Differentiation

Ye Wang
University of North Carolina at Chapel Hill

Mathematics and Statistics For Political Research
POLI783

A brief history of calculus

- Scholars since Archimedes have been interested in how to calculate the tangent line to a curve or the area under it.
- By the 17th century, these geometric questions became urgent practical problems in science and engineering.
- For example: How do I calculate the speed of a cannonball at a particular point along its path?
- How can I determine the distance a falling object has traveled over time?
- How do I compute the volume of irregular shapes, such as the hull of a ship?
- How fast is a planet moving at a specific point in its orbit?

A brief history of calculus

- The first breakthrough came from Isaac Newton.
- He was an undergraduate at Cambridge in 1665.
- Then, a plague hit, and Cambridge was under lock-down.
- Over the next two years, working in isolation, Newton made a series of discoveries:
- Calculus, laws of motion, law of universal gravitation, dispersion of light, etc.
- After the plague, the world was never the same.
- "Nature and Nature's laws lay hid in night: God said, Let Newton be! and all was light." —Alexander Pope.

A brief history of calculus

- A few years later, a German diplomat, Gottfried Wilhelm Leibniz, independently developed calculus.
- He corresponded with Newton and challenged him with a series of mathematical questions.
- Today, both Newton and Leibniz are recognized as co-founders of calculus.
- The most fundamental result in calculus is known as the Newton–Leibniz formula

Derivative

- For any function $f(x)$, we often want to what will happen to its value if x changes slightly at x_0?
- How much more a worker would earn with one more year of education?
- If a citizens lives one kilometer closer to a wind turbine, how would they vote in the next election?
- Let's define $\Delta x = x - x_0$, then what are interested in is

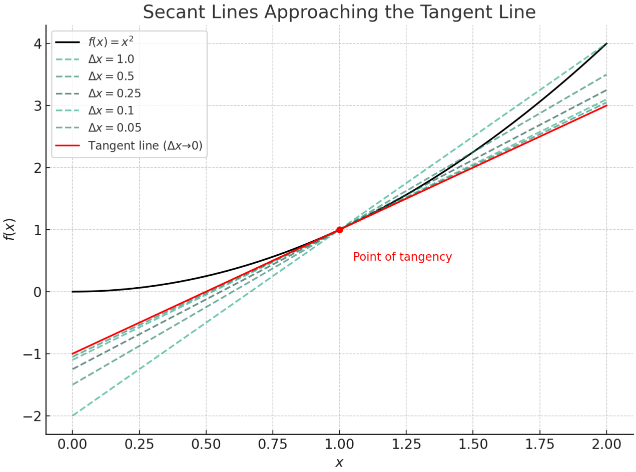
$$\frac{f(x) - f(x_0)}{x - x_0} = \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x}.$$

- This gives the slope of the secant at x_0.
- As Δx approches zero, the secant line becomes the tagent line, and the ratio in differences becomes the derivative:

$$\frac{df(x)}{dx}|_{x=x_0} = \lim_{\Delta x \to 0} \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x}.$$

Derivative

- We can visualize this process.



Derivative

- $f(x)$'s derivative at x_0 is a fixed value rather than a function.
- It captures the effect of an infinitesimal change in x at x_0 on $f(x)$.
- $\frac{df(x)}{dx}|_{x=x_0}$ is not well-defined if $f(x)$ is not continuous at x_0.
- But continuity at x_0 does not imply the existence of $\frac{df(x)}{dx}|_{x=x_0}$.
- There are functions that are continuous everywhere but have no derivative anywhere.
- If $\frac{df(x)}{dx}|_{x=x_0}$ is well-defined everywhere on $f(x)$'s domain, $\frac{df(x)}{dx}|_{x}$ can be seen as a function.
- We call it the derivative function of $f(x)$ and denote it as $f'(x)$.
- We can write $\frac{df(x)}{dx}|_{x=x_0}$ as $f'(x_0)$.
- $dy = df(x) = f'(x)dx$ is known as $f(x)$'s differential.

Differentiation

- The process of finding a function's derivative function is called differentiation.
- As a derivation is a limit, we can calculate it by definition.
- E.g., what is the derivative of x^2?

$$\lim_{\Delta x \to 0} \frac{(x_0 + \Delta x)^2 - (x_0)^2}{\Delta x} = \lim_{\Delta x \to 0} \frac{2x_0 \Delta x + (\Delta x)^2}{\Delta x}$$
$$= \lim_{\Delta x \to 0} (2x_0 + \Delta x) = 2x_0.$$

- In general, $(x^k)' = kx^{k-1}$.
- Some commonly used results: $(c)' = 0$, $(e^x)' = e^x$, $(\sin(x))' = \cos(x)$, , $(\cos(x))' = -\sin(x)$.

Rules of differentiation

- But we don't want to compute limits all the time.
- Some rules can be quite helpful.
  1. $(af(x) + bg(x))' = af'(x) + bg'(x)$.
  2. $(f(x)g(x))' = f'(x)g(x) + f(x)g'(x)$.
  3. $(f \circ g(x))' = f'(g(x))g'(x)$ (the chain rule).
  4. $(f^{-1}(y))' = \frac{1}{f'(x)}$.
- We implicitly assume that both $f(x)$ and $g(x)$ are differentiable everywhere.
- Each of these rules generalizes to multiple functions.

Rules of differentiation

- Example 1:
  $\left(\frac{1}{g(x)}\right)' = ((g(x))^{-1})' = -(g(x))^{-2}g'(x) = -\frac{g'(x)}{(g(x))^2}$.
- Example 2: $\left(\frac{f(x)}{g(x)}\right)' = \frac{f'(x)}{g(x)} - \frac{f(x)g'(x)}{(g(x))^2} = \frac{f'(x)g(x)-f(x)g'(x)}{(g(x))^2}$.
- Example 3: $(\ln(x))' = \frac{1}{e^y} = \frac{1}{x}$.
- Example 4:
  $(a^x)' = \left(e^{x\ln(a)}\right)' = e^{x\ln(a)}(x\ln a)' = e^{x\ln(a)}\ln a = a^x \ln a$.
- Exercise: find the derivative function for $e^{\sqrt{1+\cos x}}$.

Higher-order derivatives

- We can similarly define the derivative of a derivative function:

$$f''(x) = (f'(x))' = \frac{d^2y}{dx^2}.$$

- Higher-order derivatives can be similarly defined (if they exist):

$$f^{(k)}(x) = \frac{d^ky}{dx^k}.$$

- Derivatives of $f(x)$ depict its properties at each point.
- If $f'(x_0) > 0$, we expect $f(x) > f(x_0)$ when x is slightly larger than x_0, and vice versa.
- If $f'(x) > 0$ for every x in its domain, $f(x)$ is monotonically increasing:

$$\forall x' > x, f(x') > f(x),$$

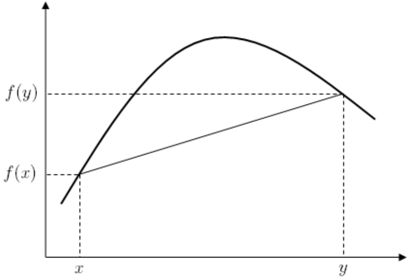
  and vice versa.

Higher-order derivatives

- The second-order derivative captures the concavity of $f(x)$.
- If $f''(x_0) < 0$, then the change of $f(x)$ becomes less drastic as we move from x_0 to $x > x_0$.
- $f'(x)$ depicts the speed and $f''(x)$ depicts acceleration.
- If If $f''(x) < 0$ for every x in its domain, $f(x)$ is a concave function:

$$\forall x' > x, f\left(\lambda x' + (1-\lambda)x\right) > \lambda f(x') + (1-\lambda)f(x),$$

for any $\lambda > 0$.



Taylor expansion

- In general, we can use a polynomial to approximate any smooth function $f(x)$:

$$
\begin{aligned}
f(x) =& f(x_0) + f'(x_0)(x - x_0) + f''(x_0)\frac{(x - x_0)^2}{2} + \dots \\
=& \sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k = \sum_{k=0}^{K} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k + R_K(x),
\end{aligned}
$$

  where $R_K(x)$ include terms whose order is higher than K.
- Some examples:

$$
\begin{aligned}
e^x &= \sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots \\
\sin x &= \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots
\end{aligned}
$$

First-order approximation

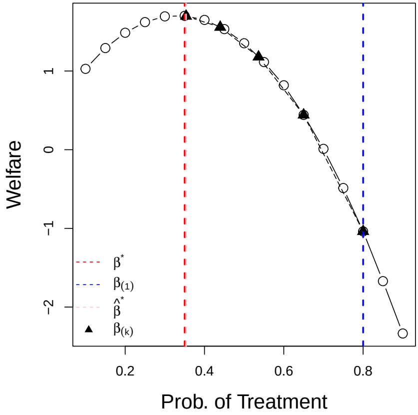
- We can approximate any smooth function with a desired degree of precision.
- The most useful one in practice is the first-order approximation.
- We can approximate $f(x)$'s value at any x using a linear function:

$$f(x) \approx f(x_0) + f'(x_0)(x - x_0).$$

- We can see that $f'(x_0)$ captures the first-order effect of x on $f(x)$ at x_0.
- This is the logical foundation of using linear models in social science research.
- E.g., linear regression, gradient descent, etc.
- Do we need more complex models in social science?

First-order approximation

- Our goal here is to find the policy that maximizes the social welfare.
- We start from a random guess and update the policy gradually along the direction of $f'(x)$.



First-order approximation

- A very useful approximation in practice: $\ln(1+x) \approx x$.
- In regression analysis, it is common to transform the outcome from y to $\ln y$ or $\ln(y+1)$.
- Then, for a small change ΔY, we have

$$\ln(Y+\Delta Y) - \ln(Y) = \ln\left(\frac{Y+\Delta Y}{Y}\right)$$
$$= \ln\left(1+\frac{\Delta Y}{Y}\right) \approx \frac{\Delta Y}{Y}.$$

- The change in $\ln(Y)$ is approximately the percentage change in Y.
- It can be understood as the growth rate of Y and is known as "elasticity" in economics.
- We can similarly show that $e^x - 1 \approx x$ and $\sin x \approx x$.

Optimization

- A function's first- and second-order derivative allows us to depict its shape.
- Based on its shape, we can easily find its extrema (maxima and minima).
- Note that extrema may not exist for some functions (e.g., $y = x$ on $\mathbb{R}$).
- If a function $f(x)$ is continuous on a closed interval $[a, b]$, then there exists points $x_{\text{min}}, x_{\text{max}} \in [a, b]$, such that

$$f(x_{\text{min}}) \leq f(x) \leq f(x_{\text{max}})$$

  for any $x \in [a, b]$ (Extreme Value Theorem).
- We call x_{min} and x_{max} point of minimum and maximum, respectively.

Optimization

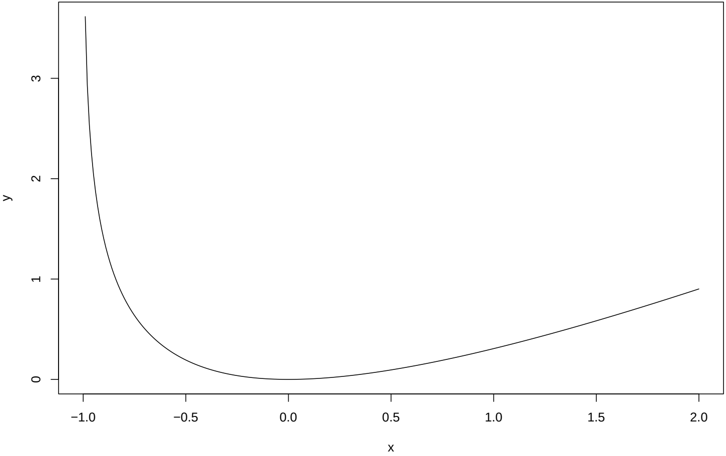
- How do we find x_{min} and x_{max}?
- Step 0: determined the domain and range of $f(x)$.
- Step 1: find all the stationary points x^* where $f'(x^*) = 0$.
- Step 2: calculate $f''(x^*)$ for each stationary point x^*.
- If $f''(x^*) < 0 (> 0)$, then $f(x)$ has a local maximum (minimum) at x^*.
- If $f''(x^*) = 0$, we need to check higher-order derivatives.
- Compare values of local minima, local maxima, and $f(x)$'s values at the boundary and discontinuous points.
- Visualization helps!

Optimization

- Consider the function $y = x - \ln(1 + x)$.
- The function is well-defined when $x > -1$.
- To understand its behavior, let's calculate its first- and second-order derivatives.
- $f'(x) = 1 - \frac{1}{1+x}$, $f''(x) = \frac{1}{(1+x)^2}$.
- We can see that $f''(x) > 0$, thus $f(x)$ is a convex function.
- $f'(x) = 0$ has a unique solution $x = 0$, and $f(0) = 0$ is a local minimum.
- $f'(x)$ is increasing when $x > 0$ and decreasing when $x < 0$.
- $f(x) \to \infty$ when $x \to -1$ or $x \to \infty$.
- Therefore, $f(x)$ has a global minimum 0 and no local or global maximum.

Optimization

- We draw x from $[-1, 2]$ and generate y:



An investment decision

- A dictator controls 50 resource units at time $t = 0$.
- He must choose how much to consume now, c_0, and how much to invest, $I = 50 - c_0$.
- Investment produces future consumption via a nonlinear return function: $c_1 = \sqrt{4I}$.
- The dictator's goal is to maximize total consumption across two periods:

$$U(c_0) = c_0 + \sqrt{4(50 - c_0)}.$$

- What will be his optimal choice of c_0?

An investment decision

- First note that there are constraints on c_0 implied by the setting.
- $0 \leq c_0, c_1 \leq 50$ and $50 - c_0 \geq 0$.
- Let's calculate the first-order condition:

$$U'(c_0) = 1 - \frac{2}{\sqrt{4(50 - c_0)}} = 0.$$

- We can solve $c_0^* = 49$.
- Is this the final answer?
- We need to check the second-order derivative:

$$U'(c_0^*) = -\frac{4}{(4(50 - c_0^*))^{3/2}} < 0.$$

- Finally, ensure that $0 \leq c_0^* \leq 50$.

Differentiation with multivariate functions

- For a multivariate function $y = f(x_1, x_2, \ldots, x_P) = f(\mathbf{x})$, we can define derivatives for each of the variables:

$$\frac{\partial f}{\partial x_p}|_{\mathbf{x}=\mathbf{x}_0} = \lim_{\Delta x \to 0} \frac{f(x_{10}, \ldots, x_{p0} + \Delta x, \ldots, x_{P0}) - f(x_{10}, \ldots, x_{P0})}{\Delta x}.$$

- This is called the partial derivative of $f(\cdot)$.
- The vector $\nabla f(\mathbf{x}) = \left(\frac{\partial f}{\partial x_1}|_{\mathbf{x}=\mathbf{x}_0}, \frac{\partial f}{\partial x_2}|_{\mathbf{x}=\mathbf{x}_0}, \ldots, \frac{\partial f}{\partial x_P}|_{\mathbf{x}=\mathbf{x}_0}\right)'$ is known as the gradient of $f(\mathbf{x})$ at $f(\mathbf{x}_0)$.
- The first-order condition for $f(\mathbf{x})$ is $\nabla f(\mathbf{x}) = \mathbf{0}$.

Taylor expansion for multivariate functions

- We can generalize Taylor expansion to multivariate functions.
- Consider a simple case where $y = f(x_1, x_2)$.
- Its first-order approximation at (x_{10}, x_{20}) equals

$$f(x_1, x_2) \approx f(x_{10}, x_{20}) + \frac{\partial f(x_{10}, x_{20})}{\partial x_1}(x_1 - x_{10}) + \frac{\partial f(x_{10}, x_{20})}{\partial x_2}(x_2 - x_{20}).$$

- The second-order Taylor expansion involves the second-order partial derivatives, $\frac{\partial f^2(x_{10}, x_{20})}{\partial x_1^2}$ and $\frac{\partial f^2(x_{10}, x_{20})}{\partial x_2^2}$, as well as the cross partial derivative $\frac{\partial f^2(x_{10}, x_{20})}{\partial x_1 \partial x_2}$.

Profit maximization

- Consider the Cobb-Douglas production function we saw before:

$$y = AK^aL^b, a + b < 1,$$

where y, K, and L represent the product, capital, and labor, respectively.

- The company aims to maximize its profit, the benefit from selling the product minus the cost of input:

$$\max_{K,L}(py - p_K K - p_L L).$$

- The first-order conditions are

$$paAK^{a-1}L^b = p_K,$$
$$pbAK^aL^{b-1} = p_L.$$

- Dividing condition 1 by condition 2, we have $\frac{aL}{bK} = \frac{p_K}{p_L}$.