

Chapter 6

Orthogonality and Least Squares

Section 6.1

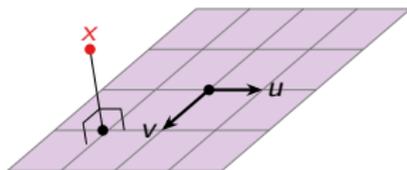
Inner Product, Length, and Orthogonality

Recall: This course is about learning to:

- ▶ Solve the matrix equation $Ax = b$
- ▶ Solve the matrix equation $Ax = \lambda x$
- ▶ Almost solve the equation $Ax = b$

We are now aiming at the last topic.

Idea: In the real world, data is imperfect. Suppose you measure a data point x which you know for theoretical reasons must lie on a plane spanned by two vectors u and v .



Due to measurement error, though, the measured x is not actually in $\text{Span}\{u, v\}$. In other words, the equation $au + bv = x$ has no solution. What do you do? The real value is probably the *closest* point to x on $\text{Span}\{u, v\}$. Which point is that?

The Dot Product

We need a notion of *angle* between two vectors, and in particular, a notion of *orthogonality* (i.e. when two vectors are perpendicular). This is the purpose of the dot product.

Definition

The **dot product** of two vectors x, y in \mathbf{R}^n is

$$x \cdot y = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \cdot \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \stackrel{\text{def}}{=} x_1y_1 + x_2y_2 + \cdots + x_ny_n.$$

Thinking of x, y as column vectors, this is the same as $x^T y$.

Example

$$\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \cdot \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix} = (1 \quad 2 \quad 3) \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix} = 1 \cdot 4 + 2 \cdot 5 + 3 \cdot 6 = 32.$$

Properties of the Dot Product

Many usual arithmetic rules hold, as long as you remember you can only dot two vectors together, and that the result is a scalar.

- ▶ $x \cdot y = y \cdot x$
- ▶ $(x + y) \cdot z = x \cdot z + y \cdot z$
- ▶ $(cx) \cdot y = c(x \cdot y)$

Dotting a vector with itself is special:

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = x_1^2 + x_2^2 + \cdots + x_n^2.$$

Hence:

- ▶ $x \cdot x \geq 0$
- ▶ $x \cdot x = 0$ if and only if $x = 0$.

Important: $x \cdot y = 0$ does *not* imply $x = 0$ or $y = 0$. For example, $\begin{pmatrix} 1 \\ 0 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} = 0$.

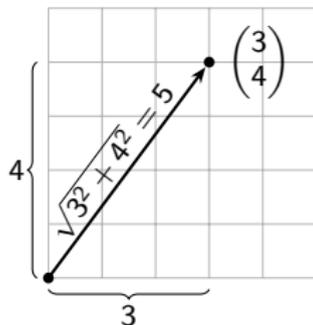
The Dot Product and Length

Definition

The **length** or **norm** of a vector x in \mathbf{R}^n is

$$\|x\| = \sqrt{x \cdot x} = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}.$$

Why is this a good definition? The Pythagorean theorem!



$$\left\| \begin{pmatrix} 3 \\ 4 \end{pmatrix} \right\| = \sqrt{3^2 + 4^2} = 5$$

Fact

If x is a vector and c is a scalar, then $\|cx\| = |c| \cdot \|x\|$.

$$\left\| \begin{pmatrix} 6 \\ 8 \end{pmatrix} \right\| = \left\| 2 \begin{pmatrix} 3 \\ 4 \end{pmatrix} \right\| = 2 \left\| \begin{pmatrix} 3 \\ 4 \end{pmatrix} \right\| = 10$$

The Dot Product and Distance

Definition

The **distance** between two points x, y in \mathbf{R}^n is

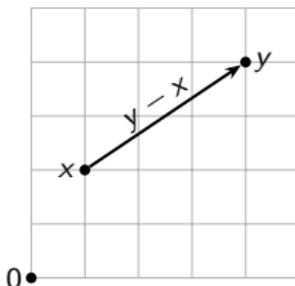
$$\text{dist}(x, y) = \|y - x\|.$$

This is just the length of the vector from x to y .

Example

Let $x = (1, 2)$ and $y = (4, 4)$. Then

$$\text{dist}(x, y) = \|y - x\| = \left\| \begin{pmatrix} 3 \\ 2 \end{pmatrix} \right\| = \sqrt{3^2 + 2^2} = \sqrt{13}.$$



Definition

A **unit vector** is a vector v with length $\|v\| = 1$.

Example

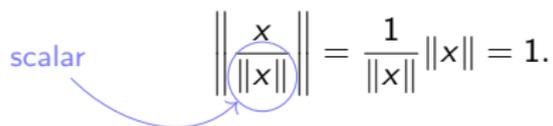
The unit coordinate vectors are unit vectors:

$$\|e_1\| = \left\| \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right\| = \sqrt{1^2 + 0^2 + 0^2} = 1$$

Definition

Let x be a nonzero vector in \mathbf{R}^n . The **unit vector in the direction of x** is the vector $\frac{x}{\|x\|}$.

This is in fact a unit vector:



A blue arrow labeled "scalar" points from the word "scalar" to the fraction $\frac{x}{\|x\|}$ inside the norm symbols of the equation $\left\| \frac{x}{\|x\|} \right\| = \frac{1}{\|x\|} \|x\| = 1$. The fraction $\frac{x}{\|x\|}$ is circled in blue.

$$\left\| \frac{x}{\|x\|} \right\| = \frac{1}{\|x\|} \|x\| = 1.$$

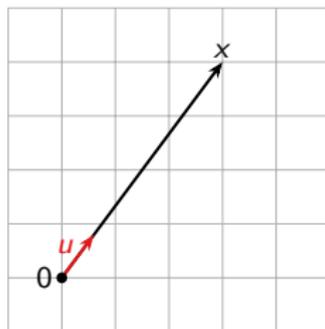
Unit Vectors

Example

Example

What is the unit vector in the direction of $x = \begin{pmatrix} 3 \\ 4 \end{pmatrix}$?

$$u = \frac{x}{\|x\|} = \frac{1}{\sqrt{3^2 + 4^2}} \begin{pmatrix} 3 \\ 4 \end{pmatrix} = \frac{1}{5} \begin{pmatrix} 3 \\ 4 \end{pmatrix}.$$

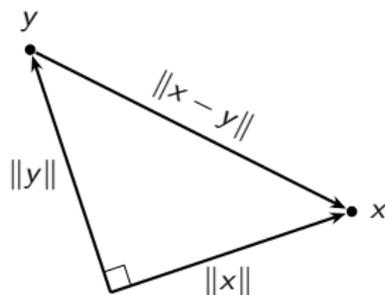


Definition

Two vectors x, y are **orthogonal** or **perpendicular** if $x \cdot y = 0$.

Notation: $x \perp y$ means $x \cdot y = 0$.

Why is this a good definition? The Pythagorean theorem / law of cosines!



$$\begin{aligned}x \text{ and } y \text{ are perpendicular} &\iff \|x\|^2 + \|y\|^2 = \|x - y\|^2 \\ &\iff x \cdot x + y \cdot y = (x - y) \cdot (x - y) \\ &\iff x \cdot x + y \cdot y = x \cdot x + y \cdot y - 2x \cdot y \\ &\iff x \cdot y = 0\end{aligned}$$

Fact: $x \perp y \iff \|x - y\|^2 = \|x\|^2 + \|y\|^2$

Orthogonality

Example

Problem: Find *all* vectors orthogonal to $v = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}$.

We have to find all vectors x such that $x \cdot v = 0$. This means solving the equation

$$0 = x \cdot v = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} = x_1 + x_2 - x_3.$$

The parametric form for the solution is $x_1 = -x_2 + x_3$, so the parametric vector form of the general solution is

$$x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = x_2 \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}.$$

For instance, $\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \perp \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}$ because $\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} = 0$.

Orthogonality

Example

Problem: Find *all* vectors orthogonal to both $v = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}$ and $w = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$.

Now we have to solve the system of two homogeneous equations

$$0 = x \cdot v = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} = x_1 + x_2 - x_3$$

$$0 = x \cdot w = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = x_1 + x_2 + x_3.$$

In matrix form:

The rows are v and $w \longrightarrow \begin{pmatrix} 1 & 1 & -1 \\ 1 & 1 & 1 \end{pmatrix} \xrightarrow{\text{rref}} \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$

The parametric vector form of the solution is

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = x_2 \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}.$$

Orthogonality

General procedure

Problem: Find all vectors orthogonal to some number of vectors v_1, v_2, \dots, v_m in \mathbf{R}^n .

This is the same as finding all vectors x such that

$$0 = v_1^T x = v_2^T x = \dots = v_m^T x.$$

Putting the *row* vectors $v_1^T, v_2^T, \dots, v_m^T$ into a matrix, this is the same as finding all x such that

$$\begin{pmatrix} -v_1^T & - \\ -v_2^T & - \\ \vdots & \vdots \\ -v_m^T & - \end{pmatrix} x = \begin{pmatrix} v_1 \cdot x \\ v_2 \cdot x \\ \vdots \\ v_m \cdot x \end{pmatrix} = 0.$$

Important

The set of all vectors orthogonal to some vectors v_1, v_2, \dots, v_m in \mathbf{R}^n is the *null space* of the $m \times n$ matrix

$$\begin{pmatrix} -v_1^T & - \\ -v_2^T & - \\ \vdots & \vdots \\ -v_m^T & - \end{pmatrix}.$$

In particular, this set is a subspace!

Orthogonal Complements

Definition

Let W be a subspace of \mathbf{R}^n . Its **orthogonal complement** is

$$W^\perp = \{v \text{ in } \mathbf{R}^n \mid v \cdot w = 0 \text{ for all } w \text{ in } W\} \quad \text{read "W perp".}$$

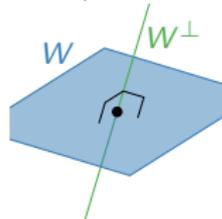
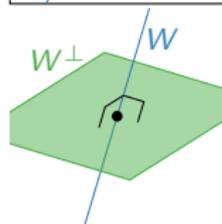
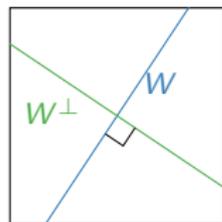
W^\perp is orthogonal complement
 A^T is transpose

Pictures:

The orthogonal complement of a **line** in \mathbf{R}^2 is the perpendicular **line**. [demo]

The orthogonal complement of a **line** in \mathbf{R}^3 is the perpendicular **plane**. [demo]

The orthogonal complement of a **plane** in \mathbf{R}^3 is the perpendicular **line**. [demo]



Poll

Let W be a plane in \mathbf{R}^4 . How would you describe W^\perp ?

- A. The zero space $\{0\}$.
- B. A line in \mathbf{R}^4 .
- C. A plane in \mathbf{R}^4 .
- D. A 3-dimensional space in \mathbf{R}^4 .
- E. All of \mathbf{R}^4 .

Orthogonal Complements

Basic properties

Let W be a subspace of \mathbf{R}^n .

Facts:

1. W^\perp is also a subspace of \mathbf{R}^n
2. $(W^\perp)^\perp = W$
3. $\dim W + \dim W^\perp = n$
4. If $W = \text{Span}\{v_1, v_2, \dots, v_m\}$, then

$$\begin{aligned}W^\perp &= \text{all vectors orthogonal to each } v_1, v_2, \dots, v_m \\ &= \{x \text{ in } \mathbf{R}^n \mid x \cdot v_i = 0 \text{ for all } i = 1, 2, \dots, m\} \\ &= \text{Nul} \begin{pmatrix} -v_1^T & - \\ -v_2^T & - \\ \vdots & \\ -v_m^T & - \end{pmatrix}.\end{aligned}$$

Let's check 1.

- ▶ Is 0 in W^\perp ? Yes: $0 \cdot w = 0$ for any w in W .
- ▶ Suppose x, y are in W^\perp . So $x \cdot w = 0$ and $y \cdot w = 0$ for all w in W . Then $(x + y) \cdot w = x \cdot w + y \cdot w = 0 + 0 = 0$ for all w in W . So $x + y$ is also in W^\perp .
- ▶ Suppose x is in W^\perp . So $x \cdot w = 0$ for all w in W . If c is a scalar, then $(cx) \cdot w = c(x \cdot w) = c(0) = 0$ for any w in W . So cx is in W^\perp .

Orthogonal Complements

Computation

Problem: if $W = \text{Span} \left\{ \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\}$, compute W^\perp .

By property 4, we have to find the null space of the matrix whose rows are $(1 \ 1 \ -1)$ and $(1 \ 1 \ 1)$, which we did before:

$$\text{Nul} \begin{pmatrix} 1 & 1 & -1 \\ 1 & 1 & 1 \end{pmatrix} = \text{Span} \left\{ \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \right\}.$$

$$\text{Span}\{v_1, v_2, \dots, v_m\}^\perp = \text{Nul} \begin{pmatrix} -v_1^T & - \\ -v_2^T & - \\ \vdots & \\ -v_m^T & - \end{pmatrix}$$

Orthogonal Complements

Row space, column space, null space

Definition

The **row space** of an $m \times n$ matrix A is the span of the *rows* of A . It is denoted $\text{Row } A$. Equivalently, it is the column span of A^T :

$$\text{Row } A = \text{Col } A^T.$$

It is a subspace of \mathbf{R}^n .

We showed before that if A has rows $v_1^T, v_2^T, \dots, v_m^T$, then

$$\text{Span}\{v_1, v_2, \dots, v_m\}^\perp = \text{Nul } A.$$

Hence we have shown:

Fact: $(\text{Row } A)^\perp = \text{Nul } A$.

Replacing A by A^T , and remembering $\text{Row } A^T = \text{Col } A$:

Fact: $(\text{Col } A)^\perp = \text{Nul } A^T$.

Using property 2 and taking the orthogonal complements of both sides, we get:

Fact: $(\text{Nul } A)^\perp = \text{Row } A$ and $\text{Col } A = (\text{Nul } A^T)^\perp$.

Orthogonal Complements of Most of the Subspaces We've Seen

For any vectors v_1, v_2, \dots, v_m :

$$\text{Span}\{v_1, v_2, \dots, v_m\}^\perp = \text{Nul} \begin{pmatrix} -v_1^T & - \\ -v_2^T & - \\ \vdots & \\ -v_m^T & - \end{pmatrix}$$

For any matrix A :

$$\text{Row } A = \text{Col } A^T$$

and

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