Math 217 Worksheet 4: Linear Transformations (§2.1)

Definition: A **linear transformation** $T : \mathbb{R}^m \to \mathbb{R}^n$ is a map (i.e., a function) satisfying both of the following:

- $T(\vec{v} + \vec{w}) = T(\vec{v}) + T(\vec{w})$ for all $\vec{v}$, $\vec{w} \in \mathbb{R}^m$ (that is, "T *respects vector addition*").
- $T(a\vec{v}) = aT(\vec{v})$ for all $a \in \mathbb{R}$ and $\vec{v} \in \mathbb{R}^m$ (that is, "T *respects scalar multiplication*").

Problem 1: Examples and Non-Examples. In (a) – (f) below, you are given a map T from $\mathbb{R}^m$ to $\mathbb{R}^n$ for some particular n and m. Determine, in each case, *using only the definition above,* whether or not the given map is a linear transformation.

(a) T is the *identity function* on $\mathbb{R}^4$, defined by $T(\vec{v}) = \vec{v}$ for all $\vec{v} \in \mathbb{R}^4$.

(b) $T : \mathbb{R}^3 \to \mathbb{R}^3$ is *dilation by* 2, defined by $T(\vec{v}) = 2\vec{v}$ for all $\vec{v} \in \mathbb{R}^3$.

(c) $T : \mathbb{R} \to \mathbb{R}$ is defined by $T(x) = x + 1$ for all $x \in \mathbb{R}$.

(d) $T : \mathbb{R}^2 \to \mathbb{R}^2$ sending each vector $\begin{bmatrix} x \\ y \end{bmatrix}$ to its reflection $\begin{bmatrix} y \\ x \end{bmatrix}$ over the line $y = x$.

(e) $T : \mathbb{R}^2 \to \mathbb{R}^2$ where $T(\begin{bmatrix} x \\ y \end{bmatrix}) = \begin{bmatrix} \pi x + \sqrt{17}y \\ -y \end{bmatrix}$

(f) $T : \mathbb{R}^2 \to \mathbb{R}$ assigns to every point in the plane its distance from the origin.

Solution: (a), (b), (d) and (e) are linear, but (c) and (f) are not. For example, (c) does not respect scalar multiplication since $T(2 \times 0) \neq 2T(0)$ and (f) does not respect addition since $T(\vec{e}_1 + (-\vec{e}_1)) \neq T(\vec{e}_1) + T(-\vec{e}_1)$.

Problem 2: Linear transformations from Matrix Multiplication. Let A be an $n \times m$ matrix. We can use matrix multiplication to define a map

$$T_A : \mathbb{R}^m \to \mathbb{R}^n \qquad \vec{x} \mapsto A\vec{x}.$$

(a) Let $A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix}$. What are the source and target of T_A? Write a formula for $T_A(\begin{bmatrix} x \\ y \\ z \end{bmatrix})$.

(b) For A as in (a), compute $T(\vec{e}_1), T(\vec{e}_2)$ and $T(\vec{e}_3)$, and compare to the columns of A.

(c) For A as in (a), prove that T_A is a linear transformation. [First Line: We must check two things...]

(d) For arbitrary $A \in \mathbb{R}^{n\times m}$, prove that T_A is a linear transformation. [Hint: Use Theorem 1.3.10.*]

*The textbook's Theorem 1.3.10 says: For all matrices $A \in \mathbb{R}^{n\times m}$, all vectors $\vec{x}, \vec{y} \in \mathbb{R}^m$, and all scalars $k \in \mathbb{R}$, we have $A(\vec{x} + \vec{y}) = A\vec{x} + A\vec{y}$ and $A(k\vec{x}) = k(A\vec{x})$.

Solution:

(a) The source is $\mathbb{R}^3$ and the target is $\mathbb{R}^2$. A formula is $T_A(\begin{bmatrix} x \\ y \\ z \end{bmatrix}) = \begin{bmatrix} x+y+z \\ y \end{bmatrix})$.

(b) $T_A(\vec{e}_j)$ is the j-th column of A!

(c) We need to show two things:

(i) $T_A(\vec{v}_1 + \vec{v}_2) = T_A(\vec{v}_1) + T_A(\vec{v}_2)$ for all $\vec{v}_1, \vec{v}_2 \in \mathbb{R}^3$.

(ii) $T_A(\lambda\vec{v}) = \lambda T_A(\vec{v})$ for all $\vec{v} \in \mathbb{R}^2$ and $\lambda \in \mathbb{R}$.

Let $\vec{v}_1 = \begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix}$ and $\vec{v}_2 = \begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix}$. Note that $\vec{v}_1 + \vec{v}_2 = \begin{bmatrix} x_1+x_2 \\ y_1+y_2 \\ z_1+z_2 \end{bmatrix}$, and $\lambda\vec{v} = \begin{bmatrix} \lambda x_1 \\ \lambda y_1 \\ \lambda z_1 \end{bmatrix}$. So using our formula in (a), what we need to check becomes

(i') $\begin{bmatrix} (x_1+x_2)+(y_1+y_2)+(z_1+z_2) \\ y_1+y_2 \end{bmatrix} = \begin{bmatrix} x_1+y_1+z_1 \\ y_1 \end{bmatrix} + \begin{bmatrix} x_2+y_2+z_2 \\ y_2 \end{bmatrix}$; and

(ii') $\begin{bmatrix} \lambda x+\lambda y+\lambda z \\ \lambda y \end{bmatrix} = \lambda \begin{bmatrix} x+y+z \\ y \end{bmatrix}$.

Both (i') and (ii') are easy to check, by inspection, using the definition of vector addition and scalar multiplication, and basic properties of arithmetic for real numbers (such as commutativity of addition, distributivity, etc). QED.

(d) We need to show two things:

(i) $T_A(\vec{x} + \vec{y}) = T_A(\vec{x}) + T_A(\vec{y})$ for all $\vec{x}, \vec{y} \in \mathbb{R}^2$.

(ii) $T_A(\lambda\vec{x}) = \lambda T_A(\vec{x})$ for all $\vec{x} \in \mathbb{R}^2$ and $\lambda \in \mathbb{R}$.

For (i):

$$\begin{aligned} T_A(\vec{x}+\vec{y}) &= A(\vec{x}+\vec{y}) \quad \text{by the definition of } T_A \\ &= A\vec{x} + A\vec{y} \quad \text{by the distributive property of matrix multiplication} \\ &= T_A(\vec{x}) + T_A(\vec{y}) \quad \text{by the definition of } T_A. \end{aligned}$$

For (ii): $T_A(\lambda\vec{x}) = A(\lambda\vec{x})$ (by definition of T_A), which is $\lambda A\vec{x}$ by a property of matrix multiplication (Theorem 1.3.10). So $T_A(\lambda\vec{x}) = \lambda T_A(\vec{x})$.

Problem 3. Linear Transformations and Linear Combinations.
Let $\mathbb{R}^2 \xrightarrow{T} \mathbb{R}^2$ be a linear transformation such that $T(\begin{bmatrix} 1 \\ 0 \end{bmatrix}) = \begin{bmatrix} 4 \\ 2 \end{bmatrix}$ and $T(\begin{bmatrix} 0 \\ 1 \end{bmatrix}) = \begin{bmatrix} -2 \\ -2 \end{bmatrix}$.

(a) Using only the *definition* of linear transformation† and the fact that $\begin{bmatrix} 0 \\ 1 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, find $T(\begin{bmatrix} 1 \\ 1 \end{bmatrix})$. Similarly, find $T(\begin{bmatrix} 2 \\ 0 \end{bmatrix})$ and then $T(\begin{bmatrix} 2 \\ 1 \end{bmatrix})$.

(b) Using only the *definition* of linear transformation above and the fact that $\begin{bmatrix} a \\ b \end{bmatrix} = a\begin{bmatrix} 1 \\ 0 \end{bmatrix} + b\begin{bmatrix} 0 \\ 1 \end{bmatrix}$,

†Note that the definition of linear transformation does NOT involve any matrices!

find a formula for $T(\begin{bmatrix} a \\ b \end{bmatrix})$, for the given transformation T.

(c) Find a 2×2 matrix A such that $T(\vec{x}) = A\vec{x}$ by examining the formula you found in (c).

(d) How do the outputs $T(\vec{e}_1)$ and $T(\vec{e}_2)$ relate to your matrix A?

Solution:

(a) Since T respects vector addition, we have

$$T\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right) = T\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) = T\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) + \left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 4 \\ 2 \end{bmatrix} + \begin{bmatrix} -2 \\ -2 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}.$$

Similarly,

$$T\left(\begin{bmatrix} 2 \\ 0 \end{bmatrix}\right) = 2T\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) = 2\begin{bmatrix} 4 \\ 2 \end{bmatrix} = \begin{bmatrix} 8 \\ 4 \end{bmatrix},$$

and $T\left(\begin{bmatrix} 2 \\ 1 \end{bmatrix}\right) = T\left(\begin{bmatrix} 2 \\ 0 \end{bmatrix}\right) + T\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 8 \\ 4 \end{bmatrix} + \begin{bmatrix} -2 \\ -2 \end{bmatrix} = \begin{bmatrix} 6 \\ 2 \end{bmatrix}$.

(b) Because T respects vector addition and scalar multiplication,

$$T\left(\begin{bmatrix} a \\ b \end{bmatrix}\right) = T\left(\begin{bmatrix} a \\ 0 \end{bmatrix}\right) + T\left(\begin{bmatrix} 0 \\ b \end{bmatrix}\right) = aT\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) + bT\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) = a\begin{bmatrix} 4 \\ 2 \end{bmatrix} + b\begin{bmatrix} -2 \\ -2 \end{bmatrix} = \begin{bmatrix} 4a - 2b \\ 2a - 2b \end{bmatrix}.$$

(c) We can take $A = \begin{bmatrix} 4 & -2 \\ 2 & -2 \end{bmatrix}$. We check: $\begin{bmatrix} 4 & -2 \\ 2 & -2 \end{bmatrix}\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 4a - 2b \\ 2a - 2b \end{bmatrix}$, which recovers our formula for $T(\begin{bmatrix} a \\ b \end{bmatrix})$.

(d) The outputs $T(\vec{e}_1)$ and $T(\vec{e}_2)$ are the first and second columns of A, respectively.

Key Theorem. Given a linear transformation $T : \mathbb{R}^d \to \mathbb{R}^n$, there exists a *unique* matrix A such that $T(\vec{x}) = A\vec{x}$ for all $\vec{x} \in \mathbb{R}^d$. Furthermore, we can find A as follows: the j-th column of A is $T(\vec{e}_j)$ where $\vec{e}_j$ is the standard unit column vector in $\mathbb{R}^d$, for each $j = 1, 2, \ldots d$.

Problem 4.

(a) What is the size of the matrix A guaranteed by the Theorem (in terms of info about T)?

(b) For the linear transformation $T : \mathbb{R}^3 \to \mathbb{R}^2$ defined by $T\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} 2x + y + 3z \\ x + y - z \end{bmatrix}$, what is the matrix A guaranteed by the Theorem?

(c) Use the Key Theorem to find a matrix A such that $T = T_A$ for each of the four linear transformations T in Problem 1.

Solution:

(a) $n \times d$.

(b) By the Key Theorem, we can compute the first column: $T(\vec{e}_1) = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$, the second column is $T(\vec{e}_1) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, and the third column is $T(\vec{e}_3) = \begin{bmatrix} 3 \\ -1 \end{bmatrix}$. So the matrix of T is $\begin{bmatrix} 2 & 1 & 3 \\ 1 & 1 & -1 \end{bmatrix}$

(c) Using the Key Theorem, we plug in each of the standard unit vectors $\vec{e}_i$ to find the columns of A:

- For (a): $A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$.
- For (b): $A = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$.
- For (d): $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$.
- For (e): $A = \begin{bmatrix} \pi & \sqrt{17} \\ 0 & -1 \end{bmatrix}$

Problem 5: Linear transformations preserve linear combinations. Let $T : \mathbb{R}^m \to \mathbb{R}^p$ be a linear transformation. Prove that for any vectors $\vec{v}_1, \ldots, \vec{v}_n \in \mathbb{R}^m$ and scalars $c_1, \ldots, c_n \in \mathbb{R}$,

$$T\left(\sum_{i=1}^{n} c_i \vec{v}_i\right) = \sum_{i=1}^{n} c_i T(\vec{v}_i).$$

[Hint: Use the *definition* and induction on n.]

Solution: For the induction base $n = 1$, we have that for any linear transformation $T : \mathbb{R}^m \to \mathbb{R}^p$ and for any vector $\vec{v}_1 \in \mathbb{R}^m$ and scalar $c_1 \in \mathbb{R}$, $T(c_1\vec{v}_1) = c_1 T(\vec{v}_1)$ by linearity of T.

For the inductive step, let $n \in \mathbb{N}$ and assume for inductive hypothesis that for every linear transformation $f : \mathbb{R}^m \to \mathbb{R}^p$ and for all vectors $\vec{v}_1, \ldots, \vec{v}_n \in \mathbb{R}^m$ and scalars $c_1, \ldots, c_n \in \mathbb{R}$,

$$T\left(\sum_{i=1}^{n} c_i \vec{v}_i\right) = \sum_{i=1}^{n} c_i T(\vec{v}_i).$$

Let $T : \mathbb{R}^m \to \mathbb{R}^p$ be linear, and let $\vec{v}_1, \ldots, \vec{v}_{n+1} \in \mathbb{R}^m$ and $c_1, \ldots, c_{n+1} \in \mathbb{R}$. Then, using linearity of T and the inductive hypothesis, we have

$$\begin{aligned} T\left(\sum_{i=1}^{n+1} c_i \vec{v}_i\right) &= T\left(c_{n+1}\vec{v}_{n+1} + \sum_{i=1}^{n} c_i \vec{v}_i\right) = T(c_{n+1}\vec{v}_{n+1}) + T\left(\sum_{i=1}^{n} c_i \vec{v}_i\right) \\ &= c_{n+1}T(\vec{v}_{n+1}) + \sum_{i=1}^{n} c_i T(\vec{v}_i) = \sum_{i=1}^{n+1} c_i T(\vec{v}_i). \end{aligned}$$

This completes the proof by induction.

Problem 6. Let $T : \mathbb{R}^3 \to \mathbb{R}^3$ be the linear transformation which rotates space one quarter turn clockwise around the z-azis (looking down from the positive part of the z-axis). Find a matrix A such that $T(\vec{v}) = A\vec{v}$ for all $\vec{v} \in \mathbb{R}^3$. What is the size of A? Can there be more than choice for A?

Solution: Using the Key theorem, We see $\vec{e}_1$ rotates to $-\vec{e}_2$, and $\vec{e}_2$ rotates to $\vec{e}_1$, while $\vec{e}_3$, being on the z-axis, is fixed. So the matrix is $A = [-\vec{e}_2\ \vec{e}_1\ \vec{e}_3] = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$. There is no other matrix A such that $T(\vec{v}) = A\vec{v}$, by the uniqueness statement in the Key theorem.

Problem 7: Proof of the Key Theorem. Let $T : \mathbb{R}^d \to \mathbb{R}^n$ be a linear transformation.

(a) Let A be the $n \times d$ matrix whose columns are $T(\vec{e}_1), \ldots T(\vec{e}_d)$. For arbitrary $\vec{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_d \end{bmatrix} \in \mathbb{R}^d$, prove that $T(\vec{x}) = A\vec{x}$. [Hint: Write arbitrary $\vec{x} \in \mathbb{R}^d$ as a linear combination of $\vec{e}_1, \ldots, \vec{e}_d$.]

(b) Suppose A and B are two $n \times d$ matrices such that $A\vec{e}_j = B\vec{e}_j$ for all standard unit vectors $\vec{e}_j \in \mathbb{R}^d$. Explain why $A = B$. [Hint: Check that $A\vec{e}j$ is the j-th column of A.]

(c) Prove the Key Theorem. [Proof Technique: To prove a theorem that states "There exists a unique...", there are two steps: (1) existence and (2) uniqueness. To prove existence in this case, write down some matrix A (you *know* what it has to be, right?) and show it has the desired properties using (a). For uniqueness, assume you have two such matrices and prove they must be the same using (b).]

Solution:

(a) For all $\vec{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_d \end{bmatrix} \in \mathbb{R}^d$, we have $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_d \end{bmatrix} = x_1\vec{e}_1 + \cdots + x_d\vec{e}_d$. So

$$T(\vec{x}) = x_1 T(\vec{e}_1) + \cdots + x_d T(\vec{e}_d).$$

On the other hand, for the matrix $A = \begin{bmatrix} | & & | \\ T(\vec{e}_1) & \cdots & T(\vec{e}_n) \\ | & & | \end{bmatrix}$, we have

$$A\vec{x} = \begin{bmatrix} T(\vec{e}_1) & \cdots & T(\vec{e}_d) \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_d \end{bmatrix} = x_1 T(\vec{e}_1) + \cdots + x_d T(\vec{e}_d).$$

So $T_A(\vec{x}) = A\vec{x}$ for all $\vec{x} \in \mathbb{R}^d$.

(b) It is straightforward to check that $A\vec{e}_j$ is the j-th column of A, for any $j = 1, 2, \ldots, d$ (called the "Unreasonable Useful Lemma" in the reading). So if $A\vec{e}_j = B\vec{e}_j$ for some j, then A and B have the same j-th column. So if this holds for all j, it means every column of A is the same as the corresponding column of B. In other words, $A = B$.

(c) Take A to be $\begin{bmatrix} | & & | \\ T(\vec{e}_1) & \cdots & T(\vec{e}_n) \\ | & & | \end{bmatrix}$. Then $T(\vec{x}) = A\vec{x}$ for all $\vec{x} \in \mathbb{R}^d$ by (a). If B is another matrix such that $T(\vec{x}) = B\vec{x}$ for all $\vec{x} \in \mathbb{R}^d$, then in particular, $T(\vec{e}_j) = A\vec{e}_j = B\vec{e}_j$ for all vectors $\vec{e}_1, \ldots, \vec{e}_d$. By (b), this says that $A = B$. So the matrix is unique. This completes the proof.