Theorem. (The Orthogonal Decomposition Theorem)
Let $W$ be a subspace of $\mathbb{R}^n$. Then each $y$ in $\mathbb{R}^n$ can be uniquely represented in the form

$$ y = \hat{y} + z $$

where $\hat{y}$ is in $W$ and $z$ is in $W^\perp$.

In fact, if $\{u_1, u_2, \ldots, u_p\}$ is any orthogonal basis for $W$, then

$$ \hat{y} = \left( \frac{y \cdot u_1}{u_1 \cdot u_1} \right) u_1 + \cdots + \left( \frac{y \cdot u_p}{u_p \cdot u_p} \right) u_p $$

and $z = y - \hat{y}$.

The vector $\hat{y}$ is called the orthogonal projection of $y$ onto $W$. 
Theorem (The Best Approximation Theorem). Let $W$ be a subspace of $\mathbb{R}^n$, $y$ any vector in $\mathbb{R}^n$, and $\hat{y}$ the orthogonal projection of $y$ onto $W$. Then $\hat{y}$ is the point in $W$ closest to $y$, in the sense that
\[ \|y - \hat{y}\| < \|y - v\| \]
for all $v$ in $W$ distinct from $\hat{y}$.

Outline of Proof.
Theorem. If \( \{u_1, u_2, \ldots, u_p\} \) is an orthonormal basis for a subspace \( W \) of \( \mathbb{R}^n \), then
\[
\text{proj}_W y = (y \cdot u_1)u_1 + \cdots + (y \cdot u_p)u_p.
\]
If \( U = [u_1 \cdots u_p] \), then
\[
\text{proj}_W y = UU^T y \quad \text{for all } y \text{ in } \mathbb{R}^n.
\]
Outline of Proof.
6.4 The Gram-Schmidt Process

**Goal:** Form an orthogonal basis for a subspace $W$.

**Example.** Suppose $W = \text{Span}\{x_1, x_2\}$ where

\[
x_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \quad x_2 = \begin{bmatrix} 2 \\ 2 \\ 3 \end{bmatrix}.
\]

Find an orthogonal basis \( \{v_1, v_2\} \) for $W$. 

Example. Suppose $W = \text{Span}\{x_1, x_2\}$ where

\[
x_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \quad x_2 = \begin{bmatrix} 2 \\ 2 \\ 3 \end{bmatrix}.
\]

Find an orthogonal basis $\{v_1, v_2\}$ for $W$. 
Example. Suppose $\{x_1, x_2, x_3\}$ is a basis for a subspace $W$ of $\mathbb{R}^4$. Derive an orthogonal basis $\{v_1, v_2, v_3\}$ for $W$.

Solution.
Theorem. (The Gram-Schmidt Process)

Given a basis \( \{x_1, x_2, \cdots, x_p\} \) for a subspace \( W \) of \( \mathbb{R}^n \), define

\[
\begin{align*}
v_1 &= x_1 \\
v_2 &= x_2 - \frac{x_2 \cdot v_1}{v_1 \cdot v_1} v_1 \\
v_3 &= x_3 - \frac{x_3 \cdot v_1}{v_1 \cdot v_1} v_1 - \frac{x_3 \cdot v_2}{v_2 \cdot v_2} v_2 \\
\vdots \\
v_p &= x_p - \frac{x_p \cdot v_1}{v_1 \cdot v_1} v_1 - \frac{x_p \cdot v_2}{v_2 \cdot v_2} v_2 - \cdots - \frac{x_p \cdot v_{p-1}}{v_{p-1} \cdot v_{p-1}} v_{p-1}.
\end{align*}
\]

Then \( \{v_1, v_2, \cdots, v_p\} \) is an orthogonal basis for \( W \) and \( \text{Span}\{x_1, \cdots, x_p\} = \text{Span}\{v_1, \cdots, v_p\} \).
Example. Suppose \( \{x_1, x_2, x_3\} \) is a basis for a subspace \( W \) of \( \mathbb{R}^4 \), where

\[
x_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 0 \end{bmatrix}, \quad x_2 = \begin{bmatrix} 1 \\ 2 \\ 0 \\ 0 \end{bmatrix}, \quad x_3 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}.
\]

Derive an orthogonal basis for \( W \).
QR Factorization

Theorem (The QR Factorization)

If \( A \) is an \( m \times n \) matrix with linearly independent columns, then \( A \) can be factored as \( A = QR \), where \( Q \) is an \( m \times n \) matrix whose columns form an orthonormal basis for \( \text{Col} \ A \) and \( R \) is an \( n \times n \) upper triangular invertible matrix with positive entries on the main diagonal.
Example. Find the QR factorization of $A = \begin{bmatrix} 1 & 2 \\ 1 & 2 \\ 0 & 3 \end{bmatrix}$.

Solution.
6.5 Least-Squares Problems

Problem: What do we do when \( Ax=b \) has no solution \( x \)?

Answer: Find \( \hat{x} \) such that \( A\hat{x} \) is as “close” as possible to \( b \). (Least Squares Problem)

If \( A \) is \( m \times n \) and \( b \) is in \( \mathbb{R}^m \), a least-squares solution of \( Ax=b \) is an \( \hat{x} \) in \( \mathbb{R}^n \) such that

\[
\| b - A\hat{x} \| \leq \| b - Ax \|
\]

for all \( x \) in \( \mathbb{R}^n \).
Let $W = \text{Col } A$ where $A$ is $m \times n$ and $A = [a_1 \ a_2 \cdots \ a_n]$. Suppose $b$ is in $\mathbb{R}^m$ and $\hat{b} = \text{proj}_W b$.

$\hat{b}$ is the point in $W = \text{Col } A$ closest to $b$. Since $b$ is in $\text{Col } A$, there is at least one $\hat{x}$ in $\mathbb{R}^n$ such that $\hat{b} = A\hat{x}$. 
Let $z = \mathbf{b} - A\hat{x}$. By the orthogonal projection theorem, $z$ is in $W^\perp$.

Thus, $z = \mathbf{b} - A\hat{x}$ is orthogonal to every column of $A$: 
Theorem. The set of least squares solutions of $Ax = b$ is the set of all solutions of the normal equations $A^T A \hat{x} = A^T b$. 
Example. Find a least squares solution to the inconsistent system $Ax = b$ where

$$A = \begin{bmatrix} 2 & 0 \\ 0 & 1 \\ 2 & 2 \end{bmatrix} \quad \text{and} \quad b = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}.$$

Solution.
When $\mathbf{A}^T \mathbf{A}$ is invertible:
Theorem. The matrix $A^T A$ is invertible if and only if the columns of $A$ are linearly independent. In this case, the equation $Ax = b$ has only one least-squares solution $\hat{x}$, and it is given by

$$\hat{x} = (A^T A)^{-1} A^T b.$$ 

Least-squares error $= \|b - A\hat{x}\|$
Calculating Least-Squares Solutions using the QR Factorization:

**Theorem.** Given an $mxn$ matrix of rank $n$, let $A=QR$ be a QR factorization of $A$. Then, for each $b$ in $\mathbb{R}^m$, the equation $Ax = b$ has a unique least squares solution, given by
Calculating Least-Squares Solutions using the QR Factorization:

**Theorem.** Given an \( mxn \) matrix of rank \( n \), let \( A=QR \) be a QR factorization of \( A \). Then, for each \( b \) in \( \mathbb{R}^m \), the equation \( Ax = b \) has a unique least squares solution, given by

\[
\hat{x} = R^{-1}Q^T b.
\]

(The solution \( \hat{x} \) may be obtained by using back substitution to solve \( R\hat{x} = Q^T b \).)
Example. Find the least-squares solution to the inconsistent system

\[
\begin{bmatrix}
1 & -1 \\
1 & 4 \\
1 & -1 \\
1 & 4
\end{bmatrix}
\begin{bmatrix}
x_1 \\
x_2
\end{bmatrix}
= 
\begin{bmatrix}
-1 \\
6 \\
5 \\
7
\end{bmatrix}.
\]