Math Camp - Day 3

Aaron Rudkin

Trinity College Dublin
Department of Political Science

rudkina@tcd.ie

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!
- $\dfrac{d}{dx}x^n = n \times x^{n-1}dx$
- 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

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

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

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:** $\lim_{x \to 0} \frac{f(x)}{g(x)} = \lim_{x \to 0} \frac{f'(x)}{g'(x)}$

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

Finding Maxima and Minima

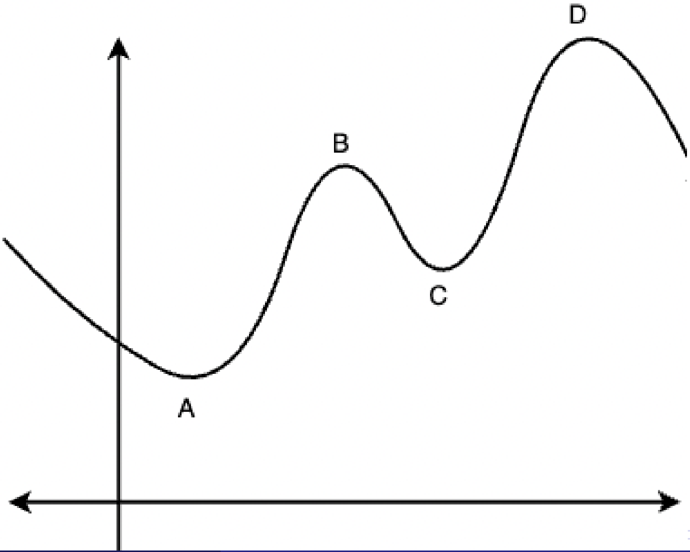
Finding Maxima and Minima

- Why else do we actually care about derivatives?
- Derivatives are useful to "optimize" functions: find the maxima (highest points) and minima (lowest points)
- Data science: The essence of fitting a model to data is finding the parameters that maximize fit or minimize error!
- Game theory and behaviour: Which response optimizes my utility?
- Regression:
  - Ordinary Least Squares **minimizes** squared error: $e'e$
  - Maximum Likelihood Estimation **maximizes** a likelihood function.

Finding Maxima and Minima

- For any function, a high point is called a maximum (plural maxima) and a low point is called a minimum (plural minima).
- Recall graphical example from yesterday
- At the extremum, drawing a tangent line (line that touches the function only at the extremum) will have a slope zero
- Intuition: A maximum happens when the slope of a function goes from positive to negative, passing through zero. A minimum happens when the slope of a function goes from negative to positive, passing through zero. The derivative is the slope of the function!
- An extremum is **local** if it is the largest or smallest value in its neighborhood (some interval of the domain of the function), and **global** if it is the largest or smallest value across all points of the function

Identifying extrema



Higher Order Derivatives

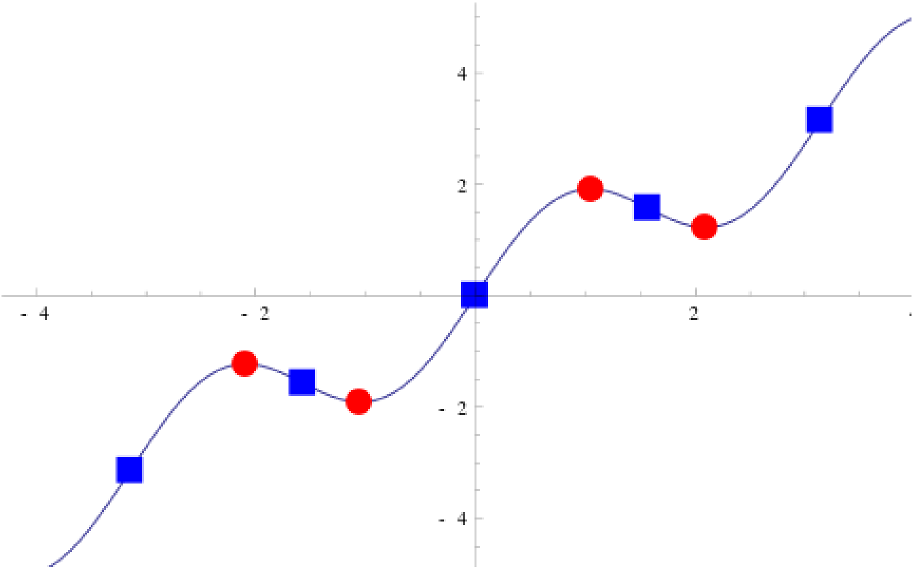
- The n^{th}-**order derivative** is the rate of change of the $n-1^{th}$-order derivative.
- The first-order derivative is the rate of change (slope) of the function. If the function is distance, the derivative is speed
- The second-order derivative is the rate of change of the instantaneous rate of change. If the function is distance, the second-order derivative is acceleration.
- The third-order derivative is the rate of change of the second-order derivative, and so on.
- Given a polynomial function, no x^n term where n is less than the order of the derivative will be in the answer!

  $f(x) = 2x^4 \rightarrow f'(x) = 8x^3 \rightarrow f''(x) = 24x^2 \rightarrow f'''(x) = 48x \rightarrow f''''(x) = 48 \rightarrow f'''''(x) = 0$

Critical Points and Inflection Points

- A **critical point** is any point x^* such that either $f'(x^*) = 0$ or $f'(x^*)$ doesn't exist
- A critical point, called a **stationary point** when the slope is flat, can be a local minimum, a local maximum, or an inflection point.
- **Inflection points** are points at which the graph of the function changes from concave (bending down) to convex (bending up) or vice versa.

Critical Points and Inflection Points



Method to Find Minima and Maxima

1. Given a function $f(x)$, we wish to find a minimum or maximum.
2. First, find $f'(x)$.
3. Set $f'(x^*) = 0$ and solve for all x^*. These are stationary points of the function. Any minimum or maximum that exists will be among these.
4. If we are unsure whether this is a minimum, maximum, or inflection point, we have a mathematical tool to check.
5. Find $f''(x)$ – the second derivative.
6. For each stationary point x^*, say a stationary point at $x = 3$, substitute x^* into $f''(x)$, e.g. $f''(3)$.
   - If $f''(x^*) < 0$, $f(x)$ has a local maximum at x^*.
   - If $f''(x^*) > 0$, $f(x)$ has a local minimum at x^*.
   - If $f''(x^*) = 0$, x^* may be an inflection point (rules for adjudicating this can be found online)

tl;dr

To find the extremum, set the first derivative to zero and solve for x. To determine if it is a maximum or minimum, take the sign of the second derivative.

Introduction to Integrals

- An integral, also called an anti-derivative, is **the Area under the Curve**.
- Basic notation: $F(x) = \int f(x)dx$
- What do these terms mean?

Defining notation

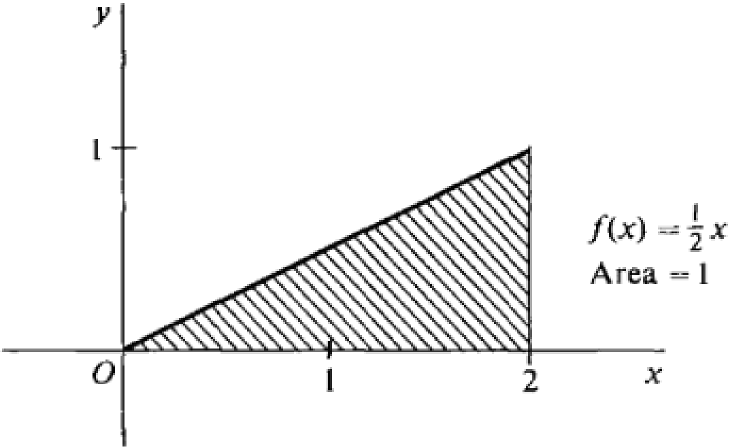
From "Calculus Made Easy" (1910):

The preliminary terror, which chokes off most [people] from even attempting to learn how to [integrate], can be abolished once for all by simply stating what is the meaning of the two principal symbols that are used in [integrating]:

(1) d which merely means "a little bit of.". Thus dx means a little bot of x. You will find that these bits may be considered to be indefinitely small.
(2) $\int$ which is merely a long S, and may be called "the sum of".

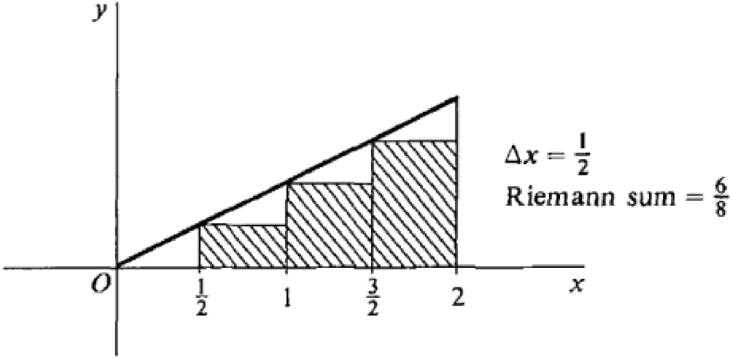
Thus, $\int dx$ means the sum of all the little bits of x. Now any fool can see that if x is considered as made up of a lot of little bits, each of which is called dx, if you add them up together you get the sum of all the dx's, (which is the same thing as the whole of x).

Taking an integral without even realizing it



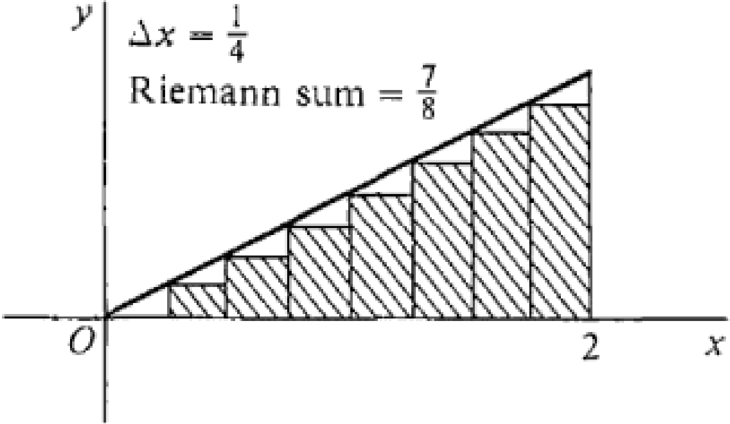
Intuition of integral

What if we didn't have an analytical formula like $area = \frac{b \times h}{2}$? Use **Riemann sums** to measure sections:



Intuition of integrals

Can't get the answer exactly right; will always count too much or too little! But by shrinking rectangles, can get closer to correct.



Reimann Sum and definite integral

- We can treat the integral as the limit of a sum of many very small rectangles
- In other words, “the sum of a little bit of x”, like the text said.
- $\int f(x)dx = \lim_{\Delta x \to 0} \sum_{i}^{n} f(x_i^*)\Delta x_i$
- Plain English: As we make the rectangles narrower and narrower, let's sum up the value of the function ($y = f(x)$) and multiply it by the width of the rectangle (δx) to get the combined area.

Definite and Indefinite Integrals

- There are two types of integrals: **definite** integrals (which measure the area under the curve in a fixed range of x) and **indefinite** integrals (which measure the area under the curve across every value of x).
- Indefinite integral: $\int f(x)dx = F(x)$
- Definite integral: $\int_a^b f(x)dx = F(b) - F(a)$
- Definite integrals are called the **fundamental theorem of calculus**.
- Typically indefinite integrals are represented as functions, not solved, while definite integrals are represented as numerical solutions.

Definite and indefinite integrals, right triangle

- Consider our right triangle example, $f(x) = x$. You can draw this function to show that it will produce a line with slope 1, which forms a right triangle with respect to the x-axis, as we saw in the images above.
- What's the area under the function? Recall that base and height are both equal to x in this example.
  $\int xdx = F(x) = \frac{bh}{2} = \frac{x^2}{2}$.
- What's the area under the function from $x = 0$ to $x = 2$?
  $F(2) = \frac{2*2}{2} = 2$
- What's the area under the function from $x = 1$ to $x = 2$? Uh oh! The area here isn't just a triangle! Take definite integral!
  $$\int_1^2 xdx = F(2) - F(1) = \frac{2 \times 2}{2} - \frac{1 \times 1}{2} = 1.5$$

Relationship between integrals and derivatives

- Integral and derivative are ALMOST inverse operations:
- $\underbrace{F(x)}_{\text{integral}} \xrightarrow{\text{derivative}} \underbrace{f(x)}_{\text{function}} \xrightarrow{\text{derivative}} \underbrace{f'(x)}_{\text{derivative}}$
- $\underbrace{f'(x)}_{\text{derivative}} \xrightarrow{\text{integrate}} \underbrace{f(x)}_{\text{function}} \xrightarrow{\text{integrate}} \underbrace{F(x)}_{\text{integral}}$
- Why? Integral is area under the curve (AUC): made by adding up little bits of the curve itself.
- Rate of change of the AUC is, thus, equal to the value of the little bits – the original function!
- Rate of change of the original function is the original function's slope!

Constant of integration

- Why ALMOST? The **constant of integration**.
- Consider the function
$$f(x) = \frac{x^2}{2} + 5$$
- Its derivative is $f'(x) = x\,dx$. Why? Derivative of constant term is 0, and so the $+5$ term drops away.
- Now, integrate this function:
$$\int f'(x) = \int xdx = \frac{x^2}{2}$$
- What happened to the $+5$ term? Why didn't it come back?

Constant of integration

- Why ALMOST? The **constant of integration**.
- Consider the function
$$f(x) = \frac{x^2}{2} + 5$$
- Its derivative is $f'(x) = x\,dx$. Why? Derivative of constant term is 0, and so the $+5$ term drops away.
- Now, integrate this function:
$$\int f'(x) = \int xdx = \frac{x^2}{2}$$
- What happened to the $+5$ term? Why didn't it come back?
- For all indefinite integrals, must add a term $+C$ to represent the unknown and unknowable constant term in order to ensure integration and derivatives are inverse operators.
- For definite integrals, we can ignore this, because that constant term is the same at $F(a)$ and $F(b)$, so $F(b) - F(a)$ will remote the constant anyway.

Computing Integrals

Power rule 1	$\int x^n dx = \frac{x^{n+1}}{n+1} + C$ if $n \neq -1$
Power rule 2	$\int x^{-1} dx = ln\|x\| + C$
Exponential rule 1	$\int e^x dx = e^x + C$
Exponential rule 2	$\int a^x dx = \frac{a^x}{ln(a)} + C$
Logarithm rule 1	$\int ln(x) dx = xln(x) - x + C$

Useful to remember first power rule and first exponential rule, but generally OK to look these up as needed!

Computing Integrals

Definite integrals 1	$\int_a^b f(x)dx = -\int_b^a f(x)dx$
Empty definite integral	$\int_a^a f(x)dx = 0$
Piecewise definite int.	$\int_a^b f(x)dx = \int_a^c f(x)dx + \int_c^b f(x)dx$
Linear rule	$\int (af(x) + bg(x))dx = a\int f(x)dx + b\int g(x)dx$
Integration by parts	$\int f(x)g'(x)dx = f(x)g(x) - \int f'(x)g(x)dx$

u-substitution

One powerful integration rule is **u-substitution**, which is the integration inverse of the derivative chain rule. Reminder:

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

The **u-substitution** rule is:

If $\int f(x)dx$ cannot be solved, choose u such that $\int f(x)dx = \int udu$, solve, and substitute x for u

u-substitution, huh?

Solve:

$$\int (x^4 + 3)^3 (x^3) dx$$

u-substitution:

$$\text{Let } u = x^4 + 3 \text{ then } \frac{du}{dx} = 4x^3 \text{ and } du = 4x^3 dx$$

So:

$$\begin{aligned}
x^3 dx &= \frac{1}{4} du \\
\int (x^4 + 3)^3 (x^3) dx &= \int u^3 (\frac{1}{4} du) \\
&= \frac{1}{4} \int u^3 du \\
&= \frac{1}{4} (\frac{u^4}{4} + C) = \frac{1}{16} (u^4 + C) \\
&= \frac{1}{16} ((x^4 + 3)^4 + C)
\end{aligned}$$

Conclusion

Tomorrow

- Scalars, Vectors, and Matrices, oh my!