Orthonormal sets

**Definition.** A set of vectors \( \{ \mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k \} \) is orthonormal if it is orthogonal and \( \mathbf{v}_i \cdot \mathbf{v}_i = 1 \) for all \( i \).

**Example.** Consider the vectors

\[
\mathbf{v}_1 = \begin{bmatrix} 2 \\ -1 \\ 2 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} -1 \\ 2 \\ 2 \end{bmatrix}, \quad \text{and} \quad \mathbf{v}_3 = \begin{bmatrix} 2 \\ 2 \\ -1 \end{bmatrix}.
\]

We can use matrices to express the fact that a set is orthonormal.
Theorem. Let $A$ be an $n \times n$ matrix. The following three conditions are equivalent.

1. $A^T = A^{-1}$
2. The columns of $A$ form an orthonormal basis of $\mathbb{R}^n$.
3. The rows of $A$ form an orthonormal basis of $\mathbb{R}^n$.

Definition. Whenever a matrix satisfies the above theorem, it is said to be an orthogonal matrix.

Example. We can use the orthonormal basis of $\mathbb{R}^3$ that we derived earlier to produce an orthogonal matrix.

Why are orthogonal matrices special?
Orthogonal projection

How do we project a vector $v$ onto a subspace $W$?

**Theorem.** (Orthogonal Decomposition Theorem)

1. Each vector $v$ in $\mathbb{R}^n$ can be written uniquely as
   \[ v = w + w^\perp, \]
   where $w$ is in $W$ and $w^\perp$ is in $W^\perp$.

2. Given an orthogonal basis \{\(w_1, \ldots, w_k\)\} of $W$, then
   \[
   w = \left( \frac{v \cdot w_1}{w_1 \cdot w_1} \right) w_1 + \ldots + \left( \frac{v \cdot w_k}{w_k \cdot w_k} \right) w_k
   \]
   and $w^\perp = v - w$.

Important consequence: If we want to find the distance of a vector $v$ to a subspace $W$, then we compute
\[
||w^\perp|| = ||v - w||.
\]
Example. Find the point closest to

\[ \mathbf{v} = \begin{bmatrix} 3 \\ -1 \\ 1 \\ 13 \end{bmatrix} \]

in the subspace \( W \) spanned by the two vectors

\[ \mathbf{w}_1 = \begin{bmatrix} 1 \\ -2 \\ -1 \\ 2 \end{bmatrix} \quad \text{and} \quad \mathbf{w}_2 = \begin{bmatrix} -4 \\ 1 \\ 0 \\ 3 \end{bmatrix} \].
Theorem. If \( \{u_1, \ldots, u_k\} \) is an orthonormal basis for a subspace \( W \), then

\[
w = (v \cdot u_1)u_1 + \ldots + (v \cdot u_k)u_k.
\]

If

\[
U = \begin{bmatrix}
    u_1 & u_2 & \ldots & u_k
\end{bmatrix},
\]

then \( w = UU^Tv \).