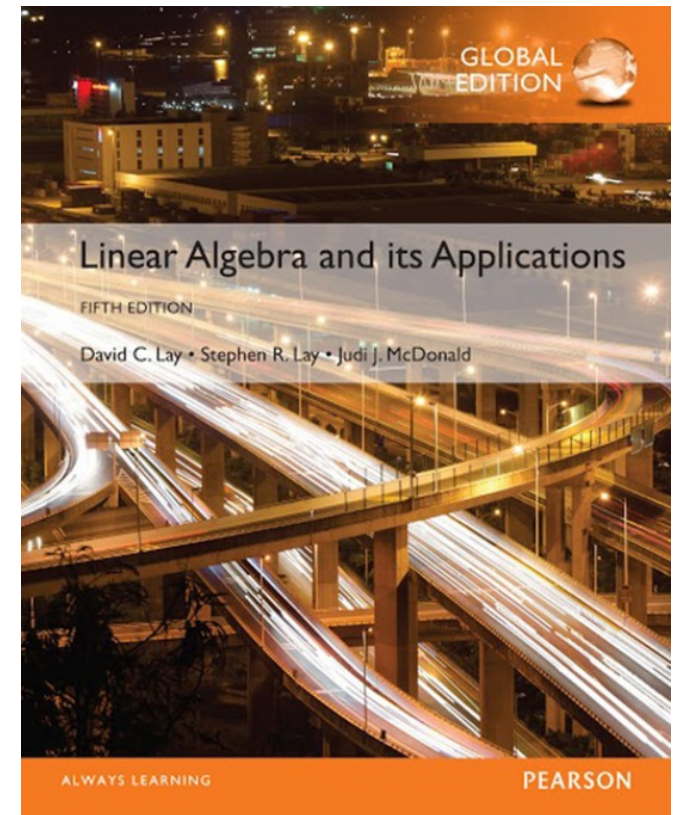


6

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 1×1 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 $u \cdot v$.
- This 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 \mathbf{u} , \mathbf{v} , and \mathbf{w} be vectors in \mathbb{R}^n , and let c be a scalar. Then
 - a.* $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$
 - b.* $(\mathbf{u} + \mathbf{v}) \cdot \mathbf{w} = \mathbf{u} \cdot \mathbf{w} + \mathbf{v} \cdot \mathbf{w}$
 - c.* $(c\mathbf{u}) \cdot \mathbf{v} = c(\mathbf{u} \cdot \mathbf{v}) = \mathbf{u} \cdot (c\mathbf{v})$
 - d.* $\mathbf{u} \cdot \mathbf{u} \geq 0$, and $\mathbf{u} \cdot \mathbf{u} = 0$ if and only if $\mathbf{u} = \mathbf{0}$
- Properties (b) and (c) can be combined several times to produce the following useful
$$\mathbf{u} \cdot (c_1\mathbf{u}_1 + \dots + c_p\mathbf{u}_p) = c_1(\mathbf{u} \cdot \mathbf{u}_1) + \dots + c_p(\mathbf{u} \cdot \mathbf{u}_p)$$

THE LENGTH OF A VECTOR

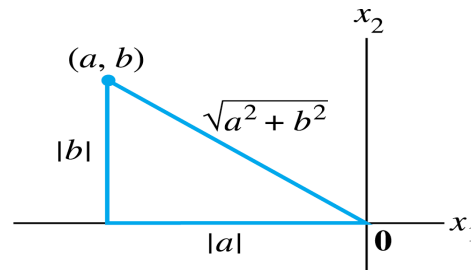
- If \mathbf{v} is in \mathbb{R}^n , with entries v_1, \dots, v_n , then the square root of $\mathbf{v} \cdot \mathbf{v}$ is defined because $\mathbf{v} \cdot \mathbf{v}$ is nonnegative.
- **Definition:** The **length** (or **norm**) of \mathbf{v} is the nonnegative scalar $\|\mathbf{v}\|$ defined by

$$\|\mathbf{v}\| = \sqrt{\mathbf{v} \cdot \mathbf{v}} = \sqrt{v_1^2 + v_2^2 + \dots + v_n^2}, \text{ and } \|\mathbf{v}\|^2 = \mathbf{v} \cdot \mathbf{v}$$

- 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.



Interpretation of $\|\mathbf{v}\|$ as length.

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

THE LENGTH OF A VECTOR

- A vector whose length is 1 is called a **unit vector**.
- If we *divide* a nonzero vector \mathbf{v} by its length—that is, multiply by $1 / \|\mathbf{v}\|$ —we obtain a unit vector \mathbf{u} because the length of \mathbf{u} is $(1 / \|\mathbf{v}\|)\|\mathbf{v}\|$.
- 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 2:** Let $\mathbf{v} = (1, -2, 2, 0)$. Find a unit vector \mathbf{u} in the same direction as \mathbf{v} .
- **Solution:** First, compute the length of \mathbf{v} :

$$\|\mathbf{v}\|^2 = \mathbf{v} \cdot \mathbf{v} = (1)^2 + (-2)^2 + (2)^2 + (0)^2 = 9$$

$$\|\mathbf{v}\| = \sqrt{9} = 3$$

- Then, multiply \mathbf{v} by $1 / \|\mathbf{v}\|$ to obtain

$$\mathbf{u} = \frac{1}{\|\mathbf{v}\|} \mathbf{v} = \frac{1}{3} \mathbf{v} = \frac{1}{3} \begin{bmatrix} 1 \\ -2 \\ 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 1/3 \\ -2/3 \\ 2/3 \\ 0 \end{bmatrix}$$

DISTANCE IN \mathbb{R}^n

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

$$\begin{aligned}\|\mathbf{u}\|^2 &= \mathbf{u} \cdot \mathbf{u} = \left(\frac{1}{3}\right)^2 + \left(-\frac{2}{3}\right)^2 + \left(\frac{2}{3}\right)^2 + (0)^2 \\ &= \frac{1}{9} + \frac{4}{9} + \frac{4}{9} + 0 = 1\end{aligned}$$

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

$$\text{dist}(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\|$$

DISTANCE IN \mathbb{R}^n

- **Example 4:** Compute the distance between the vectors $\mathbf{u} = (7, 1)$ and $\mathbf{v} = (3, 2)$.

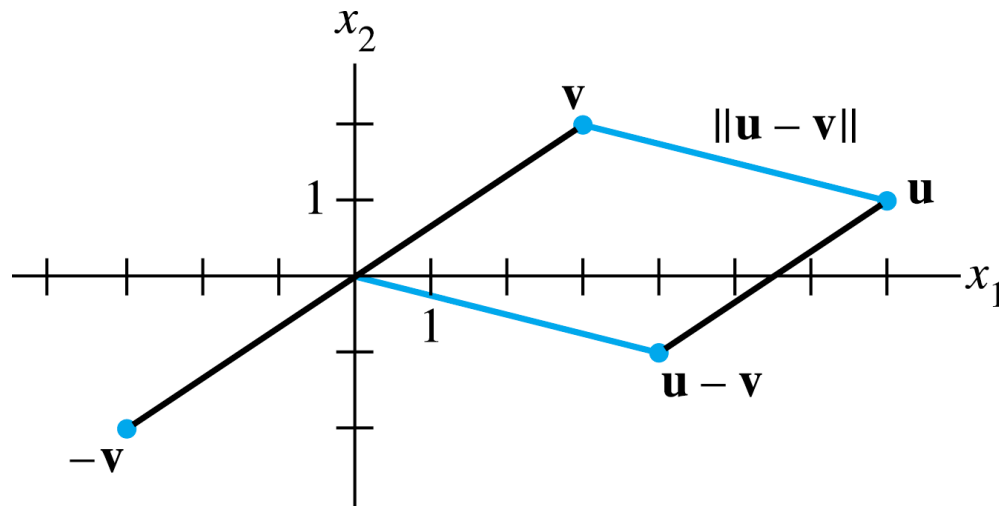
- **Solution:** Calculate

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

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

- The vectors \mathbf{u} , \mathbf{v} , and $\mathbf{u} - \mathbf{v}$ are shown in the figure on the next slide.
- When the vector $\mathbf{u} - \mathbf{v}$ is added to \mathbf{v} , the result is \mathbf{u} .

DISTANCE IN \mathbb{R}^n

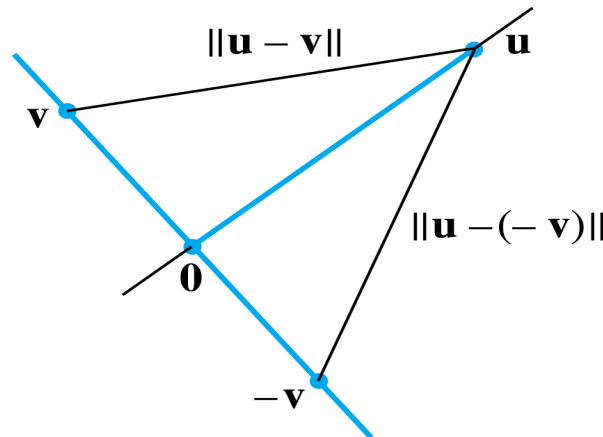


The distance between \mathbf{u} and \mathbf{v} is the length of $\mathbf{u} - \mathbf{v}$.

- 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 $\mathbf{0}$.

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

$$\begin{aligned}[dist(u, -v)]^2 &= \|u - (-v)\|^2 = \|u + v\|^2 \\ &= (u + v) \cdot (u + v)\end{aligned}$$

$$= u \cdot (u + v) + v \cdot (u + v)$$

$$= u \cdot u + u \cdot v + v \cdot u + v \cdot v$$

$$= \|u\|^2 + \|v\|^2 + 2u \cdot v$$

Theorem 1(b)

Theorem 1(a), (b)

Theorem 1(a)

- The same calculations with v and $-v$ interchanged show that

$$[dist(u, v)]^2 = \|u\|^2 + \|-v\|^2 + 2u \cdot (-v)$$

$$= \|u\|^2 + \|v\|^2 - 2u \cdot v$$

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 \mathbf{u} and \mathbf{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 \mathbf{u} and \mathbf{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 $0^T \mathbf{v} = 0$ for all \mathbf{v} .

THE PYTHOGOREAN THEOREM

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

Orthogonal Complements

- If a vector \mathbf{z} is orthogonal to every vector in a subspace W of \mathbb{R}^n , then \mathbf{z} is said to be **orthogonal to W** .
- The set of all vectors \mathbf{z} that are orthogonal to W is called the **orthogonal complement** of W and is denoted by W^\perp (and read as “ W perpendicular” or simply “ 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{R}^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 :

$$\begin{array}{ccc} & \text{and} & \\ (\text{Row } A)^\perp = \text{Nul } A & & (\text{Col } A)^\perp = \text{Nul } A^T \end{array}$$

ORTHOGONAL COMPLEMENTS

- **Proof:** The row-column rule for computing $A\mathbf{x}$ shows that if \mathbf{x} is in $\text{Nul}A$, then \mathbf{x} is orthogonal to each row of A (with the rows treated as vectors in \mathbb{R}^n).
- Since the rows of A span the row space, \mathbf{x} is orthogonal to $\text{Row } A$.
- Conversely, if \mathbf{x} is orthogonal to $\text{Row } A$, then \mathbf{x} is certainly orthogonal to each row of A , and hence $A\mathbf{x} = \mathbf{0}$.
- 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 $\text{Row } A^T = \text{Col } A$.