Math Camp - Day 2

Aaron Rudkin

Trinity College Dublin
Department of Political Science

rudkina@tcd.ie

Functions

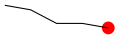
Functions

- What is a function?
  - A function is a **black box –** a device that translates from input $\rightarrow$ output
  - A function will map one or more input values (characteristics of data or an object) onto values measuring another characteristic.
  - Typically written in the style $f(x)$ where f is a label being assigned to the function and x is an input argument.
- $f(x) = y$
- $f(x) = 4x + 1 \rightarrow f(3) = 4(3) + 1 = 13$
- Arguments can be constants or themselves functions: $f(x) = x^2 + 2 \rightarrow f(y + 3z) = (y + 3z)^2 + 2$

Functions

- The **domain** of the function is the set of possible input values for which the function has a meaningful output.
  - $\frac{x+1}{x-1} \rightarrow$ Not defined when $x = 1$ (why not?)
- **Function composition** chains multiple functions, another way to use functions as inputs:
  - If $f(x) = 2x + 1$ and $g(x) = x^2 + 1$
  - Then $f \circ g(x) = 2(x^2 + 1) + 1 = 2x^2 + 3$
- **Identity function** – $f(x) = x$: Elements in domain are mapped to identical elements without transformation.
- **Inverse function** – $f^{-1}(f(x)) = x$: Function that when composed with original function returns identity function.
- Can have multiple inputs or outputs.
- In programming, used for any set of instructions.

Increasing and Decreasing Functions

- Increasing 
  As inputs increase, outputs increase.
- Decreasing 
  As inputs increase, outputs decrease.
- Strictly increasing: Increasing function that is never flat.
- Strictly decreasing: Decreasing function that is never flat.
- Weakly increasing: Increasing function that is sometimes flat.
- Weakly decreasing: Decreasing function that is sometimes flat.

Functions

Table: Interval Notation

Interval	Set	Description
(a, ∞)	$\{x \in \mathbf{R}^1 : x > a\}$	Open, bounded
$[a, \infty)$	$\{x \in \mathbf{R}^1 : x \geq a\}$	Half-open, bounded
$(-\infty, a)$	$\{x \in \mathbf{R}^1 : x < a\}$	Open, bounded
$(-\infty, a]$	$\{x \in \mathbf{R}^1 : x \leq a\}$	Half-open, bounded
$(-\infty, \infty)$	$\mathbf{R}$	Open, unbounded

Linear Functions

- Simplest linear function: $y = a + bx$, where a and b are constants.

- When b is negative, function is decreasing. When b is positive, function is increasing.
- Linear functions can also have multiple inputs: $z = a + bx + cy$
- Easily interpreted in regression format: The "effect" of a one-unit change in x in output y is the same everywhere in the domain. Givens us statements like: "One additional year of schooling is associated with 2,000 euro additional annual income".
- Slope (how steep is the line): $b = \frac{y_2 - y_1}{x_2 - x_1} \rightarrow$ "rise over run"

Nonlinear Functions

- The formal definition of a linear function is any function with the following properties:
  - Additivity (superposition): $f(x1 + x2) = f(x1) + f(x2)$
  - Scaling (homogeneity): $f(ax) = af(x)$
- Not all functions are linear. Some examples:
  - Exponents and roots: $y = x^2 + 3x + 2$
  - Logarithms
  - Exponential functions: $y = e^{x^2+1}$

Nonlinear Functions – Exponents

Recap from yesterday: multiplication of a number by itself. e.g. $a^3 = a \times a \times a$

Table: Rules of Exponents

Multiplication rule	$a^x \times a^y = a^{x+y}$
Division rule	$a^x \div a^y = a^{x-y}$
Power of a power rule	$(a^x)^y = a^{xy}$
Power of a product rule	$(ab)^x = a^x b^x$
Power of a fraction rule	$(\frac{a}{b})^x = \frac{a^x}{b^x}$
Zero exponent	$a^0 = 1$
Negative exponent	$a^{-x} = \frac{1}{a^x}$
Fractional exponent	$a^{\frac{1}{x}} = \sqrt[x]{a}$

Squared and cubic terms

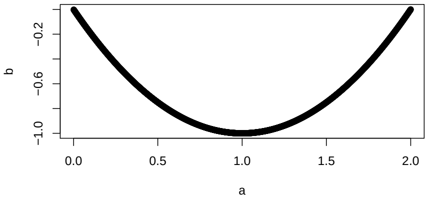


Figure: $y = -2x + x^2$

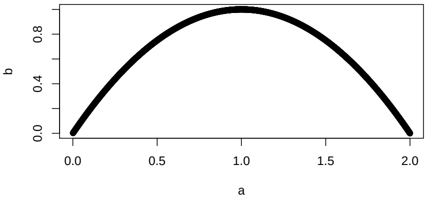


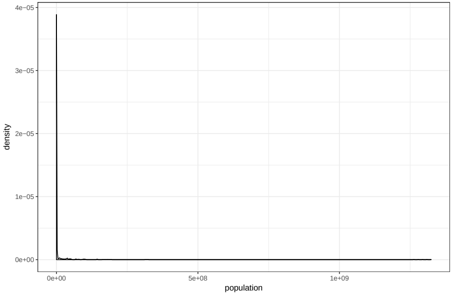
Figure: $y = 2x - x^2$

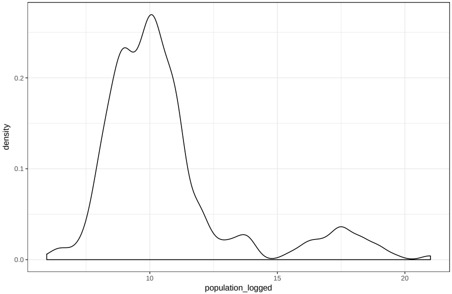
Logarithms

Logarithms

- $y = log_a z \Leftrightarrow z = a^y$
- The logarithm of z – (y) is the power to which one must raise a to yield z.
- $log_x x^{2y} = 2y$
- $log_3 27 = 3$
- $log_4 256 = 4$
- Why do we need logarithms? Useful for modelling functions where growth depends on level: increasing or diminishing marginal returns. Bank account interest, population growth, happiness from additional slices of pizza, magnitude of an earthquake.

Logarithmic transformation





Logarithms “squish” large differences, making them appear smaller, and “stretch” small differences, making them appear larger.

Properties of Logarithm, Recap

- $log(x \times y) = log(x) + log(y)$
- $log(\frac{1}{x}) = -log(x)$
- $log(\frac{x}{y}) = log(x) - log(y)$
- $logx^y = ylog(x)$
- $log(1) = 0$

Change of Base Trick

- The Trick:
  $$log_b(a) = \frac{log_x(a)}{log_x b}$$
- Example:
  $$log_2 48 = \frac{log_{10} 48}{log_{10} 2} \approx 5.58$$
- Easy to quickly convert to a base that can be solved on a calculator
- Independent study: Why does this work? How do rules of exponents lead to this?

Interpreting Log-Transformed Variables

- In a regression context, either function inputs, function outputs, or both can be log-transformed.
- Suppose we wish to study the impact of additional investment euro in a charity's programming.
- $Y = \alpha + 10.43X + \epsilon$ and let us assume X is log-transformed. (ϵ is used for the "error term": unexplained variation)

Interpreting Log-Transformed Variables

- In a regression context, either function inputs, function outputs, or both can be log-transformed.
- Suppose we wish to study the impact of additional investment euro in a charity's programming.
- $Y = \alpha + 10.43X + \epsilon$ and let us assume X is log-transformed. (ϵ is used for the "error term": unexplained variation)
- Implies a one-unit increase in X in log terms leads to 10.43 unit increase in Y.
- Because X is log-tranformed, a "one unit increase" in X is not interpretable in euro terms.
  - Depending on baseline, increase of one log unit may be larger or smaller
  - $exp(4) - exp(3) \approxeq 34.5$
  - $exp(15) - exp(14) = 2{,}066{,}413$

Quick Hack for Unpacking Log-Transformed Variables

- Given $y = \alpha + \beta X$ linear function, and say $\beta = 10.43$ as before, the expected change in Y associated with a p% increase in X (logged) can be calculated as $\beta \dot{l}og([100 + p]/100)$.
- A 10% increase in X will increase Y by $10.43 \times log(1.10) = 10.43 \times .09531 \approx 0.994$
- A 50% increase in X will increase Y by $10.43 \times log(1.50) = 10.43 \times .17609 \approx 1.83$

Limits

Limits and Functions

- For the function $y = f(x)$, a limit is the value of y that the function tends toward (approaches) given arbitrarily small movements toward a specific value of x, say $x = c$.
- Notation: $\lim_{x \to c} f(x)$
- Limits do not always exist, but they exist where $f(x)$ is smooth and continuous (no breaks in the curve or plane fit by the function).
- Limits can even exist where $f(x)$ is not defined: consider $f(x) = \frac{1}{x}$. This function is not defined at $x = 0$. But $lim_{x \to 0+} = \infty$.

Continuity

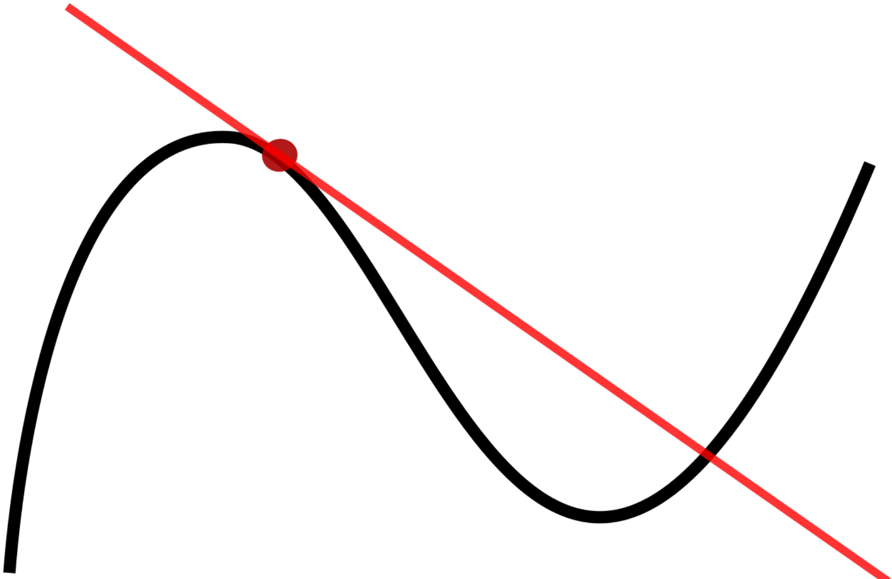
- Intuitively, a **continuous function** is a function without sudden breaks in it.
- When you draw the graph of a continuous function, you never need to lift your pencil from the page.
- $f(x)$ is continuous at $x = c$ iff $\lim_{x \to c} f(x) = f(c)$.
- A non-continuous function might have a single discontinuity (say, $f(x) = \frac{1}{(x-1)}$) or else it could be **piecewise**.

Introduction to Derivatives

Introduction to Derivatives

- Earlier, we saw that functions have slopes (how steeply they increase or decrease in a given section)
- In a linear function, the slope is constant, but in a non-linear function, the slope is sometimes steeper and sometimes less steep
- A slope is also called a "rate of change": how much does the function change as its inputs change?
- The derivative of a function at a point is the *instantaneous* rate of change. At any given point, how much is it changing?

Visual intuition of a derivative



Derivatives and Limits

- Intuition to calculate: evaluate slope near point of interest
- But the smaller the change in x over which we evaluate the slope, the closer we get to evaluate instantaneous rate of change.
- Making the difference between points smaller and smaller until if approaches to zero, we take the limit to get the derivative – "rise over run".
- But thankfully, don't need to use this approach: we have analytical solutions, rules for whole classes of functions.
- Derivatives of polynomial functions, as we'll see in rules tomorrow, lower order of polynomial: linear functions have constant derivatives, quadratic functions have linear derivatives,

$$\lim_{h \to 0} \frac{f(x+h) - f(x)}{(x+h) - x}$$

Notation

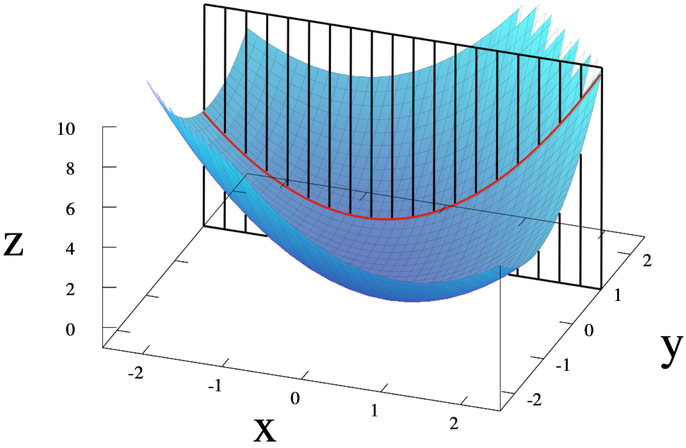
- $\frac{d}{dx}f(x) \rightarrow$ read as "derivative with respect to x of f(x)", i.e. how is the function changing as we change x?

- $\frac{dy}{dx} \rightarrow$ read as "derivative of y with respect to x of f(x)", i.e. how is y, the output of the function, changing as we change x?

- $f'(x) \rightarrow$ read as "f prime of x". Each time we take a derivative, add another prime mark.

- **All of these notations are equivalent**

Partial derivatives



Partial Derivatives

- Some functions have multiple inputs, but still have a class of derivatives.
- For $f(x, z)$, to know how y changes with x holding z constant, we need partial derivatives.
- Partial derivative keeps all but one input to function constant (taking a "slice" out of the function) and takes derivative with respect to one variable's slope.
- Notation: $\frac{\partial}{\partial x} f(x, z)$
- Note use of lowercase delta by convention.

Rules of Differentiation

Rules of Differentiation

Name of rule	Function type	Solution
Constant rule	$f(x) = a$	$f'(x) = 0$
Power rule	$f(x) = (x^n)$	$f'(x) = nx^{n-1}dx$
Sum rule	$f(x) = g(x) + h(x)$	$f'(x) = g'(x) + h'(x)dx$
Difference rule	$f(x) = g(x) - h(x)$	$f'(x) = g'(x) - h'(x)dx$
Product rule	$f(x) = g(x)h(x)$	$f'(x) = g(x) + h(x)g'(x)dx$
Quotient rule	$f(x) = \frac{g(x)}{h(x)}$	$f'(x) = \frac{g'(x)h(x)-g(x)h'(x)}{h(x)^2}dx$
Constants	$f(ax)$	$f'(ax) = af'(x)dx$
Chain rule	$f(g(x))$	$f'(g(x)) = f'(x)g'(x)dx$
Exponential rule	$f(x) = a^x$	$f'(x) = a^x ln(x)dx$
Logarithm rule	$f(x) = log_a(x)$	$f'(x) = \frac{1}{xln(a)}dx$

Power Rule

- **Power Rule**: Simple to learn and master, extremely powerful; almost all of the additive models we use are some sort of polynomial!
- $\frac{d}{dx}x^n = n \times x^{n-1}dx$
- Examples:

$$\frac{d}{dx}4x^3 = 4 \times 3x^2dx = 12x^2dx$$
$$\frac{d}{dx}\frac{x^4}{2} = \frac{1}{2}4x^3dx = 2x^3dx$$
$$\frac{d}{dx}\sqrt{x} = \frac{d}{dx}x^{\frac{1}{2}} = \frac{1}{2}x^{-\frac{1}{2}}dx = \frac{1}{2\sqrt{x}}dx$$

Chain Rule

- **Chain Rule**: If you can't find a rule to solve a function, write the function as a composite!
- $\dfrac{d}{dx}f(g(x)) = f'(x)g'(x)dx$
- Examples:

  $\dfrac{d}{dx}\sqrt{3x+2}$

  Let $f(x) = \sqrt{x}; g(x) = 3x + 2$

  Then $f'(x) = \frac{1}{2\sqrt{x}}; g'(x) = 3$

  $\dfrac{d}{dx}\sqrt{3x+2} = \dfrac{1}{2\sqrt{3x+2}}3dx = \dfrac{3}{2\sqrt{3x+2}}dx$

Approaching a derivative

- Break a complicated function down into its constituent parts using rules
- Each individual part can be solved easily
- Keep track of various functions

$$
\begin{aligned}
f(x) &= x^3 + 3x^2 - 6x + 5 \\
&= {\color{red}x^3} + {\color{blue}3x^2} - {\color{orange}6x} + {\color{purple}5} \\
f'(x) &= 3x^2 + 6x - 6 + 0\,dx
\end{aligned}
$$

$$
\begin{aligned}
{\color{red}\frac{d}{dx}x^3} &= 3x^2 dx \\
{\color{blue}\frac{d}{dx}3x^2} &= 3(2x) = 6xdx \\
{\color{orange}\frac{d}{dx}6x} &= 6x^0 = 6dx \\
{\color{purple}\frac{d}{dx}5} &= 0
\end{aligned}
$$

L'Hopital's Rule

- One immediate use of a derivative is that it allows us to solve the limits of functions that would otherwise be undefined.
- Consider $\lim_{x \to 0} \frac{e^x - 1}{3x}$
- As x approaches zero, the numerator tends to 0... and the denominator tends to 0. How to solve?
- **L'Hopital's Rule:** $\lim_{x \to 0} \frac{f(x)}{g(x)} = \lim_{x \to 0} \frac{f'(x)}{g'(x)}$

$$\lim_{x \to 0} \frac{e^x - 1}{x^2} = \lim_{x \to 0} \frac{e^x}{3} = 0$$

Conclusion

Tomorrow

- Derivatives for Optimization
- Intro and Rules for Integrals
- Intro to Matrices