Solution for Assignment 3

1. **a)** By definition, T is a linear transformation if and only if $T(\mathbf{x}+\mathbf{y}) = T(\mathbf{x}) + T(\mathbf{y})$ and $T(\lambda\mathbf{x}) = \lambda T(\mathbf{x})$ holds for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ and $\lambda \in \mathbb{R}$. Hence, we need to verify this. Let $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ and $\lambda \in \mathbb{R}$ be arbitrary. Indeed, we have

$$T(\mathbf{x}+\mathbf{y}) = \sum_{k=1}^{n} k(x_k + y_k) = \sum_{k=1}^{n}(kx_k + ky_k) = \sum_{k=1}^{n} kx_k + \sum_{k=1}^{n} ky_k = T(\mathbf{x}) + T(\mathbf{y})$$

and

$$T(\lambda\mathbf{x}) = \sum_{k=1}^{n} k(\lambda x_k) = \lambda \sum_{k=1}^{n} kx_k = \lambda T(\mathbf{x})$$

which proves that T is a linear transformation.

b) No, T is not a linear transformation. Consider the standard unit vector $\mathbf{e}_n \in \mathbb{R}^n$ and choose $\lambda = 2$. By definition of T, we get

$$T(\lambda\mathbf{e}_n) = \sum_{k=1}^{n-1} \lambda^k 0 + \lambda^n = \lambda^n$$

and also

$$T(\mathbf{e}_n) = \sum_{k=1}^{n-1} 0 + 1 = 1.$$

Hence, we have $T(\lambda\mathbf{e}_n) = \lambda^n = 2^n \neq 2 = \lambda T(\mathbf{e}_n)$ which means that T is not a linear transformation. Note that the assumption $n \geq 2$ is crucial for this last step to work.

2. We argue both directions separately.

($\Rightarrow$) Assume that T is linear, and observe that, by definition of T, we have

$$T(\mathbf{0}) = \begin{bmatrix} | & | & & | & | \\ \mathbf{v}_1 & \mathbf{v}_2 & \dots & \mathbf{v}_n & \mathbf{v}_{n+1} \\ | & | & & | & | \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} = \mathbf{v}_{n+1}.$$

By Lemma 2.28 we also have $T(\mathbf{0}) = \mathbf{0}$ which implies $\mathbf{v}_{n+1} = T(\mathbf{0}) = \mathbf{0}$, as desired.

($\Leftarrow$) Assume now that $\mathbf{v}_{n+1} = \mathbf{0}$. Observe that we have

$$T(\mathbf{x}) = \begin{bmatrix} | & | & & | & | \\ \mathbf{v}_1 & \mathbf{v}_2 & \dots & \mathbf{v}_n & \mathbf{v}_{n+1} \\ | & | & & | & | \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \\ 1 \end{bmatrix} = \begin{bmatrix} | & | & & | \\ \mathbf{v}_1 & \mathbf{v}_2 & \dots & \mathbf{v}_n \\ | & | & & | \end{bmatrix} \mathbf{x} + \mathbf{v}_{n+1}$$

for all $\mathbf{x} \in \mathbb{R}^n$. Since $\mathbf{v}_{n+1} = \mathbf{0}$, this means that $T(\mathbf{x}) = B\mathbf{x}$ with

$$B = \begin{bmatrix} | & | & & | \\ \mathbf{v}_1 & \mathbf{v}_2 & \dots & \mathbf{v}_n \\ | & | & & | \end{bmatrix}$$

for all $\mathbf{x} \in \mathbb{R}^n$. From Observation 2.26, we conclude T is a linear transformation.

3. **a)** We have to find some angle $\phi \in \mathbb{R}$ such that $A = Q(\phi)$. Since $\cos(\phi)$ should be zero, we have two candidates $\phi = \frac{1}{2}\pi$ and $\phi = \frac{3}{2}\pi$ if we restrict ourselves to $\phi \in [0, 2\pi)$. But we also need $\sin(\phi) = 1$ which is only true for $\phi = \frac{1}{2}\pi$. Indeed, for $\phi = \frac{1}{2}\pi$ we have

$$Q(\phi) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} = A.$$

b) We use the definition of the matrix product and the following addition theorems of the trigonometric functions

$$\sin(\alpha + \beta) = \sin\alpha\cos\beta + \cos\alpha\sin\beta$$
$$\cos(\alpha + \beta) = \cos\alpha\cos\beta - \sin\alpha\sin\beta$$

to get

$$\begin{aligned} Q(\phi_1) \cdot Q(\phi_2) &= \begin{bmatrix} \cos\phi_1 & -\sin\phi_1 \\ \sin\phi_1 & \cos\phi_1 \end{bmatrix} \begin{bmatrix} \cos\phi_2 & -\sin\phi_2 \\ \sin\phi_2 & \cos\phi_2 \end{bmatrix} \\ &= \begin{bmatrix} \cos\phi_1\cos\phi_2 - \sin\phi_1\sin\phi_2 & -\cos\phi_1\sin\phi_2 - \sin\phi_1\cos\phi_2 \\ \cos\phi_1\sin\phi_2 + \sin\phi_1\cos\phi_2 & \cos\phi_1\cos\phi_2 - \sin\phi_1\sin\phi_2 \end{bmatrix} \\ &= \begin{bmatrix} \cos(\phi_1 + \phi_2) & -\sin(\phi_1 + \phi_2) \\ \sin(\phi_1 + \phi_2) & \cos(\phi_1 + \phi_2) \end{bmatrix}. \end{aligned}$$

So the matrix product $Q(\phi_1)Q(\phi_2)$ is a rotation matrix $Q(\phi_1 + \phi_2)$ with the rotation angle $\phi_3 = \phi_1 + \phi_2$.

c) Since A is a rotation matrix, there exists $\phi \in \mathbb{R}$ with $A = Q(\phi)$. Choose B as the rotation matrix $B = Q(-\phi)$. By part b), we have

$$AB = Q(\phi)Q(-\phi) = Q(\phi - \phi) = Q(0) = I = Q(0) = Q(-\phi + \phi) = Q(-\phi)Q(\phi) = BA.$$

4. **a)** Let $\mathbf{a}_1, \ldots, \mathbf{a}_m$ be the rows of A and let $\mathbf{b}_1, \ldots \mathbf{b}_m$ be the columns of B. By the definition of matrix multiplication, we have

$$AB = \begin{bmatrix} — & \mathbf{a}_1 & — \\ — & \mathbf{a}_2 & — \\ \vdots & \vdots & \vdots \\ — & \mathbf{a}_m & — \end{bmatrix} \begin{bmatrix} | & | & \cdots & | \\ \mathbf{b}_1 & \mathbf{b}_2 & \cdots & \mathbf{b}_m \\ | & | & \cdots & | \end{bmatrix} = \begin{bmatrix} \mathbf{a}_1 \cdot \mathbf{b}_1 & \mathbf{a}_1 \cdot \mathbf{b}_2 & \ldots & \mathbf{a}_1 \cdot \mathbf{b}_m \\ \mathbf{a}_2 \cdot \mathbf{b}_1 & \mathbf{a}_2 \cdot \mathbf{b}_2 & \ldots & \mathbf{a}_2 \cdot \mathbf{b}_m \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{a}_m \cdot \mathbf{b}_1 & \mathbf{a}_m \cdot \mathbf{b}_2 & \ldots & \mathbf{a}_m \cdot \mathbf{b}_m \end{bmatrix}.$$

Notice that by the triangular shape of A, the last $m - i$ entries of $\mathbf{a}_i$ are zero for all $i \in [m]$. Similarly, the first $i - 1$ entries of $\mathbf{b}_i$ are zero for all $i \in [m]$. In particular, for all $i, j \in [m]$ with $i < j$, we get $\mathbf{a}_i \cdot \mathbf{b}_j = 0$. Hence, AB is indeed lower triangular.

b) Let A and B be $m \times m$ upper triangular matrices. Then $A^\top$ and $B^\top$ are lower triangular. By subtask a), this implies that $B^\top A^\top$ is lower triangular and we know $B^\top A^\top = (AB)^\top$. Hence, AB is upper triangular.

5. **a)** Recall that the 2×2 matrix

$$A' = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$$

is a rotation matrix. In particular, it rotates vectors in $\mathbb{R}^2$ by $\frac{\pi}{4}$ in counter-clockwise direction. Observe that the matrix A contains A' as a submatrix, i.e. we can obtain A' by removing the second column and second row from A.

Now consider an arbitrary vector $\mathbf{v} \in \mathbb{R}^3$ that is spanned by $\mathbf{e}_1$ and $\mathbf{e}_3$, i.e. $\mathbf{v} = v_1\mathbf{e}_1 + v_3\mathbf{e}_3$. Then applying A to $\mathbf{v}$ has no effect on the second coordinate, but it will rotate that $\mathbf{v}$ in the plane spanned by $\mathbf{e}_1$ and $\mathbf{e}_3$. A better way to say this, is that applying A to $\mathbf{v}$ corresponds to a rotation around the axis $\mathbf{e}_2$ (the vecor orthogonal to the plane spanned by $\mathbf{e}_1$ and $\mathbf{e}_3$).

Finally, observe that this is still true even if we don't have $v_2 = 0$. Since the second column of A is simply $\mathbf{e}_2$, the second coordinate of $\mathbf{v}$ remains unchanged when applying A to it.

In conclusion, the linear transformation given by matrix A is a rotation by $\frac{\pi}{4}$ around the axis spanned by $\mathbf{e}_2$.

b) Note that we did not specify a direction for the rotation in the exercise. The exercise is still well defined because we want to rotate by π and hence the direction does not matter.

We start by thinking about what such a transformation would do to the standard unit vectors $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$. The vector $\mathbf{e}_1$ should be mapped to $\mathbf{e}_2$, the vector $\mathbf{e}_2$ should be mapped to $\mathbf{e}_1$, and the vector $\mathbf{e}_3$ should be mapped to $-\mathbf{e}_3$. In particular, we want

$$A \begin{bmatrix} | & | & | \\ \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ | & | & | \end{bmatrix} \overset{!}{=} \begin{bmatrix} | & | & | \\ \mathbf{e}_2 & \mathbf{e}_1 & -\mathbf{e}_3 \\ | & | & | \end{bmatrix}.$$

From this, we conclude that A has to be the matrix

$$A = \begin{bmatrix} | & | & | \\ \mathbf{e}_2 & \mathbf{e}_1 & -\mathbf{e}_3 \\ | & | & | \end{bmatrix}.$$

6. **a)** Observe first that

$$T(\begin{bmatrix} 0 \\ 1 \end{bmatrix}) = T(\begin{bmatrix} 1 \\ 1 \end{bmatrix} - \begin{bmatrix} 1 \\ 0 \end{bmatrix}) = \begin{bmatrix} 2 \\ 3 \\ 2 \end{bmatrix} - \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}$$

by linearity of T. Therefore, we get

$$T(\begin{bmatrix} x \\ y \end{bmatrix}) = T(x\begin{bmatrix} 1 \\ 0 \end{bmatrix} + y\begin{bmatrix} 0 \\ 1 \end{bmatrix}) = x\begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} + y\begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} = \begin{bmatrix} x+y \\ x+2y \\ 2x \end{bmatrix}$$

for all $x, y \in \mathbb{R}$.

b) Recall from subtask a) that we have

$$T(\begin{bmatrix} x \\ y \end{bmatrix}) = x\begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} + y\begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}$$

for all $x, y \in \mathbb{R}$. We can write this as a matrix-vector product

$$T(\begin{bmatrix} x \\ y \end{bmatrix}) = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 2 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

and thus conclude that $T = T_A$ where

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 2 & 0 \end{bmatrix}.$$

7. We define $T(\mathbf{x}) := \mathbf{v}\mathbf{x}$ for all $\mathbf{x} \in \mathbb{R}$. By Observation 2.26, T is indeed a linear transformation. It remains to observe that we have

$$L = \{\lambda\mathbf{v} \,:\, \lambda \in \mathbb{R}\} = \{\mathbf{v}\lambda \,:\, \lambda \in \mathbb{R}\} = \{\mathbf{v}\mathbf{x} \,:\, \mathbf{x} \in \mathbb{R}\} = \{T(\mathbf{x}) \,:\, \mathbf{x} \in \mathbb{R}\}.$$

Note that most of this is just symbolic manipulation, i.e. by replacing $\lambda \in \mathbb{R}$ with $\mathbf{x} \in \mathbb{R}$ we just change the name of a variable, nothing interesting happens. Also note that we write $\mathbf{x}$ in bold to emphasize that we think of it as a vector in $\mathbb{R}^1 = \mathbb{R}$, while we think of $\lambda \in \mathbb{R}$ as a scalar (there is no mathematical difference between these two things, the different 'names' just try to hint at how we use the objects).