Chapter 6

Orthogonality and Least Squares

Section 6.1

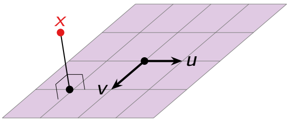
Inner Product, Length, and Orthogonality

Orientation

We are now aiming at the **last topic**.

- Almost solve the equation $Ax = b$

Problem: In the real world, *data is imperfect*.



But due to measurement error, the measured x is not actually in $\text{Span}\{u, v\}$.
But you know, *for theoretical reasons*, it must lie on that plane.

What do you do?
The real value is *probably the closest point*, on the plane, to x.

New terms: Orthogonal projection ('closest point'), orthogonal vectors, angle.

The Dot Product

The dot product encodes the notion of *angle* between two vectors. We will use it to define *orthogonality* (i.e. when two vectors are perpendicular)

Definition

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

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

This is the same as $x^T y$.

Example

$$\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \cdot \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \end{pmatrix} \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix} =$$

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, $\binom{1}{0} \cdot \binom{0}{1} = 0$.

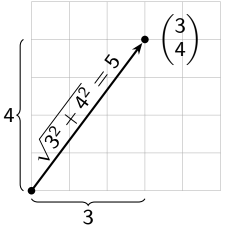
The Dot Product and Length

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 $\|cx\| = |c| \cdot \|x\|$.

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

The Dot Product and Distance

The following is just *the length* of the vector *from x to y*.

Definition

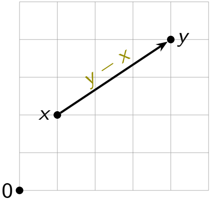
The **distance** between two points x, y in $\mathbf{R}^n$ is

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

Example

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

$$\mathrm{dist}(x, y) =$$



Unit Vectors

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 $\dfrac{x}{\|x\|}$.

Is this really a unit vector?

scalar

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

Example

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

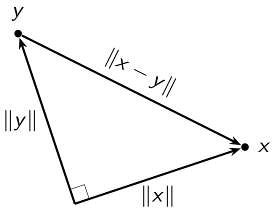
Orthogonality

Definition

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

Notation: Write it as $x \perp y$.

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



Fact: $x \perp y \iff \|x - y\|^2 = \|x\|^2 + \|y\|^2$ (Pythagorean Theorem)

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

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

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

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

Problem: Find all *vectors orthogonal* to $v_1, v_2, \ldots, v_m$ in $\mathbf{R}^n$.

This is the same as finding all vectors x such that

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

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

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

The key observation

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

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

Orthogonal Complements

Definition

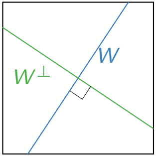
Let W be a subspace of $\mathbf{R}^n$. Its **orthogonal complement** is

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

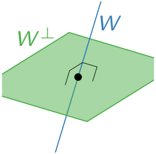
$W^{\perp}$ is orthogonal complement
A^{T} is transpose

Pictures:

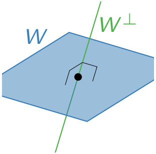
The orthogonal complement of a line in $\mathbf{R}^2$ is the perpendicular line.



The orthogonal complement of a line in $\mathbf{R}^3$ is the perpendicular plane.



The orthogonal complement of a plane in $\mathbf{R}^3$ is the perpendicular line.



Poll

Orthogonal Complements

Basic properties

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

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 $W = \mathrm{Span}\{v_1, v_2, \ldots, v_m\}$, then

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

Property 4

$$\mathrm{Span}\{v_1, v_2, \ldots, v_m\}^\perp = \mathrm{Nul}\begin{pmatrix} — v_1^T — \\ — v_2^T — \\ \vdots \\ — v_m^T — \end{pmatrix}$$

Definition

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

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

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

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

$$\operatorname{Span}\{v_1, v_2, \ldots, v_m\}^\perp = \operatorname{Nul} A.$$

Hence we have shown: $(\operatorname{Row} A)^\perp = \operatorname{Nul} A$.

Extra: Reference sheet

Orthogonal Complements of Most of the Subspaces We've Seen

For any vectors $v_1, v_2, \ldots, v_m$:

$$(\mathrm{Span}\{v_1, v_2, \ldots, v_m\})^\perp = \mathrm{Nul}\begin{pmatrix} — v_1^T — \\ — v_2^T — \\ \vdots \\ — v_m^T — \end{pmatrix}$$

For any matrix A:

$$\mathrm{Row}\, A = \mathrm{Col}\, A^T$$

thus

$$(\mathrm{Row}\, A)^\perp = \mathrm{Nul}\, A \qquad \mathrm{Row}\, A = (\mathrm{Nul}\, A)^\perp$$

$$(\mathrm{Col}\, A)^\perp = \mathrm{Nul}\, A^T \qquad \mathrm{Col}\, A = (\mathrm{Nul}\, A^T)^\perp$$

Extra: Practice proving a set is subspace and some facts