

# Math Camp

Justin Grimmer

Associate Professor  
Department of Political Science  
Stanford University

September 4th, 2015

# Where We've Been, Where We're Going

Calculus: Analyze behavior of functions on real line

- Convergence
- Differentiation
- Integration

Linear Algebra

- Data stored in matrices
- Higher dimensional spaces
  - complex world, condition on many factors
  - flood of big data, store in many dimensions
- Linear Algebra:
  - Algebra of matrices
  - Geometry of high dimensional space
  - Calculus (multivariable) in many dimensions

Very important for regression(!!!!)

# Points + Vectors

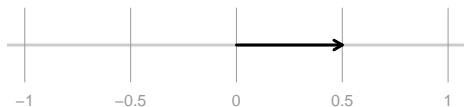
- A point in  $\mathbb{R}^1$ 
  - 1
  - $\pi$
  - $e$
- An ordered pair in  $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$ 
  - $(1, 2)$
  - $(0, 0)$
  - $(\pi, e)$
- An ordered triple in  $\mathbb{R}^3 = \mathbb{R} \times \mathbb{R} \times \mathbb{R}$ 
  - $(3.1, 4.5, 6.11132)$
- ⋮
- An ordered n-tuple in  $\mathbb{R}^n = \mathbb{R} \times \mathbb{R} \times \dots \times \mathbb{R}$ 
  - $(a_1, a_2, \dots, a_n)$

# Points and Vectors

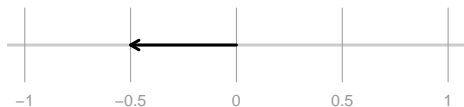
## Definition

*A point  $\mathbf{x} \in \mathbb{R}^n$  is an ordered  $n$ -tuple,  $(x_1, x_2, \dots, x_n)$ . The vector  $\mathbf{x} \in \mathbb{R}^n$  is the arrow pointing from the origin  $(0, 0, \dots, 0)$  to  $\mathbf{x}$ .*

# One Dimensional Example



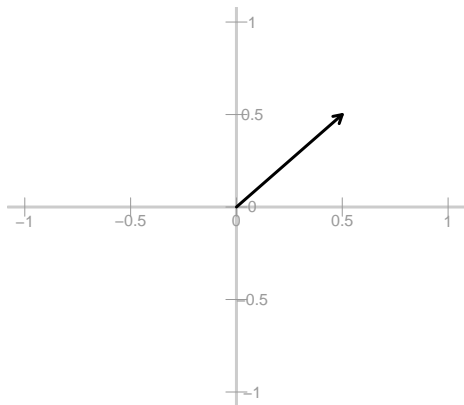
# One Dimensional Example



# One Dimensional Example

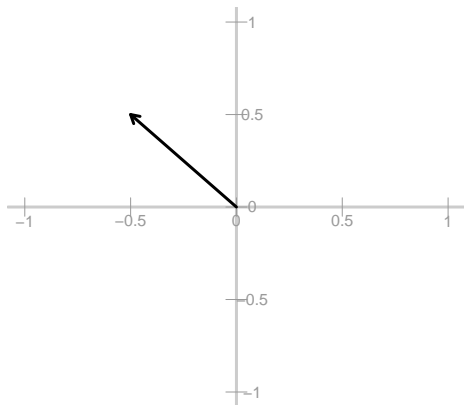


# Two Dimensional Example

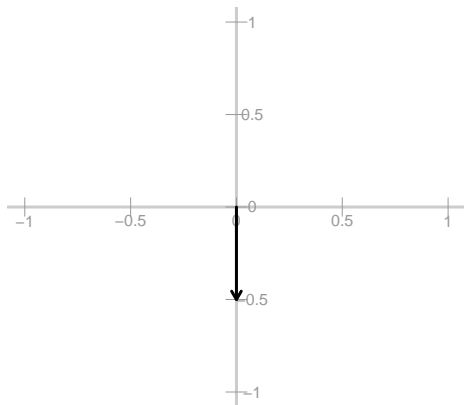




# Two Dimensional Example



# Two Dimensional Example



# Three Dimensional Example

- (Latitude, Longitude, Elevation)
- (1, 2, 3)
- (0, 1, 0)

# N-Dimensional Example

- Individual campaign donation records

$$\mathbf{x} = (1000, 0, 10, 50, 15, 4, 0, 0, 0, \dots, 24000000000)$$

- Counties have proportion of vote for Obama

$$\mathbf{y} = (0.8, 0.5, 0.6, \dots, 0.2)$$

- Run experiment, assess feeling thermometer of elected official

$$\mathbf{t} = (0, 100, 50, 70, 80, \dots, 100)$$

# Arithmetic with Vectors

## Definition

Suppose  $\mathbf{u}$  and  $\mathbf{v}$  are vectors  $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ ,

$$\mathbf{u} = (u_1, u_2, u_3, \dots, u_n)$$

$$\mathbf{v} = (v_1, v_2, v_3, \dots, v_n)$$

We will say  $\mathbf{u} = \mathbf{v}$  if  $u_1 = v_1, u_2 = v_2, \dots, u_n = v_n$

Define the **sum** of  $\mathbf{u} + \mathbf{v}$  as

$$\mathbf{u} + \mathbf{v} = (u_1 + v_1, u_2 + v_2, u_3 + v_3, \dots, u_n + v_n)$$

Suppose  $k \in \mathbb{R}$ . We will call  $k$  a **scalar**.

Define  $k\mathbf{u}$  as the **scalar product**

$$k\mathbf{u} = (ku_1, ku_2, \dots, ku_n)$$

# Examples

Suppose:

$$\mathbf{u} = (1, 2, 3, 4, 5)$$

$$\mathbf{v} = (1, 1, 1, 1, 1)$$

$$k = 2$$

Then,

$$\mathbf{u} + \mathbf{v} = (1 + 1, 2 + 1, 3 + 1, 4 + 1, 5 + 1) = (2, 3, 4, 5, 6)$$

$$k\mathbf{u} = (2 \times 1, 2 \times 2, 2 \times 3, 2 \times 4, 2 \times 5) = (2, 4, 6, 8, 10)$$

$$k\mathbf{v} = (2 \times 1, 2 \times 1, 2 \times 1, 2 \times 1, 2 \times 1) = (2, 2, 2, 2, 2)$$

# Properties of Arithmetic

**Challenge Proofs**—we can do these!

Theorem

*Suppose that  $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^n$  and  $k$  and  $l$  are scalars.*

a)  $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$

# Properties of Arithmetic

**Challenge Proofs**—we can do these!

Theorem

Suppose that  $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^n$  and  $k$  and  $l$  are scalars.

a)  $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$

Proof.

$$\begin{aligned}\mathbf{u} + \mathbf{v} &= (u_1 + v_1, u_2 + v_2, \dots, u_n + v_n) \\ &= (v_1 + u_1, v_2 + u_2, \dots, v_n + u_n) \\ &= \mathbf{v} + \mathbf{u}\end{aligned}$$





# Properties of Arithmetic

**Challenge Proofs**—we can do these!

Theorem

*Suppose that  $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^n$  and  $k$  and  $l$  are scalars.*

b)  $\mathbf{u} + \mathbf{0} = \mathbf{0} + \mathbf{u} = \mathbf{u}$

# Properties of Arithmetic

**Challenge Proofs**—we can do these!

Theorem

Suppose that  $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^n$  and  $k$  and  $l$  are scalars.

b)  $\mathbf{u} + \mathbf{0} = \mathbf{0} + \mathbf{u} = \mathbf{u}$

Proof.

$$\begin{aligned}\mathbf{u} + \mathbf{0} &= (u_1 + 0, u_2 + 0, \dots, u_n + 0) \\ &= (0 + u_1, 0 + u_2, \dots, 0 + u_n) = \mathbf{0} + \mathbf{u} \\ &= (u_1, u_2, \dots, u_n) \\ &= \mathbf{u}\end{aligned}$$



# Properties of Arithmetic

Challenge Proofs—we can do these!

Theorem

Suppose that  $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^n$  and  $k$  and  $l$  are scalars.

$$c) (l + k)\mathbf{u} = l(\mathbf{u}) + k(\mathbf{u})$$

Proof.

How can we show this?



# Challenge Proofs

- Show that  $1\mathbf{u} = \mathbf{u}$
- Show that  $\mathbf{u} + -1\mathbf{u} = \mathbf{0}$

# Inner Product

## Definition

Suppose  $\mathbf{u} \in \mathbb{R}^n$  and  $\mathbf{v} \in \mathbb{R}^n$  then define  $\mathbf{u} \cdot \mathbf{v}$ ,

$$\begin{aligned}\mathbf{u} \cdot \mathbf{v} &= u_1 v_1 + u_2 v_2 + \dots + u_n v_n \\ &= \sum_{i=1}^N u_i v_i\end{aligned}$$

# Examples

Suppose  $\mathbf{u} = (1, 2, 3)$  and  $\mathbf{v} = (2, 3, 1)$ . Then,

$$\begin{aligned}\mathbf{u} \cdot \mathbf{v} &= 1 \times 2 + 2 \times 3 + 3 \times 1 \\ &= 2 + 6 + 3 \\ &= 11\end{aligned}$$

Suppose  $\mathbf{y} = (y_1, y_2, \dots, y_N)$  and  $\mathbf{1} = (1, 1, 1, \dots, 1)$ . Then,

$$\begin{aligned}\mathbf{y} \cdot \mathbf{1} &= y_1 + y_2 + \dots + y_n \\ &= \sum_{i=1}^n y_i\end{aligned}$$

# R Code

Create a vector in R

# R Code

Create a vector in R  
`vec <- c(1, 2, 3, 4, 5)`



# R Code

```
Create a vector in R  
vec <- c(1, 2, 3, 4, 5)  
vec<- c()
```

# R Code

```
Create a vector in R  
vec <- c(1, 2, 3, 4, 5)  
vec<- c()  
vec[1]<- 1
```

# R Code

```
Create a vector in R  
vec <- c(1, 2, 3, 4, 5)  
vec<- c()  
vec[1]<- 1  
vec[2]<- 2
```

# R Code

```
Create a vector in R  
vec <- c(1, 2, 3, 4, 5)  
vec<- c()  
vec[1]<- 1  
vec[2]<- 2  
vec[3]<- 3
```

# R Code

```
Create a vector in R  
vec <- c(1, 2, 3, 4, 5)  
vec<- c()  
vec[1]<- 1  
vec[2]<- 2  
vec[3]<- 3  
vec[4]<- 4
```

# R Code

```
Create a vector in R  
vec <- c(1, 2, 3, 4, 5)  
vec<- c()  
vec[1]<- 1  
vec[2]<- 2  
vec[3]<- 3  
vec[4]<- 4  
vec[5]<- 5
```

# R Code

```
Create a vector in R
vec <- c(1, 2, 3, 4, 5)
vec<- c()
vec[1]<- 1
vec[2]<- 2
vec[3]<- 3
vec[4]<- 4
vec[5]<- 5
x1<- c(2, 2, 3, 2)
```

# R Code

```
Create a vector in R
vec <- c(1, 2, 3, 4, 5)
vec<- c()
vec[1]<- 1
vec[2]<- 2
vec[3]<- 3
vec[4]<- 4
vec[5]<- 5
x1<- c(2, 2, 3, 2)
x2<- c(5, 3, 1, 3)
```



# R Code

```
Create a vector in R
vec <- c(1, 2, 3, 4, 5)
vec<- c()
vec[1]<- 1
vec[2]<- 2
vec[3]<- 3
vec[4]<- 4
vec[5]<- 5
x1<- c(2, 2, 3, 2)
x2<- c(5, 3, 1, 3)
add <- x1 + x2
```

# R Code

```
Create a vector in R
vec <- c(1, 2, 3, 4, 5)
vec<- c()
vec[1]<- 1
vec[2]<- 2
vec[3]<- 3
vec[4]<- 4
vec[5]<- 5
x1<- c(2, 2, 3, 2)
x2<- c(5, 3, 1, 3)
add <- x1 + x2
add
[1] 7 5 4 5
```

# R Code

Create a vector in R

```
vec <- c(1, 2, 3, 4, 5)
```

```
vec<- c()
```

```
vec[1]<- 1
```

```
vec[2]<- 2
```

```
vec[3]<- 3
```

```
vec[4]<- 4
```

```
vec[5]<- 5
```

```
x1<- c(2, 2, 3, 2)
```

```
x2<- c(5, 3, 1, 3)
```

```
add <- x1 + x2
```

```
add
```

```
[1] 7 5 4 5
```

```
scalar<- 10 *x1
```

# R Code

```
Create a vector in R
vec <- c(1, 2, 3, 4, 5)
vec<- c()
vec[1]<- 1
vec[2]<- 2
vec[3]<- 3
vec[4]<- 4
vec[5]<- 5
x1<- c(2, 2, 3, 2)
x2<- c(5, 3, 1, 3)
add <- x1 + x2
add
[1] 7 5 4 5
```

```
scalar<- 10 *x1
scalar
[1] 20 20 30 20
```

# R Code

Create a vector in R  
`vec <- c(1, 2, 3, 4, 5)`

`vec<- c()`

`vec[1]<- 1`

`vec[2]<- 2`

`vec[3]<- 3`

`vec[4]<- 4`

`vec[5]<- 5`

`x1<- c(2, 2, 3, 2)`

`x2<- c(5, 3, 1, 3)`

`add <- x1 + x2`

`add`

`[1] 7 5 4 5`

`scalar<- 10 *x1`

`scalar`

`[1] 20 20 30 20`

`output<- x1 %*% x2`

# R Code

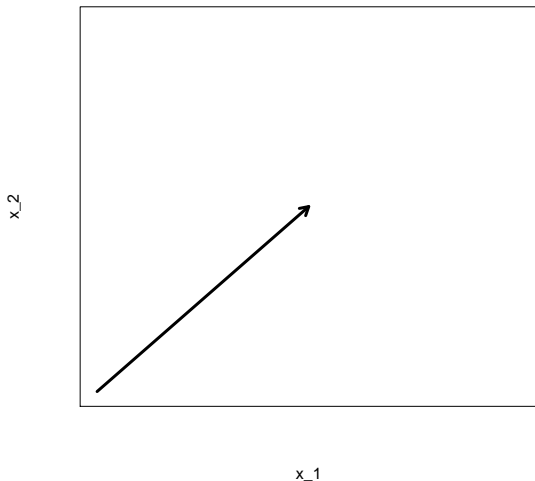
```
Create a vector in R
vec <- c(1, 2, 3, 4, 5)
vec<- c()
vec[1]<- 1
vec[2]<- 2
vec[3]<- 3
vec[4]<- 4
vec[5]<- 5
x1<- c(2, 2, 3, 2)
x2<- c(5, 3, 1, 3)
add <- x1 + x2
add
[1] 7 5 4 5
```

```
scalar<- 10 *x1
scalar
[1] 20 20 30 20
output<- x1 %*% x2
output
[,1]
[1,] 25
```

# Challenge Problems

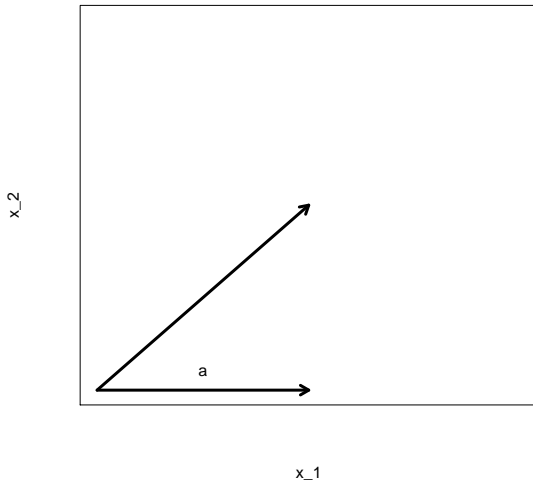
- Suppose  $\mathbf{v} = (1, 4, 1, 4)$  and  $\mathbf{w} = (4, 1, 4, 1)$ . Calculate:  $\mathbf{v} \cdot \mathbf{w}$
- Prove  $\mathbf{v} \cdot \mathbf{w} = \mathbf{w} \cdot \mathbf{v}$
- **Super hard:** prove  $\mathbf{v} \cdot \mathbf{v} \geq 0$  and  $\mathbf{v} \cdot \mathbf{v} = 0$  if and only if  $\mathbf{v} = \mathbf{0}$ .

# Vector Length



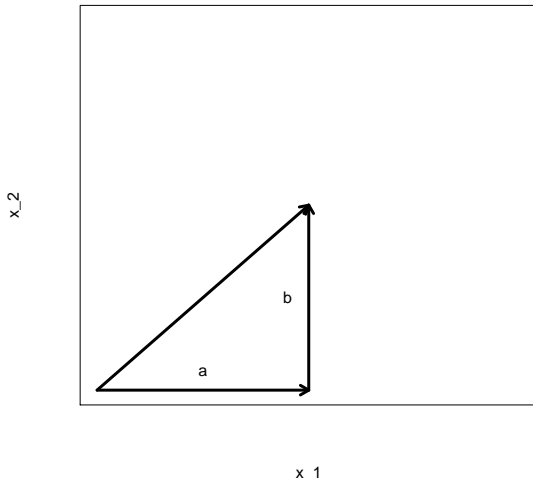


# Vector Length



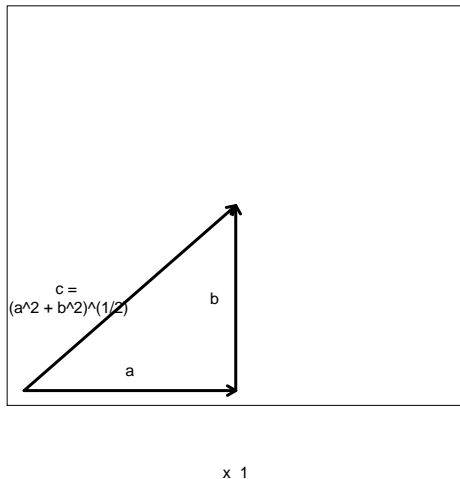
- **Pythagorean Theorem:**  
Side with length  $a$

# Vector Length



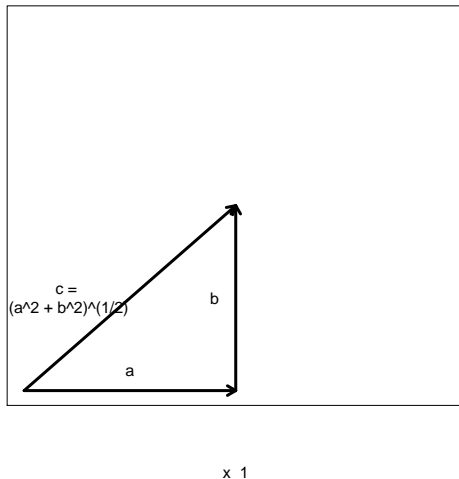
- **Pythagorean Theorem:**  
Side with length  $a$
- Side with length  $b$  and  
right triangle

# Vector Length



- **Pythagorean Theorem:**  
Side with length  $a$
- Side with length  $b$  and right triangle
- $c = \sqrt{a^2 + b^2}$

# Vector Length



- **Pythagorean Theorem:**  
Side with length  $a$
- Side with length  $b$  and  
right triangle
- $c = \sqrt{a^2 + b^2}$
- **This is generally true**

# Vector Length

## Definition

Suppose  $\mathbf{v} \in \mathbb{R}^n$ . Then, we will define its *length* as

$$\begin{aligned}\|\mathbf{v}\| &= (\mathbf{v} \cdot \mathbf{v})^{1/2} \\ &= (v_1^2 + v_2^2 + v_3^2 + \dots + v_n^2)^{1/2}\end{aligned}$$

# Calculating a Length

Example 1: suppose  $\mathbf{x} = (1, 1, 1)$  .

$$\begin{aligned}\|\mathbf{x}\| &= (\mathbf{x} \cdot \mathbf{x})^{1/2} \\ &= (1 + 1 + 1)^{1/2} \\ &= \sqrt{3}\end{aligned}$$

Example 2: R code for length function

```
len.vec<- function(x) {  
  p1<- sqrt(x%*%x)  
  return(p1)  
}  
x <- c(1,1,1)  
len.vec(x)  
[,1]  
[1,] 1.732051
```

# Coding Problem

Let's calculate the length of some vectors

- Write a function to assess the length of a vector.
- Use it to calculate the length of:
  - `y<- c(10, 20, 30, 40)`
  - `x<- seq(1, 1000*pi, len=1000)`

# Texts in Space



# Texts in Space

$$\text{Doc1} = (1, 1, 3, \dots, 5)$$

# Texts in Space

Doc1 =  $(1, 1, 3, \dots, 5)$

Doc2 =  $(2, 0, 0, \dots, 1)$

# Texts in Space

$$\text{Doc1} = (1, 1, 3, \dots, 5)$$

$$\text{Doc2} = (2, 0, 0, \dots, 1)$$

$$\mathbf{Doc1}, \mathbf{Doc2} \in \mathbb{R}^M$$

# Texts in Space

$$\text{Doc1} = (1, 1, 3, \dots, 5)$$

$$\text{Doc2} = (2, 0, 0, \dots, 1)$$

$$\mathbf{Doc1}, \mathbf{Doc2} \in \mathbb{R}^M$$

Provides **many** operations that will be useful

# Texts in Space

$$\text{Doc1} = (1, 1, 3, \dots, 5)$$

$$\text{Doc2} = (2, 0, 0, \dots, 1)$$

$$\mathbf{Doc1}, \mathbf{Doc2} \in \mathbb{R}^M$$

Provides **many** operations that will be useful

**Inner Product** between documents:

# Texts in Space

$$\text{Doc1} = (1, 1, 3, \dots, 5)$$

$$\text{Doc2} = (2, 0, 0, \dots, 1)$$

$$\mathbf{Doc1}, \mathbf{Doc2} \in \mathbb{R}^M$$

Provides **many** operations that will be useful

**Inner Product** between documents:

$$\mathbf{Doc1} \cdot \mathbf{Doc2} = (1, 1, 3, \dots, 5)' (2, 0, 0, \dots, 1)$$

# Texts in Space

$$\text{Doc1} = (1, 1, 3, \dots, 5)$$

$$\text{Doc2} = (2, 0, 0, \dots, 1)$$

$$\mathbf{Doc1}, \mathbf{Doc2} \in \mathbb{R}^M$$

Provides **many** operations that will be useful

**Inner Product** between documents:

$$\begin{aligned}\mathbf{Doc1} \cdot \mathbf{Doc2} &= (1, 1, 3, \dots, 5)' (2, 0, 0, \dots, 1) \\ &= 1 \times 2 + 1 \times 0 + 3 \times 0 + \dots + 5 \times 1\end{aligned}$$

# Texts in Space

$$\text{Doc1} = (1, 1, 3, \dots, 5)$$

$$\text{Doc2} = (2, 0, 0, \dots, 1)$$

$$\mathbf{Doc1}, \mathbf{Doc2} \in \mathbb{R}^M$$

Provides **many** operations that will be useful

**Inner Product** between documents:

$$\begin{aligned}\mathbf{Doc1} \cdot \mathbf{Doc2} &= (1, 1, 3, \dots, 5)' (2, 0, 0, \dots, 1) \\ &= 1 \times 2 + 1 \times 0 + 3 \times 0 + \dots + 5 \times 1 \\ &= 7\end{aligned}$$



Length of document:

Length of document:

$$\begin{aligned} ||\mathbf{Doc1}|| &\equiv \sqrt{\mathbf{Doc1} \cdot \mathbf{Doc1}} \\ &= \sqrt{(1, 1, 3, \dots, 5)'(1, 1, 3, \dots, 5)} \\ &= \sqrt{1^2 + 1^2 + 3^2 + 5^2} \\ &= 6 \end{aligned}$$

Length of document:

$$\begin{aligned} ||\mathbf{Doc1}|| &\equiv \sqrt{\mathbf{Doc1} \cdot \mathbf{Doc1}} \\ &= \sqrt{(1, 1, 3, \dots, 5)'(1, 1, 3, \dots, 5)} \\ &= \sqrt{1^2 + 1^2 + 3^2 + 5^2} \\ &= 6 \end{aligned}$$

Cosine of the angle between documents:

Length of document:

$$\begin{aligned} ||\mathbf{Doc1}|| &\equiv \sqrt{\mathbf{Doc1} \cdot \mathbf{Doc1}} \\ &= \sqrt{(1, 1, 3, \dots, 5)'(1, 1, 3, \dots, 5)} \\ &= \sqrt{1^2 + 1^2 + 3^2 + 5^2} \\ &= 6 \end{aligned}$$

Cosine of the angle between documents:

$$\begin{aligned} \cos \theta &\equiv \left( \frac{\mathbf{Doc1}}{||\mathbf{Doc1}||} \right) \cdot \left( \frac{\mathbf{Doc2}}{||\mathbf{Doc2}||} \right) \\ &= \frac{7}{6 \times 2.24} \\ &= 0.52 \end{aligned}$$

# Measuring Similarity

Documents in space  $\rightarrow$  measure similarity/dissimilarity

# Measuring Similarity

Documents in space  $\rightarrow$  measure similarity/dissimilarity  
What properties should similarity measure have?

# Measuring Similarity

Documents in space  $\rightarrow$  measure similarity/dissimilarity

What properties should similarity measure have?

- Maximum: document with itself

# Measuring Similarity

Documents in space  $\rightarrow$  measure similarity/dissimilarity

What properties should similarity measure have?

- Maximum: document with itself
- Minimum: documents have no words in common (orthogonal)



# Measuring Similarity

Documents in space  $\rightarrow$  measure similarity/dissimilarity

What properties should similarity measure have?

- Maximum: document with itself
- Minimum: documents have no words in common (orthogonal )
- Increasing when more of same words used

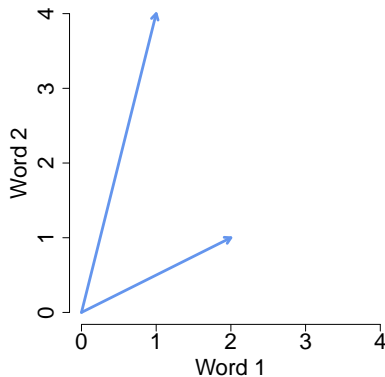
# Measuring Similarity

Documents in space  $\rightarrow$  measure similarity/dissimilarity

What properties should similarity measure have?

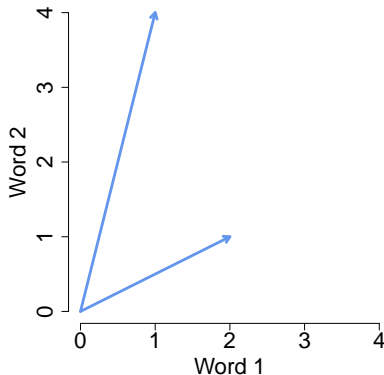
- Maximum: document with itself
- Minimum: documents have no words in common (orthogonal)
- Increasing when more of same words used
- ?  $s(a, b) = s(b, a)$ .

# Measuring Similarity



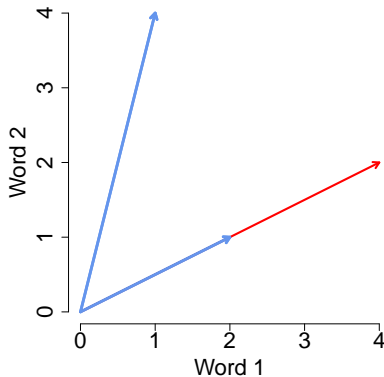
Measure 1: Inner product

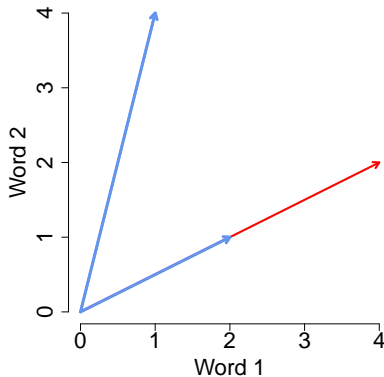
# Measuring Similarity



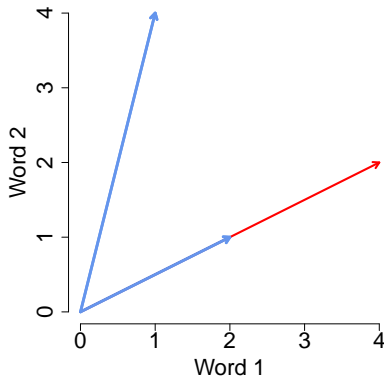
Measure 1: Inner product

$$(2, 1)' \cdot (1, 4) = 6$$



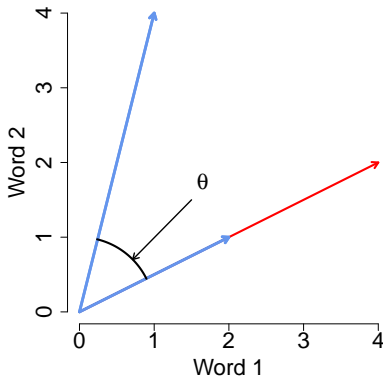


Problem(?): length dependent



**Problem(?)**: length dependent

$$(4, 2)'(1, 4) = 12$$



**Problem(?)**: length dependent

$$(4, 2)'(1, 4) = 12$$

$$a \cdot b = ||a|| \times ||b|| \times \cos \theta$$



# Cosine Similarity

$\cos \theta$ : removes document length from similarity measure

# Cosine Similarity

$\cos \theta$ : removes document length from similarity measure

$$\cos \theta = \left( \frac{a}{||a||} \right) \cdot \left( \frac{b}{||b||} \right)$$

# Cosine Similarity

$\cos \theta$ : removes document length from similarity measure

$$\cos \theta = \left( \frac{a}{||a||} \right) \cdot \left( \frac{b}{||b||} \right)$$
$$\frac{(4, 2)}{||(4, 2)||} = (0.89, 0.45)$$

# Cosine Similarity

$\cos \theta$ : removes document length from similarity measure

$$\cos \theta = \left( \frac{a}{\|a\|} \right) \cdot \left( \frac{b}{\|b\|} \right)$$

$$\frac{(4, 2)}{\|(4, 2)\|} = (0.89, 0.45)$$

$$\frac{(2, 1)}{\|(2, 1)\|} = (0.89, 0.45)$$

# Cosine Similarity

$\cos \theta$ : removes document length from similarity measure

$$\cos \theta = \left( \frac{a}{||a||} \right) \cdot \left( \frac{b}{||b||} \right)$$

$$\frac{(4, 2)}{||(4, 2)||} = (0.89, 0.45)$$

$$\frac{(2, 1)}{||(2, 1)||} = (0.89, 0.45)$$

$$\frac{(1, 4)}{||(1, 4)||} = (0.24, 0.97)$$

# Cosine Similarity

$\cos \theta$ : removes document length from similarity measure

$$\cos \theta = \left( \frac{a}{\|a\|} \right) \cdot \left( \frac{b}{\|b\|} \right)$$

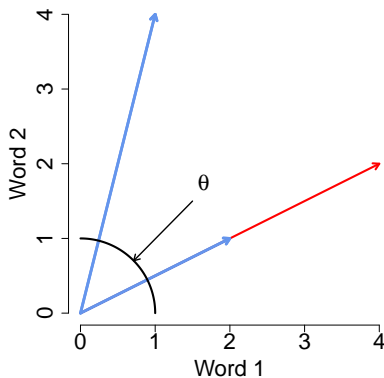
$$\frac{(4, 2)}{\|(4, 2)\|} = (0.89, 0.45)$$

$$\frac{(2, 1)}{\|(2, 1)\|} = (0.89, 0.45)$$

$$\frac{(1, 4)}{\|(1, 4)\|} = (0.24, 0.97)$$

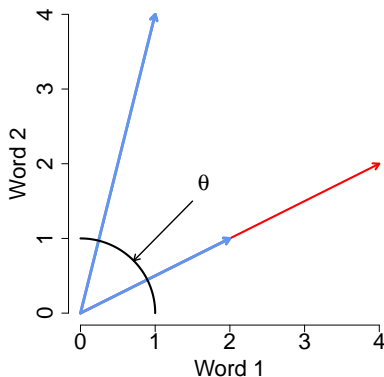
$$(0.89, 0.45)' (0.24, 0.97) = 0.65$$

# Cosine Similarity



$\cos \theta$ : removes document length from similarity measure

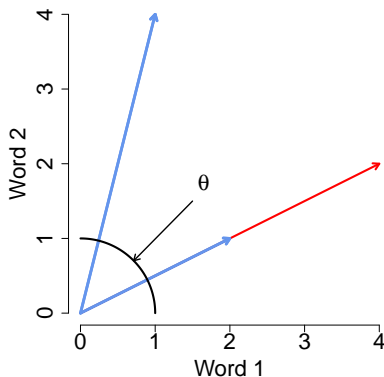
# Cosine Similarity



$\cos \theta$ : removes document length from similarity measure  
Project onto Hypersphere



# Cosine Similarity

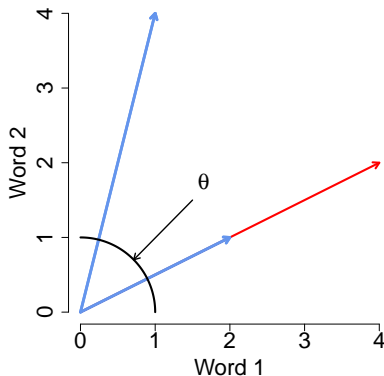


$\cos \theta$ : removes document length from similarity measure

Project onto Hypersphere

$\cos \theta \rightarrow$  Inverse distance on Hypersphere

# Cosine Similarity



$\cos \theta$ : removes document length from similarity measure

Project onto Hypersphere

$\cos \theta \rightarrow$  Inverse distance on Hypersphere

**von Mises Fisher distribution** : distribution on sphere surface

# Matrices

## Definition

A **Matrix** is a rectangular array of numbers

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

If  $\mathbf{A}$  has  $m$  rows  $n$  columns we will say that  $\mathbf{A}$  is an  $m \times n$  matrix.

Suppose  $\mathbf{X}$  and  $\mathbf{Y}$  are  $m \times n$  matrices. Then  $\mathbf{X} = \mathbf{Y}$  if  $x_{ij} = y_{ij}$  for all  $i$  and  $j$

# Simple Examples

$$\mathbf{I} = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}$$

If  $\mathbf{I}$  is an  $n \times n$  matrix we will call an **identity** matrix.

# Simple Examples

$$\mathbf{X} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 4 \end{pmatrix}$$

$\mathbf{X}$  is an  $2 \times 3$  matrix

# Matrix Algebra

## Definition

Suppose  $\mathbf{X}$  and  $\mathbf{Y}$  are  $m \times n$  matrices. Then define

$$\begin{aligned}\mathbf{X} + \mathbf{Y} &= \begin{pmatrix} x_{11} & x_{12} & \dots & x_{1n} \\ x_{21} & x_{22} & \dots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{m1} & x_{m2} & \dots & x_{mn} \end{pmatrix} + \begin{pmatrix} y_{11} & y_{12} & \dots & y_{1n} \\ y_{21} & y_{22} & \dots & y_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ y_{m1} & y_{m2} & \dots & y_{mn} \end{pmatrix} \\ &= \begin{pmatrix} x_{11} + y_{11} & x_{12} + y_{12} & \dots & x_{1n} + y_{1n} \\ x_{21} + y_{21} & x_{22} + y_{22} & \dots & x_{2n} + y_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{m1} + y_{m1} & x_{m2} + y_{m2} & \dots & x_{mn} + y_{mn} \end{pmatrix}\end{aligned}$$

# Matrix Algebra

## Definition

Suppose  $\mathbf{X}$  is an  $m \times n$  matrix and  $k \in \mathbb{R}$ . Then,

$$k\mathbf{X} = \begin{pmatrix} kx_{11} & kx_{12} & \dots & kx_{1n} \\ kx_{21} & kx_{22} & \dots & kx_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ kx_{m1} & kx_{m2} & \dots & kx_{mn} \end{pmatrix}$$

Prove theorems about this tonight

# R Code

Using **matrix** command `mat1<- matrix(NA, nrow=3, ncol=2) ##`

Creating matrix

```
mat1[1,1]<- 1
```

```
mat1[1,2]<- 2
```

```
mat1[2,1]<- 1
```

```
mat1[2,2]<- 4
```

```
mat1[3,1]<- exp(1)
```

```
mat1[3,2]<- 4
```



# R Code

Using **rbind**

```
r1<- c(1, 2)
```

```
r2<- c(1, 4)
```

```
r3<- c(exp(1) , 4)
```

```
mat1<- rbind(r1, r2, r3)
```

# R Code

Using **cbind**

```
c1<- c(1, 1, exp(1) )
```

```
c2<- c(2, 4, 4)
```

# R Code

```
dim(mat1)
[1] 3 2
mat1 + mat1
[,1] [,2]
[1,] 2.000000 4
[2,] 2.000000 8
[3,] 5.436564 8
```

# R Code

What if the matrices are of different dimension

```
mat1<- matrix(1, nrow=3, ncol=2)
```

```
mat2<- matrix(2, nrow=10, ncol=3)
```

```
mat1 + mat2
```

```
Error in mat1 + mat2 : non-conformable arrays
```

# Matrix Transpose

We will use `matrix transpose` to flip the dimensionality of a matrix

# Matrix Transpose

We will use **matrix transpose** to flip the dimensionality of a matrix

$$\mathbf{X} = \begin{pmatrix} x_{11} & x_{12} & \dots & x_{1n} \\ x_{21} & x_{22} & \dots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{m1} & x_{m2} & \dots & x_{mn} \end{pmatrix}$$

# Matrix Transpose

We will use **matrix transpose** to flip the dimensionality of a matrix

$$\mathbf{X} = \begin{pmatrix} x_{11} & x_{12} & \dots & x_{1n} \\ x_{21} & x_{22} & \dots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{m1} & x_{m2} & \dots & x_{mn} \end{pmatrix}$$
$$\mathbf{X}' = \begin{pmatrix} x_{11} \\ x_{12} \\ \vdots \\ x_{1n} \end{pmatrix}$$

# Matrix Transpose

We will use **matrix transpose** to flip the dimensionality of a matrix

$$\mathbf{X} = \begin{pmatrix} x_{11} & x_{12} & \dots & x_{1n} \\ \color{red}{x_{21}} & \color{red}{x_{22}} & \dots & \color{red}{x_{2n}} \\ \vdots & \vdots & \ddots & \vdots \\ x_{m1} & x_{m2} & \dots & x_{mn} \end{pmatrix}$$
$$\mathbf{X}' = \begin{pmatrix} x_{11} & \color{red}{x_{21}} \\ x_{12} & \color{red}{x_{22}} \\ \vdots & \vdots \\ x_{1n} & \color{red}{x_{2n}} \end{pmatrix}$$



# Matrix Transpose

We will use **matrix transpose** to flip the dimensionality of a matrix

$$\mathbf{X} = \begin{pmatrix} x_{11} & x_{12} & \dots & x_{1n} \\ x_{21} & x_{22} & \dots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{m1} & x_{m2} & \dots & x_{mn} \end{pmatrix}$$
$$\mathbf{X}' = \begin{pmatrix} x_{11} & x_{21} & \dots \\ x_{12} & x_{22} & \dots \\ \vdots & \vdots & \ddots \\ x_{1n} & x_{2n} & \dots \end{pmatrix}$$

# Matrix Transpose

We will use **matrix transpose** to flip the dimensionality of a matrix

$$\mathbf{X} = \begin{pmatrix} x_{11} & x_{12} & \dots & x_{1n} \\ x_{21} & x_{22} & \dots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{m1} & x_{m2} & \dots & x_{mn} \end{pmatrix}$$
$$\mathbf{X}' = \begin{pmatrix} x_{11} & x_{21} & \dots & x_{m1} \\ x_{12} & x_{22} & \dots & x_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ x_{1n} & x_{2n} & \dots & x_{mn} \end{pmatrix}$$

# Matrix Transpose

We will use **matrix transpose** to flip the dimensionality of a matrix

$$\mathbf{X} = \begin{pmatrix} x_{11} & x_{12} & \dots & x_{1n} \\ x_{21} & x_{22} & \dots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{m1} & x_{m2} & \dots & x_{mn} \end{pmatrix}$$
$$\mathbf{X}' = \begin{pmatrix} x_{11} & x_{21} & \dots & x_{m1} \\ x_{12} & x_{22} & \dots & x_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ x_{1n} & x_{2n} & \dots & x_{mn} \end{pmatrix}$$

If  $\mathbf{X}$  is an  $m \times n$  then  $\mathbf{X}'$  is  $n \times m$ .

# Matrix Transpose

We will use **matrix transpose** to flip the dimensionality of a matrix

$$\mathbf{X} = \begin{pmatrix} x_{11} & x_{12} & \dots & x_{1n} \\ x_{21} & x_{22} & \dots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{m1} & x_{m2} & \dots & x_{mn} \end{pmatrix}$$
$$\mathbf{X}' = \begin{pmatrix} x_{11} & x_{21} & \dots & x_{m1} \\ x_{12} & x_{22} & \dots & x_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ x_{1n} & x_{2n} & \dots & x_{mn} \end{pmatrix}$$

If  $\mathbf{X}$  is an  $m \times n$  then  $\mathbf{X}'$  is  $n \times m$ .

If  $\mathbf{X} = \mathbf{X}'$  then we say  $\mathbf{X}$  is symmetric.

# Matrix Transpose

Example 1:  $\mathbf{X} = \begin{pmatrix} 4 & 1 & 2 \\ 1 & 2 & 3 \end{pmatrix}$  then  $\mathbf{X}' = \begin{pmatrix} 4 & 1 \\ 1 & 2 \\ 2 & 3 \end{pmatrix}$

In R

```
mat1<- matrix(c(1, 2, 3), nrow=3, ncol=2)
```

```
mat2<- t(mat1)
```

```
dim(mat1)
```

```
3 2
```

```
dim(mat2)
```

```
2 3
```

# Matrix Multiplication

How do we **multiply** matrices?

# Matrix Multiplication

How do we **multiply** matrices?

Because we want to use matrix multiplication to solve equations we won't use an intuitive definition

# Matrix Multiplication

How do we **multiply** matrices?

Because we want to use matrix multiplication to solve equations we won't use an intuitive definition

Suppose we have two matrices



# Matrix Multiplication

How do we **multiply** matrices?

Because we want to use matrix multiplication to solve equations we won't use an intuitive definition

Suppose we have two matrices

$$\mathbf{X} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \mathbf{Y} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$$

# Matrix Multiplication

How do we **multiply** matrices?

Because we want to use matrix multiplication to solve equations we won't use an intuitive definition

Suppose we have two matrices

$$\mathbf{X} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \mathbf{Y} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$$

We will create a new matrix **A** by matrix multiplication:

# Matrix Multiplication

How do we **multiply** matrices?

Because we want to use matrix multiplication to solve equations we won't use an intuitive definition

Suppose we have two matrices

$$\mathbf{X} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \mathbf{Y} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$$

We will create a new matrix **A** by matrix multiplication:

$$\mathbf{A} = \mathbf{XY}$$

# Matrix Multiplication

How do we **multiply** matrices?

Because we want to use matrix multiplication to solve equations we won't use an intuitive definition

Suppose we have two matrices

$$\mathbf{X} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \mathbf{Y} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$$

We will create a new matrix  $\mathbf{A}$  by matrix multiplication:

$$\begin{aligned} \mathbf{A} &= \mathbf{XY} \\ &= \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \end{aligned}$$

# Matrix Multiplication

How do we **multiply** matrices?

Because we want to use matrix multiplication to solve equations we won't use an intuitive definition

Suppose we have two matrices

$$\mathbf{X} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \mathbf{Y} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$$

We will create a new matrix **A** by matrix multiplication:

$$\begin{aligned} \mathbf{A} &= \mathbf{XY} \\ &= \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \\ &= \begin{pmatrix} 1 \times 1 + 1 \times 3 & \\ & \end{pmatrix} \end{aligned}$$

# Matrix Multiplication

How do we **multiply** matrices?

Because we want to use matrix multiplication to solve equations we won't use an intuitive definition

Suppose we have two matrices

$$\mathbf{X} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \mathbf{Y} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$$

We will create a new matrix **A** by matrix multiplication:

$$\begin{aligned} \mathbf{A} &= \mathbf{XY} \\ &= \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \\ &= \begin{pmatrix} 1 \times 1 + 1 \times 3 & 1 \times 2 + 1 \times 4 \\ & \end{pmatrix} \end{aligned}$$

# Matrix Multiplication

How do we **multiply** matrices?

Because we want to use matrix multiplication to solve equations we won't use an intuitive definition

Suppose we have two matrices

$$\mathbf{X} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \mathbf{Y} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$$

We will create a new matrix **A** by matrix multiplication:

$$\begin{aligned} \mathbf{A} &= \mathbf{XY} \\ &= \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \\ &= \begin{pmatrix} 1 \times 1 + 1 \times 3 & 1 \times 2 + 1 \times 4 \\ 1 \times 1 + 1 \times 3 & \end{pmatrix} \end{aligned}$$

# Matrix Multiplication

How do we **multiply** matrices?

Because we want to use matrix multiplication to solve equations we won't use an intuitive definition

Suppose we have two matrices

$$\mathbf{X} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \mathbf{Y} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$$

We will create a new matrix  $\mathbf{A}$  by matrix multiplication:

$$\begin{aligned} \mathbf{A} &= \mathbf{XY} \\ &= \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \\ &= \begin{pmatrix} 1 \times 1 + 1 \times 3 & 1 \times 2 + 1 \times 4 \\ 1 \times 1 + 1 \times 3 & 1 \times 2 + 1 \times 4 \end{pmatrix} \end{aligned}$$



# Matrix Multiplication

How do we **multiply** matrices?

Because we want to use matrix multiplication to solve equations we won't use an intuitive definition

Suppose we have two matrices

$$\mathbf{X} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \mathbf{Y} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$$

We will create a new matrix **A** by matrix multiplication:

$$\begin{aligned} \mathbf{A} &= \mathbf{XY} \\ &= \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \\ &= \begin{pmatrix} 1 \times 1 + 1 \times 3 & 1 \times 2 + 1 \times 4 \\ 1 \times 1 + 1 \times 3 & 1 \times 2 + 1 \times 4 \end{pmatrix} \\ &= \begin{pmatrix} 4 & 6 \\ 4 & 6 \end{pmatrix} \end{aligned}$$

# Matrix Multiplication

## Definition

Suppose  $\mathbf{X}$  is an  $m \times n$  matrix and  $\mathbf{Y}$  is an  $n \times k$  matrix. Then define the matrix  $\mathbf{A}$  an  $m \times k$  matrix that obtains from **multiplying**  $\mathbf{X}$  and  $\mathbf{Y}$  as,

$$\mathbf{A} = \mathbf{XY}$$

$$= \begin{pmatrix} x_{11} & x_{12} & \dots & x_{1n} \\ x_{21} & x_{22} & \dots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{m1} & x_{m2} & \dots & x_{mn} \end{pmatrix} \begin{pmatrix} y_{11} & y_{12} & \dots & y_{1k} \\ y_{21} & y_{22} & \dots & y_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ y_{n1} & y_{n2} & \dots & y_{nk} \end{pmatrix}$$

$$= \begin{pmatrix} x_{11}y_{11} + x_{12}y_{21} + \dots + x_{1n}y_{n1} & \dots & x_{11}y_{1k} + x_{12}y_{2k} + \dots + x_{1n}y_{nk} \\ \vdots & \ddots & \vdots \\ x_{m1}y_{11} + x_{m2}y_{21} + \dots + x_{mn}y_{n1} & \dots & x_{m1}y_{1k} + x_{m2}y_{2k} + \dots + x_{mn}y_{nk} \end{pmatrix}$$

## Definition

Suppose  $\mathbf{X}$  is an  $m \times n$  matrix and  $\mathbf{Y}$  is an  $n \times k$  matrix. Write the **row**

vectors of  $\mathbf{X} = \begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \vdots \\ \mathbf{x}_m \end{pmatrix}$  and  $\mathbf{Y}$  as column vector  $\mathbf{Y} = (\mathbf{y}_1 \ \mathbf{y}_2 \ \dots \ \mathbf{y}_k)$ .

Then the  $m \times k$  matrix  $\mathbf{A} = \mathbf{X}\mathbf{Y}$  can be written as

$$\mathbf{A} = \begin{pmatrix} \mathbf{x}_1 \cdot \mathbf{y}_1 & \mathbf{x}_1 \cdot \mathbf{y}_2 & \dots & \mathbf{x}_1 \cdot \mathbf{y}_k \\ \mathbf{x}_2 \cdot \mathbf{y}_1 & \mathbf{x}_2 \cdot \mathbf{y}_2 & \dots & \mathbf{x}_2 \cdot \mathbf{y}_k \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{x}_m \cdot \mathbf{y}_1 & \mathbf{x}_m \cdot \mathbf{y}_2 & \dots & \mathbf{x}_m \cdot \mathbf{y}_k \end{pmatrix}$$

# Matrix Multiplication

Let's work on an example together!

$$\mathbf{X} = \begin{pmatrix} 1 & 4 & 5 \\ 10 & 2 & 3 \end{pmatrix}$$

$$\mathbf{Y} = \begin{pmatrix} 2 & 3 \\ 1 & 5 \\ 3 & 5 \end{pmatrix} \text{ What is } \mathbf{XY}?$$

# Matrix Multiplication

Let's work on an example together!

$$\mathbf{X} = \begin{pmatrix} 1 & 4 & 5 \\ 10 & 2 & 3 \end{pmatrix}$$

$$\mathbf{Y} = \begin{pmatrix} 2 & 3 \\ 1 & 5 \\ 3 & 5 \end{pmatrix} \text{ What is } \mathbf{XY}?$$

Not all matrices can be multiplied.

Matrix  $\mathbf{AB}$  exists only if the number of columns in  $\mathbf{A}$  = number of rows in  $\mathbf{B}$ . If  $\mathbf{AB}$  exists we will say the matrices are **conformable**

# Matrix Multiplication with a Vector

Suppose  $\mathbf{X} = \begin{pmatrix} 2 & 3 & 4 & 5 \\ 1 & 5 & 1 & 2 \\ 3 & 5 & 3 & 4 \end{pmatrix}$  a  $3 \times 4$  matrix and that  $\mathbf{v} = \begin{pmatrix} 3 \\ 3 \\ 4 \\ 10 \end{pmatrix}$  a  $4 \times 1$

matrix (or a **column** vector) what is

$\mathbf{X}\mathbf{v}$ ?

What is  $\mathbf{X}'\mathbf{v}$ ?

# Algebraic Properties

Suppose  $\mathbf{X}$  is an  $m \times n$  matrix and  $\mathbf{Y}$  is an  $n \times k$  matrix. Suppose that

$\mathbf{I} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}$  as the **identity** matrix and that  $k \in \mathbb{R}$ .

- $\mathbf{XY} \neq \mathbf{YX}$  **in general !!!!** (but it could)
- $\mathbf{XI} = \mathbf{X}$  (let's talk it out!)
- $(\mathbf{X}')' = \mathbf{X}$
- $(\mathbf{XY})' = \mathbf{Y}'\mathbf{X}'$
- $(k\mathbf{X})' = k\mathbf{X}'$
- $(\mathbf{X} + \mathbf{Y})' = \mathbf{X}' + \mathbf{Y}'$

# Examples, Implementing in R

## R and matrix multiplication

```
X<- matrix(NA, nrow=2, ncol=3)
```

```
Y<- matrix(NA, nrow=3, ncol=2)
```

```
X[1,<- c(1, 4, 5)
```

```
X[2,<- c(10, 2, 3)
```

```
Y[1,<- c(2, 3)
```

```
Y[2,<- c(1, 5)
```

```
Y[3,<- c(3, 5)
```

```
A<- X%*%Y
```

```
> A
```

```
 [,1] [,2]
```

```
[1,] 21 48
```

```
[2,] 31 55
```



# Matrix Inversion

Big topic: suppose  $\mathbf{X}$  is an  $n \times n$  matrix. We want to find the matrix  $\mathbf{X}^{-1}$  such that

$$\mathbf{X}^{-1}\mathbf{X} = \mathbf{X}\mathbf{X}^{-1} = \mathbf{I}$$

where  $\mathbf{I}$  is the  $n \times n$  identity matrix.

Why?

- Regression
- Solving systems of equations
- Will provide intuition about “colinearity”, “fixed effects”, “treatment designs” and what we can learn as social scientists

Calculate  $\rightsquigarrow$  Properties of Inverses  $\rightsquigarrow$  when do inverses exist  $\rightsquigarrow$

Application to regression analysis

# Some Motivating Examples

Consider the following equations:

$$\begin{aligned}x_1 + x_2 + x_3 &= 0 \\ 0x_1 + 0x_2 + x_3 &= 5\end{aligned}\tag{0.1}$$

# Some Motivating Examples

Consider the following equations:

$$x_1 + x_2 + x_3 = 0$$

$$x_1 + x_2 + 0x_3 = 0$$

$$0x_1 + x_2 + x_3 = 0$$

$$x_1 + 0x_2 + x_3 = 0$$

# Some Motivating Examples

Consider the following equations:

$$x_1 + x_2 + x_3 = 0$$

$$0x_1 + 5x_2 + 0x_3 = 5$$

$$0x_1 + 0x_2 + 3x_3 = 6$$

# Some Motivating Examples

Consider the following equations:

$$x_1 + x_2 + x_3 = 0$$

$$0x_1 + 5x_2 + 0x_3 = 5$$

$$0x_1 + 0x_2 + 3x_3 = 6$$

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 5 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

# Some Motivating Examples

Consider the following equations:

$$x_1 + x_2 + x_3 = 0$$

$$0x_1 + 5x_2 + 0x_3 = 5$$

$$0x_1 + 0x_2 + 3x_3 = 6$$

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 5 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

$$\mathbf{x} = (x_1, x_2, x_3)$$

# Some Motivating Examples

Consider the following equations:

$$x_1 + x_2 + x_3 = 0$$

$$0x_1 + 5x_2 + 0x_3 = 5$$

$$0x_1 + 0x_2 + 3x_3 = 6$$

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 5 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

$$\mathbf{x} = (x_1, x_2, x_3)$$

$$\mathbf{b} = (0, 5, 6)$$

# Some Motivating Examples

Consider the following equations:

$$x_1 + x_2 + x_3 = 0$$

$$0x_1 + 5x_2 + 0x_3 = 5$$

$$0x_1 + 0x_2 + 3x_3 = 6$$

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 5 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

$$\mathbf{x} = (x_1, x_2, x_3)$$

$$\mathbf{b} = (0, 5, 6)$$

The system of equations are now,



# Some Motivating Examples

Consider the following equations:

$$x_1 + x_2 + x_3 = 0$$

$$0x_1 + 5x_2 + 0x_3 = 5$$

$$0x_1 + 0x_2 + 3x_3 = 6$$

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 5 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

$$\mathbf{x} = (x_1, x_2, x_3)$$

$$\mathbf{b} = (0, 5, 6)$$

The system of equations are now,

$$\mathbf{Ax} = \mathbf{b}$$

# Some Motivating Examples

Consider the following equations:

$$x_1 + x_2 + x_3 = 0$$

$$0x_1 + 5x_2 + 0x_3 = 5$$

$$0x_1 + 0x_2 + 3x_3 = 6$$

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 5 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

$$\mathbf{x} = (x_1, x_2, x_3)$$

$$\mathbf{b} = (0, 5, 6)$$

The system of equations are now,

$$\mathbf{Ax} = \mathbf{b}$$

$\mathbf{A}^{-1}$  exists **if and only if**  $\mathbf{Ax} = \mathbf{b}$  has only one solution.

# Matrix Inversion, Definition

## Definition

Suppose  $\mathbf{X}$  is an  $n \times n$  matrix. We will call  $\mathbf{X}^{-1}$  the *inverse* of  $\mathbf{X}$  if

$$\mathbf{X}^{-1}\mathbf{X} = \mathbf{X}\mathbf{X}^{-1} = \mathbf{I}$$

If  $\mathbf{X}^{-1}$  exists then  $\mathbf{X}$  is invertible. If  $\mathbf{X}^{-1}$  does not exist, then we will say  $\mathbf{X}$  is *singular*.

# Matrix Inversion

You'll never invert a matrix by hand.

We're going to use R

```
X<- matrix(NA, nrow=3, ncol=3)
X[1,]<- c(2, 3, 4)
X[2,]<- c(3, 1, 3)
X[3,]<- c(2, 4, 2)
X.inv<- solve(X)
> X.inv
[,1] [,2] [,3]
[1,] -0.5 0.5 0.25
[2,] 0.0 -0.2 0.30
[3,] 0.5 -0.1 -0.35
X.inv%%X
[,1] [,2] [,3]
[1,] 1 0.000000e+00 -2.220446e-16
[2,] 0 1.000000e+00 0.000000e+00
[3,] 0 -2.220446e-16 1.000000e+00
```

# Matrix Inversion

- 1) Calculate Inverses
- 2) **Properties of Inverses**

# Matrix Inversion

- 1) Calculate Inverses
- 2) Properties of Inverses

## Theorem

*The inverse of matrix  $\mathbf{X}$ ,  $\mathbf{X}^{-1}$ , is unique*

# Matrix Inversion

- 1) Calculate Inverses
- 2) **Properties of Inverses**

## Theorem

*The inverse of matrix  $\mathbf{X}$ ,  $\mathbf{X}^{-1}$ , is unique*

Proof.

By way of contradiction, suppose not. Then there are at least two matrices  $\mathbf{A}$  and  $\mathbf{C}$  such that  $\mathbf{AX} = \mathbf{I}$  and  $\mathbf{CX} = \mathbf{I}$

This implies that,

$$\begin{aligned}\mathbf{AXC} &= (\mathbf{AX})\mathbf{C} \\ &= \mathbf{IC} \\ &= \mathbf{C}\end{aligned}$$

# Matrix Inversion

But it also implies that

$$\begin{aligned}\mathbf{A}\mathbf{X}\mathbf{C} &= \mathbf{A}(\mathbf{X}\mathbf{C}) \\ &= \mathbf{A}(\mathbf{I}) \\ &= \mathbf{A}\end{aligned}$$

So  $\mathbf{C} = \mathbf{A}\mathbf{X}\mathbf{C} = \mathbf{A}$  or  $\mathbf{C} = \mathbf{A}$  but this contradicts our assumption that there are two unique inverses.



# Matrix Inversion

## Theorem

*Suppose  $\mathbf{A}$  has inverse  $\mathbf{A}^{-1}$  and  $\mathbf{B}$  has inverse  $\mathbf{B}^{-1}$ . Then,*

$$(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$$

# Matrix Inversion

## Theorem

Suppose  $\mathbf{A}$  has inverse  $\mathbf{A}^{-1}$  and  $\mathbf{B}$  has inverse  $\mathbf{B}^{-1}$ . Then,

$$(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$$

Proof.

We need to show that  $(\mathbf{B}^{-1}\mathbf{A}^{-1})(\mathbf{AB}) = (\mathbf{AB})(\mathbf{B}^{-1}\mathbf{A}^{-1}) = \mathbf{I}$ .

$$\begin{aligned}(\mathbf{B}^{-1}\mathbf{A}^{-1})(\mathbf{AB}) &= \mathbf{B}^{-1}(\mathbf{A}^{-1}\mathbf{A})\mathbf{B} \\&= \mathbf{B}^{-1}\mathbf{I}\mathbf{B} \\&= \mathbf{B}^{-1}\mathbf{B} \\&= \mathbf{I}\end{aligned}$$

# Matrix Inversion

$$\begin{aligned}(\mathbf{AB})(\mathbf{B}^{-1}\mathbf{A}^{-1}) &= \mathbf{A}(\mathbf{BB}^{-1})\mathbf{A}^{-1} \\ &= \mathbf{A}\mathbf{I}\mathbf{A}^{-1} \\ &= \mathbf{AA}^{-1} \\ &= \mathbf{I}\end{aligned}$$

So  $\mathbf{AB}$  is invertible and  $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$ .

# Challenge Inversion Proofs

- Show that  $(\mathbf{A}^{-1})^{-1} = \mathbf{A}$ .
- Show that  $(k\mathbf{A})^{-1} = \frac{1}{k}\mathbf{A}^{-1}$

# Matrix Inversion

- 1) How to Calculate an Inverse
- 2) Inversion properties
- 3) When do inverses exist?

**Linear Independence:** not repeated information in matrix will be the key  
(for both inversion and regressions)

# Matrix Inversion: Existence

## Definition

*Suppose we have a set of vectors  $S = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_r\}$*

*And consider the system of equations*

$$k_1 \mathbf{v}_1 + k_2 \mathbf{v}_2 + \dots + k_r \mathbf{v}_r = \mathbf{0}$$

*If the only solution is  $k_1 = 0, k_2 = 0, k_3 = 0, \dots, k_r = 0$  then we say that the set is **linearly independent**. If there are other solutions, then the set is **linearly dependent**.*

# Matrix Inversion: Existence

Consider  $\mathbf{v}_1 = (1, 0, 0)$ ,  $\mathbf{v}_2 = (0, 1, 0)$ ,  $\mathbf{v}_3 = (0, 0, 1)$

Can we write this as a combination of other vectors?

# Matrix Inversion: Existence

Consider  $\mathbf{v}_1 = (1, 0, 0)$ ,  $\mathbf{v}_2 = (0, 1, 0)$ ,  $\mathbf{v}_3 = (0, 0, 1)$

Can we write this as a combination of other vectors? **no!**



# Matrix Inversion: Existence

Consider  $\mathbf{v}_1 = (1, 0, 0)$ ,  $\mathbf{v}_2 = (0, 1, 0)$ ,  $\mathbf{v}_3 = (0, 0, 1)$

Can we write this as a combination of other vectors? **no!**

Consider  $\mathbf{v}_1 = (1, 0, 0)$ ,  $\mathbf{v}_2 = (0, 1, 0)$ ,  $\mathbf{v}_3 = (0, 0, 1)$ ,  $\mathbf{v}_4 = (1, 2, 3)$ .

Can we write this as a combination of other vectors?

# Matrix Inversion: Existence

Consider  $\mathbf{v}_1 = (1, 0, 0)$ ,  $\mathbf{v}_2 = (0, 1, 0)$ ,  $\mathbf{v}_3 = (0, 0, 1)$

Can we write this as a combination of other vectors? **no!**

Consider  $\mathbf{v}_1 = (1, 0, 0)$ ,  $\mathbf{v}_2 = (0, 1, 0)$ ,  $\mathbf{v}_3 = (0, 0, 1)$ ,  $\mathbf{v}_4 = (1, 2, 3)$ .

Can we write this as a combination of other vectors?

$$\mathbf{v}_4 = \mathbf{v}_1 + 2\mathbf{v}_2 + 3\mathbf{v}_3$$

# Matrix Inversion: Existence

## Theorem

*Suppose  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_K \in \mathbb{R}^n$ . If  $K > n$  then the set is linearly dependent*

# Matrix Inversion: Existence

## Theorem

*Suppose  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_K \in \mathbb{R}^n$ . If  $K > n$  then the set is linearly dependent*

-  $\mathbf{v}_1 = (v_{11}, v_{21}, \dots, v_{n1})$

# Matrix Inversion: Existence

## Theorem

*Suppose  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_K \in \mathbb{R}^n$ . If  $K > n$  then the set is linearly dependent*

- $\mathbf{v}_1 = (v_{11}, v_{21}, \dots, v_{n1})$
- Says that if there are more vectors in the set than elements in each vector, one must be linearly dependent

# Matrix Inversion: Existence

We care because of the following theorem

Theorem

Suppose  $\mathbf{X}$  is an  $n \times n$  matrix. Recall we can write this matrix as  $\begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \vdots \\ \mathbf{x}_n \end{pmatrix}$ .

Then  $\mathbf{X}$  has an inverse *if and only if*  $S = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$  is linearly independent

If this is true, we say  $\mathbf{X}$  has full rank

# Linear Regression

# Linear Regression

In 350a you learn about linear regression. For each  $i$  (individual) we observe covariates  $x_{i1}, x_{i2}, \dots, x_{ik}$  and independent variable  $Y_i$ . Then,



# Linear Regression

In 350a you learn about linear regression. For each  $i$  (individual) we observe covariates  $x_{i1}, x_{i2}, \dots, x_{ik}$  and independent variable  $Y_i$ . Then,

$$Y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik}$$

# Linear Regression

In **350a** you learn about linear regression. For each  $i$  (individual) we observe covariates  $x_{i1}, x_{i2}, \dots, x_{ik}$  and independent variable  $Y_i$ . Then,

$$Y_1 = \beta_0 + \beta_1 x_{11} + \beta_2 x_{12} + \dots + \beta_k x_{1k}$$

$$Y_2 = \beta_0 + \beta_1 x_{21} + \beta_2 x_{22} + \dots + \beta_k x_{2k}$$

$$\vdots \quad \vdots \quad \vdots$$

$$Y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik}$$

$$\vdots \quad \vdots \quad \vdots$$

$$Y_n = \beta_0 + \beta_1 x_{n1} + \beta_2 x_{n2} + \dots + \beta_k x_{nk}$$

# Linear Regression

- Define  $\mathbf{x}_i = (1, x_{i1}, x_{i2}, \dots, x_{ik})$
- Define  $\mathbf{X} = \begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \vdots \\ \mathbf{x}_n \end{pmatrix}$
- Define  $\boldsymbol{\beta} = (\beta_0, \beta_1, \dots, \beta_k)$
- Define  $\mathbf{Y} = (Y_1, Y_2, \dots, Y_n)$ .

Then we can write

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta}$$

# Linear Regression

$$\mathbf{Y} = \mathbf{X}\beta$$

$$\mathbf{X}'\mathbf{Y} = \mathbf{X}'\mathbf{X}\beta$$

$$(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{X}\beta$$

$$(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y} = \beta$$

Big question: is  $(\mathbf{X}'\mathbf{X})^{-1}$  invertible?

We'll investigate in homework!