Orthogonal basis and Gram-Schmidt orthogonalization

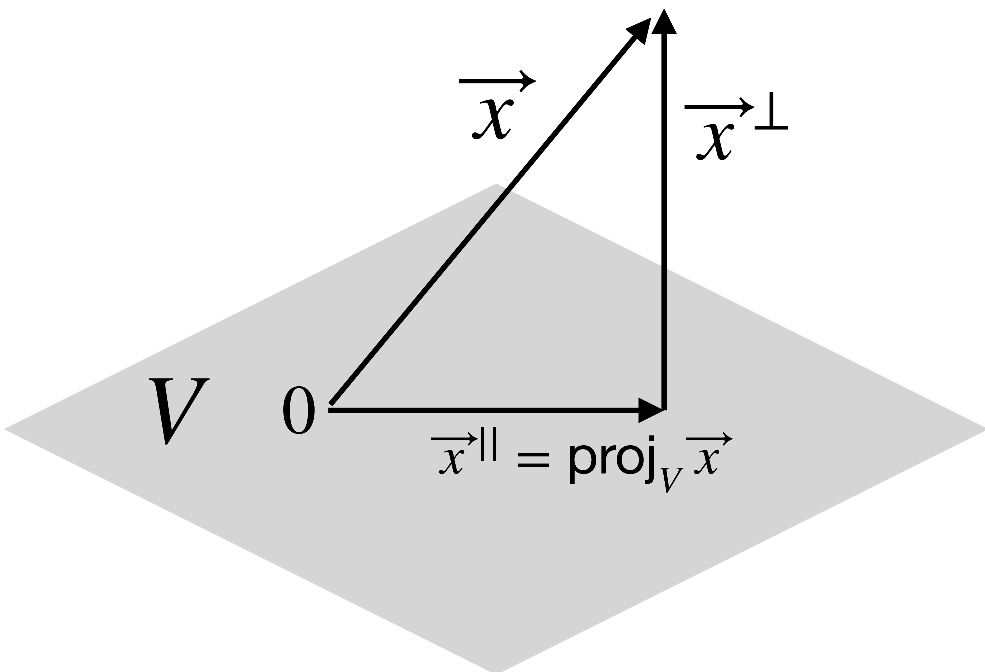
Two vectors $\overrightarrow{u}_1$ and $\overrightarrow{u}_2$ are *orthogonal* if and only if $\langle \overrightarrow{u}_1, \overrightarrow{u}_2 \rangle = 0$.
The vectors $\overrightarrow{u}_1, \overrightarrow{u}_2, \cdots, \overrightarrow{u}_m \in \mathbb{R}^n$ are *orthonormal* if and only if $\langle \overrightarrow{u}_i, \overrightarrow{u}_j \rangle = \delta_{ij}$.

Example: The vectors $\overrightarrow{e}_1, \overrightarrow{e}_2, \ldots, \overrightarrow{e}_n \in \mathbb{R}^n$ are orthonormal.

Properties of orthonormal vectors:

1. Orthonormal vectors are (automatically) linearly independent.
2. Orthonormal vectors $\overrightarrow{u}_1, \overrightarrow{u}_2, \cdots, \overrightarrow{u}_n \in \mathbb{R}^n$ form a basis in $\mathbb{R}^n$.

The shaded area denoted by V in the figure below is an **infinite** plane through the origin.



Orthogonal projection and orthogonal complement:
Let $\overrightarrow{x} \in \mathbb{R}^n$ and a subspace V of $\mathbb{R}^n$. Then we can wrtie $\overrightarrow{x} = \overrightarrow{x}^{||} + \overrightarrow{x}^{\perp}$, where $\overrightarrow{x}^{||} \in V$ and $\overrightarrow{x}^{\perp} \in V^{\perp}$. The above representation is **unique**.

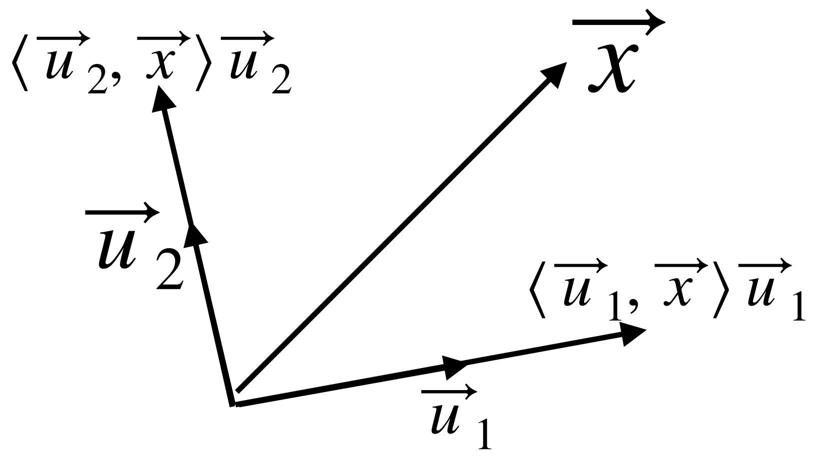
Here $V^{\perp} = \{\overrightarrow{x} \in \mathbb{R}^n : \langle \overrightarrow{v}, \overrightarrow{x} \rangle = 0, \ \ \forall \overrightarrow{v} \in V\}$. The transformation $T(\overrightarrow{x}) = \text{proj}_V \overrightarrow{x} = \overrightarrow{x}^{||}$ from $\mathbb{R}^n$ to $\mathbb{R}^n$ is linear. $V^{\perp} = \text{Ker}(T)$.

How do we compute $\overrightarrow{x}^{||}$?
Consider an orthonormal basis of V: $\overrightarrow{u}_1, \overrightarrow{u}_2, \cdots, \overrightarrow{u}_m \in V$ which is a subspace of $\mathbb{R}^n$. Then

$$\overrightarrow{x}^{||} = \langle \overrightarrow{u}_1, \overrightarrow{x} \rangle \overrightarrow{u}_1 + \cdots + \langle \overrightarrow{u}_m, \overrightarrow{x} \rangle \overrightarrow{u}_m; \quad \forall \overrightarrow{x} \in \mathbb{R}^n.$$

Consequently, consider an orthonormal basis of $\mathbb{R}^n$: $\overrightarrow{u}_1, \overrightarrow{u}_2, \cdots, \overrightarrow{u}_n$. Then any $\overrightarrow{x} \in \mathbb{R}^n$,

$$\overrightarrow{x} = \langle \overrightarrow{u}_1, \overrightarrow{x} \rangle \overrightarrow{u}_1 + \cdots + \langle \overrightarrow{u}_n, \overrightarrow{x} \rangle \overrightarrow{u}_n.$$



Properties of orthogonal complement:

Consider a subspace $V \in \mathbb{R}^n$.

1. $V^\perp$ is a subspace of $\mathbb{R}^n$.
2. $V \cap V^\perp = \{\overrightarrow{0}\}$.
3. $\dim(V) + \dim(V^\perp) = n$.
4. $(V^\perp)^\perp = V$.

Example: Consider the subspace $V = Im(A)$ of $\mathbb{R}^4$, where $A = \begin{pmatrix} 1 & 1 \\ 1 & -1 \\ 1 & -1 \\ 1 & 1 \end{pmatrix}$. Find $\overrightarrow{x}^{||}$ for $\overrightarrow{x} = \begin{pmatrix} 1 \\ 3 \\ 1 \\ 7 \end{pmatrix}$.

Solution: Recall that the column space of A is $Im(A)$. It can be easily checked that the column vectors of A are orthogonal by taking their scalar product. Thus we can construct an orthonormal basis of $Im(A)$. The basis vectors are: $\overrightarrow{u}_1 = \begin{pmatrix} 1/2 \\ 1/2 \\ 1/2 \\ 1/2 \end{pmatrix}$ and $\overrightarrow{u}_2 = \begin{pmatrix} 1/2 \\ -1/2 \\ -1/2 \\ 1/2 \end{pmatrix}$.

Then $\overrightarrow{x}^{||} = \langle \overrightarrow{u}_1, \overrightarrow{x} \rangle \overrightarrow{u}_1 + \langle \overrightarrow{u}_2, \overrightarrow{x} \rangle \overrightarrow{u}_2 = 6\overrightarrow{u}_1 + 2\overrightarrow{u}_2 = \begin{pmatrix} 3 \\ 3 \\ 3 \\ 3 \end{pmatrix} + \begin{pmatrix} 1 \\ -1 \\ -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 4 \\ 2 \\ 2 \\ 4 \end{pmatrix}$. In order to check that this answer is indeed correct, verify that $(\overrightarrow{x} - \overrightarrow{x}^{||}) \perp \overrightarrow{u}_1, \overrightarrow{u}_2$.

Why are orthonormal basis vectors useful?

1. We know that if we have some basis $\overrightarrow{v}_1, \overrightarrow{v}_2, \ldots, \overrightarrow{v}_n$ of an n-dimensional vector space W. Then any vector $\overrightarrow{x} \in W$ can be written as $\overrightarrow{x} = \alpha_1 \overrightarrow{v}_1 + \alpha_2 \overrightarrow{v}_2 + \cdots + \alpha_n \overrightarrow{v}_n$ (as a linear combination of the basis vectors) but there is no first-principles or convenient way of finding the unique coefficients $\alpha_1, \alpha_2, \ldots, \alpha_n$ except by explicit guesswork calculations. Now instead if we have an orthonormal basis set $\overrightarrow{u}_1, \overrightarrow{u}_2, \ldots, \overrightarrow{u}_n$ then any vector can be written as a linear combination of this orthonormal basis set as follows:

   $\overrightarrow{x} = \beta_1 \overrightarrow{u}_1 + \beta_2 \overrightarrow{u}_2 + \cdots + \beta_n \overrightarrow{u}_n$ where the coefficients can now be uniquely determined as $\beta_i = \langle \overrightarrow{u}_i, \overrightarrow{x} \rangle, \quad \forall i = 1,2,...,n$

2. Orthogonality guarantees linear independence.

Why are orthogonal transformations useful?

1. Orthogonal transformations are metric preserving transformations, i.e. if $T : \mathbb{R}^n \to \mathbb{R}^n$ is orthogonal, then $||T(\overrightarrow{x})|| = ||\overrightarrow{x}||, \quad \forall \overrightarrow{x} \in \mathbb{R}^n$. [1]

2. Orthogonal transformations are angle preserving transformations for orthogonal vectors. If $\overrightarrow{u} \perp \overrightarrow{w}$, then $T(\overrightarrow{u}) \perp T(\overrightarrow{w})$.

---

[1] If $T(\overrightarrow{x}) = A\overrightarrow{x}$ is an orthogonal transformation, then we say that A is an orthogonal matrix.