Linear Transformation

Recall. Let A, B be sets. A *well-defined map* is a subset $f \subset A \times B$ that satisfies the following conditions:
(WD1) For each $a \in A$, there exists a $b \in B$ such that $(a, b) \in f$.
(WD2) If $(a_1, b_1) \in f$ and $(a_2, b_2) \in f$ such that $a_1 = a_2$, then $b_1 = b_2$.

We write $f : A \to B$ to denote that f is a map from A to B. The set A is called the *domain* of f and B is called the *co-domain* of f. In the more familiar notation, we write $f(a) = b$ if $(a, b) \in f$. Thus, we recast the definition above as follows:

(WD1) For each $a \in A$, there exists a $b \in B$ such that $f(a) = b$.
(WD2) If $a_1 = a_2$ then $f(a_1) = f(a_2)$.

Recall. Let $f : A \to B$ be a map. We have "dual" notions to the well-definedness conditions from Definition 0.2.1.
(1) We call f *onto* (*surjective*) if for every $b \in B$, there exists an $a \in A$ such that $b = f(a)$. It is immediate from the definition that f is surjective if and only if $f(A) = B$.
(2) We call f *1-1* (*injective*) if $f(a_1) = f(a_2)$ impies $a_1 = a_2$ for all $a_1, a_2 \in A$.
(3) We call f *bijective* if it is both 1-1 and onto. Recall that map is invertible ($f^{-1} : B \to A$) if andonly if f is a bijection.

Definition 1. Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a map. We call T a linear transformation if the following conditions are satisfied:
(1) $T(\mathbf{x} + \mathbf{y}) = T(\mathbf{x}) + T(\mathbf{y})$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$.
(2) $T(c\mathbf{x}) = cT(\mathbf{x})$ for all $\mathbf{x} \in \mathbb{R}^n$ and $c \in \mathbb{R}$.

Examples 2.
(1) Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be given by

$$T\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x+y \\ 1 \end{pmatrix}.$$

Then T is *not* a linear transformation since

$$T\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 2 \\ 5 \end{pmatrix}\right) = T\begin{pmatrix} 2 \\ 5 \end{pmatrix} = \begin{pmatrix} 2+5 \\ 1 \end{pmatrix} = \begin{pmatrix} 7 \\ 1 \end{pmatrix}$$

while

$$T\begin{pmatrix} 0 \\ 0 \end{pmatrix} + T\begin{pmatrix} 2 \\ 5 \end{pmatrix} = \begin{pmatrix} 0+0 \\ 1 \end{pmatrix} + \begin{pmatrix} 2+5 \\ 1 \end{pmatrix} = \begin{pmatrix} 7 \\ 2 \end{pmatrix}.$$

Thus property (1) of the definition fails.

(2) Let $T : \mathbb{R}^3 \to \mathbb{R}^3$ be given by

$$T\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x+y \\ x-y \\ z \end{pmatrix}.$$

Then T is a linear transformation aswe verify

$$\begin{aligned}
T\left(\begin{pmatrix} x \\ y \\ z \end{pmatrix} + \begin{pmatrix} r \\ s \\ t \end{pmatrix}\right) &= T\begin{pmatrix} x+r \\ y+s \\ z+t \end{pmatrix} \quad \text{(Def of } + \text{ in } \mathbb{R}^3\text{)} \\
&= \begin{pmatrix} (x+r)+(y+s) \\ (x+r)-(y+s) \\ z+t \end{pmatrix} \quad \text{(Def of } T\text{)} \\
&= \begin{pmatrix} (x+y)+(r+s) \\ (x-y)+(r-s) \\ z+t \end{pmatrix} \quad \text{(usual algebra in } \mathbb{R}\text{)} \\
&= \begin{pmatrix} x+y \\ x-y \\ z \end{pmatrix} + \begin{pmatrix} r+s \\ r-s \\ t \end{pmatrix} \quad \text{(Def of } + \text{ in } \mathbb{R}^3\text{)} \\
&= T\begin{pmatrix} x \\ y \\ z \end{pmatrix} + T\begin{pmatrix} r \\ s \\ t \end{pmatrix} \quad \text{(Def of } T\text{)}
\end{aligned}$$

and

$$\begin{aligned}
T\left(c\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) &= T\begin{pmatrix} cx \\ cy \\ cz \end{pmatrix} \quad \text{(Def of } + \text{ in } \mathbb{R}^3\text{)} \\
&= \begin{pmatrix} cx+cy \\ cx-cy \\ cz \end{pmatrix} \quad \text{(Def of } T\text{)} \\
&= \begin{pmatrix} c(x+y) \\ c(x-y) \\ cz \end{pmatrix} \quad \text{(Usual algebra in } \mathbb{R}\text{)} \\
&= c\begin{pmatrix} x+y \\ x-y \\ z \end{pmatrix} \quad \text{(Def of scalar mult in } \mathbb{R}^3\text{)} \\
&= cT\begin{pmatrix} x+y \\ x-y \\ z \end{pmatrix} \quad \text{(Def of } T\text{)}
\end{aligned}$$

Facts 3. Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation.
(1) $T(\mathbf{0}) = \mathbf{0}$.
(2) $T(\sum_{i=1}^{n} c_i \mathbf{v}_i) = \sum_{i=1}^{n} c_i T(\mathbf{v}_i)$ for every $n \in \mathbb{Z}^+$.

Proof.
(1) We have that $T(\mathbf{0}) = T(\mathbf{0}+\mathbf{0}) = T(\mathbf{0}) + (\mathbf{0})$. Adding $-T(\mathbf{0})$ to both sides, we find that $\mathbf{0} = T(\mathbf{0})$.
(2) We induct on n. Let $P(n)$ be the statement "$T(\sum_{i=1}^{n} c_i \mathbf{v}_i) = \sum_{i=1}^{n} c_i T(\mathbf{v}_i)$". Since $T(\sum_{i=1}^{1} c_i \mathbf{v}_i) = T(c_1 \mathbf{v}_1) = c_1 T(\mathbf{v}_1) = \sum_{i=1}^{1} c_1 T(\mathbf{v}_1)$, we have that $P(1)$ is true. Suppose that $k \geq 1$ and that $P(k)$ is true. We must show that $P(k+1)$ is true. That is, we must show that $T(\sum_{i=1}^{k+1} c_i \mathbf{v}_i) = \sum_{i=1}^{k+1} c_i T(\mathbf{v}_i)$. But

$$
\begin{aligned}
& T(\sum_{i=1}^{k+1} c_i \mathbf{v}_i) \\
= \quad & T(c_{k+1}\mathbf{v}_{k+1} + \sum_{i=1}^{k} c_i \mathbf{v}_i) \quad \text{(Inductive defn of } \textstyle\sum \text{)} \\
= \quad & c_{k+1}T(\mathbf{v}_{k+1}) + T(\sum_{i=1}^{k} c_i \mathbf{v}_i) \quad \text{(Since } T \text{ is a linear transformation)} \\
= \quad & c_{k+1}T(\mathbf{v}_{k+1}) + \sum_{i=1}^{k} c_i T(\mathbf{v}_i) \quad \text{(Induction assumption)} \\
= \quad & \sum_{i=1}^{k+1} c_i T(\mathbf{v}_i) \quad \text{(Inductive defn of } \textstyle\sum \text{)}
\end{aligned}
$$

Definition 4. Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation.
(1) We define the *kernel* of the map T to be the set

$$\ker T = \{\mathbf{x} \in \mathbb{R}^n : T(\mathbf{x}) = \mathbf{0}\}.$$

(2) We define the *image* of the map T to be the set

$$\operatorname{Im} T = \{\mathbf{b} \in \mathbb{R}^m : T(\mathbf{x}) = \mathbf{b} \text{ for some } \mathbf{x} \in \mathbb{R}^n\}.$$

Theorem 5. Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation.
(1) $\ker T \leq \mathbb{R}^n$.
(2) $\operatorname{Im} T \leq \mathbb{R}^m$.
(3) T is 1-1 if and only if $\ker T = \{\mathbf{0}\}$.
(4) T is onto if and only if $\operatorname{Im} T = \mathbb{R}^m$.

Proof.
(1) We must check three properties. (i) Since $T(\mathbf{0}) = \mathbf{0}$ (Facts 3(1)), $\mathbf{0} \in \ker T$. (ii) If $\mathbf{x}, \mathbf{y} \in \ker T$, then $T(\mathbf{x}) = \mathbf{0}$ and $T(\mathbf{y}) = \mathbf{0}$. Adding these equalities, we obtain $T(\mathbf{x}) + T(\mathbf{y}) = \mathbf{0}$. Since T is a linear transformation, $T(\mathbf{x} + \mathbf{y}) = \mathbf{0}$. Therefore, $\mathbf{x} + \mathbf{y} \in \ker T$. (iii) If $\mathbf{x} \in \ker T$ and $c \in \mathbb{R}$, then $T(\mathbf{x}) = \mathbf{0}$. Multiplying both sides of the equality by c, we obtain $cT(\mathbf{x}) = \mathbf{0}$. Since T is a linear transformation, $T(c\mathbf{x}) = \mathbf{0}$. Therefore, $c\mathbf{x} \in \ker T$.

(2) We must check three properties. (i) Since $T(\mathbf{0}) = \mathbf{0}$ (Facts 3(1)), $\mathbf{0} \in \operatorname{Im} T$. (ii) If $\mathbf{b}, \mathbf{c} \in \operatorname{Im} T$, then $T(\mathbf{x}) = \mathbf{b}$ and $T(\mathbf{y}) = \mathbf{c}$ for some $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. Adding these equalities, we obtain $T(\mathbf{x}) + T(\mathbf{y}) = \mathbf{b} + \mathbf{c}$. Since T is a linear transformation, $T(\mathbf{x} + \mathbf{y}) = \mathbf{b} + \mathbf{c}$. Therefore, $\mathbf{b} + \mathbf{c} \in \operatorname{Im} T$ since $\mathbf{x} + \mathbf{y} \in \mathbb{R}^n$. (iii) If $\mathbf{b} \in \operatorname{Im} T$ and $c \in \mathbb{R}$, then $T(\mathbf{x}) = \mathbf{b}$ for some $\mathbf{x} \in \mathbb{R}^n$. multiplying both sides of the equality by c, we obtain $cT(\mathbf{x}) = c\mathbf{b}$. Since T is a linear transformation, $T(c\mathbf{x}) = c\mathbf{b}$. Therefore, $c\mathbf{b} \in \operatorname{Im} T$ since $c\mathbf{x} \in \mathbb{R}^n$.

(3) ($\Rightarrow$) Suppose that $\mathbf{x} \in \ker T$. Then $T(\mathbf{x}) = \mathbf{0}$, and Facts 3(1) says $T(\mathbf{0}) = \mathbf{0}$. It follows that $T(\mathbf{x}) = T(\mathbf{0})$. Since T is 1-1, $\mathbf{x} = \mathbf{0}$. Therefore, $\ker T \subseteq \{\mathbf{0}\}$ and so $T = \{\mathbf{0}\}$.

($\Leftarrow$) Suppose that $T(\mathbf{x}) = T(\mathbf{y})$. We must show that $\mathbf{x} = \mathbf{y}$. Now, $T(\mathbf{x}) = T(\mathbf{y})$ implies $T(\mathbf{x}) - T(\mathbf{y}) = \mathbf{0}$. Since T is a linear transformation, $T(\mathbf{x} - \mathbf{y}) = \mathbf{0}$. It follows that $\mathbf{x} - \mathbf{y} \in \ker T$ and the hypothesis says $\mathbf{x} - \mathbf{y} = \mathbf{0}$. Therefore, $\mathbf{x} = \mathbf{y}$ as needed.

(4) Follows almost immediately from the definition.

Theorem 6. Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. There exists a matrix $A \in \mathcal{M}_{m \times n}$ such that $T(\mathbf{x}) = A\mathbf{x}$ for every $\mathbf{x} \in \mathbb{R}^n$.

Proof. Let

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, ..., \mathbf{e}_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

be the standard basis in $\mathbb{R}^n$. Since $T(\mathbf{e}_j) \in \mathbb{R}^m$, we can write

$$T(\mathbf{e}_1) = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix}, T(\mathbf{e}_2) = \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix}, ..., T(\mathbf{e}_n) = \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix}.$$

Set

$$A = \left[\begin{array}{cccc} T(\mathbf{e}_1) & T(\mathbf{e}_2) & \cdots & T(\mathbf{e}_n) \end{array}\right].$$

Then

$$T(\mathbf{x}) = T\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = T(\sum_{j=1}^{n} x_j \mathbf{e}_j) = \sum_{j=1}^{n} x_j T(\mathbf{e}_j) = \sum_{j=1}^{n} x_j \mathbf{c}_j(A) = A\mathbf{x}.$$

Definition 7. Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. The matrix $A_T = \begin{bmatrix} T(\mathbf{e}_1) & T(\mathbf{e}_2) & \cdots & T(\mathbf{e}_n) \end{bmatrix} \in \mathcal{M}_{m\times n}$ is called the standard matrix of the the transformation T.

Theorem 8. Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation and let A_T be the standard matrix for T.
(1) T is 1-1 if and onlly if $\mathrm{Null}(A_T) = \{\mathbf{0}\}$.
(2) T is onto if and onlly if $\mathrm{Col}(A_T) = \mathbb{R}^m$.

Proof. One easily checks that $\mathrm{Null}(A_T) = \ker T$ and $\mathrm{Col}(A_T) = \mathrm{Im}\, T$.

Example 9. We find A_T for the transformation $T : \mathbb{R}^3 \to \mathbb{R}^3$ be given by

$$T\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x+y \\ x-y \\ z \end{pmatrix}.$$

We have

$$T\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, T\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, T\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

and so

$$A_T = \begin{bmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Since

$$rref(A_T) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

it follows that

$$\mathrm{Null}(A_T) = \{\mathbf{0}\} \text{ and } \mathrm{Col}(A_T) = \mathbb{R}^3.$$

Theorem 10. Let $T : \mathbb{R}^n \to \mathbb{R}^p$ and $S : \mathbb{R}^p \to \mathbb{R}^m$ be linear transformations. Then $(S \circ T) : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation. Moreover, $A_{S\circ T} = A_S A_T$.

Proof. For $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ and $c \in \mathbb{R}$, we have

$$\begin{aligned}
& (S \circ T)(\mathbf{x} + \mathbf{y}) \\
= \ & S(T(\mathbf{x} + \mathbf{y})) \quad (\text{Def of } \circ) \\
= \ & S(T(\mathbf{x}) + T(\mathbf{y})) \quad (T \text{ is LT}) \\
= \ & S(T(\mathbf{x})) + S(T(\mathbf{y})) \quad (S \text{ is LT}) \\
= \ & (S \circ T)(\mathbf{x}) + (S \circ T)(\mathbf{y}) \quad (\text{Def of } \circ)
\end{aligned}$$

and

$$\begin{aligned}
& (S \circ T)(c\mathbf{x}) \\
= \ & S(T(c\mathbf{x})) \quad (\text{Def of } \circ) \\
= \ & S(cT(\mathbf{x})) \quad (T \text{ is LT}) \\
= \ & cS(T(\mathbf{x})) \quad (S \text{ is LT}) \\
= \ & c(S \circ T)(\mathbf{x}) \quad (\text{Def of } \circ).
\end{aligned}$$

For the second part, choose any $\mathbf{x} \in \mathbb{R}^n$. We have

$$\begin{aligned}
(S \circ T)(\mathbf{x}) &= S(T(\mathbf{x})) \quad (\text{Def of } \circ) \\
A_{S \circ T}\mathbf{x} &= S(T(\mathbf{x})) \quad (\text{Theorem 6 on Left}) \\
A_{S \circ T}\mathbf{x} &= A_S T(\mathbf{x}) \quad (\text{Theorem 6 on Right}) \\
A_{S \circ T}\mathbf{x} &= A_S(A_T\mathbf{x}) \quad (\text{Theorem 6 on Right}) \\
A_{S \circ T}\mathbf{x} &= (A_S A_T)\mathbf{x} \quad (\text{Associativity}) \\
A_{S \circ T} &= A_S A_T \quad (\text{An old HW exercise})
\end{aligned}$$

Theorem 11. If $T : \mathbb{R}^n \to \mathbb{R}^n$ be a bijective a linear transformation, then $T^{-1} : \mathbb{R}^n \to \mathbb{R}^n$ is a linear transformation. Moreover, $A_{T^{-1}} = A_T^{-1}$.

Proof. Exercise.

Linear Transformations

1. Let $T : \mathbb{R}^3 \to \mathbb{R}^2$ be given by

$$T\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x - 2y \\ 2x + y \end{pmatrix}.$$

(a) Find A_T.
(b) Is T 1-1? If not, does there exist a 1-1 map $P : \mathbb{R}^3 \to \mathbb{R}^2$? Justify.
(c) Is T onto? Justify.
(d) Find $\dim(\ker T)$ and $\dim(\operatorname{Im} T)$.

2. Let $S : \mathbb{R}^2 \to \mathbb{R}^4$ be given by

$$S\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 3x \\ y - 2x \\ 2x \\ x + 2y \end{pmatrix}.$$

Find a formula for $(S \circ T) : \mathbb{R}^3 \to \mathbb{R}^4$.

3. Prove Theorem 11: If $T : \mathbb{R}^n \to \mathbb{R}^n$ be a bijective a linear transformation, then $T^{-1} : \mathbb{R}^n \to \mathbb{R}^n$ is a linear transformation. Moreover, $A_{T^{-1}} = A_T^{-1}$.

4. Let $T : \mathbb{R}^3 \to \mathbb{R}^3$ be given by

$$T\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2x + 3y + z \\ 3x + 3y + z \\ 2x + 4y + z \end{pmatrix}.$$

Prove that T is an invertible map (1-1 and onto) and find a fromula for $T^{-1} : \mathbb{R}^3 \to \mathbb{R}^3$.

5. Construct a linear transformation $\rho : \mathbb{R}^2 \to \mathbb{R}^2$ that rotates every vector through an angle of $\theta = \frac{\pi}{2}$. Find the standard matrix A_ρ of the transformation and verify that ρ really does rotate the plane through $\theta = \frac{\pi}{2}$.

6. Let $B = \{\mathbf{v}_1, ..., \mathbf{v}_n\}$ be a *basis* of $\mathbb{R}^n$. Let $\{\mathbf{w}_1, ..., \mathbf{w}_n\}$ be a *list* of n vectors in $\mathbb{R}^m$. Prove the following statements
(a) There exists a linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ such that $T(\mathbf{v}_i) = \mathbf{w}_i$ for each index $i \leq n$.
(b) If $S : \mathbb{R}^n \to \mathbb{R}^m$ is another linear transformation such that $S(\mathbf{v}_i) = \mathbf{w}_i$ for each index $i \leq n$, then $S = T$.
(c) T is onto if and only if $\{\mathbf{w}_1, ..., \mathbf{w}_n\}$ spans $\mathbb{R}^m$.

(d) T is one-to-one if and only if $\{\mathbf{w}_1, ..., \mathbf{w}_n\}$ is a linearly independent subset of $\mathbb{R}^m$.
(e) T is a bijection if and only if $\{\mathbf{w}_1, ..., \mathbf{w}_n\}$ is a basis of $\mathbb{R}^m$.

7. Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. Prove the following statements.
(a) $\dim(\operatorname{Im} T) \leq n$.
(b) $n = \dim(\ker T) + \dim(\operatorname{Im} T)$.
(c) If T is 1-1, then $n \leq m$.
(d) If T is onto, then $m \leq n$.
(e) If $n = m$, then T is onto if and only if T is a bijection if and only if T is 1-1.