Solution for Assignment 2

1. The rank of A is 2 as we will prove with the following argument.

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

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

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 on the first two. For this, note that by definition of A, we have

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

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)** We can use our decomposition of A. In particular, we want $A\mathbf{x} = (\mathbf{v}\mathbf{w}^\top)\mathbf{x} = \mathbf{v}(\mathbf{w}^\top \mathbf{x}) \overset{!}{=} \mathbf{0}$. In particular, it suffices to find a vector $\mathbf{x}$ with $\mathbf{w}^\top \mathbf{x} = \begin{bmatrix} w_1 & w_2 & w_3 \end{bmatrix} \mathbf{x} = w_1 x_1 + w_2 x_2 + w_3 x_3 = 0$. There are now many possibilities for $\mathbf{x}$. One of them is $x_1 = w_2$, $x_2 = -w_1$, and $x_3 = 0$. This is a non-zero vector because we have $w_1 \neq 0$.

b) Consider the vector $\mathbf{w} = \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix}$. We claim that $\mathcal{L}$ is exactly the set of vectors that are orthogonal to $\mathbf{w}$. In other words, we claim $\mathcal{L} = \{\mathbf{x} \in \mathbb{R}^3 : \mathbf{x} \cdot \mathbf{w} = 0\}$. Note that $\{\mathbf{x} \in \mathbb{R}^3 : \mathbf{x} \cdot \mathbf{w} = 0\}$ is a hyperplane since $\mathbf{w}$ is non-zero.

In order to prove our claim, let us first consider an arbitrary vector $\mathbf{x} \in \{\mathbf{x} \in \mathbb{R}^3 : \mathbf{x} \cdot \mathbf{w} = 0\}$. Because we have $\mathbf{x} \cdot \mathbf{w} = 0$, we also get

$$A\mathbf{x} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \begin{bmatrix} w_1 & w_2 & w_3 \end{bmatrix} \mathbf{x} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} 0 = \mathbf{0}.$$

Hence, we get $\mathbf{x} \in \mathcal{L}$.

For the reverse direction, we proceed with an indirect proof: Consider an arbitrary vector $\mathbf{x} \in \mathbb{R}^3$ that is not orthogonal to $\mathbf{w}$, i.e. $\mathbf{x} \cdot \mathbf{w} = c \neq 0$ for some $c \in \mathbb{R}$. Then we have

$$A\mathbf{x} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \begin{bmatrix} w_1 & w_2 & w_3 \end{bmatrix} \mathbf{x} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} c = \begin{bmatrix} cv_1 \\ cv_2 \\ cv_3 \end{bmatrix} \neq \mathbf{0}$$ since both $v_1 \neq 0$ and $c \neq 0$. In

particular, $\mathbf{x}$ is not in $\mathcal{L}$. This concludes the proof.

3. **a)** We start by computing the powers of A. We get $A^2 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 3 & 1 \\ 1 & 0 & 3 \end{bmatrix}$ and $A^3 = \begin{bmatrix} 1 & 0 & 3 \\ 3 & 1 & 9 \\ 0 & 3 & 1 \end{bmatrix}$.

Plugging this into the equation, we get

$$B := A^3 + xA^2 + yA + zI = \begin{bmatrix} z+1 & x & 3+y \\ 3+y & 1+3x+z & 9+x+3y \\ x & 3+y & 1+3x+z \end{bmatrix} \overset{!}{=} 0.$$

In particular, we want to choose x, y, z such that all entries will become zero. From $b_{12} = x$ we can deduce $x = 0$, from $b_{13} = 3 + y$ we can deduce $y = -3$, and from $b_{11} = z + 1$ we get $z = -1$. In fact, with this choice of x, y, z all entries of B become 0 and hence we found the unique solution.

b) Note that one could prove this in a very formal way with induction. We opted for a slightly less but still sufficiently formal proof. Let $k \in \mathbb{N}$ be arbitrary. If $k = 0$, then we have $(AB)^k = I = I^2 = A^k B^k$. Thus, assume now $k > 0$ and consider the expression

$$(AB)^k = \underbrace{AB \times AB \times \cdots \times AB}_{k \text{ times}} = \underbrace{ABAB \cdots AB}_{k \text{ repetitions of } AB}.$$

Consider now the following algorithm: While there is an appearance of BA in the above expression, replace it by AB. Note that this operation does not change the product because matrix multiplication is associative and we are additionally given $AB = BA$ in this exercise. Moreover, the algorithm must eventually terminate as the number of ways to arrange k A's and k B's in a string is finite and the algorithm will never consider the same arrangement twice (to see this, note that the sum of indices of the A's decreases in every step of the algorithm). Once the algorithm terminates, all A's must be to the left of the B's (as otherwise we could find an appearance of BA). Hence, we have $(AB)^k = A^k B^k$.

c) According to the previous question, we have $(AB)^k = A^k B^k$. We compute $(AB)^k = A^k B^k = A^k 0 = 0$. Therefore, AB is nilpotent with degree at most k.

Remark: The nilpotent degree is not necessarily equal to k. If A is nilpotent of degree 5 and we choose $B = A^2$, then B is nilpotent of degree 3. But $AB = A^3$ is nilpotent of degree 2.

d) Using distributivity of matrix multiplication, we compute:

$$\begin{aligned} (I - A)(I + A + \ldots + A^{k-1}) &= (I + A + \ldots + A^{k-1}) - A(I + A + \ldots + A^{k-1}) \\ &= I + A + \ldots + A^{k-1} - A - A^2 - \ldots - A^k \\ &= I - A^k \\ &= I \qquad\qquad (A \text{ is nilpotent of degree } k) \end{aligned}$$

2

e) As is often the case, there are many ways to prove this. Here, we will argue by induction over k that the first k columns in T^k must be zero for all $k \in \mathbb{N}$.

- **Property**: The k first columns of T^k are all zero.
- **Base case**: For $k = 1$, the property is true because by our assumptions on T, the first column of T must be zero.
- **Induction step**: Fix a natural number $1 \leq k < m$ and assume that the property is true for this k (induction hypothesis). We prove that the property is true for $k + 1$. In other words, we prove that the first $(k + 1)$ columns of T^{k+1} are all zero. For this, we start by splitting T^{k+1} into $T^{k+1} = T^k T$. By our induction hypothesis, the first k columns of T^k are zero. In particular, we have

$$T^k = \begin{bmatrix} | & | & & | & | & | & & | \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_{m-k} \\ | & | & & | & | & | & & | \end{bmatrix}$$

for some vectors $\mathbf{v}_1, \ldots, \mathbf{v}_{m-k} \in \mathbb{R}^n$. Now recall the column picture

$$AB = \begin{bmatrix} | & | & & | \\ A\mathbf{b}_1 & A\mathbf{b}_2 & \cdots & A\mathbf{b}_m \\ | & | & & | \end{bmatrix}$$

for the product of two $m \times m$ matrices A and B where $\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_m$ are the columns of B. Consider splitting T^{k+1} as

$$T^{k+1} = T^k T = \begin{bmatrix} | & | & & | \\ T^k \mathbf{t}_1 & T^k \mathbf{t}_2 & \cdots & T^k \mathbf{t}_m \\ | & | & & | \end{bmatrix}$$

where $\mathbf{t}_1, \mathbf{t}_2, \cdots, \mathbf{t}_m$ are the columns of T. We want to argue that the first $(k + 1)$ columns $T^k \mathbf{t}_1, T^k \mathbf{t}_2, \ldots, T^k \mathbf{t}_{k+1}$ of this matrix are zero. Let $i \in \{1, \ldots k + 1\}$. By definition of T, only the first $i - 1$ entries of column $\mathbf{t}_i$ can be non-zero. Moreover, we know that $T^k \mathbf{t}_i$ is a linear combination of the columns of T^k where the coefficients are the entries of vector $\mathbf{t}_i$. We just observed that at most the first $i - 1$ coefficients in this linear combination can be non-zero. At the same time, we know that the first k columns of T^k are all zero. Hence, we conclude $T^k \mathbf{t}_i = \mathbf{0}$.

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

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

In other words, any vector $\mathbf{v}$ that is perpendicular to both $\mathbf{x}$ and $\mathbf{y}$ is also perpendicular to $(\lambda \mathbf{x} + \mu \mathbf{y})$. Now consider an arbitrary row $\mathbf{u}_i$ from A. The vector $\mathbf{u}_i$ must be perpendicular 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 perpendicular 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 perpendicular to $\mathbf{x}$. Hence, $\mathbf{u}_i$ must also be perpendicular to $\lambda \mathbf{x}$ for any choice of $\lambda \in \mathbb{R}$ by subtask a). In other words, $\lambda \mathbf{x}$ is perpendicular 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) In order to determine the rank of A, we have to find out how many of its columns are linearly independent. 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 quickly see that $\mathbf{v}_2$ 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 $\mathbf{v}_1$ we fail, since there is no $\lambda \in \mathbb{R}$ such that both $1\lambda = 3$ and $-2\lambda = 0$. In other words, the set of vectors $\{\mathbf{v}_1, \mathbf{v}_3\}$ is linearly independent. We found a set of two vectors that is linearly independent and also observed that the set of all three vectors is not linearly independent. Therefore, the rank of A is 2.

b) We proceed as in subtask a) and check linear 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.

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

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

6. a) There are many possible solutions here. For example, we could choose

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, B = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$$

with

$$(AB)^\top = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}^\top = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$$

and

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

b) Yes, it is possible to find examples where both A and B are symmetric. For example, consider the matrices

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

We have

$$(AB)^\top = B^\top A^\top = BA = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 2 \\ 0 & 0 \end{bmatrix}$$

but also

$$A^\top B^\top = AB = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 2 & 0 \end{bmatrix}.$$

Note that any two symmetric 2×2 matrices that do not commute (i.e. $AB \neq BA$) work as an example here.

Multiple choice Let A be an $m_1 \times n_1$ matrix and let B be an $m_2 \times n_2$ matrix for natural numbers m_1, n_1, m_2, n_2. For each statement, determine whether it is true or not (regardless of what values m_1, n_1, m_2, n_2 take).

1. If A^2 is defined, then A must be square.

$\checkmark$ **(a)** Yes

(b) No

Explanation: For matrix multiplication to work, the number of columns of the first matrix has to equal the number of rows of the second matrix. In this case, both the first and second matrix are A which implies that A must be square.

2. If $A^2 = I$, then $A = I$.

(a) Yes

$\checkmark$ **(b)** No

Explanation: A possible counterexample is $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$.

3. If $A^3 = 0$, then $A = 0$.

(a) Yes

$\checkmark$ **(b)** No

Explanation: A possible counterexample is $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$.

4. If $A = \begin{bmatrix} 1 & a \\ 0 & 1 \end{bmatrix}$, then $A^n = \begin{bmatrix} 1 & na \\ 0 & 1 \end{bmatrix}$ for all $n \in \mathbb{N}$.

$\checkmark$ **(a)** Yes

(b) No

Explanation: A full proof could be obtained via induction. We will provide an intuitive explanation. Consider a 2×2 matrix B and consider what happens when we multiply A with B. The first row of AB will be the first row of B, plus a times the second row of B. The second row of AB will be copied from the second row of B. Now if the second row of B is $\begin{bmatrix} 0 & 1 \end{bmatrix}$, this means that left-multiplying with A corresponds to adding a to the entry in the first row and second column.

5. If $AB = B$ for some choice of B, then $A = I$.

(a) Yes

$\checkmark$ **(b)** No

Explanation: If $B = 0$ and $A \neq I$, we still have $AB = B$.

6. If both products AB and BA are defined, then A and B must be square.

(**a**) Yes

✓ (**b**) No

Explanation: A could be a 3×2 matrix while B is a 2×3 matrix.

7. If both products AB and BA are defined, then AB and BA must be square.

✓ (**a**) Yes

(**b**) No

Explanation: If both products are defined, we must have $n_1 = m_2$ and $n_2 = m_1$. Observe that AB is an $m_1 \times n_2$ matrix while BA is an $m_2 \times n_1$ matrix. Hence, both must be square.

8. If two columns of A are equal and AB is defined, the corresponding columns of AB must also be equal.

(**a**) Yes

✓ (**b**) No

Explanation: A counterexample would be $A = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}$.

9. If two columns of B are equal and AB is defined, the corresponding columns of AB must also be equal.

✓ (**a**) Yes

(**b**) No

Explanation: This can be seen from the column picture of matrix multiplication (Figure 1.9 in lecture notes).

10. If two rows of A are equal and AB is defined, the corresponding rows of AB must also be equal.

✓ (**a**) Yes

(**b**) No

Explanation: This can be seen from the row picture of matrix multiplication (Figure 1.8 in lecture notes).

11. If two rows of B are equal and AB is defined, the corresponding rows of AB must also be equal.

   (a) Yes

✓   **(b)** No

   Explanation: A counterexample would be $A = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}$.

12. If A and B are symmetric matrices and AB is defined, AB is also symmetric.

   (a) Yes

✓   **(b)** No

   Explanation: A counterexample would be $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$.