If we fix any orthonormal basis \( \{v_1, v_2, \ldots, v_k\} \) of \( V \),

then \( \text{proj}_V x = \left( \frac{x \cdot v_1}{v_1 \cdot v_1} \right) v_1 + \ldots + \left( \frac{x \cdot v_k}{v_k \cdot v_k} \right) v_k \).

Remember we wanted to use this to do approximation methods.

We need to make sure our definition have the correct properties.

Let \( V \subseteq \mathbb{R}^n \) be a subspace, the orthogonal complement of \( V \) in \( \mathbb{R}^n \) is defined as \( V^\perp = \{ x' \in \mathbb{R}^n / x' \cdot v = 0 \text{ for all } v \in V \} \).

Then \( V^\perp \) is also a subspace of \( \mathbb{R}^n \).

and

- a) \( V \cap V^\perp = \{ 0 \} \)
- b) \( \dim V + \dim V^\perp = n \)
- c) \( (V^\perp)^\perp = V \)
Pythagorean thm

If \( x, y \in \mathbb{R}^n \) are orthogonal

\[ \text{if and only if} \quad \| x + y \|^2 = \| x \|^2 + \| y \|^2 \]

Thm

If \( V \) is a subspace of \( \mathbb{R}^n \), and vector \( x \) in \( \mathbb{R}^n \) then

\[ \| \text{proj}_V x \| \leq \| x \| \]

Reason:

We know by Pythagorean thm.

\[ \| x \|^2 = \| x - \text{proj}_V x \|^2 + \| \text{proj}_V x \|^2 \]

or

\[ \| x \|^2 \geq \| \text{proj}_V x \|^2 \]

Then in order to do approximation
we need to make sure we always can have a orthogonal / orthonormal basis! (Note it is easy to go from orthogonal set to orthonormal by just dividing by their length.)
Thm: Let $x \in \mathbb{R}^n$ and $V$ be a subspace of $\mathbb{R}^n$.

Then: $||x^l - \text{proj}_V x^l|| \leq ||x^l - v||$ for all $v \in \mathbb{R}^n$.

Proof: follows from pythagorean thm.

Therefore, now if we want to imitate the approximation method in $\mathbb{R}^3$ to $\mathbb{R}^n$, we need to be able to compute $\text{proj}_V x$.

This however will work only if we know how to compute find an orthogonal basis of $V$.

* Gram-Schmidt Method is a method by which we start with any basis $E_1, \ldots, E_k$ of a subspace $V$ of $\mathbb{R}^n$ and obtain an orthonormal basis of $V$.

Then: one way to find the approximation of $Ax^l = b$ when $A$ is a $m \times n$ matrix $\iff A^T A$ is inconsistent is to solve for $x^* = \text{proj}_{\text{Im} A} b$. 
Alternately, a few observations will tell us that. We don't need to actually compute the projection at all.

Fact 1. \( \bar{V} \cdot \bar{X} = \bar{V}^T \bar{X} \)

\[
\begin{bmatrix}
1 \\
0 \\
3
\end{bmatrix} \cdot \begin{bmatrix}
2 \\
1 \\
3
\end{bmatrix} = 11 \quad \text{and} \quad \begin{bmatrix}
10 \\
3
\end{bmatrix} \begin{bmatrix}
2 \\
1 \\
3
\end{bmatrix} = 11
\]

Fact 2: \( \text{Im} \)

Definition: Let \( \bar{V} \)

Fact 2: \( (\text{Im} A)^\perp = \text{Ker} A^T \)

Why? \( (\text{Im} A)^\perp = \{ \bar{X} \in \mathbb{R}^m / \bar{X}, \bar{Y} = 0 \} \) for all \( \bar{Y} \in \text{Im} \bar{A} \)

Let \( A^T = [\bar{v}_1 \ldots \bar{v}_n] \)

Then \( (\text{Im} A)^\perp = \{ \bar{X} \in \mathbb{R}^m / \bar{v}_1^T \bar{X} = 0, \ldots, \bar{v}_n^T \bar{X} = 0 \} \)
Therefore \((\text{Im } A)^\perp = \{ \mathbf{x} \in \mathbb{R}^m \mid \begin{bmatrix} v_1^T \\ \vdots \\ v_n^T \end{bmatrix} \mathbf{x} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} \}\)

\[ = \{ \mathbf{x} \in \mathbb{R}^m \mid A^T \mathbf{x} = \mathbf{0} \}\]

\[ = \ker A^T \]

Similarly, \((\ker A)^\perp = \text{Im } A^T\)

Anyway, we were trying to solve \(A \mathbf{x}^* = \text{proj}_{\text{Im } A} \mathbf{b} = \mathbf{b}^\perp\)

\(<\{\) \((\mathbf{b} - A \mathbf{x}^*) = \mathbf{b}^\perp \in (\text{Im } A)^\perp = \ker A^T\)

\(<\{\) \(A^T (\mathbf{b} - A \mathbf{x}^*) = \mathbf{0}\)

\(<\{\) \(A^T A \mathbf{x}^* = A^T \mathbf{b}\)

is called the normal equation

\(\mathbf{x}^*\) is the least squares solution of \(A^T A\).
Moreover,

**Theorem:**

\[ \text{a) } \text{Ker } A = \text{Ker } (A^T A) \]

\[ \text{b) } \text{Ker } A = \{0\} \iff (A^T A) \text{ is invertible} \]

Therefore if \( \text{Ker } A = \{0\} \), then \( Ax = b \) has a unique least squares solution.

But in general how do I find the orthogonal projection of a vector \( b \in \mathbb{R}^n \) onto a subspace \( V \subset \mathbb{R}^n \)?

**Method 1:**

First start with any basis \( \{v_1, \ldots, v_k\} \) of \( V \). Construct a orthonormal basis \( \{u_1, \ldots, u_k\} \) of \( V \) and then

\[
\text{proj}_V b = (b, u_1) u_1 + \cdots + (b, u_k) u_k.
\]

**Method 2:**

Use the reasoning used for least squares method to prove the following. Let \( \{v_1, \ldots, v_k\} \) be a basis of \( V \).
\[ A = (v_1 \ldots v_k) \text{ is a } n \times k \text{ matrix.} \]

Then, \[ \text{proj}_V b = A (A^T A)^{-1} A^T b \]

**Method:** Let me explain the idea behind the Gram-Schmidt process using a simple example.

Let \[ \left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \\ \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \\ \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \\ \end{bmatrix} \right\} \]

be a basis of a subspace \( V \) of \( \mathbb{R}^4 \).

We want to find \( \tilde{u}_1, \tilde{u}_2, \tilde{u}_3 \) such that \( \tilde{u}_1, \tilde{u}_2, \tilde{u}_3 \) is orthonormal set and in \( V \). Then clearly it is a basis since \( \dim V = 3 \) and \( \{\tilde{u}_1, \tilde{u}_2, \tilde{u}_3\} \) is linearly independent.

We should construct \( \tilde{u}_1, \tilde{u}_2, \tilde{u}_3 \) using linear combinations of \( v_1, v_2, v_3 \), since \( \text{Span} \{v_1, v_2, v_3\} = V \).
Let \( \vec{u}_1 = \frac{\vec{v}_1}{\|\vec{v}_1\|} = \begin{bmatrix} 1/2 \\ 1/2 \\ 1/2 \\ 1/2 \end{bmatrix} \).

Now \( \vec{v}_2 \) is not normal to \( \vec{u}_1 \) but \( \vec{v}_2^\perp = \vec{v}_2 - \text{proj}_{\vec{u}_1} \).

Define \( \vec{u}_2 = \frac{\vec{v}_2^\perp}{\|\vec{v}_2^\perp\|} = \begin{bmatrix} 1/2 \\ 1/2 \\ 1/2 \\ 1/2 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} - 4 \begin{bmatrix} 1/2 \\ 1/2 \\ 1/2 \\ 1/2 \end{bmatrix} \)

\[ \frac{1}{11} \begin{bmatrix} 1/2 \\ -1/2 \\ -1/2 \\ 1/2 \end{bmatrix} \]

Now \( \{\vec{u}_1, \vec{u}_2\} \) are orthonormal.

\( \vec{v}_3 \) is not in \( \text{Span}\{\vec{u}_1, \vec{u}_2\} = \text{Span}\{\vec{v}_1, \vec{v}_2\} \) so \( \vec{v}_3^\perp \) exist and is orthogonal to \( \vec{u}_1, \vec{u}_2 \).

Define \( \vec{u}_3 = \frac{\vec{v}_3^\perp}{\|\vec{v}_3^\perp\|} = \vec{v}_3 - \text{proj}_{\text{Span}\{\vec{u}_1, \vec{u}_2\}} \).