Section 6.1: Inner Product, Length, and Orthogonality

Recall that we defined the inner product in $\mathbb{R}^n$: **Definition:** For vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^n$ we define the **inner product** of $\mathbf{u}$ and $\mathbf{v}$ (also called the **dot product**) by the **matrix product**

$$\mathbf{u}^T \mathbf{v} = \begin{bmatrix} u_1 & u_2 & \cdots & u_n \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n.$$ 

We noted that this product has several properties.
Theorem (Properties of the Inner Product)

We’ll use the notation \( u \cdot v = u^T v \).

**Theorem:** For \( u, v \) and \( w \) in \( \mathbb{R}^n \) and real scalar \( c \)

(a) \( u \cdot v = v \cdot u \)

(b) \( (u + v) \cdot w = u \cdot w + v \cdot w \)

(c) \( c(u \cdot v) = (cu) \cdot v = u \cdot (cv) \)

(d) \( u \cdot u \geq 0 \), with \( u \cdot u = 0 \) if and only if \( u = 0 \).
The Norm and Orthogonality

Definition: The norm of the vector $\mathbf{v}$ in $\mathbb{R}^n$ is the nonnegative number denoted and defined by

$$\|\mathbf{v}\| = \sqrt{\mathbf{v} \cdot \mathbf{v}} = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}$$

where $v_1, v_2, \ldots, v_n$ are the components of $\mathbf{v}$.

If $\mathbf{v}$ is any nonzero vector, the vector $\mathbf{v}/\|\mathbf{v}\|$ is a unit vector in the direction of $\mathbf{v}$.

Definition: Two vectors are $\mathbf{u}$ and $\mathbf{v}$ orthogonal if $\mathbf{u} \cdot \mathbf{v} = 0$.

In $\mathbb{R}^n$, orthogonality corresponds geometrically with begin perpendicular.
Orthogonal Complement

**Definition:** Let $W$ be a subspace of $\mathbb{R}^n$. A vector $z$ in $\mathbb{R}^n$ is said to be orthogonal to $W$ if $z$ is orthogonal to every vector in $W$.

\[ \hat{z} \text{ is orthogonal to } W \text{ if } \hat{z} \cdot \hat{w} = 0 \text{ for every } \hat{w} \text{ in } W. \]

**Definition:** Given a subspace $W$ of $\mathbb{R}^n$, the set of all vectors orthogonal to $W$ is called the **orthogonal complement** of $W$ and is denoted by $W^\perp$. 
Theorem:

$W^\perp$ is a subspace of $\mathbb{R}^n$.

Note that for any $\tilde{w}$ in $W$,

$$\tilde{w} \cdot \vec{0} = w_1 \cdot 0 + w_2 \cdot 0 + \ldots + w_n \cdot 0 = 0$$

so $\vec{0}$ is orthogonal to $W$; it is in $W^\perp$.

Suppose $\tilde{u}$ and $\tilde{v}$ are in $W^\perp$. Then $\tilde{u} \cdot \tilde{w} = 0$ and $\tilde{v} \cdot \tilde{w} = 0$ for every $\tilde{w}$ in $W$. Note

$$(\tilde{u} + \tilde{v}) \cdot \tilde{w} = \tilde{u} \cdot \tilde{w} + \tilde{v} \cdot \tilde{w} = 0 + 0 = 0$$

So $\tilde{u} + \tilde{v}$ is in $W^\perp$. $W^\perp$ is closed under vector addition.
For scalar $c$ and $\mathbf{w}$ in $W$

$$(c\mathbf{u}) \cdot \mathbf{w} = c(\mathbf{u} \cdot \mathbf{w}) = c(0) = 0$$

Hence $W^\perp$ is closed under scalar multiplication.

$W^\perp$ meets all properties required to be a subspace of $\mathbb{R}^n$. 
Example

Let \( W = \text{Span}\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\} \). Show that \( W^\perp = \text{Span}\left\{ \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\} \).

Give a geometric interpretation of \( W \) and \( W^\perp \) as subspaces of \( \mathbb{R}^3 \).

If \( \bar{w} \) is in \( W \), then \( \bar{w} = w_1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + w_3 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} w_1 \\ 0 \\ w_3 \end{bmatrix} \)

for some real numbers \( w_1 \) and \( w_3 \). For \( \bar{z} \) in \( \mathbb{R}^3 \), \( \bar{z} = \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} \), if \( \bar{z} \cdot \bar{w} = 0 \), then \( \bar{z} \) is in \( W^\perp \).

\[ 0 = \bar{z} \cdot \bar{w} = z_1 w_1 + 0 + z_3 w_3 \]

This has to hold for every \( w, w_3 \) pair. If \( w_1 = 1 \) and \( w_3 = 0 \), we
get \( z_1 \cdot 1 = 0 \Rightarrow z_1 = 0 \). Taking \( w_1 = 0, w_2 = 1 \), we get \( 0 = z_2 \cdot 1 = z_2 = 0 \). So \( \tilde{z} \) must have the form \( \tilde{z} = \begin{bmatrix} 0 \\ z_2 \\ 0 \end{bmatrix} = z_2 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \). Hence

\[
W^\perp = \text{Span} \left\{ \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\}.
\]

As for the geometry, \( \tilde{w} \) in \( W \) looks like \((w_1, 0, w_2)\). Then any points in the \( xyz \) plane,
points in \( W^\perp \) look like \((0, z_2, 0)\).
This is the y-axis.
Example

Let \( A = \begin{bmatrix} 1 & 3 & 2 \\ -2 & 0 & 4 \end{bmatrix} \). Show that if \( x \) is in \( \text{Nul}(A) \), then \( x \) is in \( [\text{Row}(A)]^\perp \).

If \( \vec{x} \) is in \( \text{Nul} \ A \), then \( A\vec{x} = \vec{0} \). For such \( \vec{x} \), \( \vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \)

\[
\begin{bmatrix} 1 & 3 & 2 \\ -2 & 0 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}
\]

\[
\begin{bmatrix} x_1 + 3x_2 + 2x_3 \\ -2x_1 + 4x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}
\]
\[ x_1 + 3x_2 + 2x_3 = \hat{x} \cdot \hat{v}_1 \] where \( \hat{v}_1 \) is the first row of \( A \)

and

\[ -2x_1 + 4x_3 = \hat{x} \cdot \hat{v}_2 \] where \( \hat{v}_2 \) is the 2nd row of \( A \).

Since \( \text{row } A = \text{Span } \{ \hat{v}_1, \hat{v}_2 \} \), any \( \hat{x} \) in \( \text{null } A \)

is orthogonal to every vector in \( \text{row } A \).
**Theorem:** Let $A$ be an $m \times n$ matrix. The orthogonal complement of the row space of $A$ is the null space of $A$. That is

$$[\text{Row}(A)]^\perp = \text{Nul}(A).$$

The orthogonal complement of the column space of $A$ is the null space of $A^T$—i.e.

$$[\text{Col}(A)]^\perp = \text{Nul}(A^T).$$
Example: Find the orthogonal complement of $\text{Col}(A)$

$$A = \begin{bmatrix} 5 & 2 & 1 \\ -3 & 3 & 0 \\ 2 & 4 & 1 \\ 2 & -2 & 9 \\ 0 & 1 & -1 \end{bmatrix}$$

$$\text{col}(A)^\perp = \text{nul}(A^T)$$

$$A^T = \begin{bmatrix} 5 & -3 & 2 & 2 & 0 \\ 2 & 3 & 4 & -2 & 1 \\ 1 & 0 & 1 & 9 & -1 \end{bmatrix}$$

$$\xrightarrow{\text{ref}} \begin{bmatrix} 1 & 0 & 0 & -54 & 7 \\ 0 & 1 & 0 & -146/3 & 19/3 \\ 0 & 0 & 1 & 63 & -8 \end{bmatrix}$$

For $\vec{x}$ in $\text{nul}(A^T)$

$$x_1 = 54x_4 - 7x_5$$

$$x_2 = \frac{146}{3}x_4 - \frac{19}{3}x_5$$

$$x_3 = -63x_4 + 8x_5$$
\[ \bar{X} = x_4 \begin{bmatrix} 54 \\ 146 \\ \frac{3}{3} \\ -63 \\ 1 \\ 0 \end{bmatrix} + x_5 \begin{bmatrix} -7 \\ -19 \\ \frac{3}{3} \\ 8 \\ 0 \\ 1 \end{bmatrix} \]

so

\[ \text{[Code A]}^T = 5 \text{pm} \]

\[ \left\{ \begin{bmatrix} 54 \\ 146 \\ \frac{3}{3} \\ -63 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -7 \\ -19 \\ \frac{3}{3} \\ 8 \\ 0 \\ 1 \end{bmatrix} \right\} \]
Remark: We know that if \( B = \{ b_1, \ldots, b_p \} \) is a basis for a subspace \( W \) of \( \mathbb{R}^n \), then each vector \( x \) in \( W \) can be realized (uniquely) as a sum

\[
x = c_1 b_2 + \cdots + c_p b_p.
\]

If \( n \) is very large, the computations needed to determine the coefficients \( c_1, \ldots, c_p \) may require a lot of time (and machine memory).

Question: Can we seek a basis whose nature simplifies this task? And what properties should such a basis possess?
Orthogonal Sets

**Definition:** An indexed set \( \{ \mathbf{u}_1, \ldots, \mathbf{u}_p \} \) in \( \mathbb{R}^n \) is said to be an **orthogonal set** provided each pair of distinct vectors in the set is orthogonal. That is, provided

\[
\mathbf{u}_i \cdot \mathbf{u}_j = 0 \quad \text{whenever} \quad i \neq j.
\]

**Example:** Show that the set \( \begin{Bmatrix} \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ -4 \\ 7 \end{bmatrix} \end{Bmatrix} \) is an orthogonal set.

Label these \( \mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3 \)

\[
\mathbf{u}_1 \cdot \mathbf{u}_2 = 3(-1) + 1(2) + 1(1) = -3 + 2 + 1 = 0
\]

\[
\mathbf{u}_1 \cdot \mathbf{u}_3 = 3(-1) + 1(-4) + 1(7) = -3 - 4 + 7 = 0
\]
\[ \vec{u}_2 \cdot \vec{u}_3 = -1(-1) + 2(-4) + 1(7) = 1 - 8 + 7 = 0. \]

So \( \vec{u}_i \cdot \vec{u}_j = 0 \) for all \( i \neq j \). The set is orthogonal.
Orthogonal Basis

**Definition:** An **orthogonal basis** for a subspace \( W \) of \( \mathbb{R}^n \) is a basis that is also an orthogonal set.

**Theorem:** Let \( \{u_1, \ldots, u_p\} \) be an orthogonal basis for a subspace \( W \) of \( \mathbb{R}^n \). Then each vector \( y \) in \( W \) can be written as the linear combination

\[
y = c_1 u_1 + c_2 u_2 + \cdots + c_p u_p,
\]

where the weights

\[
c_j = \frac{y \cdot u_j}{u_j \cdot u_j}.
\]
\[ \hat{\mathbf{u}}_j \cdot \hat{\mathbf{y}} = \hat{\mathbf{u}}_j \cdot (c_1 \hat{\mathbf{u}}_1 + c_2 \hat{\mathbf{u}}_2 + \ldots + c_j \hat{\mathbf{u}}_j + \ldots + c_p \hat{\mathbf{u}}_p) \]

\[ = c_1 \hat{\mathbf{u}}_1 \cdot \hat{\mathbf{u}}_j + c_2 \hat{\mathbf{u}}_2 \cdot \hat{\mathbf{u}}_j + \ldots + c_j \hat{\mathbf{u}}_j \cdot \hat{\mathbf{u}}_j + \ldots + c_p \hat{\mathbf{u}}_p \cdot \hat{\mathbf{u}}_j \]

\[ = 0 + 0 + \ldots + c_j \| \hat{\mathbf{u}}_j \|^2 + 0 \ldots \]

\[ c_j = \frac{\hat{\mathbf{u}}_j \cdot \hat{\mathbf{y}}}{\| \hat{\mathbf{u}}_j \|^2} \]
Example

\begin{align*}
\left\{ \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ -4 \\ 7 \end{bmatrix} \right\} \text{ is an orthogonal basis of } \mathbb{R}^3. \text{ Express the vector } y = \begin{bmatrix} -2 \\ 3 \\ 0 \end{bmatrix} \text{ as a linear combination of the basis vectors.}
\end{align*}

\begin{align*}
\| \mathbf{u}_1 \|^2 &= 9 + 1 + 1 = 11, \quad \| \mathbf{u}_2 \|^2 = 1 + 4 + 1 = 6, \quad \| \mathbf{u}_3 \|^2 = 1 + 16 + 49 = 66 \\
\mathbf{y} \cdot \mathbf{u}_1 &= -6 + 3 = -3, \quad \mathbf{y} \cdot \mathbf{u}_2 = 2 + 6 = 8, \quad \mathbf{y} \cdot \mathbf{u}_3 = 2 - 12 = -10
\end{align*}

\begin{align*}
y &= \frac{-3}{11} \mathbf{u}_1 + \frac{8}{6} \mathbf{u}_2 - \frac{10}{6} \mathbf{u}_3 \\
&= \frac{-3}{11} \mathbf{u}_1 + \frac{4}{3} \mathbf{u}_2 - \frac{5}{33} \mathbf{u}_3 = -9 \mathbf{u}_1 + 4 \mathbf{u}_2 - 5 \mathbf{u}_3
\end{align*}
Projection

Given a nonzero vector $u$, suppose we wish to decompose another nonzero vector $y$ into a sum of the form

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

in such a way that $\hat{y}$ is parallel to $u$ and $z$ is perpendicular to $u$. 
Projection

Since \( \hat{y} \) is parallel to \( u \), there is a scalar \( \alpha \) such that

\[
\hat{y} = \alpha u.
\]

From \( \hat{y} = \hat{y} + \frac{\hat{y}}{u} \cdot \frac{\hat{y}}{u} \), \( \hat{y} = \hat{y} - \frac{\hat{y}}{u} \cdot \frac{\hat{y}}{u} \).

Dot product with \( u \)

\[
\hat{y} \cdot u = \hat{u} \cdot (\hat{y} - \frac{\hat{y}}{u} \cdot \frac{\hat{y}}{u}) = \hat{u} \cdot \hat{y} - \hat{u} \cdot \frac{\hat{y}}{u} \cdot \frac{\hat{y}}{u}
\]

\( \hat{u} \cdot (\alpha \hat{u}) = \hat{u} \cdot \hat{u} \cdot 0 \)

\( \alpha (\hat{u} \cdot \hat{u}) = \hat{u} \cdot \hat{y} \Rightarrow \alpha = \frac{\hat{u} \cdot \hat{y}}{\left\| \hat{u} \right\|^2} \)