

SummR Camp

Mathematics Review: Matrices and Functions

Kat Sadikova (many thanks to Louisa H. Smith)

August 26, 2020

Part 1: Matrices and linear algebra basics

Intuition and basic tools relevant for quantitative methods in population health sciences

Scalars, vectors, matrices

$a = a$, a scalar

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \text{ a vector of length } n$$

$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix}, \text{ an } n \times p \text{ matrix}$$

Note on dimensions

We read dimensions like $n \times p$ as **rows** by **columns**. X is a 3x4 matrix:

$$\mathbf{X}_{3 \times 4} = \begin{bmatrix} x_{11} & x_{12} & x_{13} & x_{14} \\ x_{21} & x_{22} & x_{23} & x_{24} \\ x_{31} & x_{32} & x_{33} & x_{34} \end{bmatrix}_{3 \times 4}$$

Individual elements are indexed in the same *row-column* order: x_{23} is the element in the second row and third column.



Imagine this very happy Ron Swanson **Rowing a Canoe** and remember that **Rows come before Columns** when indexing elements of a matrix

Data example

Let's look at a simple example - data stored in a matrix \mathbf{X} with 10 individuals and 4 variables:

- age in years
- height in inches
- a dichotomous indicator of whether the individual likes dogs
- year of birth

$$\mathbf{X} = \begin{bmatrix} 25 & 67 & \cdots & 1995 \\ 38 & 63 & \cdots & 1982 \\ \vdots & \vdots & \ddots & \vdots \\ 41 & 59 & \cdots & 1979 \end{bmatrix}$$

$$\mathbf{age} = \begin{bmatrix} 25 \\ 38 \\ \vdots \\ 41 \end{bmatrix}$$

$$age_1 = 25$$

Data example: in R

Build the data set

```
set.seed(6789)
n <- 10
dat <- data.frame(
  age = round(runif(n, 22, 45)), height = round(rnorm(n, 66, 4)), likes_dogs = rbinom(n, 1, .53)
)
dat$yob <- 2020 - dat$age
dat <- as.matrix(dat)

dat
```

```
##      age height likes_dogs   yob
## [1,]  25     67          0 1995
## [2,]  38     63          1 1982
## [3,]  35     66          1 1985
## [4,]  36     70          1 1984
## [5,]  23     62          1 1997
## [6,]  26     73          1 1994
## [7,]  44     68          0 1976
## [8,]  23     70          1 1997
## [9,]  25     70          1 1995
## [10,] 41     59          0 1979
```

Data example: in R

What are the dimensions of this matrix?

```
# both dimensions  
dim(dat)
```

```
## [1] 10 4
```

```
# number of rows  
nrow(dat)
```

```
## [1] 10
```

```
# number of columns  
ncol(dat)
```

```
## [1] 4
```

Data example: in R

Let's inspect parts of this matrix (vectors and scalars)

```
# Extract everyone's age (column vector)
dat[,1]
```

```
## [1] 25 38 35 36 23 26 44 23 25 41
```

```
# Find out if participant 2 likes dogs (scalar)
dat[2,3]
```

```
## likes_dogs
##           1
```

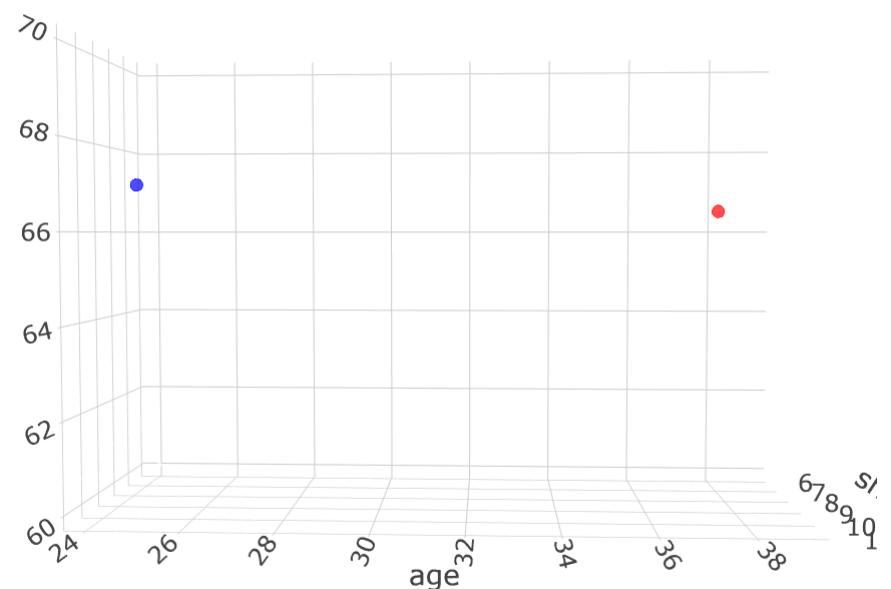
Q: How do we get all the data for participant 7?

The geometry of vectors

Let's look at the age, height and shoe size of participants 1 and 2:

$$\mathbf{x}_1 = \begin{bmatrix} 25 \\ 67 \\ 10.5 \end{bmatrix} \quad \mathbf{x}_2 = \begin{bmatrix} 38 \\ 66.5 \\ 7 \end{bmatrix}$$

The more attributes we measure about people, the higher-dimensional the space -- and the more precisely we can describe each person (they're not near anyone else in space).



Transposing vectors and matrices

We can flip a vector on its side by **transposing** it.

$$\mathbf{x}_1 = \begin{bmatrix} 25 \\ 67 \\ 10.5 \end{bmatrix} \quad \mathbf{x}_2 = \begin{bmatrix} 38 \\ 66.5 \\ 7 \end{bmatrix}$$

$$\mathbf{x}_1^T = [25 \ 67 \ 10.5] \quad \mathbf{x}_2^T = [38 \ 66.5 \ 7]$$

Note: We may also refer to transposed vectors (and matrices) with "prime" notation: $\mathbf{x}_1^T = \mathbf{x}'_1$ (Not to be confused with a derivative!)

Stack these row vectors and you have a matrix!

$$\mathbf{X} = \begin{bmatrix} \mathbf{x}_1^T \\ \mathbf{x}_2^T \end{bmatrix} = \begin{bmatrix} 25 & 67 & 10.5 \\ 38 & 66.5 & 7 \end{bmatrix}$$

Q: What are the dimensions of \mathbf{X} ?

Transposing vectors and matrices

We can also think of vectors as single-columned matrices:

- A column vector \mathbf{y} of length n has dimensions $n \times 1$.

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}_{n \times 1}$$

- A row vector \mathbf{y}' of length n has dimensions $1 \times n$.

$$[y_1 \quad y_2 \quad \cdots \quad y_n]_{1 \times n}$$

Thinking about vectors this way will be really handy when we multiply vectors and matrices!

Transposing vectors and matrices

We can also transpose matrices

$$\mathbf{X} = \begin{bmatrix} 25 & 67 & \dots & 1995 \\ 38 & 63 & \dots & 1982 \\ \vdots & \vdots & \ddots & \vdots \\ 41 & 59 & \dots & 1979 \end{bmatrix} \quad \mathbf{X}^T = \mathbf{X}' = \begin{bmatrix} 25 & 38 & \dots & 41 \\ 67 & 63 & \dots & 59 \\ \vdots & \vdots & \ddots & \vdots \\ 1995 & 1982 & \dots & 1979 \end{bmatrix}$$

The former columns (variables) are now rows. The former rows (participants) are now columns.

Transposing vectors and matrices

We can transpose a matrix using the `t()` function:

```
t(dat)
```

```
##           [,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8] [,9] [,10]
## age         25   38   35   36   23   26   44   23   25   41
## height      67   63   66   70   62   73   68   70   70   59
## likes_dogs  0    1    1    1    1    1    0    1    1    0
## yob        1995 1982 1985 1984 1997 1994 1976 1997 1995 1979
```

If we save this as another object, we can extract elements using the **opposite** indices as before:

```
dat_t <- t(dat)
dat[3, 4] # participant 3's year of birth
```

```
## yob
## 1985
```

```
dat_t[4, 3]
```

```
## yob
## 1985
```

Vector & matrix multiplication

When we multiply a scalar by a vector or a matrix, we can just do so element by element:

$$a\mathbf{y} = \begin{bmatrix} ay_1 \\ ay_2 \\ \vdots \\ ay_n \end{bmatrix} \quad a\mathbf{X} = \begin{bmatrix} ax_{11} & ax_{12} & \cdots & ax_{1p} \\ ax_{21} & ax_{22} & \cdots & ax_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ ax_{n1} & ax_{n2} & \cdots & ax_{np} \end{bmatrix}$$

But multiplying vectors and matrices requires special rules

Vector & matrix multiplication

We can only multiply vectors of the same length, and we have to transpose one before we can do so.

- Consider two vectors of length p , \mathbf{b} and \mathbf{c} :
- Think of them as matrices, each with 1 column
- To multiply them, their **inner** dimensions must match => must transpose one of them:

$$\mathbf{b}^T \mathbf{c} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_p \end{bmatrix}^T \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_p \end{bmatrix} = [b_1 \quad b_2 \quad \cdots \quad b_p] \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_p \end{bmatrix}$$
$$= b_1 c_1 + b_2 c_2 + \cdots + b_p c_p = \sum_{i=1}^p b_i c_i$$



Matrix \times matrix multiplication

Let's multiply a $m \times n$ matrix \mathbf{Y} by the $n \times p$ matrix \mathbf{X} :

$$\mathbf{Y}_{m \times n} \mathbf{X}_{n \times p} = \begin{bmatrix} \mathbf{y}_{11} & \mathbf{y}_{12} & \cdots & \mathbf{y}_{1n} \\ y_{21} & y_{22} & \cdots & y_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ y_{m1} & y_{m2} & \cdots & y_{mn} \end{bmatrix} \begin{bmatrix} \mathbf{x}_{11} & x_{12} & \cdots & x_{1p} \\ \mathbf{x}_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{x}_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix} = \begin{bmatrix} \mathbf{z}_{11} & \cdots \\ \vdots & \ddots & \vdots \\ \vdots & \cdots & \vdots \end{bmatrix}$$

where \mathbf{z}_{11} is the product of the first **row** of \mathbf{Y} and the first **column** of \mathbf{X}

$$\mathbf{z}_{11} = \sum_{i=1}^n y_{1i} x_{i1}$$

Matrix \times matrix multiplication

Now we do the same vector-vector multiplication with every pair of a row from \mathbf{Y} and a column from \mathbf{X}

$$\mathbf{Y}_{m \times n} \mathbf{X}_{n \times p} = \begin{bmatrix} \mathbf{y}_{11} & \mathbf{y}_{12} & \cdots & \mathbf{y}_{1n} \\ y_{21} & y_{22} & \cdots & y_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ y_{m1} & y_{m2} & \cdots & y_{mn} \end{bmatrix} \begin{bmatrix} x_{11} & \mathbf{x}_{12} & \cdots & x_{1p} \\ x_{21} & \mathbf{x}_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & \mathbf{x}_{n2} & \cdots & x_{np} \end{bmatrix} = \begin{bmatrix} z_{11} & \mathbf{z}_{12} & \cdots & \\ \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \vdots \end{bmatrix}$$

(You don't have to do it in any particular order, just keep multiplying until all row-column pairs have been multiplied.)

Q: What's the expression for z_{23} ?

$$z_{23} = \sum_{i=1}^n y_{2i} x_{i3}$$

(Note on dimensions)

Notice that I explicitly wrote out the dimensions when multiplying $\mathbf{Y}_{m \times n} \mathbf{X}_{n \times p}$

Just like we can only multiply vectors of the same length, dimensions of matrices must be compatible in order to be multiplied.



The **inner dimensions** must match, then the resulting product is a matrix with the **outer dimensions**.

I think of this as the inner dimensions "collapsing"

Matrix \times matrix multiplication: in R

```
X <- matrix(c(6, 3, 1, 6, 2, 3), ncol = 3)
X
```

```
##      [,1] [,2] [,3]
## [1,]     6     1     2
## [2,]     3     6     3
```

```
tX = t(X)
tX
```

```
##      [,1] [,2]
## [1,]     6     3
## [2,]     1     6
## [3,]     2     3
```

```
Z = tX%*%X
Z
```

```
##      [,1] [,2] [,3]
## [1,]    45    24    21
## [2,]    24    37    20
## [3,]    21    20    13
```

Matrix \times matrix multiplication: in R

```
Z = tX%*%X  
Z
```

```
##      [,1] [,2] [,3]  
## [1,]    45    24    21  
## [2,]    24    37    20  
## [3,]    21    20    13
```

```
tX[,1]%*%X[,1]
```

```
##      [,1]  
## [1,]    45
```

Identity matrix

We all know 1 is special. When you multiply a number by 1, you get the same number.

Matrices have their own special matrix, the identity matrix: **I**.

$$\mathbf{XI} = \mathbf{X} \text{ for any matrix } \mathbf{X}$$

For example:

$$\mathbf{QI} = \begin{bmatrix} r & s & t \\ u & v & w \end{bmatrix}_{2 \times 3} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}_{3 \times 3} = \begin{bmatrix} r & s & t \\ u & v & w \end{bmatrix}_{2 \times 3}$$

Symmetrical matrixes

The identity matrix \mathbf{I} is an example of a symmetric matrix: $\mathbf{I}_{i,j} = \mathbf{I}_{j,i}$ for any i, j where $i \neq j$

If you multiply a matrix by itself, the result is also a symmetric matrix:

$$\mathbf{Q}^T \mathbf{Q} = \begin{bmatrix} r^2 + u^2 & rs + uv & rt + uw \\ rs + uv & s^2 + v^2 & st + vw \\ rt + uw & st + vw & t^2 + w^2 \end{bmatrix}$$

Inverse of a matrix and linear independence

The inverse of a matrix \mathbf{Q} is the matrix \mathbf{Q}^{-1} such that

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

- For a scalar a , $a^{-1} = \frac{1}{a}$ is the multiplicative inverse of a : when we multiply the two together, we get 1
- For a matrix, the inverse is much more difficult to find, and does not exist if the columns are **linearly dependent**

Inverse of a matrix and linear independence: in R

We can attempt to invert a matrix using the `solve()` function:

```
mat_a
```

```
##      [,1] [,2]
## [1,]     2     1
## [2,]     6     8
```

```
solve(mat_a)
```

```
##      [,1] [,2]
## [1,]  0.8 -0.1
## [2,] -0.6  0.2
```

We can check to make sure this is the inverse:

```
mat_a_inv <- solve(mat_a)
mat_a_inv %*% mat_a
```

```
##      [,1] [,2]
## [1,]     1     0
## [2,]     0     1
```

We get the identity matrix I!

Inverse of a matrix and linear independence: in R

What if we have a matrix whose columns are not linearly independent?

```
mat_b
```

```
##      [,1] [,2]
## [1,]     2     1
## [2,]     6     3
```

```
solve(mat_b)
```

```
## Error in solve.default(mat_b): Lapack routine dgesv: system is exactly singular: U[2,2] = 0
```

Whenever you get an error message about something being "singular", that's code for a matrix not being invertible -- check for linear dependence!

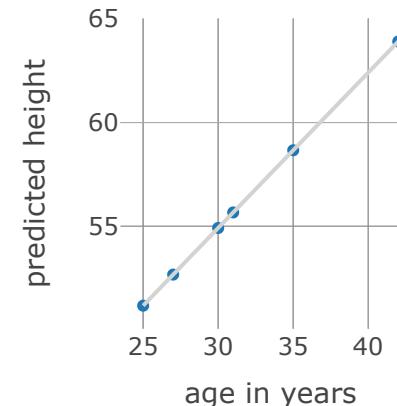
Q: How could you tell the matrix above is not invertible?

Inverse of a matrix and linear independence

Consider a situation in which we're predicting height as a linear function of several variables we have in our dataset: age in years, shoe size, and age in months.

$$\text{height} = \begin{bmatrix} 51 \\ 61 \\ 52 \\ 65 \\ 60 \\ 48 \end{bmatrix} \quad \text{other variables} = \begin{bmatrix} 30 & 7 & 360 \\ 31 & 10 & 372 \\ 25 & 9 & 300 \\ 35 & 10 & 420 \\ 42 & 6 & 504 \\ 27 & 7 & 324 \end{bmatrix}$$

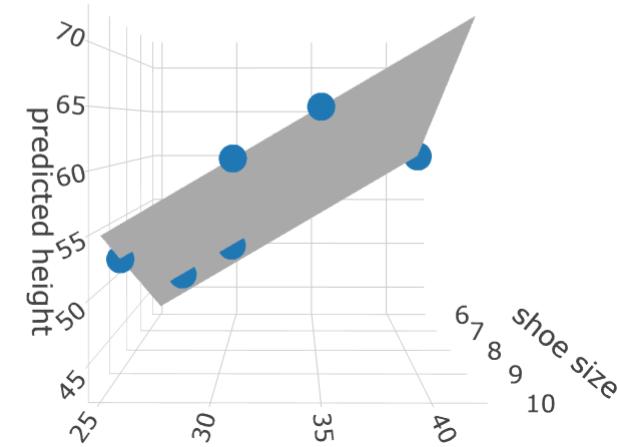
We can plot predicted height as a function of age in years alone. This is a line.



Linear independence, cont.

If we plot predicted height as a function of age in years *and* shoe size, we get a **plane**.

That is, we get more information about height from knowing someone's shoe size. If two people are the same age, but have different shoe sizes, we'll predict different heights for them.

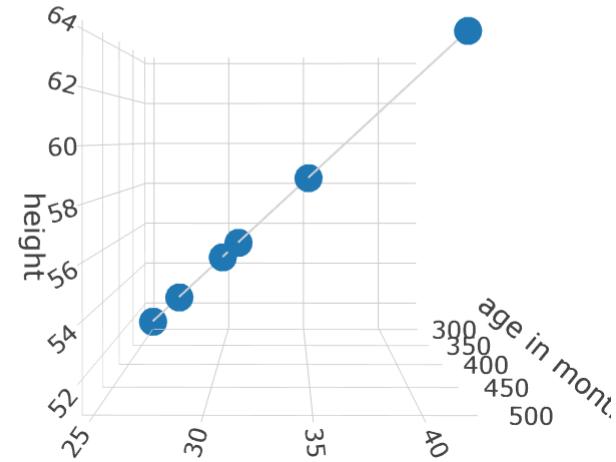


Linear dependence

But if we plot predicted height as a function of age in years and *age in months*, we get a line again, instead of a plane.

That's because age in years and age in months are **linearly dependent**: we can write one as a linear combination of the others; that is, $\text{age in months} = 12 \times \text{age in years}$.

We don't get any extra information from knowing age in months that we didn't already have from age in years.



In "real life"

Linear dependence will come into play when we cover regression. In the meantime, we can practice some regression matrix notation! You may be used to seeing a linear regression equation written out like this:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \epsilon_i$$

Well, that's the same thing as:

$$y_i = \boldsymbol{\beta}^T \mathbf{x}_i + \epsilon_i$$

where $\boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix}_{3 \times 1}$ and $\mathbf{x}_i = \begin{bmatrix} x_{i0} \\ x_{i1} \\ x_{i2} \end{bmatrix}_{3 \times 1}$ where $x_{i0} = 1$ for all i

Try multiplying it out by hand, first making sure the dimensions are compatible (notice the necessity to transpose $\boldsymbol{\beta}$)!

Stay tuned for more on this when we cover regression analysis ☺

More resources

This whole section has a lot of great information and practice with matrices.

For a more advanced introduction, work through the sections on vectors, linear combinations, and linear dependence [here](#).

You can also pick and choose from the videos [here](#), particularly those on functions and linear transformations.

Part 2: Functions, data transformations and a sprinkle of calculus

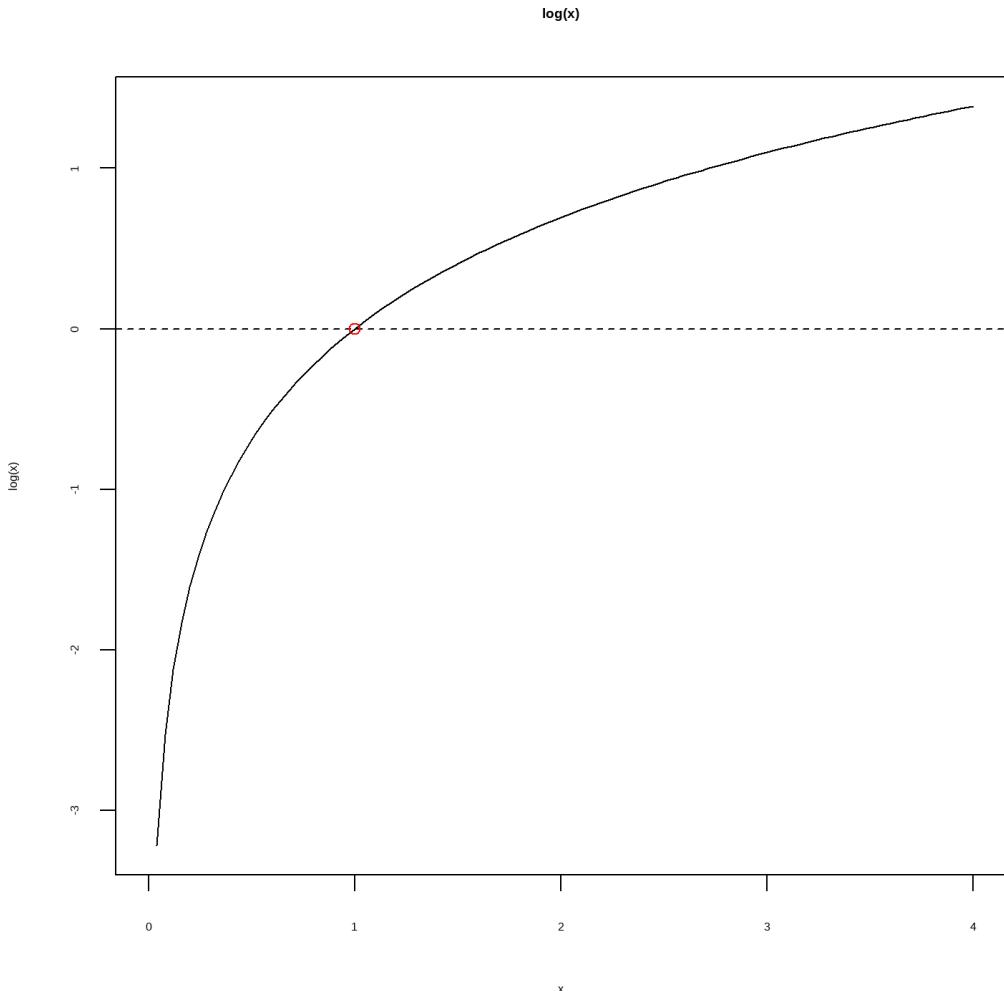
Now that we've explored how data is stored, let's look at how data can be transformed

Logarithms

If you see $\log(x)$ in this class, or basically anywhere in probability and statistics, it will refer to the natural logarithm, or $\ln(x)$.

What do you notice about the function $\log(x)$?

You can only "log" a positive number. Something like $\log(-1)$ is undefined. We can see that $\lim_{x \rightarrow 0} \log(x) = -\infty$. Importantly, $\log(1) = 0$, so $\log(x)$ for any x between 0 and 1 will give you a negative number.



Exponentiation

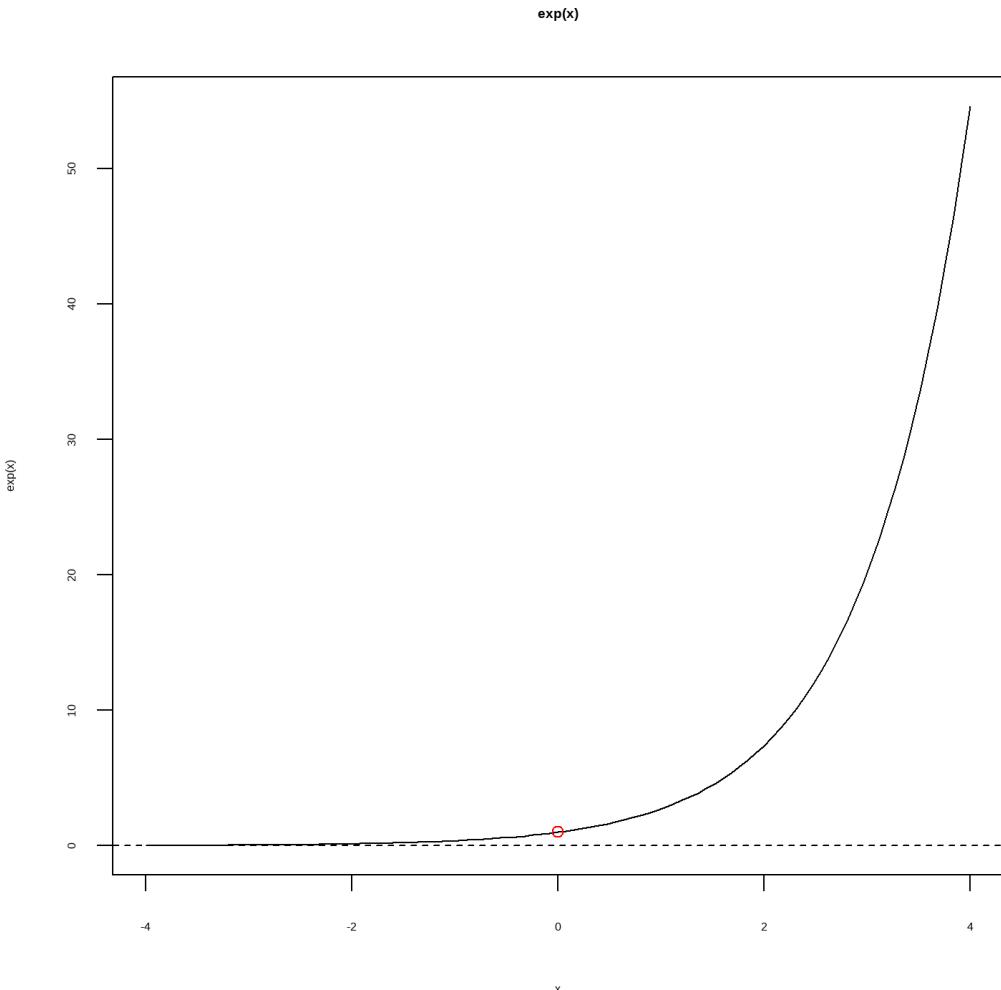
The **inverse** of the natural log is the natural exponential function e^x , which we also write as $\exp(x)$:

$$\exp(x) = y \iff x = \log(y)$$

so if one side of an equation is exponentiated, we can always "get out of it" by applying a logarithm to both sides, and vice versa.

$$\log(\exp(x)) = x \text{ and } \exp(\log(x)) = x$$

But we can't do that if x is negative! Recall that we can't take a log of a negative number, and exponentiating a number is never going to give us anything negative.



Rules to live by

When we have a sum inside an exponent, we can decompose this into the product of two exponents:

$$\exp(a + b) = \exp(a) \exp(b)$$

Similarly, a product inside a logarithm can be written as a sum of logs:

$$\log(ab) = \log(a) + \log(b)$$

Of course, the same is true of the inverses of addition and multiplication, subtraction and division:

$$\exp(a - b) = \frac{\exp(a)}{\exp(b)}$$

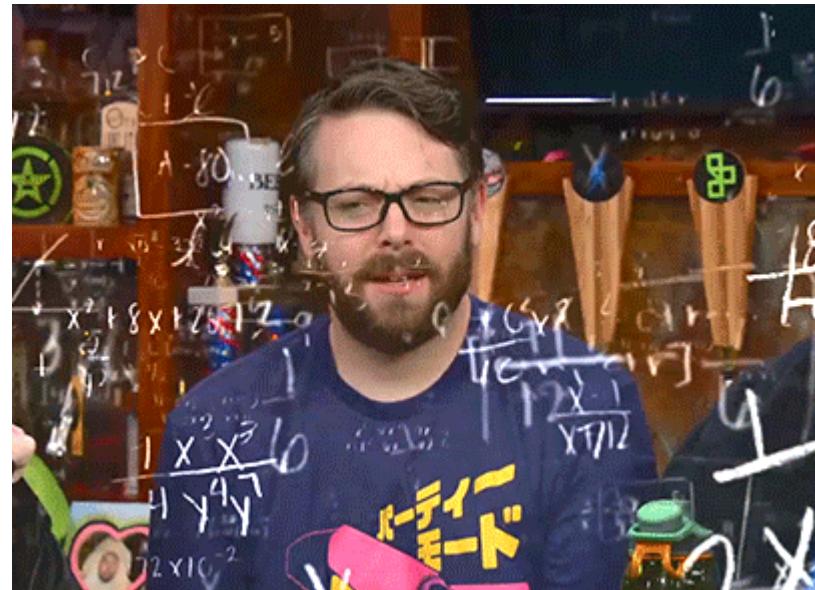
$$\log\left(\frac{a}{b}\right) = \log(a) - \log(b)$$

You will need to know these rules to understand various regression models in this class, I promise!

Why do we care?

The log function takes a number x constrained to be > 0 , and transforms it onto an unbounded number space

- Probability: ranges from 0 to 1 (strictly constrained)
- Say we know $\Pr[\text{likesdogs} = 1]$ is a function of age and height: how do we take 2 numbers, both outside the range of 0 to 1, and get a probability that is strictly between 0 and 1?



The logit function is here to help!

Odds and probabilities

Let's break it down. Think about flipping a coin lots of times: heads you win, tails you lose.

- A **probability** describes the number of successes out of the total number of trials (a proportion)
- An **odds** describes the number of successes compared to the number of failures (a ratio)

$$odds = \frac{prob}{1-prob}$$

$$prob = \frac{odds}{1+odds}$$

Let's say you get 4 heads out of 10 flips:

- probability = $\frac{4}{10}$
- odds = $\frac{4}{6}$

These are really different numbers!

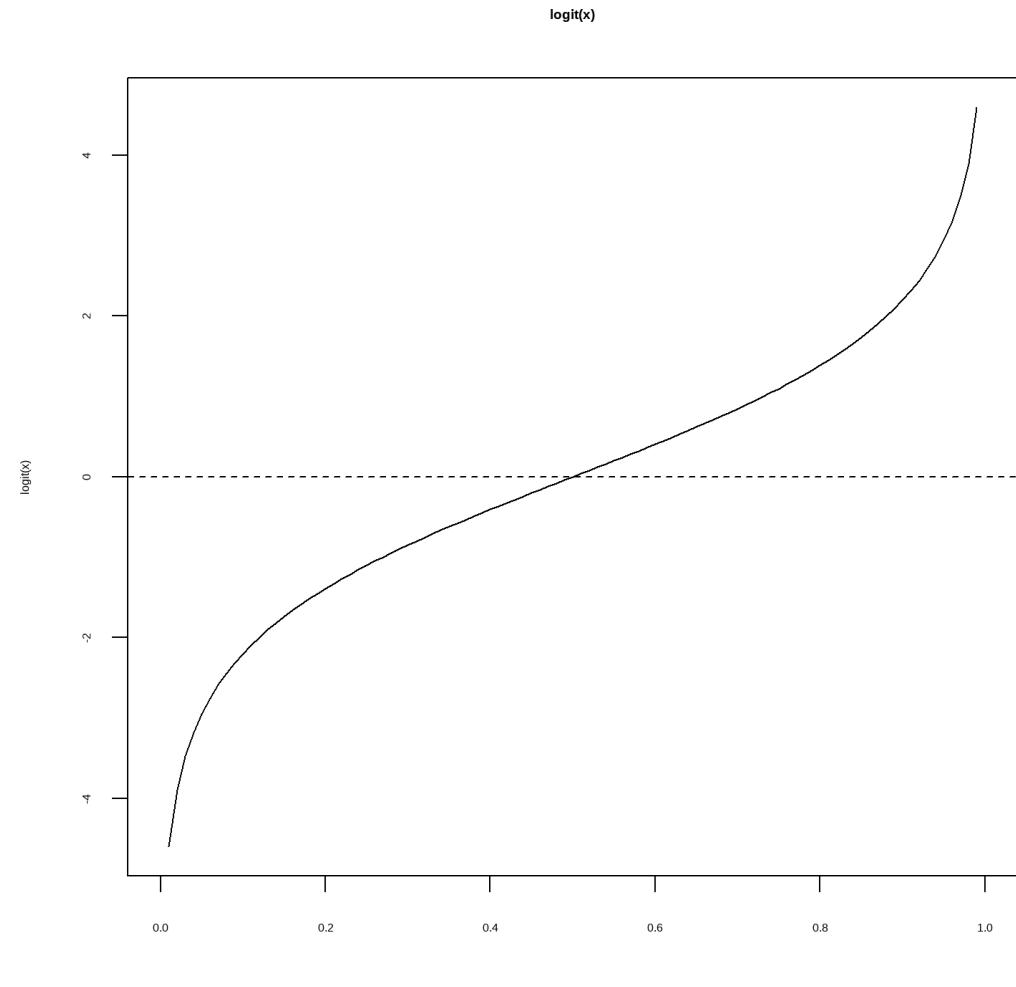
Logits and expits

The **logit** of p , aka the **log-odds** of p , can take a number between 0 and 1 (like a probability!) and transform it to a number between $-\infty$ and ∞ .

$$\text{logit}(p) = \log\left(\frac{p}{1-p}\right)$$

We can invert it to get the **expit** function, which can take any number on the real line and transform it to a value between 0 and 1:

$$\text{expit}(x) = \frac{\exp(x)}{1+\exp(x)}$$



Back to the data example:

```
##      age height likesdogs yob
## 1    25     67         0 1995
## 2    38     63         1 1982
## 3    35     66         1 1985
## 4    36     70         1 1984
## 5    23     62         1 1997
## 6    26     73         1 1994
## 7    44     68         0 1976
## 8    23     70         1 1997
## 9    25     70         1 1995
## 10   41     59         0 1979
```

Back to the data example:

Consider the following logistic regression model for the predicted probability that someone likes dogs (p_i) given their age in years (x_i):

$$\widehat{\log\left(\frac{p_i}{1-p_i}\right)} = 0.4 + 0.07 * \text{age}_i - 0.005 * \text{height}_i$$

What's the probability that person 4 likes dogs?

Log-odds: $\log\left(\frac{p_4}{1-p_4}\right) = 0.4 + 0.07 * 36 - 0.005 * 70 = 2.57$

Odds: $\frac{p_4}{1-p_4} = \exp(2.57) = 13.066$

Probability: $p_4 = \frac{13.066}{1+13.066} = \frac{\exp(0.4+0.07*36-0.005*70)}{1+\exp(0.4+0.07*36-0.005*70)} = 0.929$



Q: What's the probability of liking dogs of a person who's 20 years old and 67 inches tall?

A quick dip into calculus

Derivatives

The basic idea of a derivative is that it describes the rate of change of a function. If the function we're looking at is $g(x)$, then there are a couple of ways we usually notate the first derivative of $g(x)$, which we'll use interchangeably:

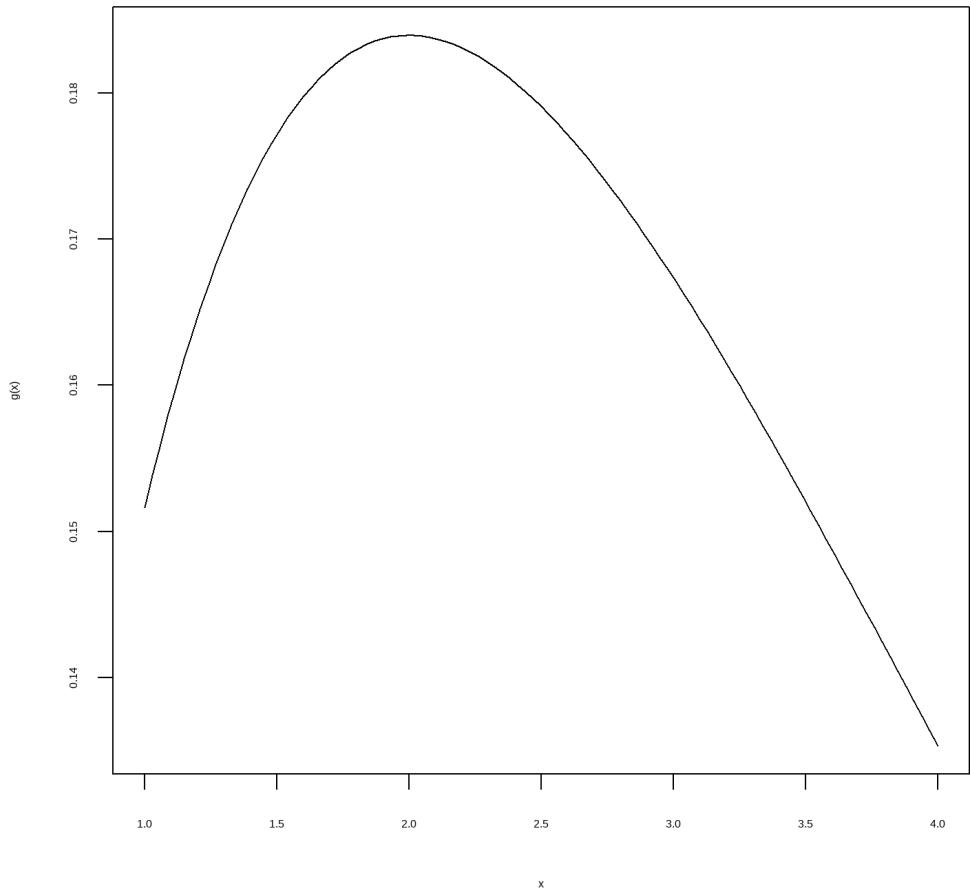
$$g'(x) = \frac{d}{dx} g(x)$$

Both equivalently tell us that we are looking at the function $g(x)$ and taking the first derivative with respect to the variable x . That means that as x changes, we want to know how much $g(x)$ changes. This is just the slope of $g(x)$ at a given value of x .

Where is $g'(x)$ greatest?

More derivatives

Where is $g'(x)$ greatest?



Where does $g'(x) = 0$?

When the first derivative is 0, the function may have reached its maximum or minimum.

So if you want to maximize a function, one way to do so is to differentiate it and then set it equal to 0.

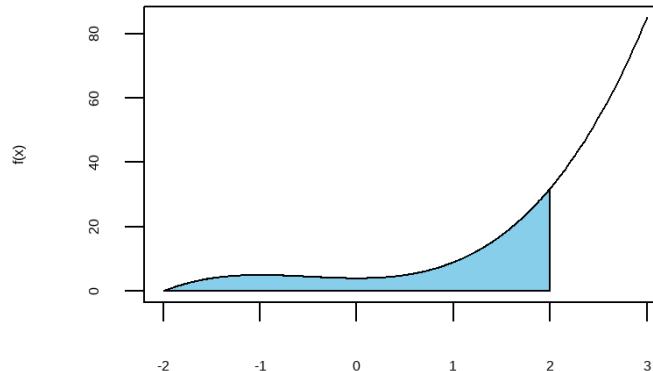
Integrals

We might want to describe a function by the area under its curve

An integral tells us how much cumulative space a function is covering (in terms of distance from the x -axis) as x gets larger.

In the graph of $f(x) = 2x^3 + 3x^2 + 4$ below, the area in blue is represented by the integral

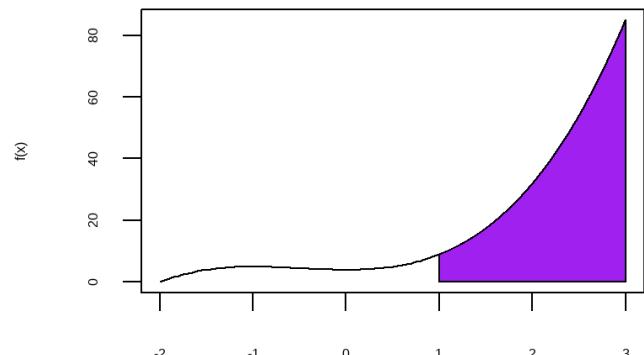
$$\int_{-2}^2 2x^3 + 3x^2 + 4 \, dx$$



Integrals

$$\int_{-2}^2 2x^3 + 3x^2 + 4 \, dx$$

The values at the top and bottom of the integral sign are those between which we're computing the integral. We could integrate over the whole function, from $-\infty$ to ∞ , or choose other limits of integration:



What limits of integration are displayed here?

The integral of a non-negative function increases with x :

$$\int_{-2}^{-1} 2x^3 + 3x^2 + 4 \, dx$$

$$\leq \int_{-2}^0 2x^3 + 3x^2 + 4 \, dx$$

$$\leq \int_{-2}^1 2x^3 + 3x^2 + 4 \, dx$$

$$\leq \int_{-2}^2 2x^3 + 3x^2 + 4 \, dx$$

and so on. On the graph, the area under the curve can only accumulate area, so the integral evaluated at greater and greater upper limits can only increase.

Data example

Let's pretend we grabbed a larger sample from the population where we are examining age, height, affinity for dogs and year of birth:

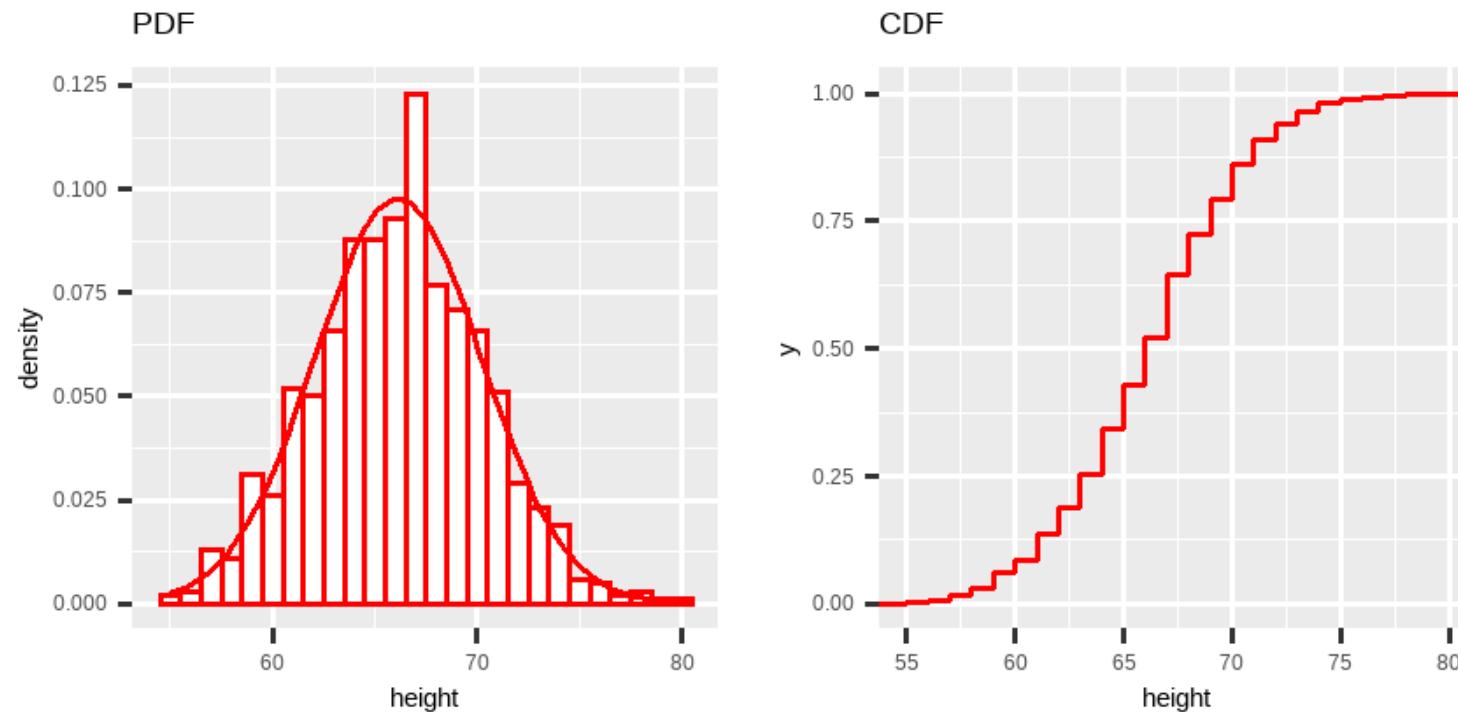
```
set.seed(6789)
n <- 1000
dat <- data.frame(
  age = round(runif(n, 22, 45)), height = round(rnorm(n, 66, 4)), likes_dogs = rbinom(n, 1, .53)
)
dat$yob <- 2020 - dat$age

summary(dat)
```

```
##      age        height     likes_dogs       yob
##  Min.   :22.00   Min.   :55.00   Min.   :0.000   Min.   :1975
##  1st Qu.:28.00   1st Qu.:63.00   1st Qu.:0.000   1st Qu.:1981
##  Median :33.00   Median :66.00   Median :1.000   Median :1987
##  Mean    :33.12   Mean    :66.13   Mean    :0.524   Mean    :1987
##  3rd Qu.:39.00   3rd Qu.:69.00   3rd Qu.:1.000   3rd Qu.:1992
##  Max.    :45.00   Max.    :80.00   Max.    :1.000   Max.    :1998
```

Data example

The distribution of height in inches:



Q: Roughly what proportion of the sample is <70 inches tall? What's an expression for this quantity in terms of the integral of the PDF of height?

More resources

Here is a **lot** of information about derivatives. You don't need more than the first few videos.

Same with **this** intro to integrals.

Watch the video on antiderivatives and indefinite integrals from **this** page and some of those on **this** page to understand the link between derivatives and integrals.