Bernd Gärtner
Robert Weismantel

Solution for Assignment 2

1. **a)** The matrices look as follows:

$$A_2 = \begin{bmatrix} 2 & 3 \\ 3 & 4 \end{bmatrix}, A_3 = \begin{bmatrix} 2 & 3 & 4 \\ 3 & 4 & 5 \\ 4 & 5 & 6 \end{bmatrix}, \text{ and } A_4 = \begin{bmatrix} 2 & 3 & 4 & 5 \\ 3 & 4 & 5 & 6 \\ 4 & 5 & 6 & 7 \\ 5 & 6 & 7 & 8 \end{bmatrix}.$$

b) The rank of A_m is 2 as we will prove with the following argument.

Let $\mathbf{v}_i$ be the i-th column of A_m, i.e.

$$\mathbf{v}_i = \begin{pmatrix} a_{1i} \\ \vdots \\ a_{mi} \end{pmatrix}$$

for all $i \in [m]$. Observe first that $\mathbf{v}_1 \neq \mathbf{0}$, so the matrix has at least rank 1. Next, observe that there is no $\lambda \in \mathbb{R}$ with $\lambda \mathbf{v}_1 = \mathbf{v}_2$: Indeed, any such λ would have to satisfy

$$2\lambda = \lambda a_{11} = a_{12} = 3$$

by looking at the first coordinates, and

$$3\lambda = \lambda a_{21} = a_{22} = 4$$

by looking at the second coordinates. But these two equations contradict each other: The first one implies $\lambda = \frac{3}{2}$ while the second one implies $\lambda = \frac{4}{3}$. We conclude that $\mathbf{v}_2$ is linearly independent from $\mathbf{v}_1$, and thus A has rank at least 2. If $m = 2$, we are done. Thus, assume now that $m \geq 3$. It remains to prove that every other column of A is dependent. We do this by showing that it is a linear combination of the first two columns. For this, note that by definition of A, we have

$$\mathbf{v}_j - \mathbf{v}_i = \begin{pmatrix} a_{1j} - a_{1i} \\ a_{2j} - a_{2i} \\ \vdots \\ a_{mj} - a_{mi} \end{pmatrix} = \begin{pmatrix} (1+j) - (1+i) \\ (2+j) - (2+i) \\ \vdots \\ (m+j) - (m+i) \end{pmatrix} = \begin{pmatrix} j-i \\ j-i \\ \vdots \\ j-i \end{pmatrix}$$

for all $j, i \in [m]$. Let now $i \in \{3, \ldots, m\}$ be arbitrary. We get

$$\mathbf{v}_i - \mathbf{v}_1 = (i-1)(\mathbf{v}_2 - \mathbf{v}_1)$$

which implies

$$\mathbf{v}_i = (2-i)\mathbf{v}_1 + (i-1)\mathbf{v}_2$$

and thus proves that $\mathbf{v}_i$ is dependent. We conclude that the rank of A is exactly 2.

2. **a)** The (column) rank is defined as the number of independent columns of A. Recall that a column is called independent if it is not a linear combination of previous columns. In the matrix A there is only one independent column, which is the first one with $\lambda_i \neq 0$. All subsequent vectors $\lambda_j \mathbf{v}$ for $j > i$ are scalar multiples of the latter:

$$\lambda_j \mathbf{v} = \left(\frac{\lambda_j}{\lambda_i}\right) \cdot \lambda_i \mathbf{v}$$

Therefore we can conclude that $\operatorname{rank}(A) = 1$.

b) The nullspace of the matrix A is the set

$$\mathbf{N}(A) = \{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} = \mathbf{0}\}$$

We claim that for $\mathbf{d} = \begin{pmatrix}\lambda_1 & \lambda_2 & \cdots & \lambda_n\end{pmatrix}^\top$ we have

$$\mathbf{N}(A) = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{d}^\top \mathbf{x} = 0\}$$

thus showing that $N(A)$ is an hyperplane through the origin in $\mathbb{R}^n$.

$\subseteq$: Let $\mathbf{v} = \begin{pmatrix}v_1 & v_2 & \cdots & v_m\end{pmatrix}^\top$, and consider an arbitrary element $\mathbf{x} \in \mathbf{N}(A)$. Then it satisfies

$$\mathbf{0} = A\mathbf{x} = \begin{bmatrix} \lambda_1 v_1 & \cdots & \lambda_n v_1 \\ \vdots & \ddots & \vdots \\ \lambda_1 v_m & \cdots & \lambda_n v_m \end{bmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} v_1 \cdot \sum_{i=1}^n \lambda_i x_i \\ \vdots \\ v_m \cdot \sum_{i=1}^n \lambda_i x_i \end{pmatrix}$$

Note that since $\mathbf{v} \neq \mathbf{0}$, this implies that

$$0 = \sum_{i=1}^n \lambda_i x_i = \mathbf{d}^\top \mathbf{x}$$

and hence $\mathbf{x} \in \{\mathbf{x} \in \mathbb{R}^n : \mathbf{d}^\top \mathbf{x} = 0\}$.

$\supseteq$: Consider an arbitrary element $\mathbf{x} \in \mathbb{R}^n$ for which we have

$$\mathbf{d}^\top \mathbf{x} = \sum_{i=1}^n \lambda_i x_i = 0$$

Then it follows that

$$A\mathbf{x} = \begin{bmatrix} \lambda_1 v_1 & \cdots & \lambda_n v_1 \\ \vdots & \ddots & \vdots \\ \lambda_1 v_m & \cdots & \lambda_n v_m \end{bmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} v_1 \cdot \sum_{i=1}^n \lambda_i x_i \\ \vdots \\ v_m \cdot \sum_{i=1}^n \lambda_i x_i \end{pmatrix} = \mathbf{0}$$

In particular, we have $\mathbf{x} \in \mathbf{N}(A)$.

3. **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 matrix 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, if

$$A = \begin{bmatrix} | & | & | \\ \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 \\ | & | & | \end{bmatrix}$$

then we want

$$\begin{aligned} \mathbf{a}_1 &= A\mathbf{e}_1 = \mathbf{e}_2 \\ \mathbf{a}_2 &= A\mathbf{e}_2 = \mathbf{e}_1 \\ \mathbf{a}_3 &= A\mathbf{e}_3 = -\mathbf{e}_3 \end{aligned}$$

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

$$A = \begin{bmatrix} | & | & | \\ \mathbf{e}_2 & \mathbf{e}_1 & -\mathbf{e}_3 \\ | & | & | \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix}.$$

4. a) We take a small detour and first consider an arbitrary vector $\mathbf{v}$ that is orthogonal to both $\mathbf{x}$ and $\mathbf{y}$. By definition, this means that we have $\mathbf{v}^\top \mathbf{x} = 0$ and $\mathbf{v}^\top \mathbf{y} = 0$. Now consider the scalar product $\mathbf{v}^\top(\lambda\mathbf{x} + \mu\mathbf{y})$. By applying Observation 1.10, we have

$$\mathbf{v}^\top(\lambda\mathbf{x} + \mu\mathbf{y}) = \lambda\, \mathbf{v}^\top \mathbf{x} + \mu\, \mathbf{v}^\top \mathbf{y} = 0.$$

In other words, any vector $\mathbf{v}$ that is orthogonal to both $\mathbf{x}$ and $\mathbf{y}$ is also orthogonal to $(\lambda\mathbf{x} + \mu\mathbf{y})$. Now consider an arbitrary row $\mathbf{u}_i$ from A. The vector $\mathbf{u}_i$ must be orthogonal to both $\mathbf{x}$ and $\mathbf{y}$ since we have $A\mathbf{x} = \mathbf{0}$ and $A\mathbf{y} = \mathbf{0}$. Hence, $\mathbf{u}_i$ must also be orthogonal to $(\lambda\mathbf{x} + \mu\mathbf{y})$.

b) No, $\mathcal{L}$ is not a finite set. Take any non-zero vector $\mathbf{x} \in \mathcal{L}$ (which must exist since $|\mathcal{L}| \geq 2$). Then every row $\mathbf{u}_i$ of A must be orthogonal to $\mathbf{x}$. Hence, $\mathbf{u}_i$ must also be orthogonal to $\lambda\mathbf{x}$ for any choice of $\lambda \in \mathbb{R}$ by subtask a). In other words, $\lambda\mathbf{x}$ is orthogonal to all rows of A. This implies that $A(\lambda\mathbf{x}) = 0$ and hence $\lambda\mathbf{x} \in \mathcal{L}$. For every choice of λ we obtain a different vector in $\mathcal{L}$. The number of choices of λ is infinite/unbounded, hence $\mathcal{L}$ is not a finite set.

5. a) To determine the rank of A, we have to find out how many of its columns are independent (not a linear combination of previous columns). Let us denote the columns of A by $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$, i.e.

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

By checking the column vectors in order we clearly see that since $\mathbf{v}_1 \neq \mathbf{0}$ the first column vector is independent. Moreover, $\mathbf{v}_2$ can be obtained from $\mathbf{v}_1$ as $\mathbf{v}_2 = -3\mathbf{v}_1$. But if we try to obtain $\mathbf{v}_3$ from $\mathbf{v}_1$ and $\mathbf{v}_2$ we fail. In fact, any linear combination of the two vectors is just a multiple of $\mathbf{v}_1$, and $\mathbf{v}_3$ is not a multiple of $\mathbf{v}_1$, since there is no $\lambda \in \mathbb{R}$ such that both $1\lambda = 3$ and $-2\lambda = 0$. In other words, the first and the third columns are independent, but the second is not. Hence, the rank of A is 2.

b) We proceed as in subtask a) and check independence of the columns of A. Let us again denote the columns of A by $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$, i.e.

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

Clearly the first column vector is independent. Then, $\mathbf{v}_2$ is also independent as there is no way of obtaining the 1 in the first coordinate of $\mathbf{v}_2$ if we only use $\mathbf{v}_1$. It remains to check whether $\mathbf{v}_3$ can be obtained as linear combination of $\mathbf{v}_2$ and $\mathbf{v}_1$. Here we observe that there is no way of obtaining the 2 in the second coordinate of $\mathbf{v}_3$ since the second entry of both $\mathbf{v}_1$ and $\mathbf{v}_2$ is 0. We conclude that the three column vectors are independent and hence A has rank 3.

6. **a)** An example is given by

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

In fact, all nonzero skew-symmetric matrices are of the form

$$\begin{bmatrix} 0 & -x \\ x & 0 \end{bmatrix}$$

for some $x \in \mathbb{R}\backslash\{0\}$ when $m = 2$.

b) From $A^\top = -A$, we obtain that $a_{ii} = -a_{ii}$ for $i \in [m]$. This implies $2a_{ii} = 0$. After dividing by two, we are left with $a_{ii} = 0$.

c) From the definition of symmetric and skew-symmetric we get that

$$A^\top = A \quad \text{and} \quad A^\top = -A$$

which implies that $A = -A$, or equivalently $2A = 0$ The only matrix satisfying this equality is only the zero matrix 0.

d) We show that there exists a nonzero vector $\mathbf{x} \in \mathbb{R}^3$ with $A\mathbf{x} = \mathbf{0}$. This implies that Observation 2.5 (ii) is not satisfied. Therefore, the columns of A are not independent and $\text{rank}(A) \leq 2$. From Part c), we know that $a_{11} = 0, a_{22} = 0$, and $a_{33} = 0$. Using that $A^\top = -A$ on the other entries, we have

$$A = \begin{bmatrix} 0 & a_{12} & a_{13} \\ -a_{12} & 0 & a_{23} \\ -a_{13} & -a_{23} & 0 \end{bmatrix}.$$

Consider the vector

$$\mathbf{x} = \begin{pmatrix} a_{23} \\ -a_{13} \\ a_{12} \end{pmatrix}.$$

If $\mathbf{x} = \mathbf{0}$, we are done since A is then a matrix with all zero entries and thus $\text{rank}(A) = 0$. So suppose that $\mathbf{x} \neq \mathbf{0}$. Then we have

$$A\mathbf{x} = \begin{pmatrix} 0 \cdot a_{23} + a_{12} \cdot (-a_{13}) + a_{13} \cdot a_{12} \\ (-a_{12}) \cdot a_{23} + 0 \cdot (-a_{13}) + a_{23} \cdot a_{12} \\ (-a_{13}) \cdot a_{23} + (-a_{23}) \cdot (-a_{13}) + 0 \cdot a_{12} \end{pmatrix} = \begin{pmatrix} -(a_{12} \cdot a_{13}) + a_{12} \cdot a_{13} \\ -(a_{12} \cdot a_{23}) + a_{12} \cdot a_{23} \\ a_{13} \cdot a_{23} - a_{13} \cdot a_{23} \end{pmatrix} = \mathbf{0}.$$

7. Let $\mathbf{v} = \begin{pmatrix} v_1 & \cdots & v_m \end{pmatrix}^\top \in \mathbb{R}^m$, then we define $T_A(\mathbf{x}) \coloneqq A\mathbf{x}$ with

$$A = \begin{bmatrix} v_1 \\ \vdots \\ v_m \end{bmatrix} \in \mathbb{R}^{m \times 1}$$

It remains to observe that we have

$$L = \{\lambda \mathbf{v} \,:\, \lambda \in \mathbb{R}\} = \{\mathbf{v}\lambda \,:\, \lambda \in \mathbb{R}\} = \{A\mathbf{x} \,:\, \mathbf{x} \in \mathbb{R}^1\} = \{T_A(\mathbf{x}) \,:\, \mathbf{x} \in \mathbb{R}^1\}.$$

Note that most of this is just symbolic manipulation, i.e. replacing $\lambda \in \mathbb{R}$ with $\mathbf{x} \in \mathbb{R}^1$ and $\mathbf{v} \in \mathbb{R}^m$ with $A \in \mathbb{R}^{m \times 1}$. Here is important to note that even though the three sets $\mathbb{R}$, $\mathbb{R}^1$, $\mathbb{R}^{1 \times 1}$ are structurally "the same" (as we will see later in the course), we don't consider them as equal. An element $\alpha \in \mathbb{R}$ is a scalar, $\begin{pmatrix} \alpha \end{pmatrix} \in \mathbb{R}^1$ is a vector and $\begin{bmatrix} \alpha \end{bmatrix} \in \mathbb{R}^{1 \times 1}$ is a matrix.