6 Orthogonality and Least Squares

6.1 INNER PRODUCT, LENGTH, AND ORTHOGONALITY
INNER PRODUCT

- If $\mathbf{u}$ and $\mathbf{v}$ are vectors in $\mathbb{R}^n$, then we regard $\mathbf{u}$ and $\mathbf{v}$ as $n \times 1$ matrices.

- The transpose $\mathbf{u}^T$ is a $1 \times n$ matrix, and the matrix product $\mathbf{u}^T \mathbf{v}$ is a $1 \times 1$ matrix, which we write as a single real number (a scalar) without brackets.

- The number $\mathbf{u}^T \mathbf{v}$ is called the inner product of $\mathbf{u}$ and $\mathbf{v}$, and it is written as $\mathbf{u} \cdot \mathbf{v}$.

- The inner product is also referred to as a dot product.
INNER PRODUCT

- If \( \mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} \) and \( \mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \),

then the inner product of \( \mathbf{u} \) and \( \mathbf{v} \) is

\[
\begin{bmatrix} u_1 & u_2 & \cdots & u_n \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n.
\]
INNER PRODUCT

- **Theorem 1:** Let \( u, v, \) and \( w \) be vectors in \( \mathbb{R}^n \), and let \( c \) be a scalar. Then
  
  a. \( u \cdot v = v \cdot u \)
  
  b. \( (u + v) \cdot w = u \cdot w + v \cdot w \)
  
  c. \( (cu) \cdot v = c(\langle u \rangle v) = u (\langle c \rangle v) \)
  
  d. \( u \cdot u \geq 0 \), and \( u \cdot u = 0 \) if and only if \( u = 0 \)

- Properties (b) and (c) can be combined several times to produce the following useful rule:
  
  \( (c_1 u_1 + \cdots + c_p u_p) \cdot w = c_1 (u_1 \cdot w) + \cdots + c_p (u_p \cdot w) \)
THE LENGTH OF A VECTOR

- If \( \mathbf{v} \) is in \( \mathbb{R}^n \), with entries \( v_1, \ldots, v_n \), then the square root of \( \mathbf{v} \mathbf{v}^\top \) is defined because \( \mathbf{v} \mathbf{v}^\top \) is nonnegative.

- **Definition:** The **length** (or **norm**) of \( \mathbf{v} \) is the nonnegative scalar \( \| \mathbf{v} \| \) defined by

  \[
  \| \mathbf{v} \| = \sqrt{\mathbf{v} \mathbf{v}^\top} = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}
  \]

  and \( \| \mathbf{v} \|^2 = \mathbf{v} \mathbf{v}^\top \)

- Suppose \( \mathbf{v} \) is in \( \mathbb{R}^2 \), say, \( \mathbf{v} = \begin{bmatrix} a \\ b \end{bmatrix} \).
THE LENGTH OF A VECTOR

- If we identify \( \mathbf{v} \) with a geometric point in the plane, as usual, then \( \| \mathbf{v} \| \) coincides with the standard notion of the length of the line segment from the origin to \( \mathbf{v} \).

- This follows from the Pythagorean Theorem applied to a triangle such as the one shown in the following figure.

- For any scalar \( c \), the length \( c \mathbf{v} \) is \( |c| \) times the length of \( \mathbf{v} \). That is,

\[
\| c \mathbf{v} \| = |c| \| \mathbf{v} \|
\]
A vector whose length is 1 is called a unit vector.

If we divide a nonzero vector \( \mathbf{v} \) by its length \( \delta \) that is, multiply by \( 1 / \| \mathbf{v} \| \delta \) we obtain a unit vector \( \mathbf{u} \) because the length of \( \mathbf{u} \) is \( (1 / \| \mathbf{v} \| \| \mathbf{v} \| ) \| \mathbf{v} \| \).

The process of creating \( \mathbf{u} \) from \( \mathbf{v} \) is sometimes called normalizing \( \mathbf{v} \), and we say that \( \mathbf{u} \) is in the same direction as \( \mathbf{v} \).
THE LENGTH OF A VECTOR

- **Example 1:** Let \( v = (1, -2, 2, 0) \). Find a unit vector \( u \) in the same direction as \( v \).

- **Solution:** First, compute the length of \( v \):

\[
\|v\|^2 = v \cdot v = (1)^2 + (-2)^2 + (2)^2 + (0)^2 = 9
\]

\[
\|v\| = \sqrt{9} = 3
\]

- Then, multiply \( v \) by \( 1 / \|v\| \) to obtain

\[
u = \frac{1}{\|v\|} v = \frac{1}{3} v = \frac{1}{3} \begin{bmatrix} 1 \\ -2 \\ 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 1/3 \\ -2/3 \\ 2/3 \\ 0 \end{bmatrix}
\]
DISTANCE IN $\mathbb{R}^n$

- To check that $\|u\| = 1$, it suffices to show that $\|u\|^2 = 1$.

$$\|u\|^2 = u \cdot u = \left(\frac{1}{3}\right)^2 + \left(-\frac{2}{3}\right) + \left(\frac{2}{3}\right)^2 + (0)^2$$

$$= \frac{1}{9} + \frac{4}{9} + \frac{4}{9} + 0 = 1$$

- **Definition:** For $u$ and $v$ in $\mathbb{R}^n$, the distance between $u$ and $v$, written as $\text{dist} (u, v)$, is the length of the vector $u - v$. That is,

$$\text{dist} (u,v) = \|u - v\|$$
DISTANCE IN $\mathbb{R}^n$

- Example 2: Compute the distance between the vectors $u = (7,1)$ and $v = (3, 2)$.

- Solution: Calculate

  $$u - v = \begin{bmatrix} 7 \\ 1 \end{bmatrix} - \begin{bmatrix} 3 \\ 2 \end{bmatrix} = \begin{bmatrix} 4 \\ -1 \end{bmatrix}$$

  $$\|u - v\| = \sqrt{4^2 + (-1)^2} = \sqrt{17}$$

- The vectors $u$, $v$, and $u - v$ are shown in the figure on the next slide.

- When the vector $u - v$ is added to $v$, the result is $u$. 
DISTANCE IN $\mathbb{R}^n$

Notice that the parallelogram in the above figure shows that the distance from $\mathbf{u}$ to $\mathbf{v}$ is the same as the distance from $\mathbf{u} - \mathbf{v}$ to $\mathbf{0}$.

The distance between $\mathbf{u}$ and $\mathbf{v}$ is the length of $\mathbf{u} - \mathbf{v}$.
ORTHOGONAL VECTORS

- Consider $\mathbb{R}^2$ or $\mathbb{R}^3$ and two lines through the origin determined by vectors $\mathbf{u}$ and $\mathbf{v}$.
- See the figure below. The two lines shown in the figure are geometrically perpendicular if and only if the distance from $\mathbf{u}$ to $\mathbf{v}$ is the same as the distance from $\mathbf{u}$ to $-\mathbf{v}$.
- This is the same as requiring the squares of the distances to be the same.
ORTHOGONAL VECTORS

- Now
\[
\left[ \text{dist}(u, -v) \right]^2 = \|u - (-v)\|^2 = \|u + v\|^2
\]
\[
= (u + v)(u + v)
\]
\[
= u\cdot(u + v) + v\cdot(u + v) \quad \text{Theorem 1(b)}
\]
\[
= u\cdot u + u\cdot v + v\cdot u + v\cdot v \quad \text{Theorem 1(a), (b)}
\]
\[
= \|u\|^2 + \|v\|^2 + 2u\cdot v \quad \text{Theorem 1(a)}
\]

- The same calculations with \(v\) and \(-v\) interchanged show that
\[
\left[ \text{dist}(u, v) \right]^2 = \|u\|^2 + \|v\|^2 + 2u\cdot(-v)
\]
\[
= \|u\|^2 + \|v\|^2 - 2u\cdot v
\]
ORTHOGONAL VECTORS

- The two squared distances are equal if and only if $2u \cdot v = -2u \cdot v$, which happens if and only if $u \cdot v = 0$.

- This calculation shows that when vectors $u$ and $v$ are identified with geometric points, the corresponding lines through the points and the origin are perpendicular if and only if $u \cdot v = 0$.

- **Definition:** Two vectors $u$ and $v$ in $\mathbb{R}^n$ are orthogonal (to each other) if $u \cdot v = 0$.

- The zero vector is orthogonal to every vector in $\mathbb{R}^n$ because $0^T v = 0$ for all $v$. 
THE PYTHAGOREAN THEOREM

**Theorem 2:** Two vectors $u$ and $v$ are orthogonal if and only if $\|u + v\|^2 = \|u\|^2 + \|v\|^2$.

**Orthogonal Complements**

- If a vector $z$ is orthogonal to every vector in a subspace $W$ of $\mathbb{R}^n$, then $z$ is said to be **orthogonal to** $W$.

- The set of all vectors $z$ that are orthogonal to $W$ is called the **orthogonal complement** of $W$ and is denoted by $W^\perp$ (and read as $\text{⊥}W$ perpendicular or simply $\text{⊥}W$ perp).
ORTHOGONAL COMPLEMENTS

1. A vector $\mathbf{x}$ is in $W^\perp$ if and only if $\mathbf{x}$ is orthogonal to every vector in a set that spans $W$.

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

- **Theorem 3:** Let $A$ be an $m \times n$ matrix. The orthogonal complement of the row space of $A$ is the null space of $A$, and the orthogonal complement of the column space of $A$ is the null space of $A^T$: $(\text{Row } A)^\perp = \text{Nul } A$ and $(\text{Col } A)^\perp = \text{Nul } A^T$
ORTHOGONAL COMPLEMENTS

- **Proof:** The row-column rule for computing $Ax$ shows that if $x$ is in Nul $A$, then $x$ is orthogonal to each row of $A$ (with the rows treated as vectors in $\mathbb{F}^n$).

  Since the rows of $A$ span the row space, $x$ is orthogonal to Row $A$.

  Conversely, if $x$ is orthogonal to Row $A$, then $x$ is certainly orthogonal to each row of $A$, and hence $Ax = 0$.

  This proves the first statement of the theorem.
ORTHOGONAL COMPLEMENTS

- Since this statement is true for any matrix, it is true for $A^T$.

- That is, the orthogonal complement of the row space of $A^T$ is the null space of $A^T$.

- This proves the second statement, because Row $A^T = \text{Col } A$. 

ANGLES IN $\mathbb{R}^2$ AND $\mathbb{R}^3$ (OPTIONAL)

- If $\mathbf{u}$ and $\mathbf{v}$ are nonzero vectors in either $\mathbb{R}^2$ or $\mathbb{R}^3$, then there is a nice connection between their inner product and the angle $\theta$ between the two line segments from the origin to the points identified with $\mathbf{u}$ and $\mathbf{v}$.

- The formula is

$$\mathbf{u} \cdot \mathbf{v} = \|\mathbf{u}\|\|\mathbf{v}\|\cos \theta \quad -----(1)$$

- To verify this formula for vectors in $\mathbb{R}^2$, consider the triangle shown in the figure on the next slide with sides of lengths $\|\mathbf{u}\|$, $\|\mathbf{v}\|$, and $\|\mathbf{u} - \mathbf{v}\|$. 

© 2012 Pearson Education, Inc.
By the law of cosines,

\[ \| \mathbf{u} - \mathbf{v} \|^2 = \| \mathbf{u} \|^2 + \| \mathbf{v} \|^2 - 2\| \mathbf{u} \|\| \mathbf{v} \| \cos \theta \]

which can be rearranged to produce the equations on the next slide.
ANGLES IN $\mathbb{R}^2$ AND $\mathbb{R}^3$ (OPTIONAL)

$\|u\| \|v\| \cos \theta = \frac{1}{2} \left[ \|u\|^2 + \|v\|^2 - \|u - v\|^2 \right]$

$= \frac{1}{2} \left[ u_1^2 + u_2^2 + v_1^2 + v_2^2 - (u_1 - v_1)^2 - (u_2 - v_2)^2 \right]$

$= u_1v_1 + u_2v_2$

$= u \cdot v$

- The verification for $\mathbb{R}^3$ is similar.
- When $n > 3$, formula (1) may be used to define the angle between two vectors in $\mathbb{R}^n$.
- In statistics, the value of $\cos \theta$ defined by (1) for suitable vectors $u$ and $v$ is called a correlation coefficient.