Math Camp

Justin Grimmer

Professor
Department of Political Science
Stanford University

September 10th, 2018

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
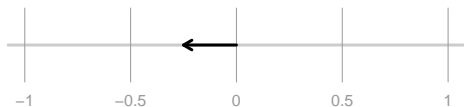
Very important for regression(!!!!)

Points + Vectors

- A point in $\Re^1$
    - 1
    - π
    - e
- An ordered pair in $\Re^2 = \Re \times \Re$
    - $(1, 2)$
    - $(0, 0)$
    - (π, e)
- An ordered triple in $\Re^3 = \Re \times \Re \times \Re$
    - $(3.1, 4.5, 6.11132)$

$\vdots$

- An ordered n-tuple in $\Re^n = \Re \times \Re \times \ldots \times \Re$
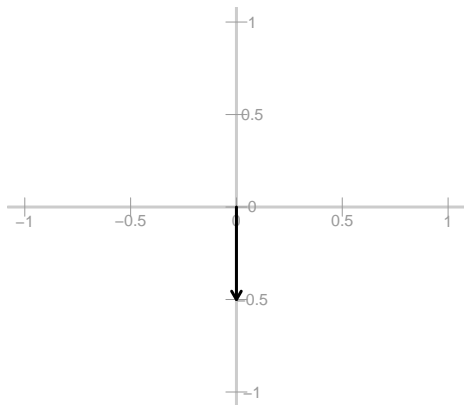   - $(a_1, a_2, \ldots, a_n)$

Points and Vectors

Definition

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

One Dimensional Example

Two Dimensional Example

Three Dimensional Example

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

N-Dimensional Example

- Individual campaign donation records

$$\boldsymbol{x} = (1000, 0, 10, 50, 15, 4, 0, 0, 0, \ldots, 2400000000)$$

- Counties have proportion of vote for Obama

$$\boldsymbol{y} = (0.8, 0.5, 0.6, \ldots, 0.2)$$

- Run experiment, assess feeling thermometer of elected official

$$\boldsymbol{t} = (0, 100, 50, 70, 80, \ldots, 100)$$

Arithmetic with Vectors

Definition

Suppose $\boldsymbol{u}$ and $\boldsymbol{v}$ are vectors $\boldsymbol{u}, \boldsymbol{v} \in \Re^n$,

$$
\begin{aligned}
\boldsymbol{u} &= (u_1, u_2, u_3, \ldots, u_n) \\
\boldsymbol{v} &= (v_1, v_2, v_3, \ldots, v_n)
\end{aligned}
$$

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

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

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

Suppose $k \in \Re$. We will call k a scalar.

Define $k\boldsymbol{u}$ as the scalar product

$$
k\boldsymbol{u} = (ku_1, ku_2, \ldots, ku_n)
$$

Examples

Suppose:

$$\begin{aligned} \boldsymbol{u} &= (1, 2, 3, 4, 5) \\ \boldsymbol{v} &= (1, 1, 1, 1, 1) \\ k &= 2 \end{aligned}$$

Then,

$$\begin{aligned} \boldsymbol{u} + \boldsymbol{v} &= (1+1, 2+1, 3+1, 4+1, 5+1) = (2, 3, 4, 5, 6) \\ k\boldsymbol{u} &= (2 \times 1, 2 \times 2, 2 \times 3, 2 \times 4, 2 \times 5) = (2, 4, 6, 8, 10) \\ k\boldsymbol{v} &= (2 \times 1, 2 \times 1, 2 \times 1, 2 \times 1, 2 \times 1) = (2, 2, 2, 2, 2) \end{aligned}$$

Properties of Arithmetic

Challenge Proofs—we can do these!

Theorem

Suppose that $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in \Re^n$ and k and l are scalars.

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

Proof.

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

$\square$

Properties of Arithmetic

Challenge Proofs—we can do these!

Theorem

Suppose that $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in \Re^n$ and k and l are scalars.

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

Proof.

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

$\square$

Properties of Arithmetic

Challenge Proofs—we can do these!

Theorem

Suppose that $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in \Re^n$ and k and l are scalars.

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

Proof.

How can we show this?

$\square$

Challenge Proofs

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

Inner Product

Definition

Suppose $\boldsymbol{u} \in \Re^n$ and $\boldsymbol{v} \in \Re^n$ then define $\boldsymbol{u} \cdot \boldsymbol{v}$,

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

Examples

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

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

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

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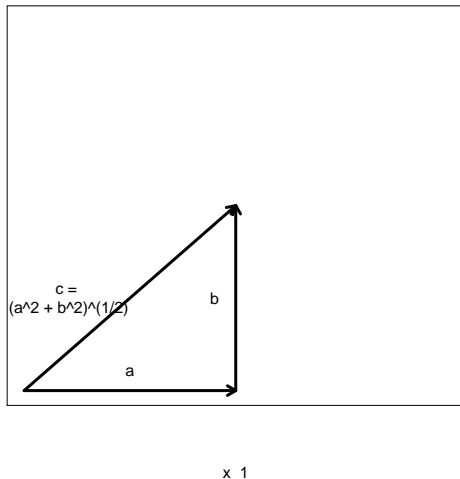&= \sum_{i=1}^{n} y_i
\end{aligned}
$$

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 $\boldsymbol{v} = (1, 4, 1, 4)$ and $\boldsymbol{w} = (4, 1, 4, 1)$. Calculate: $\boldsymbol{v} \cdot \boldsymbol{w}$
- Prove $\boldsymbol{v} \cdot \boldsymbol{w} = \boldsymbol{w} \cdot \boldsymbol{v}$
- Prove $\boldsymbol{v} \cdot \boldsymbol{v} \geq 0$ and $\boldsymbol{v} \cdot \boldsymbol{v} = 0$ if and only if $\boldsymbol{v} = \boldsymbol{0}$.

Vector Length

- Pythogorean 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 $\mathbf{v} \in \Re^n$. Then, we will define its length as

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

Calculating a Length

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

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

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

$$\begin{aligned} \text{Doc1} &= (1, 1, 3, \ldots, 5) \\ \text{Doc2} &= (2, 0, 0, \ldots, 1) \\ \mathbf{Doc1}, \mathbf{Doc2} &\in \Re^M \end{aligned}$$

Provides many operations that will be useful
Inner Product between documents:

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

Length of document:

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

Cosine of the angle between documents:

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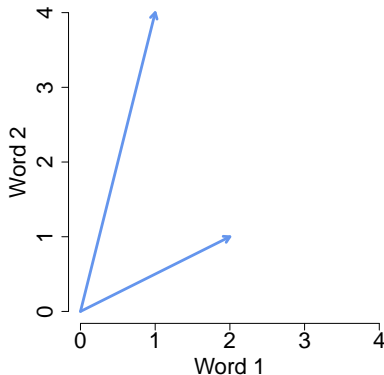&= \frac{7}{6 \times 2.24} \\
&= 0.52
\end{aligned}
$$

Measuring Similarity

Documents in space $\rightarrow$ measure similarity/dissimilarity

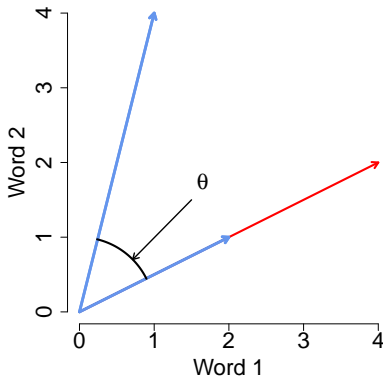What properties should similarity measure have?

- 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

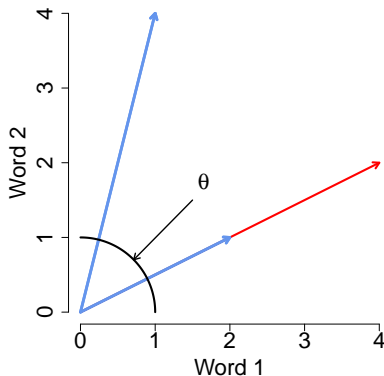$$(2,1)^{'} \cdot (1,4) \;\; = \;\; 6$$

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

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

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

If $\boldsymbol{A}$ has m rows n columns we will say that $\boldsymbol{A}$ is an $m \times n$ matrix.
Suppose $\boldsymbol{X}$ and $\boldsymbol{Y}$ are $m \times n$ matrices. Then $\boldsymbol{X} = \boldsymbol{Y}$ if $x_{ij} = y_{ij}$ for all i and j

Simple Examples

$$\boldsymbol{I} \;=\; \begin{pmatrix} 1 & 0 & 0 & \ldots & 0 \\ 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & 1 \end{pmatrix}$$

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

Simple Examples

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

$\boldsymbol{X}$ is an 2×3 matrix

Matrix Algebra

Definition

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

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

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

Matrix Algebra

Definition

Suppose $\boldsymbol{X}$ is an $m \times n$ matrix and $k \in \Re$. Then,

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

Prove theorems about this tonight

R Code

```
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

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

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

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

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

Matrix Transpose

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

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?

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

Suppose we have two matrices

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

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

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

Matrix Multiplication

Definition

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

$$
\begin{aligned}
\boldsymbol{A} &= \boldsymbol{X}\boldsymbol{Y} \\
&= \begin{pmatrix} x_{11} & x_{12} & \ldots & x_{1n} \\ x_{21} & x_{22} & \ldots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{m1} & x_{m2} & \ldots & x_{mn} \end{pmatrix} \begin{pmatrix} y_{11} & y_{12} & \ldots & y_{1k} \\ y_{21} & y_{22} & \ldots & y_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ y_{n1} & y_{n2} & \ldots & y_{nk} \end{pmatrix} \\
&= \begin{pmatrix} x_{11}y_{11} + x_{12}y_{21} + \ldots + x_{1n}y_{n1} & \ldots & x_{11}y_{1k} + x_{12}y_{2k} + \ldots + x_{1n}y_{nk} \\ \vdots & \ddots & \vdots \\ x_{m1}y_{11} + x_{m2}y_{21} + \ldots + x_{mn}y_{n1} & \ldots & x_{m1}y_{11} + x_{m2}y_{12} + \ldots + x_{mn}y_{nk} \end{pmatrix}
\end{aligned}
$$

Definition

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

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

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

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

Matrix Multiplication

Let's work on an example together!

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

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

Not all matrices can be multiplied.

Matrix AB exists only if the number of columns in A = number of rows in B. If AB exists we will say the matrices are conformable

Matrix Multiplication with a Vector

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

matrix (or a column vector) what is

$\boldsymbol{Xv}$?

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

Algebraic Properties

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

$$\boldsymbol{I} = \begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 1 \end{pmatrix} \text{ as the identity matrix and that } k \in Re.$$

- $\boldsymbol{X}\boldsymbol{Y} \neq \boldsymbol{Y}\boldsymbol{X}$ in general !!!! (but it could)
- $\boldsymbol{X}\boldsymbol{I} = \boldsymbol{X}$ (let's talk it out!)
- $(\boldsymbol{X}^{'})^{'} = \boldsymbol{X}$
- $(\boldsymbol{X}\boldsymbol{Y})^{'} = \boldsymbol{Y}^{'}\boldsymbol{X}^{'}$
- $(k\boldsymbol{X})^{'} = k\boldsymbol{X}^{'}$
- $(\boldsymbol{X} + \boldsymbol{Y})^{'} = \boldsymbol{X}^{'} + \boldsymbol{Y}^{'}$

Examples, Implenting 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 $\boldsymbol{X}$ is an $n \times n$ matrix. We want to find the matrix X^{-1} such that

$$\boldsymbol{X}^{-1}\boldsymbol{X} \;\; = \;\; \boldsymbol{X}\boldsymbol{X}^{-1} = \boldsymbol{I}$$

where $\boldsymbol{I}$ 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 ⇝ Properties of Inverses ⇝ when do inverses exist ⇝ Application to regression analysis

Some Motivating Examples

Consider the following equations:

$$\begin{aligned}
x_1 + x_2 + x_3 &= 0 \\
0x_1 + 5x_2 + 0x_3 &= 5 \\
0x_1 + 0x_2 + 3x_3 &= 6
\end{aligned}$$

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

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

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

The system of equations are now,

$$\boldsymbol{Ax} = \boldsymbol{b}$$

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

Matrix Inversion, Definition

Definition

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

$$X^{-1}X \;\; = \;\; XX^{-1} = I$$

If X^{-1} exists then X is invertible. If X^{-1} does not exist, then we will say 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

Theorem

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

Proof.

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

This implies that,

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

Matrix Inversion

But it also implies that

$$
\begin{aligned}
\boldsymbol{AXC} &= \boldsymbol{A}(\boldsymbol{XC}) \\
&= \boldsymbol{A}(\boldsymbol{I}) \\
&= \boldsymbol{A}
\end{aligned}
$$

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

Matrix Inversion

Theorem

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

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

Proof.

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

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

Matrix Inversion

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

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

Challenge Inversion Proofs

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

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 $S = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_r\}$
And consider the system of equations

$$k_1 \boldsymbol{v}_1 + k_2 \boldsymbol{v}_2 + \ldots + k_r \boldsymbol{v}_r = \boldsymbol{0}$$

If the only solution is $k_1 = 0, k_2 = 0, k_3 = 0, \ldots, k_r = 0$ 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? 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 $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_K \in \Re^n$. If $K > n$ then the set is linearly dependent

- $\boldsymbol{v}_1 = (v_{11}, v_{21}, \ldots, v_{n1})$
- 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 $\boldsymbol{X}$ is an $n \times n$ matrix. Recall we can write this matrix as $\begin{pmatrix} \boldsymbol{x}_1 \\ \boldsymbol{x}_2 \\ \vdots \\ \boldsymbol{x}_n \end{pmatrix}$.

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

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

Linear Regression

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

$$
\begin{aligned}
Y_1 &= \beta_0 + \beta_1 x_{11} + \beta_2 x_{12} + \ldots + \beta_k x_{1k} \\
Y_2 &= \beta_0 + \beta_1 x_{21} + \beta_2 x_{22} + \ldots + \beta_k x_{2k} \\
&\;\;\vdots \quad \vdots \quad \vdots \\
Y_i &= \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \ldots + \beta_k x_{ik} \\
&\;\;\vdots \quad \vdots \quad \vdots \\
Y_n &= \beta_0 + \beta_1 x_{n1} + \beta_2 x_{n2} + \ldots + \beta_k x_{nk}
\end{aligned}
$$

Linear Regression

- Define $\boldsymbol{x}_i = (1, x_{i1}, x_{i2}, \ldots, x_{ik})$

- Define $\boldsymbol{X} = \begin{pmatrix} \boldsymbol{x}_1 \\ \boldsymbol{x}_2 \\ \vdots \\ \boldsymbol{x}_n \end{pmatrix}$

- Define $\boldsymbol{\beta} = (\beta_0, \beta_1, \ldots, \beta_k)$

- Define $\boldsymbol{Y} = (Y_1, Y_2, \ldots, Y_n)$.

Then we can write

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

Linear Regression

$$
\begin{aligned}
\boldsymbol{Y} &= \boldsymbol{X\beta} \\
\boldsymbol{X}'\boldsymbol{Y} &= \boldsymbol{X}'\boldsymbol{X\beta} \\
(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{Y} &= (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{X\beta} \\
(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{Y} &= \boldsymbol{\beta}
\end{aligned}
$$

Big question: is $(\boldsymbol{X}'\boldsymbol{X})^{-1}$ invertible?
We'll investigate in homework!

An Introduction to Eigenvectors, Values, and Diagonalization

Definition

Suppose $\boldsymbol{A}$ is an $N \times N$ matrix and λ is a scalar.
If

$$\boldsymbol{Ax} \;\; = \;\; \lambda \boldsymbol{x}$$

Then $\boldsymbol{x}$ is an eigenvector and λ is the associated eigenvalue

- $\boldsymbol{A}$ stretches the eigenvector $\boldsymbol{x}$
- $\boldsymbol{A}$ stretches $\boldsymbol{x}$ by λ
- To find eigenvectors/values: (`eigen` in R)
    - Find λ that solves $\det(\boldsymbol{A} - \lambda \boldsymbol{I}) = 0$
    - Find vectors in null space of:

$$(\boldsymbol{A} - \lambda \boldsymbol{I}) \;\; = \;\; 0$$

Definition

Suppose A is an N × N matrix and A has N linearly independent eigenvectors. Then, we can write A as

$$\boldsymbol{A} = \boldsymbol{W}^{'} \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_N \end{pmatrix} \boldsymbol{W}$$

Where $\lambda_1, \lambda_2, \dots, \lambda_N$ are the eigenvalues and $\boldsymbol{w}$ is a matrix of the eigenvectors.

Definition

*Suppose **X** is an $N \times J$ matrix. Then **X** can be written as:*

$$\boldsymbol{X} = \underbrace{\boldsymbol{U}}_{N \times N} \underbrace{\boldsymbol{S}}_{N \times J} \underbrace{\boldsymbol{V}'}_{J \times J}$$

Where:

$$\boldsymbol{U}'\boldsymbol{U} = \boldsymbol{I}_N$$
$$\boldsymbol{V}'\boldsymbol{V} = \boldsymbol{V}\boldsymbol{V}' = \boldsymbol{I}_J$$

S *contains* $\min(N, J)$ *singular values,* $\sqrt{\lambda_j} \geq 0$ *down the diagonal and then 0's for the remaining entries*