**Math 4242 Fall 2016 (Darij Grinberg): midterm 2 with solutions**
**Mon, 7 Nov 2016, in class (75 minutes). Proofs are NOT required.**

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

**(b)** The list $\mathbf{b} = \left((1,0,1)^T\right)$ is linearly independent. Extend this list to a basis of $\mathbb{R}^3$ by appending to it some elements from the list $\mathbf{a}$. [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[1]: We scan the list $\mathbf{a}$ from left to right. Each time we read an entry of $\mathbf{a}$, we check if this entry is a linear combination of the entries before it. If it is, then we remove this entry from $\mathbf{a}$ and start from scratch with the new (shorter) $\mathbf{a}$. 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 $\mathbf{a}$ 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 $\overrightarrow{0}$. Our first entry $(2,-4,2)^T$ is not $\overrightarrow{0}$. 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[2]), 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[3]. Thus, we remove the entry from $\mathbf{a}$, and start from scratch with the new (shorter) $\mathbf{a}$.

[1] In this algorithm, we treat $\mathbf{a}$ as a mutable variable.

[2] 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.

[3] And, again, we can tell this immediately without Gaussian elimination as well, by observing that $(1,-2,1)^T = \frac{1}{2}(2,-4,2)^T$.

- We scan the new list $\mathbf{a} = \left((2,-4,2)^T, (-3,6,3)^T, (3,1,3)^T, (2,1,0)^T\right)$ (the result of removing $(1,-2,1)^T$ from the old list $\mathbf{a}$) 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 $\mathbf{a}$. 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 $\mathbf{a}$ 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 $\mathbf{a}$, and start from scratch with the new (shorter) $\mathbf{a}$.

- We scan the new list $\mathbf{a} = \left((2,-4,2)^T, (-3,6,3)^T, (3,1,3)^T\right)$ 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 $\mathbf{a}$.

**(b)** We solve this using the following algorithm[4]: We scan the list $\mathbf{a}$ from left to right. Each time we read an entry of $\mathbf{a}$, we check if this entry is a linear combination of the (current) entries of $\mathbf{b}$. If it isn't, then we append this entry to $\mathbf{b}$. In either case, we proceed to the next entry. By the time we have scanned all entries of $\mathbf{a}$, the list $\mathbf{b}$ has become a basis of $\mathbb{R}^3$.

[4]In this algorithm, we treat $\mathbf{b}$ as a mutable variable.

In order to simplify our life, we use not the original list

$$\mathbf{a} = \left( (2, -4, 2)^T, (-3, 6, 3)^T, (1, -2, 1)^T, (3, 1, 3)^T, (2, 1, 0)^T \right),$$

but the shorter list

$$\mathbf{a} = \left( (2, -4, 2)^T, (-3, 6, 3)^T, (3, 1, 3)^T \right)$$

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 $\mathbf{b} = \left( (1, 0, 1)^T, (2, -4, 2)^T, (-3, 6, 3)^T \right)$. This list is therefore a basis of $\mathbb{R}^3$ obtained by appending some elements from **a** to the (old) list **b**. □

[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 $A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 5 \end{pmatrix}$. [6 points]

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

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

[*Proof:* Since $\operatorname{Col} A$ is the span of the columns of $A$, we have

$$\operatorname{Col} A = \operatorname{span}\left((1,2,3)^T, (2,3,4)^T, (3,4,5)^T\right).$$

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 $\operatorname{Col} A = \operatorname{span}\left((1,2,3)^T, (2,3,4)^T\right)$. The vectors $(1,2,3)^T$ and $(2,3,4)^T$ are fairly obviously linearly independent, and so they form a basis of $\operatorname{Col} A$.]

*Row space:* A basis of $\operatorname{Row} A$ is $((1,2,3), (2,3,4))$.

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

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

[*Proof:* We have $\operatorname{rank} A = \dim(\operatorname{Col} A) = 2$ (since we have found a basis of $\operatorname{Col} A$, and this basis has size 2). Now, the rank-nullity theorem yields $\operatorname{rank} A + \dim(\operatorname{Ker} A) = 3$, so that $\dim(\operatorname{Ker} A) = 3 - \underbrace{\operatorname{rank} A}_{=2} = 3 - 2 = 1$.

But it is easy to see that $(1,-2,1)^T \in \operatorname{Ker} A$ (indeed, this is a restatement of the linear dependency relation

$$1 \cdot \operatorname{col}_1 A + (-2) \cdot \operatorname{col}_2 A + 1 \cdot \operatorname{col}_3 A = \overrightarrow{0},$$

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 $\operatorname{Ker} A$. Since $\dim(\operatorname{Ker} A) = 1$, this shows that this list is a basis of $\operatorname{Ker} A$.

(Of course, we could have found this using Gaussian elimination as well.)]

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

[*Proof:* Again, recall that $A = A^T$. Thus, $\left(\operatorname{Ker}\underbrace{\left(A^T\right)}_{=A}\right)^T = (\operatorname{Ker} A)^T$. Thus, knowing a basis of $\operatorname{Ker} A$, we can immediately obtain a basis of $\left(\operatorname{Ker}\left(A^T\right)\right)^T$.]

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

[*Proof:* Since $\operatorname{Col} B$ is the span of the columns of $B$, we have

$$\operatorname{Col} B = \operatorname{span}\left((1,2,3,4)^T, (2,3,4,5)^T, (3,4,5,6)^T, (4,5,6,7)^T\right).$$

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 $\operatorname{Col} B = \operatorname{span}\left((1,2,3,4)^T, (2,3,4,5)^T\right)$. 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 $\operatorname{Col} B$.] □

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

[7 points]

**(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

$$a_1 + b_1 + c_1 = 0, \qquad a_2 + b_2 + c_2 = 0, \qquad a_3 + b_3 + c_3 = 0, \tag{1}$$
$$a_1 + a_2 + a_3 = 0, \qquad b_1 + b_2 + b_3 = 0, \qquad c_1 + c_2 + c_3 = 0 \tag{2}$$

(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 $\mathbb{R}^{3\times 3}$. Find a basis of this subspace. [7 points]

*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 $\mathfrak{Z}$ denote the subspace of $\mathbb{R}^{3\times 3}$ consisting of the zero-sum $3 \times 3$-matrices.

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

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

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

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

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

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

Next, we claim that $\mathfrak{Z} \subseteq \operatorname{span}(M_1, M_2, M_3, M_4)$. Indeed, let $A \in \mathfrak{Z}$. 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:

$$\begin{aligned}
c_1 &= -a_1 - b_1 && \text{(by the first equality of (1))};\\
c_2 &= -a_2 - b_2 && \text{(by the second equality of (1))};\\
a_3 &= -a_1 - a_2 && \text{(by the first equality of (2))};\\
b_3 &= -b_1 - b_2 && \text{(by the second equality of (2))};\\
c_3 &= -\underbrace{c_1}_{=-a_1-b_1} - \underbrace{c_2}_{=-a_2-b_2} && \text{(by the third equality of (2))}\\
&= -(-a_1 - b_1) - (-a_2 - b_2) = a_1 + b_1 + a_2 + b_2.
\end{aligned}$$

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

$$\begin{aligned}
A &= \begin{pmatrix} a_1 & b_1 & -a_1 - b_1 \\ a_2 & b_2 & -a_2 - b_2 \\ -a_1 - a_2 & -b_1 - b_2 & a_1 + b_1 + a_2 + b_2 \end{pmatrix} \\
&= \underbrace{\begin{pmatrix} a_1 & 0 & -a_1 \\ 0 & 0 & 0 \\ -a_1 & 0 & a_1 \end{pmatrix}}_{\substack{= a_1 \begin{pmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{pmatrix} = a_1 M_1 \\ \text{(by (3))}}} + \underbrace{\begin{pmatrix} 0 & b_1 & -b_1 \\ 0 & 0 & 0 \\ 0 & -b_1 & b_1 \end{pmatrix}}_{\substack{= b_1 \begin{pmatrix} 0 & 1 & -1 \\ 0 & 0 & 0 \\ 0 & -1 & 1 \end{pmatrix} = b_1 M_2 \\ \text{(by (3))}}} \\
&\qquad + \underbrace{\begin{pmatrix} 0 & 0 & 0 \\ a_2 & 0 & -a_2 \\ -a_2 & 0 & a_2 \end{pmatrix}}_{\substack{= a_2 \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & -1 \\ -1 & 0 & 1 \end{pmatrix} = a_2 M_3 \\ \text{(by (5))}}} + \underbrace{\begin{pmatrix} 0 & 0 & 0 \\ 0 & b_2 & -b_2 \\ 0 & -b_2 & b_2 \end{pmatrix}}_{\substack{= b_2 \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{pmatrix} = b_2 M_4 \\ \text{(by (6))}}} \\
&\qquad \left( \begin{array}{c} \text{here, we have played the usual game of splitting our matrix} \\ \text{into a sum of many matrices, each of which isolates exactly} \\ \text{one of the four variables } a_1, b_1, a_2, b_2 \end{array} \right) \\
&= a_1 M_1 + b_1 M_2 + a_2 M_3 + b_2 M_4 \in \operatorname{span}(M_1, M_2, M_3, M_4).
\end{aligned}$$

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

Hence, the list $(M_1, M_2, M_3, M_4)$ spans $\mathfrak{Z}$. In order to prove that this list is a basis of $\mathfrak{Z}$, 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

$$\begin{aligned}
\overrightarrow{0} &= \lambda_1 M_1 + \lambda_2 M_2 + \lambda_3 M_3 + \lambda_4 M_4 \\
&= \lambda_1 \begin{pmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{pmatrix} + \lambda_2 \begin{pmatrix} 0 & 1 & -1 \\ 0 & 0 & 0 \\ 0 & -1 & 1 \end{pmatrix} \\
&\qquad + \lambda_3 \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & -1 \\ -1 & 0 & 1 \end{pmatrix} + \lambda_4 \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{pmatrix} \\
&\qquad \text{(by the equalities (3), (4), (5) and (6))} \\
&= \begin{pmatrix} \lambda_1 & \lambda_2 & -\lambda_1 - \lambda_2 \\ \lambda_3 & \lambda_4 & -\lambda_3 - \lambda_4 \\ -\lambda_1 - \lambda_3 & -\lambda_2 - \lambda_4 & \lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 \end{pmatrix}.
\end{aligned}$$

Hence,

$$\begin{pmatrix} \lambda_1 & \lambda_2 & -\lambda_1 - \lambda_2 \\ \lambda_3 & \lambda_4 & -\lambda_3 - \lambda_4 \\ -\lambda_1 - \lambda_3 & -\lambda_2 - \lambda_4 & \lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 \end{pmatrix} = \overrightarrow{0} = 0_{3\times 3}.$$

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

$$\begin{pmatrix} 2 & 17 \\ 4 & 13 \\ 8 & 5 \end{pmatrix} = \underbrace{\begin{pmatrix} \frac{1}{\sqrt{21}} & \frac{2}{\sqrt{6}} \\ \frac{2}{\sqrt{21}} & \frac{1}{\sqrt{6}} \\ \frac{4}{\sqrt{21}} & \frac{-1}{\sqrt{6}} \end{pmatrix}}_{\text{this is the } Q} \underbrace{\begin{pmatrix} 2\sqrt{21} & 3\sqrt{21} \\ 0 & 7\sqrt{6} \end{pmatrix}}_{\text{this is the } R}.$$

**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,

$$w_1 = (3,4,0)^T, \qquad w_2 = (0,5,4)^T, \qquad w_3 = (0,0,3)^T.$$

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 $\lambda_{2,1} = \dfrac{\langle w_2, u_1\rangle}{\langle u_1, u_1\rangle}$. We compute these explicitly:

$$\lambda_{2,1} = \frac{\langle w_2, u_1\rangle}{\langle u_1, u_1\rangle} = \frac{\left\langle (0,5,4)^T, (3,4,0)^T \right\rangle}{\left\langle (3,4,0)^T, (3,4,0)^T \right\rangle} = \frac{20}{25} = \frac{4}{5}$$

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 $\lambda_{3,1} = \dfrac{\langle w_3, u_1\rangle}{\langle u_1, u_1\rangle}$ and $\lambda_{3,2} = \dfrac{\langle w_3, u_2\rangle}{\langle u_2, u_2\rangle}$. We compute these explicitly:

$$\lambda_{3,1} = \frac{\langle w_3, u_1\rangle}{\langle u_1, u_1\rangle} = \frac{\left\langle (0,0,3)^T, (3,4,0)^T \right\rangle}{\left\langle (3,4,0)^T, (3,4,0)^T \right\rangle} = \frac{0}{25} = 0$$

and

$$\lambda_{3,2} = \frac{\langle w_3, u_2 \rangle}{\langle u_2, u_2 \rangle} = \frac{\left\langle (0,0,3)^T, \left(-\frac{12}{5}, \frac{9}{5}, 4\right)^T \right\rangle}{\left\langle \left(-\frac{12}{5}, \frac{9}{5}, 4\right)^T, \left(-\frac{12}{5}, \frac{9}{5}, 4\right)^T \right\rangle} = \frac{12}{25}$$

and thus

$$\begin{aligned} u_3 &= w_3 - \lambda_{3,1} u_1 - \lambda_{3,2} u_2 \\ &= (0,0,3)^T - 0\,(3,4,0)^T - \frac{12}{25}\left(-\frac{12}{5}, \frac{9}{5}, 4\right)^T \\ &= \left(\frac{144}{125}, -\frac{108}{125}, \frac{27}{25}\right)^T. \end{aligned}$$

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:

$$\begin{aligned} q_1 &= \frac{1}{\|u_1\|} u_1 = \frac{1}{5}(3,4,0)^T \qquad \left(\text{since } \|u_1\| = \sqrt{\langle u_1, u_1 \rangle} = 5\right) \\ &= \left(\frac{3}{5}, \frac{4}{5}, 0\right)^T, \end{aligned}$$

$$\begin{aligned} q_2 &= \frac{1}{\|u_2\|} u_2 = \frac{1}{5}\left(-\frac{12}{5}, \frac{9}{5}, 4\right)^T \qquad \left(\text{since } \|u_2\| = \sqrt{\langle u_2, u_2 \rangle} = 5\right) \\ &= \left(-\frac{12}{25}, \frac{9}{25}, \frac{4}{5}\right)^T, \end{aligned}$$

$$\begin{aligned} q_3 &= \frac{1}{\|u_3\|} u_3 = \frac{1}{\left(\frac{9}{5}\right)}\left(\frac{144}{125}, -\frac{108}{125}, \frac{27}{25}\right)^T \qquad \left(\text{since } \|u_3\| = \sqrt{\langle u_3, u_3 \rangle} = \frac{9}{5}\right) \\ &= \left(\frac{16}{25}, -\frac{12}{25}, \frac{3}{5}\right)^T. \end{aligned}$$

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

$$Q = \begin{pmatrix} \frac{3}{5} & -\frac{12}{25} & \frac{16}{25} \\ \frac{4}{5} & \frac{9}{25} & -\frac{12}{25} \\ 0 & \frac{4}{5} & \frac{3}{5} \end{pmatrix}.$$

- The matrix $R$ will be the $3 \times 3$-matrix whose $(i,j)$-th entry (for all $i$ and $j$) is

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

In other words,

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

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

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

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 $u_3 = \left(\dfrac{144}{125}, -\dfrac{108}{125}, \dfrac{27}{25}\right)^T$.]

**(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 $\operatorname{Col} A$ (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 $\operatorname{Col} A$. [3 points]

**(d)** Compute $QQ^T b$. [3 points]

**(e)** What do you observe? [1 point]
**(f)** Find the $x \in \mathbb{R}^1$ for which $||Ax - b||$ is minimum. [3 points]

*Solution.* **(a)** The column space $\operatorname{Col} A$ of $A$ is spanned by the columns of $A$. In our situation, this means that $\operatorname{Col} A = \operatorname{span}\left((3,4)^T\right)$. Thus, the list $\left((3,4)^T\right)$ is a basis of $\operatorname{Col} A$ (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 $Q = \frac{1}{5}A = \begin{pmatrix} \frac{3}{5} \\ \frac{4}{5} \end{pmatrix}$ and $R = \begin{pmatrix} 5 \end{pmatrix}$. (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 $n = 2$, $U = \operatorname{Col} A$, $k = 1$ and $(u_1, u_2, \ldots, u_k) = \left((3,4)^T\right)$, we conclude that the projection of $b$ on $\operatorname{Col} A$ is

$$\frac{\left\langle b, (3,4)^T \right\rangle}{\left|\left|(3,4)^T\right|\right|^2} (3,4)^T = \frac{\left\langle (1,1)^T, (3,4)^T \right\rangle}{\left|\left|(3,4)^T\right|\right|^2} (3,4)^T \qquad \left(\text{since } b = \begin{pmatrix} 1 \\ 1 \end{pmatrix} = (1,1)^T\right)$$

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

**(d)** Straightforward computations show

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

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

*Proof of Proposition 0.1.* Set $u_0 = QQ^T b$ and $u_0' = b - u_0$. We are going to prove that $u_0 \in \operatorname{Col} A$, $u_0' \in (\operatorname{Col} A)^\perp$ 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 $\operatorname{Col} A$ and a vector in $(\operatorname{Col} A)^\perp$; but this will obviously imply that the projection of $b$ on $\operatorname{Col} A$ is $u_0 = QQ^T b$. 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,

$$u_0 = QQ^T b = \underbrace{QR}_{=A} R^{-1} Q^T b = AR^{-1} Q^T b \in A\mathbb{R}^k = \operatorname{Col} A.$$

Furthermore, $b = u_0 + u_0'$ follows directly from $u_0' = b - u_0$. Hence, out of the three claims $u_0 \in \operatorname{Col} A$, $u_0' \in (\operatorname{Col} A)^\perp$ 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 $u_0' \in (\operatorname{Col} A)^\perp$.

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

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

$$\begin{aligned}
\left\langle x, u_0' \right\rangle &= x^T u_0' \qquad \text{(by the definition of the inner product)} \\
&= \left( \underbrace{A}_{=QR} y \right)^T \underbrace{u_0'}_{=b-u_0} \qquad \text{(since } x = Ay\text{)} \\
&= \underbrace{(QRy)^T}_{=y^T R^T Q^T} \left( b - \underbrace{u_0}_{=QQ^T b} \right) = y^T R^T Q^T \left( b - QQ^T b \right) \\
&= y^T R^T Q^T b - y^T R^T \underbrace{Q^T Q}_{=I_k} Q^T b = y^T R^T Q^T b - y^T R^T Q^T b = 0.
\end{aligned}$$

In other words, $u_0' \perp x$.

Now, we have proven that $u_0' \perp x$ for every $x \in \operatorname{Col} A$. In other words, $u_0' \in (\operatorname{Col} A)^\perp$. This proves the one claim that remained to be proven. Thus, the proof of Proposition 0.1 is complete. □

]

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

We follow the usual method: We set $K = A^T A$ and $f = A^T b$, and then $x = K^{-1} f$. This works because the columns of $A$ are linearly independent.

Here are the computations:

$$K = A^T A = \begin{pmatrix} 3 \\ 4 \end{pmatrix}^T \begin{pmatrix} 3 \\ 4 \end{pmatrix} = \begin{pmatrix} 25 \end{pmatrix}$$

and

$$f = A^T b = \begin{pmatrix} 3 \\ 4 \end{pmatrix}^T \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 7 \end{pmatrix}.$$

Thus,

$$x = K^{-1} f = \begin{pmatrix} 25 \end{pmatrix}^{-1} \begin{pmatrix} 7 \end{pmatrix} = \begin{pmatrix} \frac{7}{25} \end{pmatrix}.$$

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

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

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

**(l)** 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)** T If $A$ is an $n \times m$-matrix, then $\operatorname{rank} A + \dim(\operatorname{Ker} A) = m$.

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

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

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

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

[**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

$$\left(A^T A\right)^T = A^T \underbrace{\left(A^T\right)^T}_{=A} = A^T A.$$

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 $\mathbf{b}$ of size $n$. This basis $\mathbf{b}$ clearly is a spanning list of $V$. Thus, if $\mathbf{a}$ is a linearly independent list of vectors in $V$, then $\mathbf{a}$ must be at most as long as this list $\mathbf{b}$ (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 $v = (v_1, v_2, \ldots, v_n)^T$ and $w = (w_1, w_2, \ldots, w_n)^T$. Then,

$$\langle v, w \rangle = v_1 w_1 + v_2 w_2 + \cdots + v_n w_n \qquad \text{and}$$
$$\langle w, v \rangle = w_1 v_1 + w_2 v_2 + \cdots + w_n v_n,$$

which are clearly equal (because $v_i w_i = w_i v_i$ 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 $n = 1$, $U = \left\{ \overrightarrow{0} \right\}$ 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 $\operatorname{rank}(AB) = n = n$. But Proposition 0.2 **(b)** on homework set #4 (applied to $k$ and $n$ instead of $m$ and $p$) yields $\operatorname{rank}(AB) \leq \operatorname{rank} A$. Finally, the equality (15) on homework set #4 (applied to $k$ instead of $m$) yields $\operatorname{rank} A \leq \min\{n, k\}$. Thus, $n = \operatorname{rank}(AB) \leq \operatorname{rank} A \leq \min\{n, k\} \leq k < n$, 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 $U = \operatorname{Col} A$.

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

$$\left\{\overrightarrow{0}\right\}^\perp = \left\{x \in \mathbb{R}^n \mid x \perp u \text{ for all } u \in \left\{\overrightarrow{0}\right\}\right\} = \left\{x \in \mathbb{R}^n \mid x \perp \overrightarrow{0}\right\} \tag{7}$$

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

**(r)** This is **true**.

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

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

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 $\operatorname{span}(u_1, u_2, \ldots, u_k, v_1, v_2, \ldots, v_\ell)$ (obviously). Hence, a basis of $\operatorname{span}(u_1, u_2, \ldots, u_k, v_1, v_2, \ldots, v_\ell)$ 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,

$$\dim(\operatorname{span}(u_1, u_2, \ldots, u_k, v_1, v_2, \ldots, v_\ell)) \leq k + \ell. \tag{8}$$

Now, I claim that

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

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, $a \in \operatorname{Col} A = \operatorname{span}(u_1, u_2, \ldots, u_k)$ (since $(u_1, u_2, \ldots, u_k)$ is a basis of $\operatorname{Col} A$). 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

$$\begin{aligned} a + b &= (\lambda_1 u_1 + \lambda_2 u_2 + \cdots + \lambda_k u_k) + (\mu_1 v_1 + \mu_2 v_2 + \cdots + \mu_\ell v_\ell) \\ &\in \operatorname{span}(u_1, u_2, \ldots, u_k, v_1, v_2, \ldots, v_\ell). \end{aligned}$$

Thus,

$$w = a + b \in \operatorname{span}(u_1, u_2, \ldots, u_k, v_1, v_2, \ldots, v_\ell).$$

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

$$\operatorname{Col}(A+B) \subseteq \operatorname{span}(u_1, u_2, \ldots, u_k, v_1, v_2, \ldots, v_\ell). \tag{10}$$

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

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

$$\begin{aligned}\dim(\operatorname{Col}(A+B)) &\leq \dim(\operatorname{span}(u_1, u_2, \ldots, u_k, v_1, v_2, \ldots, v_\ell)) \\ &\leq \underbrace{k}_{=\operatorname{rank} A} + \underbrace{\ell}_{=\operatorname{rank} B} \qquad \text{(by (8))} \\ &= \operatorname{rank} A + \operatorname{rank} B.\end{aligned}$$

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