Announcements

November 14

- WeBWorK assignment 5.5 is due on Friday at 6am.
- Midterm 3 will take place in recitation on Friday, 11/18.
  - It covers §§5.1, 5.2, 5.3, 5.5, and the material on stochastic matrices (Perron–Frobenius theorem).
- A practice exam has been posted on the website.
  - I'll post the solutions later this week.
- There are midterm details and study tips on Piazza.
- Post requests for Wednesday's review session on Piazza.
- **Triple office hours this week**: today 2–4pm, Wednesday 1–3pm, Thursday 2:30–4:30pm, and by appointment, in Skiles 221.
  - As always, TAs' office hours are posted on the website.
  - Math Lab is also a good place to visit.

Chapter 6

Orthogonality and Least Squares

Section 6.1

Inner Product, Length, and Orthogonality

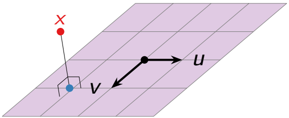
Orientation

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 for 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 $\mathrm{Span}\{u, v\}$. 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 $\mathrm{Span}\{u, v\}$. 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

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

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

Example

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

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

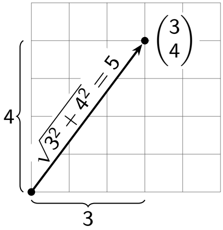
The Dot Product and Length

Definition

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

$$\|x\| = \sqrt{x \cdot x}.$$

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

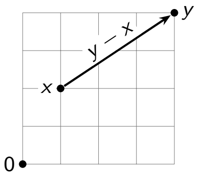
$$\text{dist}(x, y) = \|y - x\|.$$

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

Example

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

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



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

This is in fact a unit vector:

$$\left\| \frac{x}{\underbrace{\|x\|}_{\text{scalar}}} \right\| = \frac{1}{\|x\|}\|x\| = 1.$$

Example

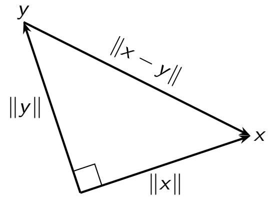
What is the unit vector in the direction of $x = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}$?

$$\frac{x}{\|x\|} = \frac{1}{\sqrt{1^2 + 2^2 + 3^2 + 4^2}} \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix} = \frac{1}{\sqrt{30}} \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}.$$

Orthogonality

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!



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

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

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 $\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} = 0$.

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

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

In matrix form:

$$\begin{pmatrix} 1 & 1 & -1 \\ 1 & 1 & 1 \end{pmatrix} \xrightarrow{\text{rref}} \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

The parametric vector form of the solution is

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

Problem: find all vectors orthogonal to some number of vectors $v_1, v_2, \ldots, v_m$ 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

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

Important

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

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

In particular, this set is a subspace!

Orthogonal Complements

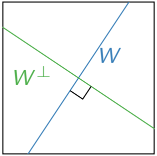
Definition

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

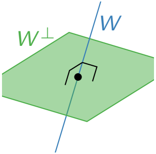
$$W^\perp = \{ v \text{ in } \mathbf{R}^n \mid v \cdot w = 0 \text{ for all } w \text{ in } W \}.$$

Pictures:

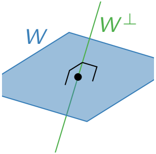
The orthogonal complement of a line in $\mathbf{R}^2$ is the perpendicular line.



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

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 $\mathbf{R}^4$.

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

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

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 $(cx) \cdot w = c(x \cdot 0) = c(0) = 0$ for any w in W. So cx is in $W^\perp$.

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

By property 4, we just have to find all vectors that are orthogonal to both $(1,1,-1)$ and $(1,1,1)$. We know how to do this: it is the null space of the matrix whose *rows* are $(1,1,-1)$ and $(1,1,1)$. We computed it before:

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

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

Definition

The **row space** of an $m \times n$ matrix A is the span of the *rows* of A. It is denoted 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 $v_1^T, v_2^T, \ldots, v_m^T$, then

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

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