Answers/Solutions of Exercise 7

1. (a) T_1 is a linear transformation with standard matrix $\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$.

   (b) T_2 is not a linear transformation.

   (c) T_3 is a linear transformation with standard matrix $\begin{pmatrix} 1 & 1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$.

   (d) T_4 is not a linear transformation.

   (e) T_5 is a linear transformation with standard matrix $\begin{pmatrix} y_1 & y_2 & \cdots & y_n \end{pmatrix}$.

   (f) T_6 is not a linear transformation.

2. (a) There is enough information.

   $$T\left(\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) = \begin{pmatrix} x+2y \\ 3x+2y+4z \\ -y+z \\ x+4y+6z \end{pmatrix} \quad \text{for } \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3.$$

   The standard matrix is $\begin{pmatrix} 1 & 2 & 0 \\ 3 & 2 & 4 \\ 0 & -1 & 1 \\ 1 & 4 & 6 \end{pmatrix}$.

   (b) The information is not enough because the two vectors do not form a basis for $\mathbb{R}^2$.

   (c) There is enough information.

   $$T\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \begin{pmatrix} x-y \\ x+y \end{pmatrix} \quad \text{for } \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2.$$

   The standard matrix is $\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$.

   (d) There is enough information.

   $$T\left(\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) = \frac{1}{5}\begin{pmatrix} x+17y-8z \\ x+22y-8z \end{pmatrix} \quad \text{for } \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3.$$

   The standard matrix is $\begin{pmatrix} \frac{1}{5} & \frac{17}{5} & \frac{-8}{5} \\ \frac{1}{5} & \frac{22}{5} & \frac{-8}{5} \end{pmatrix}$.

(e) There is enough information.

$$T\left(\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) = \begin{pmatrix} y+z \\ z \\ x+z \end{pmatrix} \quad \text{for} \quad \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3.$$

The standard matrix is $\begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix}$.

(f) The information is not enough because the three vectors do not form a basis for $\mathbb{R}^3$.

3. (a) $(S \circ T)\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \begin{pmatrix} 2x \\ 2y \\ x+y \end{pmatrix}$.

$T \circ S$ is not defined.

(b) $(S \circ T)\left(\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) = \begin{pmatrix} -2x-y+3z \\ -x-y+3z \\ -3x-2y+6z \end{pmatrix}$.

$(T \circ S)\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \begin{pmatrix} 2x+2y \\ 2x+y \end{pmatrix}$.

4. ($\Rightarrow$) It is a particular case of Theorem 7.1.4.2.

($\Leftarrow$) Suppose

$$T(c\boldsymbol{u} + d\boldsymbol{v}) = c\,T(\boldsymbol{u}) + d\,T(\boldsymbol{v}) \quad \text{for all } \boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^n \text{ and } c, d \in \mathbb{R}. \qquad (*)$$

Let $\{\boldsymbol{e_1}, \boldsymbol{e_2}, \dots, \boldsymbol{e_n}\}$ be the standard basis for $\mathbb{R}^n$ and let $\boldsymbol{A}$ be the $m \times n$ matrix $\begin{pmatrix} T(\boldsymbol{e_1}) & T(\boldsymbol{e_2}) & \cdots & T(\boldsymbol{e_n}) \end{pmatrix}$.

For any $\boldsymbol{u} = (u_1, u_2, \dots, u_n)^{\mathrm{T}} \in \mathbb{R}^n$, $\boldsymbol{u} = u_1\boldsymbol{e_1} + u_2\boldsymbol{e_2} + \cdots + u_n\boldsymbol{e_n}$. By applying $(*)$ repeatedly, we have

$$\begin{aligned} T(\boldsymbol{u}) &= u_1 T(\boldsymbol{e_1}) + u_2 T(\boldsymbol{e_2}) + \cdots + u_n T(\boldsymbol{e_n}) \\ &= \begin{pmatrix} T(\boldsymbol{e_1}) & T(\boldsymbol{e_2}) & \cdots & T(\boldsymbol{e_n}) \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} \\ &= \boldsymbol{A}\boldsymbol{u}. \end{aligned}$$

Thus T is a linear transformation.

5. (a) For any $\boldsymbol{u} \in \mathbb{R}^n$, $(T_1 + T_2)(\boldsymbol{u}) = T_1(\boldsymbol{u}) + T_2(\boldsymbol{u}) = \boldsymbol{A}\boldsymbol{u} + \boldsymbol{B}\boldsymbol{u} = (\boldsymbol{A} + \boldsymbol{B})\boldsymbol{u}$. So $T_1 + T_2$ is a linear transformation and the standard matrix for $T_1 + T_2$ is $\boldsymbol{A} + \boldsymbol{B}$.

   (b) For any $\boldsymbol{u} \in \mathbb{R}^n$, $(\lambda T)(\boldsymbol{u}) = \lambda T(\boldsymbol{u}) = \lambda \boldsymbol{A}\boldsymbol{u} = (\lambda \boldsymbol{A})\boldsymbol{u}$. So λT is a linear transformation and the standard matrix for λT is $\lambda \boldsymbol{A}$.

6. (a) (i) T is invertible and the inverse of T is T itself.

   (ii) T is not invertible. Assume there exists an inverse $S : \mathbb{R}^2 \to \mathbb{R}^2$. Then $(1,0)^{\mathrm{T}} = S \circ T((1,0)^{\mathrm{T}}) = S((1,0)^{\mathrm{T}}) = S \circ T((0,1)^{\mathrm{T}}) = (0,1)^{\mathrm{T}}$, a contradiction.

   (b) $\boldsymbol{A}^{-1}$.

7. (a) Note that $(\boldsymbol{n} \cdot \boldsymbol{x})\boldsymbol{n} = \boldsymbol{n}\boldsymbol{n}^{\mathrm{T}}\boldsymbol{x}$ where LHS is the scalar $\boldsymbol{n} \cdot \boldsymbol{x}$ multiplied to the vector $\boldsymbol{n}$ while all operations on RHS are matrix multiplications. (To verify the equation, let $\boldsymbol{n} = (a_1, \ldots, a_n)^{\mathrm{T}}$ and $\boldsymbol{x} = (x_1, \ldots, x_n)^{\mathrm{T}}$ and then check that both sides give us the same vector.)

   For any $\boldsymbol{x} \in \mathbb{R}^n$, $P(\boldsymbol{x}) = \boldsymbol{x} - (\boldsymbol{n} \cdot \boldsymbol{x})\boldsymbol{n} = \boldsymbol{I}\boldsymbol{x} - \boldsymbol{n}\boldsymbol{n}^{\mathrm{T}}\boldsymbol{x} = (\boldsymbol{I} - \boldsymbol{n}\boldsymbol{n}^{\mathrm{T}})\boldsymbol{x}$. So P is a linear transformation and the standard matrix for P is $\boldsymbol{I} - \boldsymbol{n}\boldsymbol{n}^{\mathrm{T}}$.

   (b) Since for all $\boldsymbol{x} \in \mathbb{R}^n$,

$$
\begin{aligned}
(P \circ P)(\boldsymbol{x}) = P(P(\boldsymbol{x})) &= P(\boldsymbol{x} - (\boldsymbol{n} \cdot \boldsymbol{x})\boldsymbol{n}) \\
&= \boldsymbol{x} - (\boldsymbol{n} \cdot \boldsymbol{x})\boldsymbol{n} - \{\boldsymbol{n} \cdot [\boldsymbol{x} - (\boldsymbol{n} \cdot \boldsymbol{x})\boldsymbol{n}]\}\boldsymbol{n} \\
&= \boldsymbol{x} - (\boldsymbol{n} \cdot \boldsymbol{x})\boldsymbol{n} - \{(\boldsymbol{n} \cdot \boldsymbol{x}) - (\boldsymbol{n} \cdot \boldsymbol{x})(\boldsymbol{n} \cdot \boldsymbol{n})\}\boldsymbol{n} \\
&= \boldsymbol{x} - (\boldsymbol{n} \cdot \boldsymbol{x})\boldsymbol{n} = P(\boldsymbol{x}),
\end{aligned}
$$

   $P \circ P = P$.

   Alternatively, since $\boldsymbol{n}$ is a unit vector, $\boldsymbol{n}^{\mathrm{T}}\boldsymbol{n} = \boldsymbol{n} \cdot \boldsymbol{n} = 1$. Thus

$$
(\boldsymbol{I} - \boldsymbol{n}\boldsymbol{n}^{\mathrm{T}})^2 = (\boldsymbol{I} - \boldsymbol{n}\boldsymbol{n}^{\mathrm{T}})(\boldsymbol{I} - \boldsymbol{n}\boldsymbol{n}^{\mathrm{T}}) = \boldsymbol{I} - 2\boldsymbol{n}\boldsymbol{n}^{\mathrm{T}} + \boldsymbol{n}\boldsymbol{n}^{\mathrm{T}}\boldsymbol{n}\boldsymbol{n}^{\mathrm{T}} = \boldsymbol{I} - \boldsymbol{n}\boldsymbol{n}^{\mathrm{T}}.
$$

   By Theorem 7.1.11, $P \circ P = P$.

8. (a) Suppose T is not the zero transformation. So there exists $\boldsymbol{x} \in \mathbb{R}^n$ such that $T(\boldsymbol{x}) \neq 0$. Define $\boldsymbol{u} = T(\boldsymbol{x})$. Then $\boldsymbol{u}$ is a nonzero vector and

$$
T(\boldsymbol{u}) = T(T(\boldsymbol{x})) = (T \circ T)(\boldsymbol{x}) = T(\boldsymbol{x}) = \boldsymbol{u}.
$$

   (b) Suppose T is not the identity transformation. So there exists $\boldsymbol{y} \in \mathbb{R}^n$ such that $T(\boldsymbol{y}) \neq \boldsymbol{y}$. Define $\boldsymbol{v} = T(\boldsymbol{y}) - \boldsymbol{y}$. Then $\boldsymbol{v}$ is a nonzero vector and

$$
T(\boldsymbol{v}) = T(T(\boldsymbol{y}) - \boldsymbol{y}) = (T \circ T)(\boldsymbol{y}) - T(\boldsymbol{y}) = T(\boldsymbol{y}) - T(\boldsymbol{y}) = \boldsymbol{0}.
$$

(c) Let $\boldsymbol{A}$ be the standard matrix for T. If T is not the zero transformation and the identity transformation, then by (a) and (b), 1 and 0 are the eigenvalues of $\boldsymbol{A}$. So by the result of Question 6.4,

$$\boldsymbol{A} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} r & s \\ t & 1-r \end{pmatrix} \quad \text{where } st = r(1-r).$$

9. (a) Similar to Question 7.7, for any $\boldsymbol{x} \in \mathbb{R}^n$, $F(\boldsymbol{x}) = \boldsymbol{x} - 2(\boldsymbol{n} \cdot \boldsymbol{x})\boldsymbol{n} = \boldsymbol{I}\boldsymbol{x} - 2\boldsymbol{n}\boldsymbol{n}^{\mathrm{T}}\boldsymbol{x} = (\boldsymbol{I} - 2\boldsymbol{n}\boldsymbol{n}^{\mathrm{T}})\boldsymbol{x}$. So F is a linear transformation and the standard matrix for F is $\boldsymbol{I} - 2\boldsymbol{n}\boldsymbol{n}^{\mathrm{T}}$.

(b) Since for all $\boldsymbol{x} \in \mathbb{R}^n$,

$$\begin{aligned}(F \circ F)(\boldsymbol{x}) = F(F(\boldsymbol{x})) &= F(\boldsymbol{x} - 2(\boldsymbol{n} \cdot \boldsymbol{x})\boldsymbol{n}) \\ &= \boldsymbol{x} - 2(\boldsymbol{n} \cdot \boldsymbol{x})\boldsymbol{n} - 2\{\boldsymbol{n} \cdot [\boldsymbol{x} - 2(\boldsymbol{n} \cdot \boldsymbol{x})\boldsymbol{n}]\}\boldsymbol{n} \\ &= \boldsymbol{x} - 2(\boldsymbol{n} \cdot \boldsymbol{x})\boldsymbol{n} - 2\{(\boldsymbol{n} \cdot \boldsymbol{x}) - 2(\boldsymbol{n} \cdot \boldsymbol{x})(\boldsymbol{n} \cdot \boldsymbol{n})\}\boldsymbol{n} \\ &= \boldsymbol{x} - 2(\boldsymbol{n} \cdot \boldsymbol{x})\boldsymbol{n} - 2\{-(\boldsymbol{n} \cdot \boldsymbol{x})\} = \boldsymbol{x},\end{aligned}$$

$F \circ F$ is the identity transformation.

Alternatively,

$$(\boldsymbol{I} - 2\boldsymbol{n}\boldsymbol{n}^{\mathrm{T}})^2 = (\boldsymbol{I} - 2\boldsymbol{n}\boldsymbol{n}^{\mathrm{T}})(\boldsymbol{I} - 2\boldsymbol{n}\boldsymbol{n}^{\mathrm{T}}) = \boldsymbol{I} - 4\boldsymbol{n}\boldsymbol{n}^{\mathrm{T}} + 4\boldsymbol{n}\boldsymbol{n}^{\mathrm{T}}\boldsymbol{n}\boldsymbol{n}^{\mathrm{T}} = \boldsymbol{I}.$$

By Theorem 7.1.11, $F \circ F$ is the identity transformation.

(c) Note that $(\boldsymbol{I} - 2\boldsymbol{n}\boldsymbol{n}^{\mathrm{T}})^{\mathrm{T}} = \boldsymbol{I} - 2(\boldsymbol{n}\boldsymbol{n}^{\mathrm{T}})^{\mathrm{T}} = \boldsymbol{I} - 2\boldsymbol{n}\boldsymbol{n}^{\mathrm{T}}$. Thus

$$(\boldsymbol{I} - 2\boldsymbol{n}\boldsymbol{n}^{\mathrm{T}})(\boldsymbol{I} - 2\boldsymbol{n}\boldsymbol{n}^{\mathrm{T}})^{\mathrm{T}} = (\boldsymbol{I} - 2\boldsymbol{n}\boldsymbol{n}^{\mathrm{T}})^2 = \boldsymbol{I}$$

by (b). The standard matrix is an orthogonal matrix.

10. (a) By Theorem 7.1.4.2,

$$\begin{aligned}T(\boldsymbol{u}+\boldsymbol{v}) \cdot T(\boldsymbol{u}+\boldsymbol{v}) &= (T(\boldsymbol{u}) + T(\boldsymbol{v})) \cdot (T(\boldsymbol{u}) + T(\boldsymbol{v})) \\ &= T(\boldsymbol{u}) \cdot T(\boldsymbol{u}) + 2(T(\boldsymbol{u}) \cdot T(\boldsymbol{v})) + T(\boldsymbol{v}) \cdot T(\boldsymbol{v}) \\ &= ||T(\boldsymbol{u})||^2 + ||T(\boldsymbol{v})||^2 + 2(T(\boldsymbol{u}) \cdot T(\boldsymbol{v})) \\ &= ||\boldsymbol{u}||^2 + ||\boldsymbol{v}||^2 + 2(T(\boldsymbol{u}) \cdot T(\boldsymbol{v})). \qquad (1)\end{aligned}$$

On the other hand,

$$\begin{aligned}T(\boldsymbol{u}+\boldsymbol{v}) \cdot T(\boldsymbol{u}+\boldsymbol{v}) &= ||T(\boldsymbol{u}+\boldsymbol{v})||^2 \\ &= ||\boldsymbol{u}+\boldsymbol{v}||^2 \\ &= (\boldsymbol{u}+\boldsymbol{v}) \cdot (\boldsymbol{u}+\boldsymbol{v}) \\ &= \boldsymbol{u} \cdot \boldsymbol{u} + 2(\boldsymbol{u} \cdot \boldsymbol{v}) + \boldsymbol{v} \cdot \boldsymbol{v} \\ &= ||\boldsymbol{u}||^2 + ||\boldsymbol{v}||^2 + 2(\boldsymbol{u} \cdot \boldsymbol{v}). \qquad (2)\end{aligned}$$

Thus (1) and (2) imply $T(\boldsymbol{u}) \cdot T(\boldsymbol{v}) = \boldsymbol{u} \cdot \boldsymbol{v}$.

(b) ($\Leftarrow$) Suppose $\boldsymbol{A}$ is an orthogonal matrix of order n. Then by Question 5.32, for all $\boldsymbol{u} \in \mathbb{R}^n$,

$$||T(\boldsymbol{u})|| = ||\boldsymbol{A}\boldsymbol{u}|| = ||\boldsymbol{u}||.$$

So T is an isometry.

($\Rightarrow$) Suppose T is an isometry on $\mathbb{R}^n$. Let $\{\boldsymbol{e_1}, \boldsymbol{e_2}, \ldots, \boldsymbol{e_n}\}$ be the standard basis for $\mathbb{R}^n$. Then

$$\begin{aligned}(\boldsymbol{A}\boldsymbol{e_i}) \cdot (\boldsymbol{A}\boldsymbol{e_j}) = (\boldsymbol{A}\boldsymbol{e_i})^{\mathrm{T}}\boldsymbol{A}\boldsymbol{e_j} &= \boldsymbol{e_i}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{e_j} \\ &= \text{the } (i,j)\text{-entry of } \boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}. \end{aligned} \tag{3}$$

On the other hand, by (a),

$$(\boldsymbol{A}\boldsymbol{e_i}) \cdot (\boldsymbol{A}\boldsymbol{e_j}) = T(\boldsymbol{e_i}) \cdot T(\boldsymbol{e_j}) = \boldsymbol{e_i} \cdot \boldsymbol{e_j} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j. \end{cases} \tag{4}$$

By (3) and (4), $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A} = \boldsymbol{I}$. By Remark 5.4.4, $\boldsymbol{A}$ is an orthogonal matrix.

(c) All isometries on $\mathbb{R}^2$ are of the form

$$T\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \begin{pmatrix} x\cos(\theta) + \delta y \sin(\theta) \\ x\sin(\theta) - \delta y\cos(\theta) \end{pmatrix} \quad \text{for } \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2$$

where $\delta = \pm 1$ and $0 \leq \theta < 2\pi$.

11. The standard matrix of T is $\begin{pmatrix} 2 & 1 & 0 \\ 1 & -1 & 1 \end{pmatrix}$.

$$\begin{pmatrix} 2 & 1 & 0 \\ 1 & -1 & 1 \end{pmatrix} \xrightarrow[\text{Elimination}]{\text{Gauss-Jordan}} \begin{pmatrix} 1 & 0 & \frac{1}{3} \\ 0 & 1 & -\frac{2}{3} \end{pmatrix}$$

(a) $\{(2,1)^{\mathrm{T}}, (1,-1)^{\mathrm{T}}\}$ is a basis for $\mathrm{R}(T)$. (For this example, any two linearly independent vectors in $\mathbb{R}^2$ is a basis for $\mathrm{R}(T)$. Why?)

(b) $\{(-\frac{1}{3}, \frac{2}{3}, 1)^{\mathrm{T}}\}$ is a basis for $\mathrm{Ker}(T)$.

(c) $\mathrm{rank}(T) + \mathrm{nullity}(T) = \dim(\mathrm{R}(T)) + \dim(\mathrm{Ker}(T)) = 2 + 1 = 3 = \dim(\mathbb{R}^3)$.

(d) For example, $\{(-\frac{1}{3}, \frac{2}{3}, 1)^{\mathrm{T}}, (0,1,0)^{\mathrm{T}}, (0,0,1)^{\mathrm{T}}\}$ is a basis for $\mathbb{R}^3$.

12. $\begin{pmatrix} 3 & -1 & 2 & 7 \\ 1 & 2 & 3 & 0 \\ 0 & 1 & 1 & -1 \end{pmatrix} \xrightarrow[\text{Elimination}]{\text{Gauss-Jordan}} \begin{pmatrix} 1 & 0 & 1 & 2 \\ 0 & 1 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$

(a) $\{(3,1,0)^{\mathrm{T}},\ (-1,2,1)^{\mathrm{T}}\}$ is a basis for $\mathrm{R}(T)$.

(b) $\{(-1,-1,1,0)^{\mathrm{T}},\ (-2,1,0,1)^{\mathrm{T}}\}$ is a basis for $\mathrm{Ker}(T)$.

(c) $\mathrm{rank}(T)+\mathrm{nullity}(T)=\dim(\mathrm{R}(T))+\dim(\mathrm{Ker}(T))=2+2=4=\dim(\mathbb{R}^4)$.

13. (a) 2. (b) 2. (c) 2.

14. (a) $\{\mathbf{0}\}$. (b) $\mathbb{R}^n$.

15. (a) Let $\{\boldsymbol{v_1},\ \boldsymbol{v_2},\ \ldots,\ \boldsymbol{v_k}\}$ be an orthonormal basis for V. By Theorem 5.2.15,

$$\begin{aligned} P(\boldsymbol{u}) &= (\boldsymbol{u}\cdot\boldsymbol{v_1})\boldsymbol{v_1}+(\boldsymbol{u}\cdot\boldsymbol{v_2})\boldsymbol{v_2}+\cdots+(\boldsymbol{u}\cdot\boldsymbol{v_k})\boldsymbol{v_k} \\ &= \boldsymbol{v_1}\boldsymbol{v_1}^{\mathrm{T}}\boldsymbol{u}+\boldsymbol{v_2}\boldsymbol{v_2}^{\mathrm{T}}\boldsymbol{u}+\cdots+\boldsymbol{v_k}\boldsymbol{v_k}^{\mathrm{T}}\boldsymbol{u} \\ &= (\boldsymbol{v_1}\boldsymbol{v_1}^{\mathrm{T}}+\boldsymbol{v_2}\boldsymbol{v_2}^{\mathrm{T}}+\cdots+\boldsymbol{v_k}\boldsymbol{v_k}^{\mathrm{T}})\boldsymbol{u} \end{aligned}$$

Note that $\boldsymbol{v_1}\boldsymbol{v_1}^{\mathrm{T}}+\boldsymbol{v_2}\boldsymbol{v_2}^{\mathrm{T}}+\cdots+\boldsymbol{v_k}\boldsymbol{v_k}^{\mathrm{T}}$ is an $n\times n$ matrix. So P is a linear transformation.

(b) $\mathrm{Ker}(P)=\mathrm{span}\{(a,b,c)\}$ and $\mathrm{R}(P)=V$.

16. ($\Rightarrow$) Let $\boldsymbol{u},\boldsymbol{v}\in\mathbb{R}^n$ such that $T(\boldsymbol{u})=T(\boldsymbol{v})$. Then $T(\boldsymbol{u}-\boldsymbol{v})=T(\boldsymbol{u})-T(\boldsymbol{v})=\mathbf{0}$ and hence $\boldsymbol{u}-\boldsymbol{v}\in\mathrm{Ker}(T)$. Since $\mathrm{Ker}(T)=\{\mathbf{0}\}$, $\boldsymbol{u}-\boldsymbol{v}=\mathbf{0}$, i.e. $\boldsymbol{u}=\boldsymbol{v}$. Thus T is one-to-one.

($\Leftarrow$) By Theorem 7.1.4.1, $T(\mathbf{0})=\mathbf{0}$. Since T is one-to-one, for all $\boldsymbol{v}\in\mathbb{R}^n$, if $\boldsymbol{v}\neq\mathbf{0}$, $T(\boldsymbol{v})\neq T(\mathbf{0})=\mathbf{0}$. Thus $\mathrm{Ker}(T)=\{\mathbf{0}\}$.

17. (a) Let $\boldsymbol{u}\in\mathrm{Ker}(S)$, i.e. $S(\boldsymbol{u})=\mathbf{0}$. Then $T\circ S(\boldsymbol{u})=T(S(\boldsymbol{u}))=T(\mathbf{0})=\mathbf{0}$ and hence $\boldsymbol{u}\in\mathrm{Ker}(T\circ S)$.
Thus $\mathrm{Ker}(S)\subseteq\mathrm{Ker}(T\circ S)$.

(b) Let $\boldsymbol{v}\in\mathrm{R}(T\circ S)$, i.e. there exists $\boldsymbol{u}\in\mathbb{R}^n$ such that $\boldsymbol{v}=T\circ S(\boldsymbol{u})$. Put $\boldsymbol{w}=S(\boldsymbol{u})\in\mathbb{R}^m$. Then $\boldsymbol{v}=T(S(\boldsymbol{u}))=T(\boldsymbol{w})$. This means that $\boldsymbol{v}\in\mathrm{R}(T)$. Thus $\mathrm{R}(T\circ S)\subseteq\mathrm{R}(T)$.