf{U} : W \to V$ for which $\mathsf{T} \circ \mathsf{U} = I_W$:

- *Hint for Proof #1*: Use Problem NB 10.3(b).
- *Hint for Proof #2*: Show that for any sets X, Y and functions $f : X \to Y$ and $g : Y \to X$, if $g \circ f = I_X$ (the identity map of X), then f is one-to-one and g is onto. Then use what we proved several weeks ago about non-existence of linear maps from V to W with a certain property, and linear maps from W to V with a certain property, under the given assumption that $\dim(V) < \dim(W)$.

NB 10.5. Let V and W be finite-dimensional vector spaces **of equal dimension** n. Suppose that $\mathsf{T} : V \to W$ and $\mathsf{U} : W \to V$ are linear transformations for which $\mathsf{U} \circ \mathsf{T} = I_V$. Using the steps below, prove that $\mathsf{T} \circ \mathsf{U} = I_W$.

(a) Show that $\operatorname{rank}(\mathsf{U}) = \operatorname{rank}(\mathsf{T}) = n$. (*Hint*: Problem NB 10.1.) In particular, T is onto.

(b) Show that $\mathsf{T} \circ \mathsf{U} \circ \mathsf{T} = \mathsf{T}$. (Recall that composition of composable functions is associative: if $f : X \to Y$, $g : Y \to Z$, and $h : Z \to W$ are functions, then $(h \circ g) \circ f = h \circ (g \circ f)$. Hence both sides of the latter equation can unambiguously be denoted $h \circ g \circ f$.)

(c) Now use parts (a) and (b) to show that $\mathsf{T} \circ \mathsf{U} = I_W$.

Note: Problem NB 10.4(d) shows that **in order for "$\mathsf{U} \circ \mathsf{T} = I_V$" to imply "$\mathsf{T} \circ \mathsf{U} = I_W$," (non-vacuously), the assumption that** $\dim(V) = \dim(W)$ **is *crucial*!** (The "non-vacuously" means here that there exist *some* linear $\mathsf{T} : V \to W$ and $\mathsf{U} : W \to V$ for which $\mathsf{U} \circ \mathsf{T} = I_V$.)

(d) Apply the preceding to show that if A and B are $n \times n$ matrices and $AB = I_{n \times n}$, then

$BA = I_{n\times n}$. (Take U to be the map $\mathsf{L}_A : \mathbf{R}^n \to \mathbf{R}^n$, take U to be the map $\mathsf{L}_B : \mathbf{R}^n \to \mathbf{R}^n$, and proceed from there.)

NB 10.6. (a) Let $V = \{(a_1, a_2, a_3, 0, 0) : a_1, a_2, a_3 \in \mathbf{R}\}$. Show that V is a subspace of $\mathbf{R}^5$ and is isomorphic to $\mathbf{R}^3$.

Note that V is **NOT** $\mathbf{R}^3$. Every element of V is an ordered *quintuple*; every element of $\mathbf{R}^3$ is an ordered *triple*. This difference helps to illustrate what the idea of "isomorphic" is getting at. Two isomorphic vector spaces may be very different *sets*; they are just "*essentially* the same" as far as their vector-space properties are concerned.

(b) Let n and k be integers, with $n > k > 0$. Generalize part (a) to

$$V = \{(a_1, \ldots, a_k, 0, \ldots, 0) \in \mathbf{R}^n\} = \{(a_1, \ldots, a_n) \in \mathbf{R}^n : a_i = 0 \text{ for all } i > k\} \subseteq \mathbf{R}^n.$$

(c) With n and k as in (b), generalize further: Let $i_1 < i_2 < \cdots < i_{n-k}$ be integers (where i_2 is present only if $n - k \geq 2$) and let

$$V = \{(a_1, \ldots, a_n) \in \mathbf{R}^n : a_{i_1} = a_{i_2} = \cdots = a_{i_{n-k}} = 0\} \subseteq \mathbf{R}^n.$$

Show that V is a subspace of $\mathbf{R}^n$ isomorphic to $\mathbf{R}^k$.

NB 10.7. Let $J_n = \{1, 2, \ldots, n\}$ and let $V = \mathcal{F}(J_n, \mathbf{R})$. Define $\mathsf{T} : V \to \mathbf{R}^n$ by $\mathsf{T}(f) = (f(1), f(2), \ldots, f(n))$. Show that T is an isomorphism, and hence that $\mathcal{F}(J_n, \mathbf{R}) \cong \mathbf{R}^n$.

NB 10.8. Let A and B be diagonal $n \times n$ matrices. Show that AB is also a diagonal matrix, and that if the diagonal entries of A are $\lambda_1, \ldots, \lambda_n$ and the diagonal entries of B are $\mu_1, \ldots, \mu_n$ (i.e. $\lambda_i = A_{ii}$ and $\mu_i = B_{ii}$, $1 \leq i \leq n$), then the diagonal entries of AB are simply the products $\lambda_1\mu_1, \ldots, \lambda_n\mu_n$.

NB 10.9. (a) Let $A = \begin{pmatrix} 0 & 1 & 2 & 3 \\ 0 & 0 & 4 & 5 \\ 0 & 0 & 0 & 6 \\ 0 & 0 & 0 & 0 \end{pmatrix}$. Compute A^2, A^3, and A^4.

(b) In part (a), you should have found that A^4 is a very simple matrix—so simple that you can immediately tell what all higher powers of A would be. You should also have noticed a pattern in location of the nonzero above-diagonal entries in the sequence A, A^2, A^3. Does the value of A^4, or the pattern you noticed in the sequence A, A^2, A^3, depend at all on the values of the above-diagonal entries of A?

(c) For $n \times n$ matrices with $n \geq 2$, conjecture how your observations in part (b) would generalize.

(d) Try to prove the conjecture you made in part (c).

NB 10.10. In FIS 1.3/16 you showed that $C^2(\mathbf{R})$, the space of twice-differentiable functions from $\mathbf{R}$ to $\mathbf{R}$ with continuous second derivative, is a subspace of $\mathcal{F}(\mathbf{R},\mathbf{R})$. In particular, $C^2(\mathbf{R})$ is a vector space. Its zero element is the constant function 0_{fcn} (the function f for which $f(t) = 0$ for all $t \in \mathbf{R}$). A standard theorem taught in a first course on differential equations is:

> **Theorem 1**. *Let $a, b, c \in \mathbf{R}$ with $a \neq 0$. Then for any $x_0 \in \mathbf{R}$ and $k_1, k_2 \in \mathbf{R}$, there is a unique function $y \in C^2(\mathbf{R})$ satisfying*
>
> $$ay'' + by' + cy = 0_{\text{fcn}}, \quad y(x_0) = k_1, \quad \text{and} \quad y'(x_0) = k_2$$
>
> (where "prime" denotes differentiation).

Assume the truth of Theorem 1 below.

(a) Let $a, b, c \in \mathbf{R}$ with $a \neq 0$ and define $\mathsf{L} : C^2(\mathbf{R}) \to \mathcal{F}(\mathbf{R},\mathbf{R})$ by

$$\mathsf{L}(f) = af'' + bf' + cf.$$

Show that L is linear.

(What we call the *general solution* of the *homogeneous, linear differential equation* $\mathsf{L}(y) = 0$ is simply the null space $\mathsf{N}(\mathsf{L})$. What we call a *fundamental set of solutions* for the equation $\mathsf{L}(y) = 0$, in an intro DE course, is simply a basis of $\mathsf{N}(\mathsf{L})$.)

(b) For each $x_0 \in \mathbf{R}$, define a map $\mathsf{E}_{x_0} : \mathsf{N}(\mathsf{L}) \to \mathbf{R}^2$ by

$$\mathsf{E}_{x_0}(f) = \big(f(x_0), f'(x_0)\big)$$

Using Theorem 1 above, show that for each $x_0 \in \mathbf{R}$, the map E_{x_0} is an isomorphism.

It follows that for any $x_0, x_1 \in \mathbf{R}$, the map $\mathsf{E}_{x_1} \circ (\mathsf{E}_{x_0})^{-1}$ is an isomorphism from $\mathbf{R}^2$ to $\mathbf{R}^2$. Each 2nd-order, constant coefficient, linear differential equation $\mathsf{L}(y) = 0$ determines such a "transport isomorphism" from the *space of initial conditions at* x_0 ($\mathbf{R}^2$, identified via E_{x_0} with $\big\{\big(y(x_0), y'(x_0)\big) : y \in \mathsf{N}(\mathsf{L})\big\}$)—to the space of initial conditions at any other point x_1 (similarly identified with $\mathbf{R}^2$ via E_{x_1}).

NB 10.11. Let $A, C \in M_{n\times n}(\mathbf{R})$ and assume that C is invertible. Show that for any integer $k \geq 1$,

$$(C^{-1}AC)^k = C^{-1}A^k\,C$$

and similarly

$$(CAC^{-1})^k = C\,A^k\,C^{-1}.$$

NB 10.12. (a) Let V_1, V_2, and W be vector spaces. Assume $V_1 \cong V_2$ and let $\mathsf{T} : V_1 \to V_2$ be an isomorphism. (**Note that you do not see any assumption of finite-dimensionality here.**)

(i) Define $F : \mathcal{L}(W, V_1) \to \mathcal{L}(W, V_2)$ by $F(\mathsf{U}) = \mathsf{T} \circ \mathsf{U}$ (see diagram below).

$$W \xrightarrow[\mathsf{T}]{} V_1 \xrightarrow[\mathsf{U}]{} V_2$$

(T is fixed, U is variable, and F is the map that carries U to $\mathsf{T} \circ \mathsf{U}$.)

Show that F is an isomorphism.

(ii) Define $G : \mathcal{L}(V_2, W) \to \mathcal{L}(V_1, W)$ by $G(\mathsf{U}) = \mathsf{U} \circ \mathsf{T}$. (see diagram below).

$$W \xrightarrow[\mathsf{T}]{} V_1 \xrightarrow[\mathsf{U}]{} V_2$$

(T is fixed, U is variable, and G is the map that carries U to $\mathsf{U} \circ \mathsf{T}$.)

Show that G is an isomorphism.

(iii) Define $H : \mathcal{L}(V_1, V_1) \to \mathcal{L}(V_2, V_2)$ by $H(\mathsf{U}) = \mathsf{T}^{-1} \circ \mathsf{U} \circ \mathsf{T}$ (see diagram below).

$$V_2 \xrightarrow[\mathsf{T}]{} V_1 \xrightarrow[\mathsf{U}]{} V_1 \xrightarrow[\mathsf{T}^{-1}]{} V_2$$

(T is fixed, U is variable, and H is the map that carries U to $\mathsf{T}^{-1} \circ \mathsf{U} \circ \mathsf{T}$.)

Show that H is an isomorphism.

(b) Let $n, m > 0$, let A be an invertible $m \times m$ matrix, and let B be an invertible $n \times n$ matrix. Define $\tilde{\mathsf{L}}_A : M_{m\times n}(\mathbf{R}) \to M_{m\times n}(\mathbf{R})$ and $\tilde{\mathsf{R}}_B : M_{m\times n}(\mathbf{R}) \to M_{m\times n}(\mathbf{R})$ by $\tilde{\mathsf{L}}_A(C) = AC$ and $\tilde{\mathsf{R}}_B(C) = CB$.

(i) Show, just by using matrix operations and their properties, that $\tilde{\mathsf{L}}_A$ and $\tilde{\mathsf{R}}_B$ are isomorphisms.

(ii) Show a second way that $\tilde{\mathsf{L}}_A$ is an isomorphism. Specifically, letting $\Phi : \mathcal{L}(\mathbf{R}^n, \mathbf{R}^m) \to M_{m\times n}(\mathbf{R})$ be the map that carries $\mathsf{T} \in \mathcal{L}(\mathbf{R}^n, \mathbf{R}^m)$ to $[\mathsf{T}]^{\mathrm{std},m}_{\mathrm{std},n}$ (the matrix of T with respect to the standard bases of $\mathbf{R}^n$ and $\mathbf{R}^m$), show that $\tilde{\mathsf{L}}_A = \Phi^{-1} \circ \mathsf{L}_A \circ \Phi$, and apply an earlier part of this problem.

(c) In part (a) of this problem, the vector spaces V_1, V_2, and W were not assumed to be finite-dimensional. However, when these spaces *do* have finite, positive dimension, every part of (a) has a matrix analog (for one of these parts, that's what (b)(ii) is saying). Which assigned FIS exercise is the matrix analog of part (a)(iii) of this problem?