Tutorial 3

1. Which of the following functions are linear transformations?

\( (i) \) \( T: \mathbb{R}^2 \to \mathbb{R}^2 \) defined by
\[
T(x, y) = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}
\]

\( (ii) \) \( S: \mathbb{R}^2 \to \mathbb{R}^2 \) defined by
\[
S(x, y) = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}
\]

\( (iii) \) \( g: \mathbb{R}^2 \to \mathbb{R}^3 \) defined by
\[
g(x, y) = \begin{pmatrix} 2x + y \\ y \\ x - y \end{pmatrix}
\]

\( (iv) \) \( f: \mathbb{R} \to \mathbb{R}^2 \) defined by
\[
f(x) = \begin{pmatrix} x \\ x + 1 \end{pmatrix}
\]

Solution.

\( (i) \) This function is linear. To prove this we must show that \( T(a + b) = T(a) + T(b) \) and \( T(\lambda a) = \lambda T(a) \) for all \( a, b \in \mathbb{R}^2 \) and all \( \lambda \in \mathbb{R} \). So, let \( a, b \in \mathbb{R}^2 \), \( \lambda \in \mathbb{R} \). Then \( a = \begin{pmatrix} x \\ y \end{pmatrix} \) and \( b = \begin{pmatrix} u \\ v \end{pmatrix} \) for some \( x, y, u, v \in \mathbb{R} \), and we have

\[
T(a + b) = T \left( \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} u \\ v \end{pmatrix} \right) = T \begin{pmatrix} x + u \\ y + v \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} x + u \\ y + v \end{pmatrix} = \begin{pmatrix} x + 2u + y \\ 2x + y + u + 2v \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = T(a) + T(b).
\]

Similarly

\[
T(\lambda a) = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} \lambda x \\ \lambda y \end{pmatrix} = \begin{pmatrix} \lambda x + 2\lambda y \\ 2\lambda x + \lambda y \end{pmatrix} = \lambda \begin{pmatrix} x + 2y \\ 2x + y \end{pmatrix} = \lambda \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \lambda T(a).
\]

\( (ii) \) This function is also linear, by exactly the same reasoning as in \( (i) \) above. Indeed, the same would work for any \( 2 \times 2 \) matrix.

\( (iii) \) This function is also linear, since

\[
g \left( \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} u \\ v \end{pmatrix} \right) = g \begin{pmatrix} x + u \\ y + v \end{pmatrix} = \begin{pmatrix} 2(x + u) + (y + v) \\ y + v \\ (x + u) - (y + v) \end{pmatrix} = \begin{pmatrix} 2(x + u) + (y + v) \\ y \\ x - y \end{pmatrix} + \begin{pmatrix} 2u + v \\ v \\ u - v \end{pmatrix} = g \begin{pmatrix} x \\ y \end{pmatrix} + g \begin{pmatrix} u \\ v \end{pmatrix}
\]
and similarly
\[ g\left( \lambda \begin{pmatrix} x \\ y \end{pmatrix} \right) = g\left( \frac{\lambda x}{\lambda y} \right) = \begin{pmatrix} 2\lambda x + \lambda y \\ \lambda x - \lambda y \end{pmatrix} = \lambda g\left( \begin{pmatrix} x \\ y \end{pmatrix} \right). \]

(iv) This function is not linear, since (for instance)
\[ f(0 + 0) = f(0) = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \neq \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} = f(0) + f(0). \]

2. Let \( \mathcal{A} \) be the set of all 2-component column vectors whose entries are differentiable functions from \( \mathbb{R} \) to \( \mathbb{R} \). Thus, for example, if \( h \) and \( k \) are the functions defined by \( h(t) = \cos t \) and \( k(t) = t^2 + 1 \) for all \( t \in \mathbb{R} \) then \( \begin{pmatrix} h \\ k \end{pmatrix} \) is an element of \( \mathcal{A} \).

(i) How should addition and scalar multiplication be defined so that \( \mathcal{A} \) becomes a vector space over \( \mathbb{R} \)?

(ii) If \( f \) and \( g \) are real-valued functions on \( \mathbb{R} \) then their pointwise product is the function \( f \cdot g \) defined by \( (f \cdot g)(t) = f(t)g(t) \) for all \( t \in \mathbb{R} \). Prove that
\[ \begin{pmatrix} f \\ g \end{pmatrix} \mapsto h \cdot f + g' \]
(where \( h \) is as above and \( g' \) is the derivative of \( g \)) defines a linear transformation from \( \mathcal{A} \) to the space of all real-valued functions on \( \mathbb{R} \).

Solution.

(i) Let \( a, b \in \mathcal{A} \) and \( \lambda \in \mathbb{R} \). Then
\[ a = \begin{pmatrix} \phi \\ \psi \end{pmatrix}, \quad b = \begin{pmatrix} \chi \\ \theta \end{pmatrix} \]
for some differentiable functions \( \phi, \psi, \chi \) and \( \theta \) from \( \mathbb{R} \) to \( \mathbb{R} \). We define \( a + b \) and \( \lambda a \) by
\[ a + b = \begin{pmatrix} \phi + \chi \\ \psi + \theta \end{pmatrix}, \quad \lambda a = \begin{pmatrix} \lambda \phi \\ \lambda \psi \end{pmatrix} \]
where addition and scalar multiplication for functions is defined in the usual way. That is, \( \phi + \chi \) is the function from \( \mathbb{R} \) to \( \mathbb{R} \) defined by \( (\phi + \chi)(t) = \phi(t) + \chi(t) \) for all \( t \in \mathbb{R} \), and \( \lambda \phi \) is the function from \( \mathbb{R} \) to \( \mathbb{R} \) defined by \( (\lambda \phi)(t) = \lambda(\phi(t)) \) for all \( t \in \mathbb{R} \) (and similarly for \( \psi + \theta \) and \( \lambda \psi \)).

Since addition on \( \mathcal{A} \) is meant to be a function from \( \mathcal{A} \times \mathcal{A} \) to \( \mathcal{A} \), we should check that if \( a, b \in \mathcal{A} \) then \( a + b \), as defined above, is also in \( \mathcal{A} \). Now \( a + b \) will be in \( \mathcal{A} \) if and only if both components of \( a + b \) are differentiable functions from \( \mathbb{R} \) to \( \mathbb{R} \); that is, our definition of addition will only be satisfactory if \( \phi + \chi \) and \( \psi + \theta \) are differentiable functions whenever \( \phi, \psi, \chi \) and \( \theta \) are differentiable functions. Fortunately, this is an elementary theorem of calculus. Similarly, to justify our definition of scalar multiplication we must note that if \( \lambda \in \mathbb{R} \) and \( \phi, \psi \) are differentiable functions then \( \lambda \phi \) and \( \lambda \psi \) are also differentiable functions.

Showing that these definitions of addition and scalar multiplication make \( \mathcal{A} \) into a vector space over \( \mathbb{R} \) would be a matter of checking that the eight axioms in Definition 3.2 are satisfied. This is more tedious than difficult. The first step is to observe that the set \( S \)
of all functions from \(\mathbb{R}\) to \(\mathbb{R}\) is a vector space over \(\mathbb{R}\) (by \#6, p. 54). Now let \(a, b, c \in A\) and \(\lambda, \mu \in \mathbb{R}\). Then

\[
a = \left(\begin{array}{c}
\phi \\
\psi
\end{array}\right), \quad b = \left(\begin{array}{c}
\chi \\
\theta
\end{array}\right), \quad c = \left(\begin{array}{c}
\zeta \\
\eta
\end{array}\right),
\]

where \(\phi, \psi\) etc. are differentiable functions from \(\mathbb{R}\) to \(\mathbb{R}\). Since \(S\) is a vector space we know that addition of functions is associative (vector space axiom (i)), and therefore

\[
(a + b) + c = \left(\begin{array}{c}
(\phi + \chi) + \zeta \\
(\psi + \theta) + \eta
\end{array}\right) = \left(\begin{array}{c}
\phi + \chi + \zeta \\
\psi + \theta + \eta
\end{array}\right) = a + (b + c).
\]

Similarly, since \(S\) satisfies vector space axiom (vi) it follows that

\[
\lambda(\mu a) = \left(\begin{array}{c}
\lambda(\mu \phi) \\
\lambda(\mu \psi)
\end{array}\right) = \left(\begin{array}{c}
(\lambda \mu) \phi \\
(\lambda \mu) \psi
\end{array}\right) = (\lambda \mu) a.
\]

Thus \(A\) satisfies vector space axioms (i) and (vi). Totally analogous proofs work for all the other axioms. Note that the zero element of \(A\) is \(\left(\begin{array}{c}
\xi \\
z
\end{array}\right)\), where \(z\) is the zero function (defined by \(z(t) = 0\) for all \(t\)).

Observe that we could alternatively use Exercise 13 on p. 80 of the book. In the notation of that exercise, \(A = D^2\), where \(D\) is the set of all differentiable functions from \(\mathbb{R}\) to \(\mathbb{R}\). Since \(D\) is nonempty (containing the zero function) and closed under addition and scalar multiplication (by elementary calculus, as observed above) it is a subspace of \(S\), and therefore a vector space itself. The result of Exercise 13 then shows that \(D^2\) is a vector space.

(ii) Let \(\Phi: A \rightarrow S\) be the given function; that is, if \(a = \left(\begin{array}{c}
\phi \\
\psi
\end{array}\right) \in A\) then \(\Phi(a) = h \cdot \phi + \psi'\). Recall that \(S\) is the set of all functions from \(\mathbb{R}\) to \(\mathbb{R}\), so that \(h \cdot \phi + \psi'\) is certainly an element of \(S\).

Let \(a, b \in A\) and \(\lambda \in \mathbb{R}\). As above, let \(a = \left(\begin{array}{c}
\phi \\
\psi
\end{array}\right)\) and \(b = \left(\begin{array}{c}
\chi \\
\theta
\end{array}\right)\). Then

\[
\Phi(a + b) = \Phi\left(\begin{array}{c}
\phi + \chi \\
\psi + \theta
\end{array}\right) = h \cdot (\phi + \chi) + (\psi + \theta)' = (h \cdot \phi + h \cdot \chi) + (\psi' + \theta')
\]

since elementary calculus tells us that the derivative of \(\psi + \theta\) is the sum of the derivatives of \(\psi\) and \(\theta\), while the definitions of sum and pointwise product of functions give (for all \(t \in \mathbb{R}\))

\[
(h \cdot (\psi + \chi))(t) = h(t)(\phi + \chi)(t) = h(t)(\phi(t) + \chi(t)) = h(t)\phi(t) + h(t)\chi(t) = (h \cdot \phi)(t) + (h \cdot \chi)(t).
\]

By commutativity and associativity of addition of functions it follows that

\[
\Phi(a + b) = (h \cdot \phi + \psi') + (h \cdot \chi + \theta') = \Phi(a) + \Phi(b).
\]

In a similar fashion,

\[
\phi(\lambda a) = \Phi\left(\begin{array}{c}
\lambda \phi \\
\lambda \psi
\end{array}\right) = h \cdot (\lambda \phi) + (\lambda \psi)' = \lambda(h \cdot \phi + \lambda \psi') = \lambda \Phi(a).
\]

So \(\Phi\) preserves addition and scalar multiplication; that is, \(\Phi\) is a linear transformation.
3. Let $V$ be a vector space and let $S$ and $T$ be subspaces of $V$.

(i) Prove that $S \cap T$ is a subspace of $V$.

(ii) Let $S + T = \{ x + y \mid x \in S \text{ and } y \in T \}$. Prove that $S + T$ is a subspace of $V$.

Solution.

(i) Let $u, v \in S \cap T$, $\lambda$ a scalar. Since $u, v \in S$ and $S$ is closed under addition and scalar multiplication it follows that $u + v, \lambda u \in S$, and similarly $u + v, \lambda u \in T$. So $u + v, \lambda u \in S \cap T$, and therefore $S \cap T$ is closed under addition and scalar multiplication.

Since $0 \in S$ and $0 \in T$ it follows that $0 \in S \cap T$, and so $S \cap T \neq \emptyset$.

(ii) Let $u, v \in S + T$, $\lambda$ a scalar. Then $u = x + y, v = x' + y'$ for some $x, x' \in S, y, y' \in T$, and by closure of $S$ and $T$,

$$u + v = (x + y) + (x' + y') = (x + x') + (y + y') \in S + T$$

so that $S + T$ is closed also. And $S + T \neq \emptyset$ since $0 = 0 + 0 \in S + T$.

4. Let $V$ be a vector space over the field $F$ and let $v_1, v_2, \ldots, v_n$ be arbitrary elements of $V$. Prove that the span of $\{v_1, v_2, \ldots, v_n\}$

$$\text{Span}(v_1, v_2, \ldots, v_n) = \{ \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n \mid \lambda_1, \lambda_2, \ldots, \lambda_n \in F \}$$

is a subspace of $V$.

Solution.

Let $x, y \in \text{Span}(v_1, v_2, \ldots, v_n)$ and let $\alpha$ be a scalar. Then

$$x = \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n$$

$$y = \mu_1 v_1 + \mu_2 v_2 + \cdots + \mu_n v_n$$

for some scalars $\lambda_i$ and $\mu_i$, and so

$$x + y = (\lambda_1 + \mu_1) v_1 + (\lambda_2 + \mu_2) v_2 + \cdots + (\lambda_n + \mu_n) v_n$$

and

$$\alpha x = \alpha \lambda_1 v_1 + \alpha \lambda_2 v_2 + \cdots + \alpha \lambda_n v_n$$

are both in $\text{Span}(v_1, v_2, \ldots, v_n)$. Furthermore, $\emptyset = \sum_{i=1}^n 0 v_i \in \text{Span}(v_1, v_2, \ldots, v_n)$, which is therefore nonempty.

5. Let $A$ and $B$ be $n \times n$ matrices over the field $F$. We say that $B$ is similar to $A$ if there exists a nonsingular matrix $T$ such that $B = T^{-1} A T$. Prove

(i) every $n \times n$ matrix is similar to itself,

(ii) if $B$ is similar to $A$ then $A$ is similar to $B$,

(iii) if $C$ is similar to $B$ and $B$ is similar to $A$ then $C$ is similar to $A$.

Solution.

For all $A$ we have $I^{-1} A I = A$, and so $A$ is similar to itself. (In the terminology of §1c, this says that similarity is a reflexive relation.)

Suppose that $B$ is similar to $A$. Then there exists a nonsingular $T$ with $B = T^{-1} A T$, and rearranging this equation slightly gives $A = U^{-1} B U$, where $U = T^{-1}$. We deduce that $A$ is similar to $B$ whenever $B$ is similar to $A$. (Similarity is a symmetric relation.)

Suppose that $C$ is similar to $B$ and $B$ is similar to $A$. Then there exist $U$ and $T$ with $C = U^{-1} B U$ and $B = T^{-1} A T$, and it follows that

$$C = U^{-1} B U = U^{-1} T^{-1} A T U = (T U)^{-1} A (T U),$$

whence $C$ is similar to $A$. (Thus similarity is also transitive, and hence is an equivalence relation.)