1. Prove the following:

**Theorem (Gram-Schmidt Process).** Given a basis \( \{v_1, \ldots, v_p\} \) for a nonzero subspace \( E \) of \( \mathbb{R}^n \), define

\[
\begin{align*}
    w_1 &= v_1 \\
    w_2 &= v_2 - \left( \frac{v_2 \cdot w_1}{w_1 \cdot w_1} \right) w_1 \\
    w_3 &= v_3 - \left( \frac{v_3 \cdot w_1}{w_1 \cdot w_1} \right) w_1 - \left( \frac{v_3 \cdot w_2}{w_2 \cdot w_2} \right) w_2 \\
    &\quad \ldots \\
    w_p &= v_p - \sum_{k=1}^{p-1} \left( \frac{v_p \cdot w_k}{w_k \cdot w_k} \right) w_k
\end{align*}
\]

Then, \( \{w_1, \ldots, w_p\} \) is an orthogonal basis for \( E \).

[Remarks: i) Notice that \( \text{span}(w_1, \ldots, w_k) = \text{span}(v_1, \ldots, v_k) \) for each \( k, 1 \leq k \leq p \). ii) It follows from this theorem that every subspace of \( \mathbb{R}^n \) has an orthogonal basis.]

2. Prove the following:

**Theorem (Orthogonal Decomposition).** Let \( E \) be a subspace of \( \mathbb{R}^n \). Then each \( y \in \mathbb{R}^n \) can be written uniquely in the form

\[
y = \hat{y} + z
\]

where \( y \in E \) and \( z \in E^\perp = \{v \in \mathbb{R}^n \mid v \cdot w = 0 \text{ for every } w \in E\} \). In fact, if \( \{w_1, \ldots, w_p\} \) is an orthogonal basis for \( E \), then

\[
\hat{y} = \left( \frac{y \cdot w_1}{w_1 \cdot w_1} \right) w_1 + \cdots + \left( \frac{y \cdot w_p}{w_p \cdot w_p} \right) w_p
\]

and \( z = y - \hat{y} \).

[Remarks: i) Be sure to show that this decomposition is unique. ii) \( E^\perp \) is called the orthogonal complement of \( E \). iii) This theorem justifies our claim that, given a system \( Ax = b \) with \( b \notin \text{img}(A) \), there is a unique \( \hat{b} \in \text{img}(A) \) with \( b - \hat{b} \in \text{img}(A)^\perp \).]
3. Prove the following:

**Theorem (Best Approximation).** Let $E$ be a subspace of $\mathbb{R}^n$, let $y$ be any vector in $\mathbb{R}^n$, and let $\hat{y} \in E$ such that $y - \hat{y} \in E^\perp$. Then, $|y - \hat{y}| < |y - v|$ for every $v \in E$ with $v \neq \hat{y}$.

[Remarks: i) Recall that $E^\perp = \{ v \in \mathbb{R}^n \mid v \cdot w = 0 \text{ for every } w \in E \}$ is called the orthogonal complement of $E$. ii) This theorem justifies our claim that a solution to $A\hat{x} = \hat{b}$, where $\hat{b} \in \text{img}(A)$ and $b - \hat{b} \perp \text{img}(A)$, is a best approximation to a solution of $Ax = b$.]

4. Let $E = \text{span}\left(\begin{bmatrix} -1 \\ 3 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ -4 \\ -1 \\ -2 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix}\right)$. Find an orthogonal basis for $E$.

Normalize the vectors in your basis to have integer components.

5. Find a least-squares solution of $Ax = b$ for $A = \begin{bmatrix} 1 & 3 & 5 \\ 1 & 1 & 0 \\ 1 & 1 & 2 \\ 1 & 3 & 3 \end{bmatrix}$, $b = \begin{bmatrix} 3 \\ 5 \\ 7 \\ -3 \end{bmatrix}$.

6. Describe all least-squares solutions of $Ax = b$ for $A = \begin{bmatrix} 1 & 3 & 5 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 3 & 3 \end{bmatrix}$, $b = \begin{bmatrix} 3 \\ 5 \\ 7 \\ -3 \end{bmatrix}$.

Why isn’t there a unique least-squares solution?

7. (a) Let $A$ be a $m \times n$ matrix. Prove that $\ker(A) = \ker(A^\top A)$. (Note: This requires you to show that both $\ker(A) \subset \ker(A^\top A)$ and $\ker(A^\top A) \subset \ker(A)$.)

(b) Prove the $\ker(A) = \{0\}$ if and only if $A^\top A$ is invertible.

[Remark: This shows that $Ax = b$ has a unique least-squares solution for every $b \in \mathbb{R}^m$ if and only if $A$ has linearly independent columns (i.e., $\ker(A) = \{0\}$).]