MATH 217 - W24 - LINEAR ALGEBRA
HOMEWORK 8, SOLUTIONS

Part A (15 points)

Solve the following problems from the book:

Section 5.1: 45

Solution.

45. The vector $\text{proj}_V(\vec{v}_1)$ is in $\text{Span}(\vec{v}_2, \vec{v}_3)$, so we can write it as $\text{proj}_V(\vec{v}_1) = a\vec{v}_2 + b\vec{v}_3$ for some $a, b \in \mathbb{R}$. By the definition of proj_V, $\vec{v}_1 - \text{proj}_V(\vec{v}_1) \in V^\perp$. Hence,

$$\begin{cases} (\vec{v}_1 - a\vec{v}_2 - b\vec{v}_3) \cdot \vec{v}_2 = 0 \\ (\vec{v}_1 - a\vec{v}_2 - b\vec{v}_3) \cdot \vec{v}_3 = 0 \end{cases},$$

or equivalently

$$\begin{cases} \vec{v}_1 \cdot \vec{v}_2 - a\vec{v}_2 \cdot \vec{v}_2 - b\vec{v}_3 \cdot \vec{v}_2 = 0 \\ \vec{v}_1 \cdot \vec{v}_3 - a\vec{v}_2 \cdot \vec{v}_3 - b\vec{v}_3 \cdot \vec{v}_3 = 0 \end{cases}.$$

Substituting $\vec{v}_1 \cdot \vec{v}_2 = 5$, $\vec{v}_1 \cdot \vec{v}_3 = 11$, $\vec{v}_2 \cdot \vec{v}_2 = 9$, $\vec{v}_3 \cdot \vec{v}_2 = 20$ and $\vec{v}_3 \cdot \vec{v}_3 = 49$, which we know from the definition of the matrix A, we get

$$\begin{cases} 5 - 9a - 20b = 0 \\ 11 - 20a - 49b = 0 \end{cases}.$$

which has the solution $a = \frac{25}{41}$ and $b = -\frac{1}{41}$. Hence,

$$\text{proj}_V(\vec{v}_1) = \frac{25}{41}\vec{v}_2 - \frac{1}{41}\vec{v}_3.$$

Section 5.2: 14, 26

Solution.

14. Let $\vec{v}_1$, $\vec{v}_2$ and $\vec{v}_3$ be the three vectors given. Then $\|\vec{v}_1\| = 10$, so $\vec{u}_1 = \frac{1}{10}\vec{v}_1$. Next, find $\vec{v}_2^\perp$: $\vec{v}_2^\perp = \vec{v}_2 - (\vec{u}_1 \cdot \vec{v}_2)\vec{u}_1$, and we note $\vec{u}_1 \cdot \vec{v}_2 = 10$, so we have

$$\vec{v}_2^\perp = \begin{bmatrix} 0 & 7 & 2 & 7 \end{bmatrix}^T - \begin{bmatrix} 1 & 7 & 1 & 7 \end{bmatrix}^T = \begin{bmatrix} -1 & 0 & 1 & 0 \end{bmatrix}^T.$$

This has length $\|\vec{v}_2^\perp\| = \sqrt{2}$, so $\vec{u}_2 = \frac{1}{\sqrt{2}}\begin{bmatrix} -1 & 0 & 1 & 0 \end{bmatrix}^T$. Finally, we find $\vec{v}_3^\perp = \vec{v}_3 - (\vec{u}_1 \cdot \vec{v}_3)\vec{u}_1 - (\vec{u}_2 \cdot \vec{v}_3)\vec{u}_2$. We note that $\vec{u}_1 \cdot \vec{v}_3 = 10$ and $\vec{u}_2 \cdot \vec{v}_3 = 0$, so

$$\vec{v}_3^\perp = \begin{bmatrix} 1 & 8 & 1 & 6 \end{bmatrix}^T - \begin{bmatrix} 1 & 7 & 1 & 7 \end{bmatrix}^T = \begin{bmatrix} 0 & 1 & 0 & -1 \end{bmatrix}^T.$$

Again, $\|\vec{v}_3\| = \sqrt{2}$, so $\vec{u}_3 = \frac{1}{\sqrt{2}}\begin{bmatrix} 0 & 1 & 0 & -1 \end{bmatrix}^T$.

26. Let $\vec{v}_1 = \begin{bmatrix} 2 \\ 3 \\ 0 \\ 6 \end{bmatrix}$, $\vec{v}_2 = \begin{bmatrix} 4 \\ 4 \\ 2 \\ 13 \end{bmatrix}$ be the columns of the given matrix M. The sequence of orthonormal vectors produced by the Gram-Schmidt process is:

$$\vec{u}_1 = \frac{\vec{v}_1}{\|\vec{v}_1\|} = \frac{1}{7}\begin{bmatrix} 2 \\ 3 \\ 0 \\ 6 \end{bmatrix}, \quad \vec{v}_2^{\perp} = \vec{v}_2 - (\vec{v}_2 \cdot \vec{u}_1)\vec{u}_1 = \vec{v}_2 - (14)\vec{u}_1 = \begin{bmatrix} 0 \\ -2 \\ 2 \\ 1 \end{bmatrix},$$

$$\vec{u}_2 = \frac{\vec{v}_2^{\perp}}{\|\vec{v}_2^{\perp}\|} = \frac{1}{3}\begin{bmatrix} 0 \\ -2 \\ 2 \\ 1 \end{bmatrix}.$$

Hence, the QR factorization of M is given by:

$$M = \begin{bmatrix} 2 & 4 \\ 3 & 4 \\ 0 & 2 \\ 6 & 13 \end{bmatrix} = \begin{bmatrix} 2/7 & 0 \\ 3/7 & -2/3 \\ 0 & 2/3 \\ 6/7 & 1/3 \end{bmatrix} \begin{bmatrix} 7 & 14 \\ 0 & 3 \end{bmatrix} = QR.$$

Section 5.3: 36

Solution.
36. In order to find an orthogonal matrix of the form

$$\begin{bmatrix} 2/3 & 1/\sqrt{2} & a \\ 2/3 & -1/\sqrt{2} & b \\ 1/3 & 0 & c \end{bmatrix},$$

we need to find a unit vector that is orthogonal to each of the first two columns. We can do this using the cross product:

$$\begin{bmatrix} 2/3 \\ 2/3 \\ 1/3 \end{bmatrix} \times \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \\ 0 \end{bmatrix} = \begin{bmatrix} \sqrt{2}/6 \\ \sqrt{2}/6 \\ -2\sqrt{2}/3 \end{bmatrix}.$$

Note: Since the first two columns span a two-dimensional subspace V in $\mathbb{R}^3$, its orthogonal complement $V^{\perp}$ has dimension one. This means there are only two possible choices for a unit vector in $V^{\perp}$, and hence only two possible choices for the third column in the matrix. That is, the only two possible *orthogonal* matrices of the given form are:

$$\begin{bmatrix} 2/3 & 1/\sqrt{2} & \sqrt{2}/6 \\ 2/3 & -1/\sqrt{2} & \sqrt{2}/6 \\ 1/3 & 0 & -2\sqrt{2}/3 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 2/3 & 1/\sqrt{2} & -\sqrt{2}/6 \\ 2/3 & -1/\sqrt{2} & -\sqrt{2}/6 \\ 1/3 & 0 & 2\sqrt{2}/3 \end{bmatrix}.$$

Section 5.4: 26, 32.

Solution.

26. The least squares solutions of

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \vec{x} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

are the solutions to the (consistent) system

$$\begin{bmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \vec{x} = \begin{bmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

which we can re-write as

$$\begin{bmatrix} 66 & 78 & 90 \\ 78 & 93 & 108 \\ 90 & 108 & 126 \end{bmatrix} \vec{x} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

and solve by row-reducing the augmented matrix

$$\begin{bmatrix} 66 & 78 & 90 & 1 \\ 78 & 93 & 108 & 2 \\ 90 & 108 & 126 & 3 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 0 & -1 & -7/6 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Thus we obtain:

$$\left\{ \begin{bmatrix} -7/6 \\ 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} t \ : \ t \in \mathbb{R} \right\}.$$

32. Fit a quadratic polynomial to the data points $(0, 27)$, $(1, 0)$, $(2, 0)$, $(3, 0)$ using least squares. Sketch a solution. If we name our quadratic polynomial $y = ax^2 + bx + c$, we are looking to least squares solutions to

$$\begin{bmatrix} 0 & 0 & 1 \\ 1 & 1 & 1 \\ 4 & 2 & 1 \\ 9 & 3 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 27 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

Really, we need solutions to

$$\begin{bmatrix} 0 & 1 & 4 & 9 \\ 0 & 1 & 2 & 3 \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ 1 & 1 & 1 \\ 4 & 2 & 1 \\ 9 & 3 & 1 \end{bmatrix} \vec{x} = \begin{bmatrix} 0 & 1 & 4 & 9 \\ 0 & 1 & 2 & 3 \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 27 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

which amounts to row-reducing the augmented matrix

$$\begin{bmatrix} 98 & 36 & 14 & 0 \\ 36 & 14 & 6 & 0 \\ 14 & 6 & 4 & 27 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 0 & 0 & 27/4 \\ 0 & 1 & 0 & -567/20 \\ 0 & 0 & 1 & 513/20 \end{bmatrix}$$

Thus, $y = \frac{27}{4}x^2 - \frac{567}{20}x + \frac{513}{20}$ or equivalently $y = 6.75x^2 - 28.35x + 25.65$ is the quadratic polynomial which best fits the given data points via least-squares.

Part B (25 points)

Problem 1. Let W be a subspace of $\mathbb{R}^n$ and let $\mathcal{B} = (\vec{v}_1, \ldots, \vec{v}_d)$ be a basis for W. Consider the transformation $\mathbb{R}^n \xrightarrow{\pi} \mathbb{R}^n$ defined by

$$\pi(\vec{v}) = \sum_{i=1}^{d} \frac{\vec{v} \cdot \vec{v}_i}{\vec{v}_i \cdot \vec{v}_i}\ \vec{v}_i.$$

(a) Show that if $\vec{v}_i \cdot \vec{v}_j = 0$ for all $1 \leq i \neq j \leq d$, then the transformation π is the orthogonal projection onto W. (Note: this is almost, *but not quite*, the way we defined orthogonal projection. Make sure you understand how our definition is different from this before you start trying to prove it!)

(b) Give a counterexample to show that if the basis vectors in $\mathcal{B}$ are *not* perpendicular to each other, then the linear transformation π defined above π is *not* orthogonal projection onto W .

Solution.

(a) We introduce a second basis $\mathcal{C} = (\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_d)$ for W by defining $\vec{u}_i = \frac{\vec{v}_i}{||\vec{v}_i||}$. Then each $\vec{u}_i$ is automatically a unit vector, and this combined with the condition $\vec{v}_i \cdot \vec{v}_j = 0$ implies that $\mathcal{C}$ is an orthonormal basis. By definition, then, the orthogonal projection onto W is

$$\begin{aligned}
\pi_W(\vec{v}) &= \sum_{i=1}^{d} (\vec{v} \cdot \vec{u}_i)\vec{u}_i \\
&= \sum_{i=1}^{d} \left(\vec{v} \cdot \frac{\vec{v}_i}{||\vec{v}_i||}\right) \frac{\vec{v}_i}{||\vec{v}_i||} \\
&= \sum_{i=1}^{d} \left(\frac{\vec{v} \cdot \vec{v}_i}{||\vec{v}_i||^2}\right) \vec{v}_i \\
&= \sum_{i=1}^{d} \frac{\vec{v} \cdot \vec{v}_i}{\vec{v}_i \cdot \vec{v}_i}\ \vec{v}_i \\
&= \pi(\vec{v})
\end{aligned}$$

(b) Let $W \subseteq \mathbb{R}^3$ be the subspace spanned by $\{\vec{e}_1, \vec{e}_2\}$, and choose the non-orthonormal basis $\mathcal{B} = (\vec{e}_1, \vec{e}_1 + \vec{e}_2)$. Then for any $\vec{v} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$, we have $\pi_W(\vec{v}) = \begin{bmatrix} a \\ b \\ 0 \end{bmatrix}$. However,

$$\pi_W\left(\begin{bmatrix} a \\ b \\ c \end{bmatrix}\right) = \frac{a}{1}\vec{e}_1 + \frac{a+b}{2}(\vec{e}_1 + \vec{e}_2) = \begin{bmatrix} \frac{3a+b}{2} \\ \frac{b}{2} \\ 0 \end{bmatrix}$$

and these are obviously not the same; $a = b = 1$ provides an explicit counterexample.

Problem 2. Let $\mathcal{O}_n \subseteq \mathbb{R}^{n \times n}$ denote the set of orthogonal $n \times n$ matrices. Determine whether each of the following statements is True or False, and provide a short proof (or a counter-example) of your claim.

(a) $\mathcal{O}_n$ is a subspace of $\mathbb{R}^{n\times n}$.
(b) If $A, B \in \mathcal{O}_n$, then $AB \in \mathcal{O}_n$.
(c) If $A \in \mathcal{O}_n$, then $A^2 \in \mathcal{O}_n$.
(d) If $A^2 \in \mathcal{O}_n$, then $A \in \mathcal{O}_n$.
(e) If $A \in \mathcal{O}_n$ and A^2 is the identity matrix, then A is symmetric.

Solution.

(a) FALSE. $\mathcal{O}_n$ does not contain the zero matrix; nor is it closed under addition or scalar multiplication. (For an explicit example: I_n is an orthogonal matrix, and $-I_n$ is an orthogonal matrix, but their sum is the zero matrix, which is not orthogonal. SImilarly $2I_n$ is not an orthogonal matrix.)

(b) TRUE. Suppose A and B are orthogonal $n \times n$ matrices, so A and B are invertible and $A^{-1} = A^\top$ and $B^{-1} = B^\top$. Then AB is also invertible, since products of invertible matrices are invertible, and we have $(AB)^{-1} = B^{-1}A^{-1} = B^\top A^\top = (AB)^\top$. This shows that AB is also an orthogonal matrix.

(c) TRUE. This is a straightforward consequence of (b).

(d) FALSE: A^2 could be orthogonal without A being orthogonal. To see this, let $A = \begin{bmatrix} 0 & \frac{1}{2} \\ 2 & 0 \end{bmatrix}$. Then A is not orthogonal since its columns are not unit vectors, but $A^2 = I_2$, which is orthogonal.

(e) TRUE. Let A be an orthogonal $n \times n$ matrix such that $A^2 = I_n$. Then

$$A^\top \;=\; A^\top(A^2) \;=\; A^{-1}AA \;=\; I_nA \;=\; A,$$

so A is symmetric.

Problem 3. (a) Suppose that $\mathcal{B} = (\vec{b}_1, \ldots, \vec{b}_r)$ is an orthonormal basis of a subspace V of $\mathbb{R}^n$. Prove that for all $\vec{v}, \vec{w} \in V$, $[\vec{v}]_\mathcal{B} \cdot [\vec{w}]_\mathcal{B} = \vec{v} \cdot \vec{w}$.

(b) Prove that if $\mathcal{B} = (\vec{b}_1, \ldots, \vec{b}_r)$ and $\mathcal{C} = (\vec{c}_1, \ldots, \vec{c}_r)$ are two orthonormal bases of V, then $S_{\mathcal{B}\to\mathcal{C}}$ is an orthogonal $r \times r$ matrix.

Solution.

(a) Fix $\vec{v}, \vec{w} \in V$ and let $\vec{v} = c_1\vec{b}_1 + \cdots + c_r\vec{b}_r$ and $\vec{w} = d_1\vec{b}_1 + \cdots + d_r\vec{b}_r$, where $c_1, \ldots, c_r, d_1, \ldots d_r \in \mathbb{R}$. Then:

$$[\vec{v}]_\mathcal{B} = \begin{bmatrix} c_1 \\ \vdots \\ c_r \end{bmatrix} \qquad \text{and} \qquad [\vec{w}]_\mathcal{B} = \begin{bmatrix} d_1 \\ \vdots \\ d_r \end{bmatrix}.$$

It follows from the formula for the dot product that:

$$[\vec{v}]_\mathcal{B} \cdot [\vec{w}]_\mathcal{B} = c_1d_1 + \cdots + c_rd_r.$$

Now we'll compute $\vec{v}\cdot\vec{w}$ using the basis expansion above:

$$\begin{aligned}\vec{v}\cdot\vec{w} &= (c_1\vec{b}_1+\cdots+c_r\vec{b}_r)\cdot(d_1\vec{b}_1+\cdots+d_r\vec{b}_r)\\ &= \sum_{i=1}^{r} c_id_i(\vec{b}_i\cdot\vec{b}_i)+\sum_{k\neq j} c_kd_j(\vec{b}_k\cdot\vec{b}_j)\\ &= \sum_{i=1}^{r} c_id_i(\vec{b}_i\cdot\vec{b}_i)\\ &= c_1d_1+\cdots+c_rd_r,\end{aligned}$$

where in the second equality we used the linearity of the dot product, the third equality used the fact that the $\vec{b}_i$ are orthogonal to one another, and the fourth equality used the fact that each $\vec{b}_i$ has length one. Thus, $\vec{v}\cdot\vec{w}=[\vec{v}]_{\mathcal{B}}\cdot[\vec{w}]_{\mathcal{B}}$.

(b) Let $L_{\mathcal{B}}\colon V\to\mathbb{R}^r$ be the coordinate isomorphism for the $\mathcal{B}$-coordinates and similarly let $L_{\mathcal{C}}\colon V\to\mathbb{R}^r$ be the $\mathcal{C}$-coordinate isomorphism. Then $S_{\mathcal{B}\to\mathcal{C}}$ is the matrix corresponding to the linear transformation $L_{\mathcal{C}}\circ L_{\mathcal{B}}^{-1}\colon\mathbb{R}^r\to\mathbb{R}^r$. Moreover, by part (a) we have $[\vec{v}]_{\mathcal{B}}\cdot[\vec{w}]_{\mathcal{B}}=\vec{v}\cdot\vec{w}=[\vec{v}]_{\mathcal{C}}\cdot[\vec{w}]_{\mathcal{C}}$ for every pair $\vec{v},\vec{w}\in V$. Now fix $\vec{x},\vec{y}\in\mathbb{R}^r$. Then there exist $\vec{v},\vec{w}\in V$ such that $[\vec{v}]_{\mathcal{B}}=\vec{x}$ and $[\vec{w}]_{\mathcal{B}}=\vec{y}$. Therefore, $\vec{x}\cdot\vec{y}=\vec{v}\cdot\vec{w}=[\vec{v}]_{\mathcal{C}}\cdot[\vec{w}]_{\mathcal{C}}$. Finally, we have:

$$\begin{aligned}S_{\mathcal{B}\to\mathcal{C}}\vec{x}\cdot S_{\mathcal{B}\to\mathcal{C}}\vec{y} &= [\vec{v}]_{\mathcal{C}}\cdot[\vec{w}]_{\mathcal{C}}\\ &= \vec{x}\cdot\vec{y}.\end{aligned}$$

It follows from Summary 5.3.8 that $S_{\mathcal{B}\to\mathcal{C}}$ is orthogonal.

Problem 4. Let A be an $n\times m$ matrix. Prove or disprove each of the following statements:

(a) $(\ker A)^{\perp}=\operatorname{im}A^{\top}$.
(b) $\operatorname{Rank}(A)=\operatorname{Rank}(A^{\top}A)$.
(c) $\operatorname{Rank}(A)=\operatorname{Rank}(A^{\top})$.
(d) $\operatorname{Rank}(A^{\top}A)=\operatorname{Rank}(AA^{\top})$.
(e) $\ker A=\ker AA^{\top}$.

Solution.

(a) True. By Theorem 5.4.1, we have $(\ker A^{\top})=\operatorname{im}A^{\perp}$ for all matrices. Apply this to the matrix $A^{\top}$ to get $(\ker(A^{\top})^{\top})=(\operatorname{im}A^{\top})^{\perp}$, or equivalently, $\ker A=(\operatorname{im}A^{\top})^{\perp}$. Now we can "perp" both sides (Proved on Homework 7) to get $(\ker A)^{\perp}=(\operatorname{im}A^{\top})$.

(b) True! By Theorem 5.4.2, $\ker A=\ker A^{\top}A$. Both these matrices/transformations have m-dimensional sources, so by Rank-Nullity, again they have the same rank.

(c) True. By Rank-nullity, we know rank $(A)=m-\dim\ker A$. This is also the dimension of $(\ker A)^{\perp}=\operatorname{im}A^{\top}$. So the rank of $A^{\top}$ is $\dim\operatorname{im}A^{\top}=m-\dim\ker A=\operatorname{rank}(A)$.

(d) True! Apply (b) to $A^{\top}$ to see that Rank $(A^{\top})$ = Rank $(AA^{\top})$. So by (c) this is Rank (A) and by (b) again it is Rank $(A^{\top}A)$.

(e) False! The sources of A and $A^{\top}A$ are $\mathbb{R}^m$ and $\mathbb{R}^n$ respectively, so the kernels live in different spaces and can not be equal.

For Problem 5, you will need the following definitions:

Definition. If A and B are two subsets of $\mathbb{R}^n$, then we say $A \perp B$ if for all $\vec{x} \in A$ and for all $\vec{y} \in B$, $\vec{x} \cdot \vec{y} = 0$. (Note that in this definition that A and B do not need to be subspaces, just subsets.)

Definition. A subset $A \subseteq \mathbb{R}^n$ is called *pairwise orthogonal* if any two elements $\vec{x}, \vec{y} \in A$ are orthogonal. Such a pairwise orthogonal subset $A \subseteq \mathbb{R}^n$ is called *maximally pairwise orthogonal* if it is not possible to enlarge set A to obtain a pairwise orthogonal subset $A' \subseteq \mathbb{R}^n$ that strictly contains A.

Problem 5. Let $n \in \mathbb{N}$. We consider the vector space $\mathbb{R}^n$.

(a) Prove that for all $X, Y \subseteq \mathbb{R}^n$, if $X \perp Y$ then $\mathrm{Span}(X) \perp \mathrm{Span}(Y)$.
(b) Let X and Y each be a linearly independent subset of $\mathbb{R}^n$. Prove that if $X \perp Y$, then $X \cup Y$ is linearly independent.
(c) Prove that every maximally pairwise orthogonal set of vectors in $\mathbb{R}^n$ has $n+1$ elements.

Solution.

(a) For every $\vec{v} \in \mathrm{Span}(X)$, there are $c_1, \ldots, c_k \in \mathbb{R}$ and $\vec{v}_1, \ldots, \vec{v}_k \in X$ such that $\vec{v} = c_1\vec{v}_1 + \cdots + c_k\vec{v}_k$. For every $\vec{w} \in \mathrm{Span}(Y)$, there are $d_1, \ldots, d_m \in \mathbb{R}$ and $\vec{w}_1, \ldots, \vec{w}_m \in Y$ such that $\vec{w} = d_1\vec{w}_1 + \cdots + d_m\vec{w}_m$. It follows that $\vec{v} \cdot \vec{w} = \sum_{j,l} c_j d_j (\vec{v}_j \cdot \vec{w}_l) = 0$, wehre the last equality follows from $X \perp Y$. Hence, $\mathrm{Span}(X) \perp \mathrm{Span}(Y)$.

(b) Suppose that $X = \{\vec{v}_1, \ldots, \vec{v}_k\}$ and $Y = \{\vec{w}_1, \ldots, \vec{w}_m\}$. Suppose $c_1, \ldots, c_k \in \mathbb{R}$ and $d_1, \ldots, d_m \in \mathbb{R}$, such that

$$c_1\vec{v}_1 + \cdots + c_k\vec{v}_k + d_1\vec{w}_1 + \cdots + d_m\vec{w}_m = \vec{0}.$$

By (a), we have that $\vec{v} := c_1\vec{v}_1 + \cdots + c_k\vec{v}_k$ is perpendicular to $\vec{w} = d_1\vec{w}_1 + \cdots + d_m\vec{w}_m$. Since $\vec{v} + \vec{w} = \vec{0}$ we have $\vec{w} = -\vec{v}$, and since $\vec{v} \cdot \vec{w} = 0$ we conclude $-\vec{v} \cdot \vec{v} = 0$. It follows that $\vec{v} = \vec{0}$ (and therefore $\vec{w} = \vec{0}$ as well).

Now we have $\vec{v} = c_1\vec{v}_1 + \cdots + c_k\vec{v}_k = \vec{0}$, which implies that $c_1 = \cdots = c_k = 0$, because X is linearly independent. Similarly, we have $\vec{w} = d_1\vec{w}_1 + \cdots + d_m\vec{w}_m = \vec{0}$, and therefore $d_1 = \cdots = d_m = 0$, because Y is linearly independent. So the original relation is trivial, and therefore $X \cup Y$ is linearly independent.

(c) Let Z be a maximally pairwise orthogonal set of vectors in $\mathbb{R}^n$. First, we note that $\vec{0} \in \mathbb{R}^n$, otherwise Z will not be maximal. Suppose that Z has $k+1$ vectors. It follows that Z has k non-zero vectors. Applying (b) inductively, we conclude that $Z - \{\vec{0}\}$ is linearly independent. Thus, $\mathrm{Span}(Z)$ has k dimensions. If $k < n$ then $\mathrm{Span}(Z)^{\perp}$, dimension $n - k > 0$, contains a nonzero vector $\vec{v}$ that is perpendicular to all vectors in Z. Hence $Z \cup \{\vec{v}\}$ is a pairwise orthogonal set of vectors that properly contains Z. This contradicts with Z being maximal. Hence $k = n$.

Problem 6. Let A be an $n \times m$ matrix, with $m \leq n$.

(a) If $\mathrm{rank}(A) = m$, prove that it is always possible to write $A = QL$, where Q is an $n \times m$ matrix with orthonormal columns and L is a **lower** triangular $m \times m$ matrix with positive diagonal entries.
(b) Prove that if $\mathrm{rank}(A) < m$, it is still possible to obtain such a decomposition if we allow some diagonal entries to be zero.

Solution.

(a) Recalling that the QR factorization of an $n \times m$ matrix A with $\operatorname{rank}(A) = m$ (that is, m linearly independent columns) is obtained by applying the Gram-Schmidt algorithm successively to the column vectors $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_m$ of A, we apply Gram-Schmidt to the columns of A starting with $\vec{v}_m$. Let $\mathfrak{B}$ be the basis of a subspace of $\mathbb{R}^n$ spanned by the column vectors of A, and $\mathfrak{U}$ be the orthonormal basis we obtain as follows.

Normalizing $\vec{v}_m$, we have $\vec{u}_m = \vec{v}_m/\|\vec{v}_m\|$. Then the component of $\vec{v}_{m-1}$ that is perpendicular to this is $\vec{v}_{m-1}^{\perp} = \vec{v}_{m-1} - (\vec{u}_m \cdot \vec{v}_{m-1})\vec{u}_m$, and $\vec{u}_{m-1} = \vec{v}_{m-1}^{\perp}/\|\vec{v}_{m-1}^{\perp}\|$. We may proceed similarly to obtain the remaining orthonormal vectors $\vec{u}_{m-1}, \ldots, \vec{u}_1$.

Letting $Q = \begin{bmatrix} \vec{u}_1 & \cdots & \vec{u}_m \end{bmatrix}$, theorem 4.3.4 ($B = AS$ if $B = [$ basis 1 $]$ and $A = [$ basis 2 $]$) shows that $A = QL$, where L is the change of basis matrix between the bases $\mathfrak{B}$ and $\mathfrak{U}$. It remains to be shown that L is lower triangular with positive diagonal entries. Note that the jth column of L is the coordinate vector $[\vec{v}_j]_{\mathfrak{U}}$. Then, considering the projection of $\vec{v}_j$ onto the subspace $U = \operatorname{span}(\vec{u}_{j+1}, \ldots, \vec{u}_m)$ and using the fact from our derivation that $\vec{u}_j = \vec{v}_j^{\perp}/\|\vec{v}_j^{\perp}\|$, we have

$$\begin{aligned} \vec{v}_j &= \vec{v}_j^{\perp} + \vec{v}_j^{\|} \\ &= \|\vec{v}_j^{\perp}\|\vec{u}_j + \big((\vec{u}_{j+1} \cdot \vec{v}_j)\vec{u}_{j+1} + \cdots + (\vec{u}_m \cdot \vec{v}_j)\vec{u}_m\big). \end{aligned}$$

Thus the jth column of L is $\begin{bmatrix} 0 \\ \vdots \\ 0 \\ \|\vec{v}_j^{\perp}\| \\ \vec{u}_{j+1} \cdot \vec{v}_j \\ \vdots \\ \vec{u}_m \cdot \vec{u}_j \end{bmatrix}$, and our result is obtained. □

(b) Suppose that $\operatorname{rank}(A) = k < m$. Suppose for convenience that the last k columns of A are linearly independent; we show below that the following argument applies when this is not the case.

Applying the Gram-Schmidt process as in **(a)**, we find an orthonormal set $\vec{u}_{m-k+1}, \ldots, \vec{u}_m$. Because $m < n$, we can extend this to an orthonormal set $\mathfrak{U} = \{\vec{u}_1, \ldots, \vec{u}_m\}$, which spans a subspace U of $\mathbb{R}^n$. Clearly each $\vec{v}_j \in \operatorname{span}(u_{m-k+1}, \ldots, \vec{u}_m) \subseteq \operatorname{span}(u_1, \ldots, \vec{u}_m)$. Letting $Q = \begin{bmatrix} u_1 & \cdots & u_m \end{bmatrix}$, our definition of coordinates requires that $\vec{v}_j = Q[\vec{v}_j]_{\mathfrak{U}}$, and thus (invoking Theorem 2.3.2, columns of the product AB) that $A = QL$, where L is the matrix $L = \begin{bmatrix} [\vec{v}_1]_{\mathfrak{U}} & \cdots & [\vec{v}_m]_{\mathfrak{U}} \end{bmatrix}$.

From our work in **(a)**, the last k columns of L are clearly lower-triangular. Then, because all $\vec{v}_j \in \operatorname{span}(u_{m-k+1}, \ldots, \vec{u}_m)$, the top $k-1$ entries of the first $m-k$ columns L must be zero, and L must in fact be lower-triangular.

Finally, note that this argument works for any ordering of the vectors $\vec{v}_1, \ldots, \vec{v}_m$. Letting $\vec{u}_{n-k}, \ldots, \vec{u}_m$ be the orthonormalization of the last k linearly independent vectors guarantees that L remains lower-triangular. □