Part A (15 points)

Solve the following problems from the book:

Section 5.1: 45
Section 5.2: 14, 26
Section 5.3: 36
Section 5.4: 26, 32.

5-1

In Exercises 40 through 46, consider vectors $\vec{v}_1, \vec{v}_2, \vec{v}_3$ in $\mathbb{R}^4$; we are told that $\vec{v}_i \cdot \vec{v}_j$ is the entry a_{ij} of matrix A.

$$A = \begin{bmatrix} 3 & 5 & 11 \\ 5 & 9 & 20 \\ 11 & 20 & 49 \end{bmatrix}$$

45. Find $\text{proj}_V(\vec{v}_1)$, where $V = \text{span}(\vec{v}_2, \vec{v}_3)$. Express your answer as a linear combination of $\vec{v}_2$ and $\vec{v}_3$.

Sol. Let $\text{proj}_V(\vec{v_1}) = C_2\vec{V_2} + C_3\vec{V_3}$

So $(\vec{V_1} - \text{proj}_V(\vec{v_1})) \in V^\perp$

Therefore $\begin{cases} (\vec{V_1} - C_2\vec{V_2} - C_3\vec{V_3}) \cdot \vec{V_2} = 0 & ① \\ (\vec{V_1} - C_2\vec{V_2} - C_3\vec{V_3}) \cdot \vec{V_3} = 0 & ② \end{cases}$

$① \Rightarrow C_2 \cdot a_{22} + C_3 a_{32} = 5 \Rightarrow 9C_2 + 20C_3 = 5$

$② \Rightarrow C_2 \cdot a_{23} + C_3 a_{33} = 11 \Rightarrow 20C_2 + 49C_3 = 11$

$\Rightarrow \begin{cases} C_2 = \frac{25}{41} \\ C_3 = \frac{1}{-41} \end{cases} \Rightarrow \text{proj}_V(\vec{V_1}) = \frac{25}{41}\vec{V_2} - \frac{1}{41}\vec{V_3}$

5-2

Using paper and pencil, perform the Gram–Schmidt process on the sequences of vectors given in Exercises 1 through 14.

14. $\begin{bmatrix}1\\7\\1\\7\end{bmatrix}, \begin{bmatrix}0\\7\\2\\7\end{bmatrix}, \begin{bmatrix}1\\8\\1\\6\end{bmatrix}$

<u>Sol</u> $\vec{u_1} = \frac{1}{\sqrt{98+2}}\vec{v_1} = \begin{bmatrix}\frac{1}{10}\\ \frac{7}{10}\\ \frac{1}{10}\\ \frac{7}{10}\end{bmatrix}$

$$\vec{u_2} = \frac{\vec{v_2}^{\perp}}{\|\vec{v_2}^{\perp}\|} = \frac{\vec{v_2}-(\vec{v_2}\cdot\vec{u_1})\vec{u_1}}{\|\vec{v_2}-(\vec{v_2}\cdot\vec{u_1})\vec{u_1}\|} = \frac{1}{\sqrt{2}}\begin{bmatrix}-1\\0\\1\\0\end{bmatrix} = \begin{bmatrix}-\frac{\sqrt{2}}{2}\\0\\ \frac{\sqrt{2}}{2}\\0\end{bmatrix}$$

$$\vec{u_3} = \frac{\vec{v_3}^{\perp}}{\|\vec{v_3}^{\perp}\|} = \frac{\vec{v_3}-(\vec{v_3}\cdot\vec{u_1})\vec{u_1}-(\vec{v_3}\cdot\vec{u_2})\vec{u_2}}{\|\vec{v_3}-(\vec{v_3}\cdot\vec{u_1})\vec{u_1}-(\vec{v_3}\cdot\vec{u_2})\vec{u_2}\|} = \frac{1}{\sqrt{2}}\begin{bmatrix}0\\1\\0\\-1\end{bmatrix} = \begin{bmatrix}0\\ \frac{\sqrt{2}}{2}\\0\\-\frac{\sqrt{2}}{2}\end{bmatrix}$$

$\Rightarrow$ orthonormal basis $(\vec{u_1}, \vec{u_2}, \vec{u_3})$

$$= \left(\begin{bmatrix}\frac{1}{10}\\ \frac{7}{10}\\ \frac{1}{10}\\ \frac{7}{10}\end{bmatrix}, \begin{bmatrix}-\frac{\sqrt{2}}{2}\\0\\ \frac{\sqrt{2}}{2}\\0\end{bmatrix}, \begin{bmatrix}0\\ \frac{\sqrt{2}}{2}\\0\\-\frac{\sqrt{2}}{2}\end{bmatrix}\right)$$

Using paper and pencil, find the QR factorizations of the matrices in Exercises 15 through 28. Compare with Exercises 1 through 14.

26. $\begin{array}{cc} \vec{m_1} & \vec{m_2} \\ \end{array}$
$$\begin{bmatrix} 2 & 4 \\ 3 & 4 \\ 0 & 2 \\ 6 & 13 \end{bmatrix}$$

Sol

$$\vec{u_1} = \frac{1}{7}\begin{bmatrix} 2 \\ 3 \\ 0 \\ 6 \end{bmatrix} = \begin{bmatrix} \frac{2}{7} \\ \frac{3}{7} \\ 0 \\ \frac{6}{7} \end{bmatrix}$$

$$\vec{u_2} = \frac{\vec{v_2} - (\vec{v_2} \cdot \vec{u_1})\vec{u_1}}{\| \vec{v_2} - (\vec{v_2} \cdot \vec{u_1})\vec{u_1} \|} = \frac{1}{3}\begin{bmatrix} 0 \\ -2 \\ 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ -\frac{2}{3} \\ \frac{2}{3} \\ \frac{1}{3} \end{bmatrix}$$

So $$Q = \begin{bmatrix} \frac{2}{7} & 0 \\ \frac{3}{7} & -\frac{2}{3} \\ 0 & \frac{2}{3} \\ \frac{6}{7} & \frac{1}{3} \end{bmatrix}$$

$$R = S_{M \to Q} = \begin{bmatrix} \vec{m_1} \cdot \vec{u_1} & \vec{m_1} \cdot \vec{u_2} \\ 0 & \vec{m_2} \cdot \vec{u_2} \end{bmatrix} = \begin{bmatrix} 7 & 14 \\ 0 & 3 \end{bmatrix}$$

5-3

36. Find an orthogonal matrix of the form

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

By WS 18, the roles of an orthonormal matrix are orthonormal

So $\frac{4}{9} - \frac{1}{2} + ab = 0$

$\frac{2}{9} + ac = 0$

$\frac{2}{9} + bc = 0$

So $a = b = \sqrt{\frac{1}{18}} = \frac{\sqrt{2}}{6}$

$\Rightarrow c = -\frac{2}{9} \cdot \frac{6}{\sqrt{2}} = -\frac{2\sqrt{2}}{3}$

5-4 **26.** Find the least-squares solutions $\vec{x}^*$ of the system $A\vec{x} = \vec{b}$, where

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \quad \text{and} \quad \vec{b} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}.$$

normal equation: $A^T A \vec{x} = A^T \vec{b}$

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

$$\left[\begin{array}{ccc:c} 66 & 78 & 90 & 1 \\ 78 & 93 & 108 & 2 \\ 90 & 108 & 126 & 3 \end{array}\right] \xrightarrow{rref} \left[\begin{array}{ccc:c} 1 & 0 & -1 & -\frac{7}{6} \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

$$\Rightarrow \vec{x}^* = \begin{bmatrix} -\frac{7}{6}+t \\ 1-2t \\ t \end{bmatrix} = \begin{bmatrix} -\frac{7}{6} \\ 1 \\ 0 \end{bmatrix} + t\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$$

32. Fit a quadratic polynomial to the data points $(0, 27)$, $(1, 0)$, $(2, 0)$, $(3, 0)$, using least squares. Sketch the solution.

Sol $f(x) = C_0 + C_1 t + C_2 t^2$

let $\beta = (1, t, t^2)$ be a basis

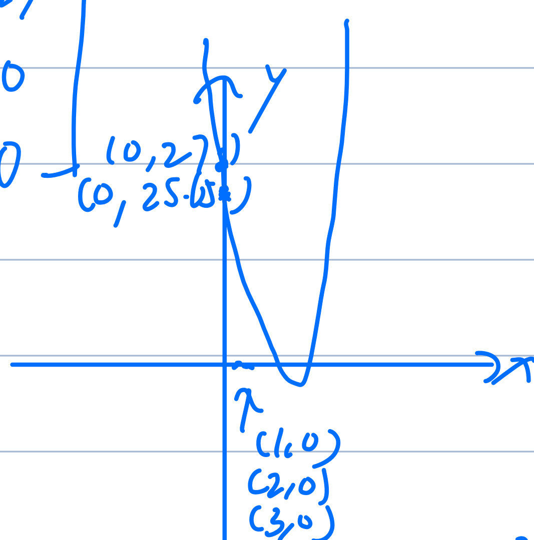
So $[f(x)]_\beta = \begin{bmatrix} C_0 \\ C_1 \\ C_2 \end{bmatrix}$

So $A = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \end{bmatrix}$, $\vec{b} = \begin{bmatrix} 27 \\ 0 \\ 0 \\ 0 \end{bmatrix}$

normal equation: $A^T A \vec{x} = A^T \vec{b}$

$$\begin{bmatrix} 4 & 6 & 14 \\ 6 & 14 & 36 \\ 14 & 36 & 98 \end{bmatrix} \vec{x} = \begin{bmatrix} 27 \\ 0 \\ 0 \end{bmatrix}$$

$$\Rightarrow \vec{x}^* = \begin{bmatrix} \frac{513}{20} \\ \frac{-567}{20} \\ \frac{27}{4} \end{bmatrix}$$



So $f^*(x) = \frac{573}{20} + \frac{-567}{20} x + \frac{27}{4} x^2 = 25.65 - 28.35x + 6.75x^2$

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.

(a) Assume $\forall\, 1 \leq i \neq j \leq d$, $\vec{v_i} \cdot \vec{v_j} = 0$

Then $(\vec{v_1}, \ldots, \vec{v_d})$ are orthogonal

So $\left(\frac{\vec{v_1}}{\|\vec{v_1}\|}, \frac{\vec{v_2}}{\|\vec{v_2}\|}, \ldots, \frac{\vec{v_d}}{\|\vec{v_d}\|}\right)$ is an orthonormal basis by definition

$\forall\, 1 \leq i \leq d$, denote $\frac{\vec{v_i}}{\|\vec{v_i}\|}$ by $\vec{u_i}$, so $(\vec{u_1}, \ldots, \vec{u_d})$ is an orthonormal basis of W

Therefore $\pi(\vec{v}) = \sum_{i=1}^{d} \frac{\vec{v} \cdot \vec{v_i}}{\|\vec{v_i}\|^2} \cdot \vec{v_i}$

$= \sum_{i=1}^{d} \left(\vec{v} \cdot \frac{\vec{v_i}}{\|\vec{v_i}\|}\right) \frac{\vec{v_i}}{\|\vec{v_i}\|} = \sum_{i=1}^{d} (\vec{v} \cdot \vec{u_i}) \vec{u_i}$

So π is the orthogonal projection onto W by definition.

(b) Assume the basis vector are not perpendicular.

Consider $\underbrace{Proj_W(\vec{v_i}) = \vec{v_i}}$

but $\pi(\vec{v_i}) = \sum_{i=1}^{d} \left(\vec{v_i} \cdot \frac{\vec{v_i}}{\|\vec{v_i}\|}\right) \frac{\vec{v_i}}{\|\vec{v_i}\|}$

Since the basis vector are not perpendicular,

$\exists$ ~~some~~ at least vector $\vec{v_j}$ such that $\vec{v_i} \cdot \vec{v_j} = a$ for some $a \neq 0$

So $\pi(\vec{v_i}) = \vec{v_i} + \dots + \frac{a}{\|\vec{v_j}\|} \vec{v_j} + \dots$

$\neq \vec{v_i}$ since these vectors are linearly independent

Therefore π is not $proj_W$.

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.

(a) False. Counter example:

$I_n \in O_n$ since $I_n^T = I_n^{-1} = I_n$

But $(I_n + I_n) \notin O_n$ since $(I_n + I_n)^T = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$

$(I_n + I_n)^{-1} = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{bmatrix}$

$(I_n + I_n)^T \neq (I_n + I_n)^{-1}$, so $I_n + I_n \notin O_n$

Therefore not a subspace

(b) True.

Proof. Assume $A, B \in O_n \Rightarrow T_A, T_B$ are orthogonal

As we have proved in WS 18, $T_A \circ T_B$ whose standard matrix is AB is orthogonal

And by WS 18. Thm B, AB is orthogonal.

(c) True

Proof. Same reason with (b): $T_A \circ T_A$ is orthogonal $\Rightarrow A^2$ is orthogonal.

(d) True

Proof. Assume A^2 is orthogonal but A is not orthogonal

Then $\|T_A(\vec{x})\| \neq \|\vec{x}\|$

Let $\|T_A(\vec{x})\| = a\|\vec{x}\|$ where $a \neq 1$

Then $\|T_A \circ T_A(\vec{x})\| = a^2\|\vec{x}\|$, $a^2 \neq 1$

So $T_A \circ T_A$ is not orthogonal, therefore A^2 is not orthogonal by WS 18. Thm B, contradicting with A^2 is orthogonal

Therefore A^2 is orthogonal $\Rightarrow A$ is orthogonal.

(e) True.

Proof. $A \in O_n \Rightarrow A^T = A^{-1}$ ①

Since $A^2 = I_n$

$AA = I_n$

$(AA)A^{-1} = I_n \cdot A^{-1}$

$A = A^{-1}$ ②

①② $\Rightarrow A = A^T$, 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.

(a) Proof

Assume $[\vec{v}]_{\mathcal{B}} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_r \end{bmatrix}$, $[\vec{w}]_{\mathcal{B}} = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_r \end{bmatrix}$

So $\vec{v} = v_1\vec{b_1} + \ldots + v_r\vec{b_r}$, $\vec{w} = w_1\vec{b_1} + \ldots + w_r\vec{b_r}$

Therefore $\vec{v} \cdot \vec{w} = (v_1\vec{b_1} + \ldots + v_r\vec{b_r}) \cdot (w_1\vec{b_1} + \ldots + w_r\vec{b_r})$

Since $\mathcal{B}$ is an orthonormal basis, $= \left(\sum_{i=1}^{n} v_i\vec{b_i}\right)\left(\sum_{j=1}^{n} w_j\vec{b_j}\right)$

$b_i \cdot b_j = \begin{cases} 0, & i \neq j \\ 1, & i = j \end{cases}$

So $\vec{v} \cdot \vec{w} = \sum_{i=1}^{n} v_i w_i = [\vec{v}]_{\mathcal{B}} \cdot [\vec{w}]_{\mathcal{B}}$.

(b) $$S_{C\to\mathcal{B}} = \begin{bmatrix} | & | & & | \\ [\vec{c_1}]_{\mathcal{B}} & [\vec{c_2}]_{\mathcal{B}} & \cdots & [\vec{c_r}]_{\mathcal{B}} \\ | & | & & | \end{bmatrix}$$

Since $\mathcal{B}$ is an orthonormal basis,

by what we proved on WS16,

$$S_{C\to\mathcal{B}} = \begin{bmatrix} \vec{c_1}\cdot\vec{b_1} & \vec{c_2}\cdot\vec{b_1} & \cdots & \vec{c_r}\cdot\vec{b_1} \\ \vec{c_1}\cdot\vec{b_2} & \ddots & & \vdots \\ \vdots & & \ddots & \\ \vec{c_1}\cdot\vec{b_r} & \cdots & \cdots & \vec{c_r}\cdot\vec{b_r} \end{bmatrix}$$

① So the ij^{th} entry of $S_{C\to\mathcal{B}}$ is $\vec{b_i}\cdot\vec{c_j}$

Similarly since C is an orthonormal basis,

$$S_{\mathcal{B}\to C} = \begin{bmatrix} \vec{b_1}\cdot\vec{c_1} & \vec{b_2}\cdot\vec{c_1} & \cdots & \vec{b_r}\cdot\vec{c_1} \\ \vec{b_1}\cdot\vec{c_2} & \ddots & & \vdots \\ \vdots & & \ddots & \\ \vec{b_1}\cdot\vec{c_r} & \cdots & \cdots & \vec{b_r}\cdot\vec{c_r} \end{bmatrix}$$

② So the ij^{th} entry of $S_{C\to\mathcal{B}}$ is $\vec{c_i}\cdot\vec{b_j}$

By ①②, $S_{C\to\mathcal{B}} = S_{\mathcal{B}\to C}^{T}$

Since $S_{C\to\mathcal{B}} = S_{\mathcal{B}\to C}^{-1}$, $S_{\mathcal{B}\to C}^{T} = S_{\mathcal{B}\to C}^{-1}$

Therefore $S_{\mathcal{B}\to C}$ is orthogonal by definition.

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) $\ker A = \ker AA^{\top}$.

(a) True.

Proof. By WS19 Thm 2, $\ker(A^T) = (\operatorname{im} A)^{\perp}$

Since $(A^T)^T = A$, we have $\ker(A) = (\operatorname{im} A^T)^{\perp}$

So $(\ker A)^{\perp} = ((\operatorname{im} A^T)^{\perp})^{\perp}$

$= \operatorname{im} A^T$ by theorem 5.1.8.

(b) True

Proof $T_A : \mathbb{R}^m \longrightarrow \mathbb{R}^n$

$T_{A^TA} : \mathbb{R}^m \longrightarrow \mathbb{R}^m$

By thm 5.4.2, $\ker(A) = \ker(A^TA)$

By rank-nullity theorem, $\ker(A) + \operatorname{rank}(A) = m$

$\ker(A^TA) + \operatorname{rank}(A^TA) = m$

$\Rightarrow \operatorname{rank}(A) = \operatorname{rank}(A^TA)$

(c) True

Proof. By WS19 Thm 2, $\ker(A^T) = (\operatorname{im} A)^{\perp} \Rightarrow \dim(\ker A^T) = \dim((\operatorname{im} A)^{\perp})$ ①

By WS16 Thm C, ② $\dim(\operatorname{im} A)^{\perp} + \dim \operatorname{im} A = m$

Since ③ $\dim(\ker A^T) + \dim(\operatorname{im} A^T) = m$ by rank-nullity thm,

By ①②③ $\Rightarrow \dim(\operatorname{im} A^T) = \dim(\operatorname{im} A) \Rightarrow \operatorname{rank}(A) = \operatorname{rank}(A^T)$ by definition

(d) True

Proof Since $\operatorname{rank}(A) = \operatorname{rank}(A^T A)$

and $(A^T)^T = A$

so $\operatorname{rank}(A^T) = \operatorname{rank}(AA^T)$

Since $\operatorname{rank}(A) = \operatorname{rank}(A^T)$,

$\operatorname{rank}(A) = \operatorname{rank}(AA^T)$

since $\operatorname{rank}(A) = \operatorname{rank}(A^T A)$,

$\operatorname{rank}(A^T A) = \operatorname{rank}(AA^T)$

(e) False. By rank-nullity theorem,

$\dim(\ker A) = m - \operatorname{rank}(A)$

$\dim(\ker AA^T) = n - \operatorname{rank}(AA^T)$

$= n - \operatorname{rank}(A)$ since $\operatorname{rank}(A) = \operatorname{rank}(AA^T)$ by (b), (d)

whenever $n \neq m$,

So $\dim(\ker A) \neq \dim(\ker AA^T)$

Then $\ker A \neq \ker AA^T$.

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 $\operatorname{Span}(X) \perp \operatorname{Span}(Y)$.

(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.

(a) Proof Select arbitrary $X = \{\vec{x_1}, \ldots, \vec{x_r}\} \subseteq \mathbb{R}^n$ and $Y = \{\vec{y_1}, \ldots, \vec{y_s}\} \subseteq \mathbb{R}^n$.

Assume $X \perp Y$.

Then $\forall\ 1 \leq i \leq r,\ 1 \leq j \leq s,\ \vec{x_i} \cdot \vec{y_j} = 0$.

Let $\vec{x} \in \operatorname{span} X$ and $\vec{y} \in \operatorname{span} Y$ be arbitrary

Then $\vec{x} = \sum_{i=1}^{r} a_i \vec{x_i}$ for some $a_1, \ldots, a_r \in \mathbb{R}$

$\vec{y} = \sum_{j=1}^{s} b_j \vec{y_j}$ for some $b_1, \ldots, b_s \in \mathbb{R}$

So $\vec{x} \cdot \vec{y} = \left(\sum_{i=1}^{r} a_i \vec{x_i}\right)\left(\sum_{j=1}^{s} b_j \vec{y_j}\right) = 0$ since every pair of $\vec{x_i} \cdot \vec{y_j} = 0$

Therefore $\operatorname{Span} X \perp \operatorname{Span} Y$ by definition

(b) Proof. Assume $X \perp Y$

let $\sum_{i=1}^{r} a_i \vec{x_i} + \sum_{j=1}^{s} b_j \vec{y_j} = 0$ be a relation on $X \cup Y$.

Assume for contradiction that some $a_{\alpha_1}, a_{\alpha_2}, \dots, a_{\alpha_n}$ and $b_{\beta_1}, b_{\beta_2}, \dots, b_{\beta_m}$ are not zero

Then $\sum_{\alpha} a_\alpha \vec{x_\alpha} = -\sum_{\beta} b_\beta \vec{y_\beta} \neq 0$, so

$\text{span}(X) \cap \text{span}(Y) \neq \{0\}$ ①

By WS16: $\text{Span}(X) \cap (\text{Span}(X))^{\perp} = \{0\}$

Since $X \perp Y$, by (a), $\text{Span}(X) \perp \text{Span}(Y)$

So $\text{Span}(Y) \subseteq \text{Span}(X)^{\perp}$

$\Longrightarrow \text{Span}(Y) \cap \text{Span}(X) = \{0\}$ ②

① ② contradicts, so assumption fails: all $a_1, \dots, a_r$ and $b_1, \dots, b_s$ are 0

Therefore $X \cup Y$ is linearly independent by definition.

(c) Claim ① Any set ($\subseteq \mathbb{R}^n$) of pairwise orthogonal vectors that have less than $n+1$ elements is not maximal.

let X be a set of pairwise orthogonal vectors $= \{\vec{x_1}, \dots, \vec{x_r}\}$ with $|X| < n-1$

So by WS 17, $X \setminus \{\vec{0}\}$ is linearly independent.

Since $\dim(\mathbb{R}^n) = n$, by performing Gram-Schmidt process, we can always find an orthonormal basis of $\mathbb{R}^n$ using $X \setminus \{\vec{0}\}$. Denote it by X'

Note that $\left(X' \setminus \left\{\frac{\vec{x_1}}{\|\vec{x_1}\|}, \dots, \frac{\vec{x_r}}{\|\vec{x_r}\|}\right\}\right) \cup (X \setminus \{0\})$ is a basis of $\mathbb{R}^n$ that have n pairwise orthogonal elements.

Now we add $\{0\}$ to it:

$\left(X' \setminus \left\{\frac{\vec{x_1}}{\|\vec{x_1}\|}, \dots, \frac{\vec{x_r}}{\|\vec{x_r}\|}\right\}\right) \cup (X \setminus \{0\}) \cup \{0\}$ has $n+1$ pairwise orthogonal elements. Note that

$$X \subsetneq \left(X' \setminus \left\{\frac{\vec{x_1}}{\|\vec{x_1}\|}, \dots, \frac{\vec{x_r}}{\|\vec{x_r}\|}\right\}\right) \cup (X \setminus \{0\}) \cup \{0\}$$

So we have proved claim ①.

Claim ②: Any set $\subseteq \mathbb{R}^n$ of more than $n+1$ element is not pairwise orthogonal.

Assume for contradiction that $Y \subseteq \mathbb{R}^n$ is pairwise orthogonal with $|Y| > n+1$ $\quad Y = \{y_1, y_2, \dots, y_m\}$

So $|Y \setminus \{0\}| > n$

Since $\dim(\mathbb{R}^n) = n$, $Y \setminus \{0\}$ is not a basis of $\mathbb{R}^n$. Therefore $\exists\ y_i \in Y/\{0\}$ s.t. y_i is a linear combination of some $y_{j1}, y_{j2}, \dots, y_{js} \in Y/\{0\}$

So $y_i = a_1 y_{j1} + a_2 y_{j2} + \dots + a_s y_{js}$ for some $a_1, \dots, a_s \neq 0$

Then $\|y_i\|^2 = y_i \cdot (a_1 y_{j1} + a_2 y_{j2} + \dots + a_s y_{js})$

Since $\|y_j\| > 0$, at least some $y_i \cdot y_{jr} \neq 0$

So $Y \setminus \{\vec{0}\}$ is not pairwise orthogonal

So Y is not pairwise orthogonal

By claim ①②, any maximally pairwise orthogonal set $\subseteq \mathbb{R}^n$ has exactly $n+1$ elements.

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

(a) If $\text{rank}(A) = m$, 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 $\text{rank}(A) < m$, it is still possible to obtain such a decomposition if we allow some diagonal entries to be zero.

Pf (a) Since $\text{rank}(A) = m$, $\dim(\text{im} A) = m$

So by Gram-Schmdmit Process, we can get orthonormal basis obtained $Q = (q_1, q_2, \dots, q_m)$ from the columns of A as another basis $\mathcal{B}$.

Now let $\mathcal{B}' = (\vec{b_m}, \vec{b_{m-1}}, \dots, \vec{b_2}, \vec{b_1})$

$$Q' = \begin{bmatrix} | & | & & | & | \\ q_m & q_{m-1} & \cdots & q_2 & q_1 \\ | & | & & | & | \end{bmatrix}$$

So By the QR factorization thm, $A = Q'R'$,

where $R' = S_{\mathcal{B}' \to \mathcal{Q}'} = \begin{bmatrix} | & | & & | \\ [b_m]_{\mathcal{Q}} & [b_{m-1}]_{\mathcal{Q}} & \cdots & [b_1]_{\mathcal{Q}} \\ | & | & & | \end{bmatrix}$

$$= \begin{bmatrix} b_m \cdot q_m & b_{m-1} \cdot q_m & \cdots & b_1 \cdot q_m \\ & b_{m-1} \cdot q_{m-1} & & \vdots \\ & & \ddots & \\ & & & b_1 \cdot q_1 \end{bmatrix}$$

Now consider $Q = \begin{bmatrix} | & | & & | & | \\ q_1 & q_2 & \cdots & q_{m-1} & q_m \\ | & | & & | & | \end{bmatrix}$, $L = R'^T$

so $A = Q'R' = QL$