Exercise 1. (a) The list \( a = \left((2, -4, 2)^T, (-3, 6, 3)^T, (1, -2, 1)^T, (3, 1, 3)^T, (2, 1, 0)^T\right) \) spans \( \mathbb{R}^3 \). Shrink this list to a basis of \( \mathbb{R}^3 \) by removing some redundant elements. [5 points]

(b) The list \( b = \left((1, 0, 1)^T\right) \) is linearly independent. Extend this list to a basis of \( \mathbb{R}^3 \) by appending to it some elements from the list \( a \). [5 points]

Solution. The method for this exercise is the same as for Exercise 1 on the midterm #2 practice sheet; thus, I’ll be briefer this time.

(a) We proceed using the standard algorithm:\footnote{In this algorithm, we treat \( a \) as a mutable variable.} We scan the list \( a \) from left to right. Each time we read an entry of \( a \), we check if this entry is a linear combination of the entries before it. If it is, then we remove this entry from \( a \) and start from scratch with the new (shorter) \( a \). If it is not, then we proceed to the next entry. If we have arrived at the end of the list, then our list has no redundant entries, and thus is a basis of \( \mathbb{R}^3 \).

Let us execute this algorithm step by step:

- We scan the list \( a \) from left to right. Thus, we begin at its first entry, which is \((2, -4, 2)^T\).

- Is this first entry \((2, -4, 2)^T\) a linear combination of the entries before it? There are no entries before it, and thus the only linear combination of the entries before it is \(0^T\). Our first entry \((2, -4, 2)^T\) is not \(0^T\). Thus, the answer is “No”. We proceed to the second entry.

- Is this second entry \((-3, 6, 3)^T\) a linear combination of the entries before it? There is only one entry before it, namely \((2, -4, 2)^T\). Hence, we are asking whether \((-3, 6, 3)^T\) is a linear combination of the vector \((2, -4, 2)^T\). Using Gaussian elimination (or just common sense\footnote{Namely, observe that every linear combination of the vector \((2, -4, 2)^T\) has its first entry equal to its third entry, but the vector \((-3, 6, 3)^T\) does not have this property.}), we see that it is not. Thus, the answer is “No”. Hence, we proceed to the third entry.

- Is this third entry \((1, -2, 1)^T\) a linear combination of the entries before it? The entries before it are \((2, -4, 2)^T\) and \((-3, 6, 3)^T\). Hence, we are asking whether \((1, -2, 1)^T\) is a linear combination of the vectors \((2, -4, 2)^T\) and \((-3, 6, 3)^T\). Once again, we can use Gaussian elimination to arrive at the answer, which is “Yes” this time\footnote{And, again, we can tell this immediately without Gaussian elimination as well, by observing that \((1, -2, 1)^T = \frac{1}{2} (2, -4, 2)^T\).}. Thus, we remove the entry from \( a \), and start from scratch with the new (shorter) \( a \).
We scan the new list $a = (2, -4, 2)^T, (-3, 6, 3)^T, (3, 1, 3)^T, (2, 1, 0)^T$ (the result of removing $(1, -2, 1)^T$ from the old list $a$) from left to right. Thus, we begin at its first entry, which is $(2, -4, 2)^T$.

Is this first entry $(1, 2, -1)^T$ a linear combination of the entries before it? Once again, this is a question we have already answered the last time we encountered this entry; the answer is “no”, and so we proceed to the second entry.

Is this second entry $(1, 1, 0)^T$ a linear combination of the entries before it? Again, we already have answered this question, and the answer is “no”. We proceed to the third entry.

Is this third entry $(3, 1, 3)^T$ a linear combination of entries before it? The entries before it are $(2, -4, 2)^T$ and $(-3, 6, 3)^T$. Hence, we are asking whether $(3, 1, 3)^T$ is a linear combination of the vectors $(2, -4, 2)^T$ and $(-3, 6, 3)^T$. As usual, we can answer this using Gaussian elimination; the answer is “no”. We thus proceed to the fourth entry.

Is this fourth entry $(2, 1, 0)^T$ a linear combination of entries before it? Again, we can use Gaussian elimination to answer this; but there is also a more obvious reason why the answer is “Yes”: Namely, let us once again take a look at the first three entries of $a$. There are no redundant entries among these (because any redundant entries would have already been removed in the previous steps); thus, they are linearly independent. But any 3 linearly independent vectors in $\mathbb{R}^3$ must form a basis of $\mathbb{R}^3$. Hence, the first three entries of $a$ form a basis of $\mathbb{R}^3$. Therefore, the fourth entry $(2, 1, 0)^T$ must (like any vector in $\mathbb{R}^3$) be a linear combination of these first three entries. Thus, we remove the entry from $a$, and start from scratch with the new (shorter) $a$.

We scan the new list $a = (2, -4, 2)^T, (-3, 6, 3)^T, (3, 1, 3)^T$ from left to right. Thus, we ask again whether the first entry is a linear combination of the entries before it, and then the same question for the second and the third entries. All of these questions have already been answered with a “no”, and so we arrive at the end of the list.

We have thus ended up with the list $\left( (2, -4, 2)^T, (-3, 6, 3)^T, (3, 1, 3)^T \right)$. This list is therefore a basis of $\mathbb{R}^3$ obtained by shrinking our (old) list $a$.

(b) We solve this using the following algorithm

We scan the list $a$ from left to right. Each time we read an entry of $a$, we check if this entry is a linear combination of the (current) entries of $b$. If it isn’t, then we append this entry to $b$. In either case, we proceed to the next entry. By the time we have scanned all entries of $a$, the list $b$ has become a basis of $\mathbb{R}^3$.

In this algorithm, we treat $b$ as a mutable variable.
In order to simplify our life, we use not the original list
\[ a = \left( (2, -4, 2)^T, (-3, 6, 3)^T, (1, -2, 1)^T, (3, 1, 3)^T, (2, 1, 0)^T \right), \]
but the shorter list
\[ a = \left( (2, -4, 2)^T, (-3, 6, 3)^T, (3, 1, 3)^T \right) \]
obtained at the end of the shrinking process in part (a) of the problem. Indeed, this shorter list works just as well (it is a basis of \( \mathbb{R}^3 \) and thus spans \( \mathbb{R}^3 \)), and clearly its elements are elements of the original list \( a \) as well.

Let us now execute our algorithm step by step:

- We scan the list \( a \) from left to right. Thus, we begin at its first entry, which is \( (2, -4, 2)^T \).
- Is this first entry \( (2, -4, 2)^T \) a linear combination of the entries of \( b \)? The entries of \( b \) are \( (1, 0, 1)^T \). Hence, we are asking whether \( (2, -4, 2)^T \) is a linear combination of the vector \( (1, 0, 1)^T \). This can be answered by Gaussian elimination\(^5\). The answer is “no”. Thus, we append this entry \( (2, -4, 2)^T \) to \( b \), so that \( b \) becomes \( \left( (1, 0, 1)^T, (2, -4, 2)^T \right) \). We now proceed to the second entry of \( a \).
- Is this second entry \( (-3, 6, 3)^T \) a linear combination of the entries of \( b \)? The entries of \( b \) are \( (1, 0, 1)^T \) and \( (2, -4, 2)^T \) (keep in mind that \( b \) has changed in the previous step!). We can answer this using Gaussian elimination; the answer is “no”. Hence, we append this entry \( (-3, 6, 3)^T \) to \( b \), so that \( b \) becomes \( \left( (1, 0, 1)^T, (2, -4, 2)^T, (-3, 6, 3)^T \right) \). We now proceed to the third entry of \( a \).
- Is this third entry \( (3, 1, 3)^T \) a linear combination of the entries of \( b \)? The answer is “yes”; this can be checked either by Gaussian elimination or by observing that the list \( b \) has three linearly independent elements and thus spans \( \mathbb{R}^3 \) (we have done something very similar in part (a) above). Thus, the list \( b \) does not change at this step. We have now arrived at the end of the list \( a \).

We have thus ended up with the list \( b = \left( (1, 0, 1)^T, (2, -4, 2)^T, (-3, 6, 3)^T \right) \). This list is therefore a basis of \( \mathbb{R}^3 \) obtained by appending some elements from \( a \) to the (old) list \( b \). \( \square \)

\(^5\) or by a quick glance at the second entry, which has to be 0 for any linear combination of \( (1, 0, 1)^T \) but fails to be 0 for \( (2, -4, 2)^T \)
Exercise 2. (a) Find bases of the four subspaces of the $3 \times 3$-matrix $A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 5 \end{pmatrix}$. 

(b) Find a basis of the column space of the $4 \times 4$-matrix $B = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 5 \\ 3 & 4 & 5 & 6 \\ 4 & 5 & 6 & 7 \end{pmatrix}$.

Solution. (a) Column space: A basis of $\text{Col } A$ is $\left( (1,2,3)^T, (2,3,4)^T \right)$. 

[Proof: Since $\text{Col } A$ is the span of the columns of $A$, we have $\text{Col } A = \text{span} \left( (1,2,3)^T, (2,3,4)^T, (3,4,5)^T \right)$. But the vector $(3,4,5)^T$ is redundant in this list, since it is a linear combination of the previous vectors (namely, it is $2 \cdot (2,3,4)^T - (1,2,3)^T$). Thus, we can throw it out and obtain $\text{Col } A = \text{span} \left( (1,2,3)^T, (2,3,4)^T \right)$. The vectors $(1,2,3)^T$ and $(2,3,4)^T$ are fairly obviously linearly independent, and so they form a basis of $\text{Col } A$.]

Row space: A basis of $\text{Row } A$ is $\left( (1,2,3)^T, (2,3,4)^T \right)$. 

[Proof: The matrix $A$ is symmetric, i.e., we have $A = A^T$. But recall that $\text{Row } A = \text{Col } (A^T) = (\text{Col } A)^T$. Thus, knowing a basis of $\text{Col } A$, we can immediately obtain a basis of $\text{Row } A$.]

Kernel: A basis of $\text{Ker } A$ is $\left( (1,-2,1)^T \right)$. 

[Proof: We have $\text{rank } A = \text{dim } (\text{Col } A) = 2$ (since we have found a basis of $\text{Col } A$, and this basis has size 2). Now, the rank-nullity theorem yields $\text{rank } A + \text{dim } (\text{Ker } A) = 3$, so that $\text{dim } (\text{Ker } A) = 3 - \text{rank } A = 3 - 2 = 1$. But it is easy to see that $(1,-2,1)^T \in \text{Ker } A$ (indeed, this is a restatement of the linear dependency relation $1 \cdot \text{col}_1 A + (-2) \cdot \text{col}_2 A + 1 \cdot \text{col}_3 A = \vec{0}$, which in turn is equivalent to our old observation that the third column $(3,4,5)^T$ of $A$ is the linear combination $2 \cdot (2,3,4)^T - (1,2,3)^T$ of the previous two columns). Hence, the linearly independent list $\left( (1,-2,1)^T \right)$ consists of 1 element of $\text{Ker } A$. Since $\text{dim } (\text{Ker } A) = 1$, this shows that this list is a basis of $\text{Ker } A$. (Of course, we could have found this using Gaussian elimination as well.)]

Left kernel: A basis of $(\text{Ker } (A^T))^T$ is $\left( (1,-2,1)^T \right)$. 


Proof: Again, recall that \( A = A^T \). Thus, \( \left( \frac{\text{Ker} \left( A^T \right)}{=A} \right)^T = (\text{Ker} \ A)^T \). Thus, knowing a basis of \( \text{Ker} \ A \), we can immediately obtain a basis of \( (\text{Ker} \ (A^T))^T \).

(b) Column space: A basis of \( \text{Col} \ B \) is \( \left( (1,2,3,4)^T, (2,3,4,5)^T \right) \).

Proof: Since \( \text{Col} \ B \) is the span of the columns of \( B \), we have
\[
\text{Col} \ B = \text{span} \left( (1,2,3,4)^T, (2,3,4,5)^T, (3,4,5,6)^T, (4,5,6,7)^T \right).
\]
But the vector \( (4,5,6,7)^T \) is redundant in this list, since it is a linear combination of the previous vectors (namely, it is \( 2 \cdot (3,4,5,6)^T - (2,3,4,5)^T \)). Similarly, the vector \( (3,4,5,6)^T \) is redundant. Thus, we can throw them both out, and obtain \( \text{Col} \ B = \text{span} \left( (1,2,3,4)^T, (2,3,4,5)^T \right) \). The vectors \( (1,2,3,4)^T \) and \( (2,3,4,5)^T \) are fairly obviously linearly independent, and so they form a basis of \( \text{Col} \ B \). \( \Box \)

Exercise 3. (a) Find a basis of the vector space of all upper-triangular \( 3 \times 3 \)-matrices.

(b) A \( 3 \times 3 \)-matrix \[
\begin{pmatrix}
  a_1 & b_1 & c_1 \\
  a_2 & b_2 & c_2 \\
  a_3 & b_3 & c_3
\end{pmatrix}
\]
is said to be zero-sum if it satisfies the equalities
\[
a_1 + b_1 + c_1 = 0, \quad a_2 + b_2 + c_2 = 0, \quad a_3 + b_3 + c_3 = 0, \quad (1)
\]
\[
a_1 + a_2 + a_3 = 0, \quad b_1 + b_2 + b_3 = 0, \quad c_1 + c_2 + c_3 = 0 \quad (2)
\]
(in other words: each row sums to 0, and each column sums to 0).

The zero-sum \( 3 \times 3 \)-matrices form a subspace of \( \mathbb{R}^{3\times3} \). Find a basis of this subspace.

Solution. (a) The list \( (E_{1,1}, E_{1,2}, E_{1,3}, E_{2,2}, E_{2,3}, E_{3,3}) \) is a basis of this space.

The proof is very similar to the proofs in the solution of Exercise 1 on homework set #4; therefore, we leave it to the reader.
(b) Let $Z$ denote the subspace of $\mathbb{R}^{3\times3}$ consisting of the zero-sum $3 \times 3$-matrices.

$$M_1 = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{pmatrix} = E_{1,1} - E_{1,3} - E_{3,1} + E_{3,3}, \quad (3)$$

$$M_2 = \begin{pmatrix} 0 & 1 & -1 \\ 0 & 0 & 0 \\ 0 & -1 & 1 \end{pmatrix} = E_{1,2} - E_{1,3} - E_{3,2} + E_{3,3}, \quad (4)$$

$$M_3 = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & -1 \\ -1 & 0 & 1 \end{pmatrix} = E_{2,1} - E_{2,3} - E_{3,1} + E_{3,3}, \quad (5)$$

$$M_4 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{pmatrix} = E_{2,2} - E_{2,3} - E_{3,2} + E_{3,3}. \quad (6)$$

Then, the list $(M_1, M_2, M_3, M_4)$ is a basis of $Z$.

[Proof: First, it is straightforward to see that all four matrices $M_1, M_2, M_3, M_4$ belong to $Z$. Hence, $\text{span}(M_1, M_2, M_3, M_4) \subseteq Z$ (since $Z$ is a subspace of $\mathbb{R}^{3\times3}$).

Next, we claim that $Z \subseteq \text{span}(M_1, M_2, M_3, M_4)$. Indeed, let $A \in Z$. Then, $A$ is a zero-sum $3 \times 3$-matrix. Write $A$ in the form $A = \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix}$. Then, the equalities (1) and (2) hold (since $A$ is zero-sum). We can use these equalities to write the five entries $c_1, c_2, a_3, b_3, c_3$ in terms of the four entries $a_1, b_1, a_2, b_2$ of our matrix:

$$c_1 = -a_1 - b_1 \quad \text{(by the first equality of (1))};$$

$$c_2 = -a_2 - b_2 \quad \text{(by the second equality of (1))};$$

$$a_3 = -a_1 - a_2 \quad \text{(by the first equality of (2))};$$

$$b_3 = -b_1 - b_2 \quad \text{(by the second equality of (2))};$$

$$c_3 = -c_1 - c_2 \quad \text{(by the third equality of (2))}$$

$$= -a_1 - b_1 - a_2 - b_2 = a_1 + b_1 + a_2 + b_2.$$
In light of these five equalities, we can rewrite $A = \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix}$ as

\[
A = \begin{pmatrix}
  a_1 & b_1 & -a_1 - b_1 \\
  a_2 & b_2 & -a_2 - b_2 \\
  a_3 & b_3 & a_1 + b_1 + a_2 + b_2
\end{pmatrix}
\]

\[
= \begin{pmatrix}
  a_1 & 0 & -a_1 \\
  0 & 0 & 0 \\
  -a_1 & 0 & a_1
\end{pmatrix} + \begin{pmatrix}
  0 & b_1 & -b_1 \\
  0 & 0 & 0 \\
  0 & -b_1 & b_1
\end{pmatrix}
\]

\[
= a_1 \begin{pmatrix}
  1 & 0 & -1 \\
  0 & 0 & 0 \\
  -1 & 0 & 1
\end{pmatrix} = a_1 M_1
\]

\[
= b_1 \begin{pmatrix}
  0 & 1 & -1 \\
  0 & 0 & 0 \\
  0 & -1 & 1
\end{pmatrix} = b_1 M_2
\]

\[
+ \begin{pmatrix}
  0 & 0 & 0 \\
  a_2 & 0 & -a_2 \\
  -a_2 & 0 & a_2
\end{pmatrix} + \begin{pmatrix}
  0 & 0 & 0 \\
  0 & b_2 & -b_2 \\
  0 & -b_2 & b_2
\end{pmatrix}
\]

\[
= a_2 \begin{pmatrix}
  0 & 0 & 0 \\
  1 & 0 & -1 \\
  -1 & 0 & 1
\end{pmatrix} = a_2 M_3
\]

\[
+ b_2 \begin{pmatrix}
  0 & 0 & 0 \\
  0 & 1 & -1 \\
  0 & -1 & 1
\end{pmatrix} = b_2 M_4
\]

\[
\text{(by (3))}
\]

\[
\text{(by (3))}
\]

\[
\text{(by (4))}
\]

\[
\text{(by (5))}
\]

\[
\text{(by (6))}
\]

\[
\text{here, we have played the usual game of splitting our matrix into a sum of many matrices, each of which isolates exactly one of the four variables $a_1, b_1, a_2, b_2$}
\]

\[
= a_1 M_1 + b_1 M_2 + a_2 M_3 + b_2 M_4 \in \text{span} \left( M_1, M_2, M_3, M_4 \right).
\]

Thus, we have proven that $A \in \text{span} \left( M_1, M_2, M_3, M_4 \right)$ for each $A \in \mathcal{Z}$. In other words, $\mathcal{Z} \subseteq \text{span} \left( M_1, M_2, M_3, M_4 \right)$. Combined with $\text{span} \left( M_1, M_2, M_3, M_4 \right) \subseteq \mathcal{Z}$, this yields $\mathcal{Z} = \text{span} \left( M_1, M_2, M_3, M_4 \right)$.

Hence, the list $(M_1, M_2, M_3, M_4)$ spans $\mathcal{Z}$. In order to prove that this list is a basis of $\mathcal{Z}$, we therefore only need to check that this list is linearly independent. Let us do this now:

Let $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ be reals such that $\lambda_1 M_1 + \lambda_2 M_2 + \lambda_3 M_3 + \lambda_4 M_4 = \overrightarrow{0}$. (Of course, $\overrightarrow{0}$ is the zero vector of $\mathbb{R}^{3 \times 3}$ here, i.e., the zero matrix $0_{3 \times 3}$.) We must show that $\lambda_1 = \lambda_2 = \lambda_3 = \lambda_4 = 0$. 

---


We have
\[ \vec{0} = \lambda_1 M_1 + \lambda_2 M_2 + \lambda_3 M_3 + \lambda_4 M_4 \]
\[ = \lambda_1 \begin{pmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{pmatrix} + \lambda_2 \begin{pmatrix} 0 & 1 & -1 \\ 0 & 0 & 0 \\ 0 & -1 & 1 \end{pmatrix} \]
\[ + \lambda_3 \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & -1 \\ -1 & 0 & 1 \end{pmatrix} + \lambda_4 \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{pmatrix} \]
(by the equalities (3), (4), (5) and (6))
\[ = \begin{pmatrix} \lambda_1 & \lambda_2 & -\lambda_1 - \lambda_2 \\ \lambda_3 & \lambda_4 & -\lambda_3 - \lambda_4 \\ -\lambda_1 - \lambda_3 & -\lambda_2 - \lambda_4 & \lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 \end{pmatrix}. \]
Hence,
\[ \begin{pmatrix} \lambda_1 & \lambda_2 & -\lambda_1 - \lambda_2 \\ \lambda_3 & \lambda_4 & -\lambda_3 - \lambda_4 \\ -\lambda_1 - \lambda_3 & -\lambda_2 - \lambda_4 & \lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 \end{pmatrix} = \vec{0} = 0_{3 \times 3}. \]
In other words, each entry of the matrix
\[ \begin{pmatrix} \lambda_1 & \lambda_2 & -\lambda_1 - \lambda_2 \\ \lambda_3 & \lambda_4 & -\lambda_3 - \lambda_4 \\ -\lambda_1 - \lambda_3 & -\lambda_2 - \lambda_4 & \lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 \end{pmatrix} \]
must be 0. In particular, all of \( \lambda_1, \lambda_2, \lambda_3, \lambda_4 \) are 0. In other words, \( \lambda_1 = \lambda_2 = \lambda_3 = \lambda_4 = 0 \). This completes our proof. 

Let me now remind you of how I understand the words “QR decomposition” (your favorite book might define it differently!). If \( A \) is an \( n \times k \)-matrix whose columns are linearly independent, then a QR decomposition of \( A \) means a way to write \( A \) in the form \( A = QR \), where:

- \( Q \) is an \( n \times k \)-matrix with orthonormal columns (this is equivalent to saying that \( Q \) is an \( n \times k \)-matrix satisfying \( Q^T Q = I_k \));
- \( R \) is an upper-triangular \( k \times k \)-matrix with nonzero diagonal entries.

For example, a QR decomposition of
\[ \begin{pmatrix} 2 & 17 \\ 4 & 13 \\ 8 & 5 \end{pmatrix} \]
is
\[ \begin{pmatrix} 2 & 17 \\ 4 & 13 \\ 8 & 5 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{21}} & \frac{2}{\sqrt{21}} \\ \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ \frac{\sqrt{21}}{4} & -\frac{1}{\sqrt{6}} \end{pmatrix} \begin{pmatrix} 2 \sqrt{21} & 3 \sqrt{21} \\ 0 & 7 \sqrt{6} \end{pmatrix}. \]
this is the \( Q \)
this is the \( R \)
Exercise 4. (a) Find a QR decomposition of the matrix \[
\begin{pmatrix}
3 & 0 & 0 \\
4 & 5 & 0 \\
0 & 4 & 3
\end{pmatrix}
\]. [4 points]
(b) Find a QR decomposition of the matrix \[
\begin{pmatrix}
1 & 2 & 3 \\
0 & 1 & 2 \\
0 & 0 & 1
\end{pmatrix}
\]. [4 points]
(c) Find a QR decomposition of the matrix \[
\begin{pmatrix}
0 & 0 & 1 \\
-1 & 0 & 0 \\
0 & 1 & 0
\end{pmatrix}
\]. [4 points]

Solution. This is solved exactly like Exercise 3 on the midterm #2 practice sheet (except that I have picked the matrix in part (a) specifically to avoid square roots appearing in its QR decomposition).

(a) Let A be our matrix \[
\begin{pmatrix}
3 & 0 & 0 \\
4 & 5 & 0 \\
0 & 4 & 3
\end{pmatrix}
\]. Let \( w_1, w_2, w_3 \) be the three columns of A; thus,
\[
\begin{align*}
w_1 &= (3, 4, 0)^T, \\
w_2 &= (0, 5, 4)^T, \\
w_3 &= (0, 0, 3)^T.
\end{align*}
\]
Now, we apply the Gram-Schmidt process to \( w_1, w_2, w_3 \):

1. At the first step, we set \( u_1 = w_1 \). Thus,
\[
u_1 = w_1 = (3, 4, 0)^T.
\]

2. At the second step, we set \( u_2 = w_2 - \lambda_{2,1}u_1 \), where \( \lambda_{2,1} = \frac{\langle w_2, u_1 \rangle}{\langle u_1, u_1 \rangle} \). We compute these explicitly:
\[
\lambda_{2,1} = \frac{\langle w_2, u_1 \rangle}{\langle u_1, u_1 \rangle} = \frac{\langle (0, 5, 4)^T, (3, 4, 0)^T \rangle}{\langle (3, 4, 0)^T, (3, 4, 0)^T \rangle} = \frac{20}{25} = \frac{4}{5}
\]
and thus
\[
u_2 = w_2 - \lambda_{2,1}u_1 = (0, 5, 4)^T - \frac{4}{5}(3, 4, 0)^T = \left(-\frac{12}{5}, \frac{9}{5}, 4\right)^T.
\]

3. At the third step, we set \( u_3 = w_3 - \lambda_{3,1}u_1 - \lambda_{3,2}u_2 \), where \( \lambda_{3,1} = \frac{\langle w_3, u_1 \rangle}{\langle u_1, u_1 \rangle} \) and \( \lambda_{3,2} = \frac{\langle w_3, u_2 \rangle}{\langle u_2, u_2 \rangle} \). We compute these explicitly:
\[
\lambda_{3,1} = \frac{\langle w_3, u_1 \rangle}{\langle u_1, u_1 \rangle} = \frac{\langle (0, 0, 3)^T, (3, 4, 0)^T \rangle}{\langle (3, 4, 0)^T, (3, 4, 0)^T \rangle} = \frac{0}{25} = 0
\]
and
\[
\lambda_{3,2} = \frac{\langle w_3, u_2 \rangle}{\langle u_2, u_2 \rangle} = \frac{\langle (0,0,3)^T, \left(\frac{-12}{5}, \frac{9}{5}, 4\right)^T \rangle}{\langle \left(\frac{-12}{5}, \frac{9}{5}, 4\right)^T, \left(\frac{-12}{5}, \frac{9}{5}, 4\right)^T \rangle} = \frac{12}{25}
\]
and thus
\[
u_3 = w_3 - \lambda_{3,1}u_1 - \lambda_{3,2}u_2
\]
\[
= (0,0,3)^T - 0 (3,4,0)^T - \frac{12}{25} \left(\frac{-12}{5}, \frac{9}{5}, 4\right)^T
\]
\[
= \left(\frac{144}{125}, \frac{-108}{125}, \frac{27}{25}\right)^T.
\]

Next, we normalize the vectors \(u_1, u_2, u_3\) – that is, we divide them by their lengths so they become orthonormal and not just orthogonal. The resulting vectors will be called \(q_1, q_2, q_3\). Explicitly:
\[
q_1 = \frac{1}{\|u_1\|}u_1 = \frac{1}{5} (3,4,0)^T \quad \text{(since } \|u_1\| = \sqrt{\langle u_1, u_1 \rangle} = 5 \text{)}
\]
\[
= \left(\frac{3}{5}, \frac{4}{5}, 0\right)^T,
\]
\[
q_2 = \frac{1}{\|u_2\|}u_2 = \frac{1}{5} \left(\frac{-12}{5}, \frac{9}{5}, 4\right)^T \quad \text{(since } \|u_2\| = \sqrt{\langle u_2, u_2 \rangle} = 5 \text{)}
\]
\[
= \left(\frac{-12}{25}, \frac{9}{25}, \frac{4}{5}\right)^T,
\]
\[
q_3 = \frac{1}{\|u_3\|}u_3 = \frac{1}{\frac{9}{5}} \left(\frac{144}{125}, \frac{-108}{125}, \frac{27}{25}\right)^T \quad \text{(since } \|u_3\| = \sqrt{\langle u_3, u_3 \rangle} = \frac{9}{5} \text{)}
\]
\[
= \left(\frac{16}{25}, \frac{-12}{25}, \frac{3}{5}\right)^T.
\]

Now, the \(Q\) and \(R\) in the QR decomposition \(A = QR\) of \(A\) can be determined as follows:

- The matrix \(Q\) will be the \(3 \times 3\)-matrix with columns \(q_1, q_2, q_3\). Plugging in the values of \(q_1, q_2, q_3\) already computed, we thus find
\[
Q = \begin{pmatrix}
3 & 12 & 16 \\
\frac{5}{25} & \frac{-25}{25} & \frac{25}{25} \\
\frac{5}{25} & \frac{9}{12} & \frac{-25}{25} \\
0 & \frac{4}{3} & \frac{3}{5} \\
\end{pmatrix}.
\]
The matrix $R$ will be the $3 \times 3$-matrix whose $(i,j)$-th entry (for all $i$ and $j$) is

$$
R_{i,j} = \begin{cases} 
\lambda_{j,i} ||u_i||, & \text{if } i < j; \\
||u_j||, & \text{if } i = j; \\
0, & \text{if } i > j.
\end{cases}
$$

In other words,

$$
R = \begin{pmatrix}
||u_1|| & \lambda_{2,1} ||u_1|| & \lambda_{3,1} ||u_1|| \\
0 & ||u_2|| & \lambda_{3,2} ||u_2|| \\
0 & 0 & ||u_3||
\end{pmatrix}.
$$

Plugging in the values of $||u_i||$ and $\lambda_{j,i}$ (which have already been computed), we obtain

$$
R = \begin{pmatrix}
5 & 4 & 0 \\
0 & 5 & 12 \\
0 & 0 & 9
\end{pmatrix}.
$$

Thus, $Q$ and $R$ have both been found.

**Remark:** I am sorry for this problem, which turned out much more laborious than I wanted it to be. I had a computer find the QR decomposition, and saw that it looked nice enough; I didn’t anticipate that the process that leads to the result would require computations such as finding the length of $u_3 = \begin{pmatrix} 144/125 & -108/125 & 27/25 \end{pmatrix}^T$.

**b)** We can use the same algorithm as in **a**. But we can also save ourselves the hassle and read off the answer from the problem: Namely, set $A = \begin{pmatrix} 1 & 2 & 3 \\
0 & 1 & 2 \\
0 & 0 & 1 \end{pmatrix}$.

Then, the matrix $A$ itself is upper-triangular. Hence, setting $Q = I_3$ and $R = A$ yields a QR decomposition $A = QR$ of $A$.

**c)** Once again, the answer can be read off from the problem: Namely, set $A = \begin{pmatrix} 0 & 0 & 1 \\
-1 & 0 & 0 \\
0 & 1 & 0 \end{pmatrix}$. Then, the matrix $A$ has orthonormal columns (in fact, its columns are distinct standard basis vectors scaled by 1 or $-1$). Hence, setting $Q = A$ and $R = I_3$ yields a QR decomposition $A = QR$ of $A$.

**Exercise 5.** Consider the $2 \times 1$-matrix $A = \begin{pmatrix} 3 \\
4 \end{pmatrix}$.

**a)** Find an orthogonal basis of Col $A$ (the column space of $A$). [3 points]

**b)** Find a QR decomposition $A = QR$ of $A$. [3 points]

**c)** Let $b = \begin{pmatrix} 1 \\
1 \end{pmatrix}$. Compute the projection of $b$ onto Col $A$. [3 points]

**d)** Compute $QQ^Tb$. [3 points]
What do you observe? [1 point]

Find the $x \in \mathbb{R}^1$ for which $\|Ax - b\|$ is minimum. [3 points]

Solution. (a) The column space $\text{Col} A$ of $A$ is spanned by the columns of $A$. In our situation, this means that $\text{Col} A = \text{span} \left( (3, 4)^T \right)$. Thus, the list $\left( (3, 4)^T \right)$ is a basis of $\text{Col} A$ (since this list is clearly linearly independent). This basis is clearly orthogonal (because it has only one entry, whereas orthogonality makes no requirements on single entries).

(b) We can just take $Q = \frac{1}{5} A = \begin{pmatrix} 3 \\ 5 \\ 4 \\ 5 \end{pmatrix}$ and $R = \begin{pmatrix} 5 \\ 0 \end{pmatrix}$. (Indeed, the matrix $\frac{1}{5} A$ has orthonormal columns, whereas the matrix $R$ is upper-triangular.)

(c) Recall the general formula that says that if $b$ is a vector in $\mathbb{R}^n$, and if $(u_1, u_2, \ldots, u_k)$ is an orthogonal basis of a subspace $U$ of $\mathbb{R}^n$, then the projection of $b$ on $U$ is

$$\frac{\langle b, u_1 \rangle}{\|u_1\|^2} u_1 + \frac{\langle b, u_2 \rangle}{\|u_2\|^2} u_2 + \cdots + \frac{\langle b, u_k \rangle}{\|u_k\|^2} u_k.$$.

Applying this to $n = 2, U = \text{Col} A, k = 1$ and $(u_1, u_2, \ldots, u_k) = \left( (3, 4)^T \right)$, we conclude that the projection of $b$ on $\text{Col} A$ is

$$\frac{\langle b, (3, 4)^T \rangle}{\|(3, 4)^T\|^2} (3, 4)^T = \frac{\langle (1, 1)^T, (3, 4)^T \rangle}{\|(3, 4)^T\|^2} (3, 4)^T$$

(since $b = \begin{pmatrix} 1 \\ 1 \end{pmatrix} = (1, 1)^T$)

$$\frac{7}{25} (3, 4)^T = \begin{pmatrix} 21 \\ 25 \\ 28 \\ 25 \end{pmatrix}^T = \frac{7}{25} (3, 4)^T = \begin{pmatrix} 7/25 \\ 28/25 \\ 25/25 \\ 25/25 \end{pmatrix} = \frac{7}{25} (3, 4)^T = \begin{pmatrix} 21 \\ 25 \\ 28 \\ 25 \end{pmatrix}.$$

(d) Straightforward computations show

$$QQ^T b = \begin{pmatrix} 3/5 \\ 4/5 \end{pmatrix} \begin{pmatrix} 3 \\ 5 \\ 4 \\ 5 \end{pmatrix}^T = \begin{pmatrix} 3/5 \\ 4/5 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 21/25 \\ 28/25 \end{pmatrix} = \begin{pmatrix} 7/25 \\ 28/25 \end{pmatrix}.$$

(e) You observe that the projection of $b$ on $\text{Col} A$ is $QQ^T b$.

[Remark: This is no coincidence. Indeed, the following general result holds:

**Proposition 0.1.** Let $A$ be an $n \times k$-matrix whose columns are linearly independent. Let $A = QR$ be the QR decomposition of $A$. Let $b \in \mathbb{R}^n$. Then, the projection of $b$ on $\text{Col} A$ is $QQ^T b$.**}
Proof of Proposition 0.1. Set $u_0 = QQ^Tb$ and $u'_0 = b - u_0$. We are going to prove that $u_0 \in \text{Col } A$, $u'_0 \in (\text{Col } A)^\perp$, and $b = u_0 + u'_0$. Once this is shown, it will follow that $b = u_0 + u'_0$ is the (unique, as we know) decomposition of $b$ into a vector in $\text{Col } A$ and a vector in $(\text{Col } A)^\perp$; but this will obviously imply that the projection of $b$ on $\text{Col } A$ is $u_0 = QQ^Tb$. So we will be done.

The matrix $R$ is upper-triangular, and its diagonal entries are nonzero. Thus, in the parlance of my lecture notes (specifically, Definition 3.30 (b)), it is invertibly upper-triangular. Hence, Theorem 3.99 in my lecture notes shows that it is invertible, and its inverse $R^{-1}$ is also invertibly upper-triangular. Now,

$$u_0 = QQ^Tb = QR R^{-1}Q^Tb = AR^{-1}Q^Tb \in A\mathbb{R}^k = \text{Col } A.$$  

Furthermore, $b = u_0 + u'_0$ follows directly from $u'_0 = b - u_0$. Hence, out of the three claims $u_0 \in \text{Col } A$, $u'_0 \in (\text{Col } A)^\perp$ and $b = u_0 + u'_0$, we have already proven the first and the third. It remains to prove the second, i.e., the claim that $u'_0 \in (\text{Col } A)^\perp$.

Let $x \in \text{Col } A$. Then, $x \in \text{Col } A = A\mathbb{R}^k$. Thus, there exists some $y \in \mathbb{R}^k$ such that $x = Ay$. Consider this $y$.

But recall that the entries of the matrix $Q^TQ$ are the inner products between the columns of the matrix $Q$. Thus, $Q^TQ = I_k$ (since the columns of the matrix $Q$ are orthonormal). Now,

$$\langle x, u'_0 \rangle = x^T u'_0 \quad \text{(by the definition of the inner product)}$$

$$= \left( \begin{array}{c} A \cdot y \\ =QR \\ \end{array} \right)^T \left( \begin{array}{c} u'_0 \\ =b-u_0 \\ \end{array} \right) \quad \text{(since } x = Ay)$$

$$= (QRy)^T \left( \begin{array}{c} b \quad u_0 \\ =Q^Tb \\ =0 \end{array} \right) = y^TR^TQ^T \left( b - QQ^Tb \right)$$

$$= y^TR^TQ^Tb - y^TR^TQ^TQ^Tb = y^TR^TQ^Tb - y^TR^TQ^Tb = 0.$$

In other words, $u'_0 \perp x$.

Now, we have proven that $u'_0 \perp x$ for every $x \in \text{Col } A$. In other words, $u'_0 \in (\text{Col } A)^\perp$. This proves the one claim that remained to be proven. Thus, the proof of Proposition 0.1 is complete. \(\square\)

1. (f) We are looking for the least-squares solution of $Ax = b$.

We follow the usual method: We set $K = A^TA$ and $f = A^Tb$, and then $x = K^{-1}f$.

This works because the columns of $A$ are linearly independent.

Here are the computations:

$$K = A^TA = \begin{pmatrix} 3 \\ 4 \end{pmatrix}^T \begin{pmatrix} 3 \\ 4 \end{pmatrix} = (25)$$
and
\[ f = A^T b = \begin{pmatrix} 3 \\ 4 \end{pmatrix}^T \begin{pmatrix} 1 \\ 1 \end{pmatrix} = (7). \]

Thus,
\[ x = K^{-1} f = (25)^{-1} (7) = \begin{pmatrix} 7 \\ 25 \end{pmatrix}. \]

[Remark: Unsurprisingly, this x satisfies \( Ax = \begin{pmatrix} 21 \\ 25 \\ 28 \\ 25 \end{pmatrix} \), which is the projection of \( b \) on \( \text{Col} \, A \). This is an example of the standard connection between the least-squares problem and the closest-point problem.]

Let me recall a few definitions:

- We denote by \( \mathbb{R}^n \) the vector space \( \mathbb{R}^{n \times 1} \). It consists of column vectors of size \( n \).
- If \( v \in \mathbb{R}^n \) and \( w \in \mathbb{R}^n \), then \( \langle v, w \rangle = v^T w \). (This is a \( 1 \times 1 \)-matrix, but we regard it as a number, just by taking its single entry and “dropping the parentheses around it”.)
- If \( v \in \mathbb{R}^n \), then \( ||v|| = \sqrt{\langle v, v \rangle} \).
- If \( v \in \mathbb{R}^n \) and \( w \in \mathbb{R}^n \), then we write \( v \perp w \) when we have \( \langle v, w \rangle = 0 \).
- If \( U \) is a subspace of \( \mathbb{R}^n \), then \( U^\perp \) denotes the subset \( \{ x \in \mathbb{R}^n \mid x \perp u \text{ for all } u \in U \} \) of \( \mathbb{R}^n \). This subset \( U^\perp \) is itself a subspace of \( \mathbb{R}^n \), and is called the orthogonal complement (or orthogonal subspace) of \( U \).

Exercise 6. Which of the following claims are true, and which are false? (Please write a “T” into the box for “True”, or an “F” for “False”.)

[2 points for each of the 18 claims]

(a) F If \( A \) is an \( n \times n \)-matrix, then the matrix \( A^T A \) is diagonal.
(b) F If \( A \) is an \( n \times n \)-matrix, then the matrix \( A^T A \) is invertible.
(c) T If \( A \) is an \( n \times n \)-matrix, then the matrix \( A^T A \) is symmetric.
(d) T If \( A \) is a lower-triangular \( n \times n \)-matrix, then \( A^T \) is an upper-triangular \( n \times n \)-matrix.
(e) T If \( a \) is a linearly independent list of vectors in an \( n \)-dimensional vector space, then \( a \) contains at most \( n \) vectors.
(f) F If a linearly independent list of vectors in a vector space \( V \) and a list of vectors that spans \( V \) have the same size, then these two lists are equal.
(g) T If \( v \) and \( w \) are two vectors in \( \mathbb{R}^n \), then \( \langle v, w \rangle = \langle w, v \rangle \).
(h) F If \( v \) and \( w \) are two vectors in \( \mathbb{R}^n \), then \( \langle 2v, 2w \rangle = 2 \langle v, w \rangle \).
(i)  \(\boxed{F}\) If \(v\) and \(w\) are two vectors in \(\mathbb{R}^n\) satisfying \(\langle v, w \rangle = 0\), then \(v = \overrightarrow{0}\) or \(w = \overrightarrow{0}\).

(j)  \(\boxed{F}\) If \(v\) and \(w\) are two vectors in \(\mathbb{R}^n\), then \(||v|| + ||w|| \leq ||v + w||\).

(k)  \(\boxed{F}\) If \(U\) and \(V\) are two subspaces of \(\mathbb{R}^n\) satisfying \(U \subseteq V\), then \(U^\perp \subseteq V^\perp\).

(l)  \(\boxed{T}\) If \(A\) is an \(n \times k\)-matrix and \(B\) is a \(k \times n\)-matrix with \(k < n\), then \(AB\) can never be invertible.

(m)  \(\boxed{T}\) If \(A\) is an \(n \times m\)-matrix, then \(\text{rank} A + \dim (\text{Ker} A) = m\).

(n)  \(\boxed{F}\) If \(A\) is an \(n \times m\)-matrix, then \(\text{rank} A + \dim (\text{Ker} A) = n\).

(o)  \(\boxed{F}\) If \(A\) is an \(n \times m\)-matrix and \(b\) is a column vector of size \(n\), then there exists a unique \(x \in \mathbb{R}^m\) for which \(||Ax - b|||\) is minimum.

(p)  \(\boxed{T}\) If \(A\) is an \(n \times m\)-matrix and \(b\) is a column vector of size \(n\), then there exists a unique \(u \in \text{Col} A\) for which \(||u - b|||\) is minimum.

(q)  \(\boxed{T}\) The orthogonal complement of the subspace \(\{ \overrightarrow{0} \}\) of \(\mathbb{R}^n\) is \(\{ \overrightarrow{0} \}^\perp = \mathbb{R}^n\).

(r)  \(\boxed{T}\) If \(A\) and \(B\) are two \(n \times m\)-matrices, then \(\text{rank} (A + B) \leq \text{rank} A + \text{rank} B\).

[Hint: Many false statements are easy to refute. Sometimes, stupid things like taking \(A = 0_{n \times m}\) or \(n = 0\) or \(n = 1\) or \(v = \overrightarrow{0}\) suffice to obtain a counterexample.]

Solution. I have entered the answers in the boxes above, but let me also comment on why the answers are the right ones:

(a) This is false. For a counterexample, take \(A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}\), and observe that \(A^T A = \begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix}\) is not diagonal.

(b) This is false. For a counterexample, take \(A = 0_{n \times n}\), and observe that \(A^T A = 0_{n \times n}\) is not invertible unless \(n = 0\).

(c) This is true. The simplest way to see that the matrix \(A^T A\) is symmetric is to show that it equals its own transpose: Since the transpose of a product of two matrices is the product of their transposes in reverse order, we have

\[
\left( A^T A \right)^T = A^T \left( A^T \right)^T = A^T A.
\]

This shows that \(A^T A\) is symmetric.

(d) This is true. And it is obvious when you look at the forms of lower- and upper-triangular matrices and recall that transposition “reflects a matrix in its diagonal”. A formal proof is also easy to make.

(e) This is true. If a vector space \(V\) is \(n\)-dimensional, then \(V\) has a basis \(b\) of size \(n\). This basis \(b\) clearly is a spanning list of \(V\). Thus, if \(a\) is a linearly independent list of vectors in \(V\), then \(a\) must be at most as long as this list \(b\) (because a linearly independent list must be at most as long as a spanning list), hence contain at most \(n\) vectors.
(f) This is false. For example, the two lists \((e_1,e_2)\) and \((e_2,e_1)\) in \(\mathbb{R}^2\) are not equal. One correct statement that we could make here instead is that both lists must be bases of \(V\).

(g) This is true. The quickest way to see this is to write \(v\) and \(w\) in the forms 
\[
\langle v, w \rangle = v_1w_1 + v_2w_2 + \cdots + v_nw_n
\]
and 
\[
\langle w, v \rangle = w_1v_1 + w_2v_2 + \cdots + w_nv_n,
\]
which are clearly equal (because \(v_iw_i = w_iv_i\) for any \(i\)).

(h) This is false. Instead, we have \(\langle 2v, 2w \rangle = 4 \langle v, w \rangle\), but \(\langle 2v, w \rangle = \langle v, 2w \rangle = 2 \langle v, w \rangle\).

More generally, \(\langle \lambda v, \mu w \rangle = \lambda \mu \langle v, w \rangle\) for any reals \(\lambda\) and \(\mu\).

(i) This is false. Any two nonzero vectors \(v\) and \(w\) orthogonal to each other (for example, \(v = e_1\) and \(w = e_2\) in \(\mathbb{R}^2\) provide a counterexample.

(j) This is false. The triangle inequality says \(||v|| + ||w|| \geq ||v + w||\) (visually, the detour is longer than the direct path). If \(v\) and \(w\) are linearly independent (again, take \(v = e_1\) and \(w = e_2\) in \(\mathbb{R}^2\) for a concrete counterexample), then the \(\geq\) sign here actually becomes a strict \(>\) sign, and so the claim \(||v|| + ||w|| \leq ||v + w||\) cannot hold.

(k) This is false. Instead, we have \(V^\perp \subseteq U^\perp\). In fact, \(U \subseteq V\) shows that every \(u \in U\) is also an element of \(V\). Therefore, if a vector \(x \in \mathbb{R}^n\) satisfies \(x \perp u\) for all \(u \in V\), then this \(x\) also satisfies \(x \perp u\) for all \(u \in U\). In other words, the set of all \(x \in \mathbb{R}^n\) satisfying \(x \perp u\) for all \(u \in V\) is a subset of the set of all \(x \in \mathbb{R}^n\) satisfying \(x \perp u\) for all \(u \in U\). In other words,
\[
\{ x \in \mathbb{R}^n \mid x \perp u \text{ for all } u \in V \} \subseteq \{ x \in \mathbb{R}^n \mid x \perp u \text{ for all } u \in U \}.
\]
Now, the definition of \(V^\perp\) shows that 
\[
V^\perp = \{ x \in \mathbb{R}^n \mid x \perp u \text{ for all } u \in V \}
\]
\[
\subseteq \{ x \in \mathbb{R}^n \mid x \perp u \text{ for all } u \in U \} = U^\perp,
\]
qed.

To provide a concrete counterexample to the false claim that \(U^\perp \subseteq V^\perp\), try \(n = 1\), \(U = \{ \hat{0} \}\) and \(V = \mathbb{R}^1\).

(l) This is true.

Proof. Assume the contrary. Thus, \(AB\) is invertible. Hence, Proposition 0.5 \(c\) on homework set #4 (applied to \(n\) and \(AB\) instead of \(m\) and \(A\)) shows that \(\text{rank}(AB) = n = n\). But Proposition 0.2 \(b\) on homework set #4 (applied to \(k\) and \(n\) instead of \(m\) and \(p\)) yields \(\text{rank}(AB) \leq \text{rank} A\). Finally, the equality (15) on homework set #4 (applied to \(k\) instead of \(m\)) yields \(\text{rank} A \leq \min \{n, k\}\). Thus, \(n = \text{rank}(AB) \leq \text{rank} A \leq \min \{n, k\} \leq k < n\), which is absurd. Hence, we have a contradiction. This shows that our assumption was wrong, and the proof is complete.
(m) This is true. It is just the rank-nullity theorem.

(n) This is false. It contradicts the rank-nullity theorem whenever \( n \neq m \).

(o) This is false. For a quick counterexample, take \( n = 1 \) and \( m = 1 \) and \( A = 0_{1 \times 1} \), in which case \( Ax \) does not depend on \( x \) at all (in fact, \( Ax = 0_{1 \times 1} \) no matter what \( x \) is) and therefore \( \|Ax - b\| \) attains its minimum value at any \( x \).

(p) This is true. Recall the fact that if \( U \) is a subspace of \( \mathbb{R}^n \), then there exists a unique \( u \in U \) for which \( \|u - b\| \) is minimum (namely, this \( u \) is the projection of \( b \) on \( U \)). The claim of (p) follows by applying this fact to \( U = \text{Col} A \).

(q) This is true. The definition of \( \{ \overrightarrow{0} \}^\perp \) yields

\[
\{ \overrightarrow{0} \}^\perp = \{ x \in \mathbb{R}^n \mid x \perp u \text{ for all } u \in \{ \overrightarrow{0} \} \} = \{ x \in \mathbb{R}^n \mid x \perp \overrightarrow{0} \} \quad (7)
\]

(since the only \( u \in \{ \overrightarrow{0} \} \) is \( \overrightarrow{0} \)). But every \( x \in \mathbb{R}^n \) satisfies \( x \perp \overrightarrow{0} \) (since \( \langle x, \overrightarrow{0} \rangle = x^T \overrightarrow{0} = 0 \)). Therefore, \( \{ x \in \mathbb{R}^n \mid x \perp \overrightarrow{0} \} = \mathbb{R}^n \). Hence, (7) rewrites as \( \{ \overrightarrow{0} \}^\perp = \mathbb{R}^n \).

(r) This is true.

Proof. Let \((u_1, u_2, \ldots, u_k)\) be a basis of \( \text{Col} A \). Then, \( k = \dim (\text{Col} A) = \text{rank} A \).

Let \((v_1, v_2, \ldots, v_\ell)\) be a basis of \( \text{Col} B \). Then, \( \ell = \dim (\text{Col} B) = \text{rank} B \).

The list \((u_1, u_2, \ldots, u_k, v_1, v_2, \ldots, v_\ell)\) might not be a basis of anything, but it spans the vector space \( \text{span} (u_1, u_2, \ldots, u_k, v_1, v_2, \ldots, v_\ell) \) (obviously). Hence, a basis of \( \text{span} (u_1, u_2, \ldots, u_k, v_1, v_2, \ldots, v_\ell) \) can be found by shrinking this list (i.e., by removing redundant elements from it). Thus, this basis will have size \( \leq k + \ell \) (because the list \((u_1, u_2, \ldots, u_k, v_1, v_2, \ldots, v_\ell)\) has size \( k + \ell \)). In other words,

\[
\dim (\text{span} (u_1, u_2, \ldots, u_k, v_1, v_2, \ldots, v_\ell)) \leq k + \ell. \quad (8)
\]

Now, I claim that

\[ \text{Col} (A + B) \text{ is a subspace of } \text{span} (u_1, u_2, \ldots, u_k, v_1, v_2, \ldots, v_\ell). \quad (9) \]

Indeed, let me show this. Let \( w \) be a column of \( A + B \). Then, \( w \) is the sum of some column of \( A \) with the respective column of \( B \). In other words, \( w = a + b \) for some column \( a \) of \( A \) and some column \( b \) of \( B \). Consider these \( a \) and \( b \). The vector \( a \) (being a column of \( A \)) must lie in the span of the columns of \( A \). In other words, \( a \in \text{Col} A = \text{span} (u_1, u_2, \ldots, u_k) \) (since \((u_1, u_2, \ldots, u_k)\) is a basis of \( \text{Col} A \)). In other words, \( a = \lambda_1 u_1 + \lambda_2 u_2 + \cdots + \lambda_k u_k \) for some reals \( \lambda_1, \lambda_2, \ldots, \lambda_k \). Similarly, \( b = \mu_1 v_1 + \mu_2 v_2 + \cdots + \mu_\ell v_\ell \) for some reals \( \mu_1, \mu_2, \ldots, \mu_\ell \). Adding the two equalities \( a = \lambda_1 u_1 + \lambda_2 u_2 + \cdots + \lambda_k u_k \) and \( b = \mu_1 v_1 + \mu_2 v_2 + \cdots + \mu_\ell v_\ell \), we obtain

\[
a + b = (\lambda_1 u_1 + \lambda_2 u_2 + \cdots + \lambda_k u_k) + (\mu_1 v_1 + \mu_2 v_2 + \cdots + \mu_\ell v_\ell)
\in \text{span} (u_1, u_2, \ldots, u_k, v_1, v_2, \ldots, v_\ell).
\]

Thus,

\[
w = a + b \in \text{span} (u_1, u_2, \ldots, u_k, v_1, v_2, \ldots, v_\ell).
\]
Thus, we have shown that every column $w$ of $A + B$ satisfies $w \in \text{span} \,(u_1, u_2, \ldots, u_k, v_1, v_2, \ldots, v_\ell)$. Therefore, the span of all columns of $A + B$ is a subset of $\text{span} \,(u_1, u_2, \ldots, u_k, v_1, v_2, \ldots, v_\ell)$. In other words,

$$\text{Col} \,(A + B) \subseteq \text{span} \,(u_1, u_2, \ldots, u_k, v_1, v_2, \ldots, v_\ell).$$

(10)

Moreover, $\text{Col} \,(A + B)$ contains the zero vector and is closed under addition and scaling (since $\text{Col} \,(A + B)$ is a subspace of $\mathbb{R}^n$). Hence, (10) shows that $\text{Col} \,(A + B)$ is actually a subspace of $\text{span} \,(u_1, u_2, \ldots, u_k, v_1, v_2, \ldots, v_\ell)$. This proves (9).

Hence, Proposition 0.1 (b) on homework set #4 shows that

$$\dim \,(\text{Col} \,(A + B)) \leq \dim \,(\text{span} \,(u_1, u_2, \ldots, u_k, v_1, v_2, \ldots, v_\ell))$$

$$\leq k + \ell \quad \text{(by (8))}$$

$$= \text{rank} \,A + \text{rank} \,B$$

Since $\text{rank} \,(A + B) = \dim \,(\text{Col} \,(A + B))$, this rewrites as $\text{rank} \,(A + B) \leq \text{rank} \,A + \text{rank} \,B$. The proof is complete. \hfill \Box