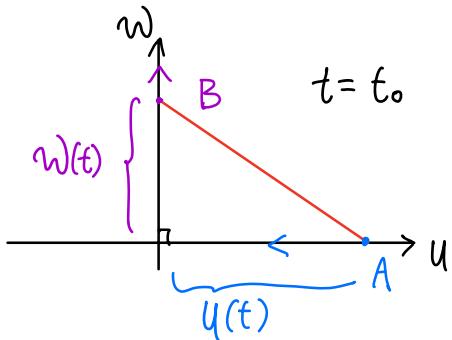


Lecture 6, Thursday, September 21/2023

Outline

- Chain rule (3.6)
- Implicit differentiation (3.7)
- Linearization and differentials (3.9)

Chain Rule (Differentiation Rule for Composite Functions)



Suppose that :

- Car A is heading west along the u -axis
- Car B is heading north along the w -axis

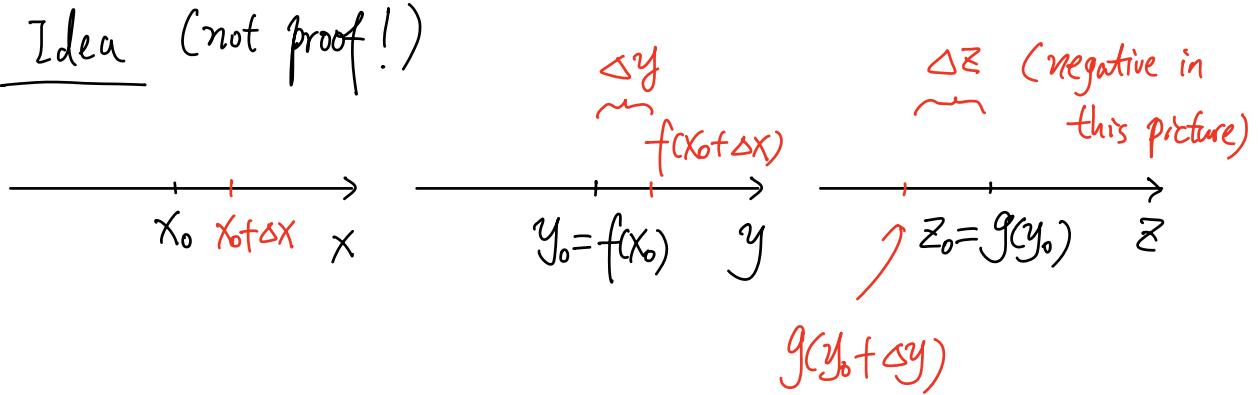
Q : How fast is the distance $D = D(A, B)$ between A and B changing when $t = t_0$?

A :

More generally, suppose that three quantities x , y and z satisfy $y = f(x)$ and $z = g(y)$, where f and g are differentiable.

Q : $\frac{dz}{dx} = ?$, i.e., $(g \circ f)'(x) = ?$

Idea (not proof!)



- Fix $x = x_0$.
- Consider a very "small" change in x -value, say Δx .
- This creates a very "small" change Δy and Δz .
- When Δx is small,

$$(g \circ f)'(x_0) = \frac{dz}{dx} \Big|_{x=x_0} \approx \frac{\Delta z}{\Delta x} = \frac{\Delta z}{\Delta y} \frac{\Delta y}{\Delta x} \quad (\text{if } \Delta y \neq 0).$$

- Since $\frac{\Delta z}{\Delta y} \approx \frac{dz}{dy} \Big|_{y=y_0} = g'(y_0) = g'(f(x_0))$ and
 $\frac{\Delta y}{\Delta x} \approx \frac{dy}{dx} \Big|_{x=x_0} = f'(x_0)$, it suggests the follow rule.
(but did not prove)

Chain Rule If $y = f(x)$ is differentiable at $x = x_0$, and $z = g(y)$ is differentiable at $y = y_0 = f(x_0)$, then $g \circ f$ is differentiable at $x = x_0$, and

$$(g \circ f)'(x_0) = g'(f(x_0)) f'(x_0).$$

Proof: After we discuss linearization (3.9).

Remark

- Another way to write the chain rule is

$$\frac{dz}{dx} \Big|_{x=x_0} = \left(\frac{ds}{dy} \Big|_{y=f(x_0)} \right) \left(\frac{dy}{dx} \Big|_{x=x_0} \right).$$

- An even shorter notation:

$$\frac{dz}{dx} = \frac{dz}{dy} \frac{dy}{dx}.$$

e.g.1 Position of an object at time t is $s = \cos(t^2 + 1)$.

Find its velocity at time t .

“Outside”: \cos .

“Inside”: $t^2 + 1$.

Sol: $v(t) = \frac{ds}{dt} = s'(t) = \cos'(t^2 + 1) \cdot (t^2 + 1)'$
 $= -\sin(t^2 + 1) \cdot 2t.$

e.g.2 Find $f'(x)$, where $f(x) = \left(\frac{x-2}{\cos(x^2)+2} \right)^9$.

Can also use the following notation.

Sol. Let $z = y^9$, $y = \frac{x-2}{\cos(x^2)+2}$. Then

$$f'(x) = \frac{dz}{dx} = \frac{dz}{dy} \cdot \frac{dy}{dx} = qy^8 \cdot \frac{(x-2)(\cos(x^2)+2)' - (\cos(x^2)+2)(x-2)'}{(\cos(x^2)+2)^2}$$

$$\begin{aligned}
 &= qy^8 \cdot \frac{(\cos(x^2)+z) \cdot 1 - (x-z)(-\sin(x^2) \cdot 2x)}{(\cos(x^2)+z)^2} \\
 &= q \left(\frac{x-z}{\cos(x^2)+z} \right)^8 \cdot \frac{\cos(x^2)+z + 2x(x-z)\sin(x^2)}{(\cos(x^2)+z)^2}.
 \end{aligned}$$

Implicit Differentiation

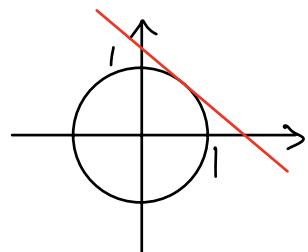
Q: Find the slope of the tangent line to the curve $x^2+y^2=1$ at $(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$.

A: • Near the point $(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$, curve

is $y = \sqrt{1-x^2}$ (since $y>0$).

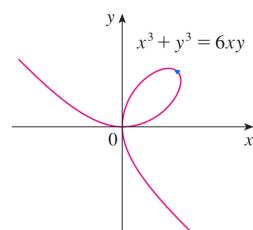
$\underbrace{y=f(x)}$, explicit function

$$\text{• Slope} = \left. \frac{df}{dx} \right|_{x=\frac{1}{\sqrt{2}}} = \left. \frac{1}{2} \frac{-2x}{\sqrt{1-x^2}} \right|_{x=\frac{1}{\sqrt{2}}} = -1.$$



Q: Find the slope of the tangent line to the curve $x^3+y^3=6xy$ at the point $(3, 3)$.

• It is not convenient to write $y=f(x)$ explicitly (as in the previous case).



- But from the curve, it can be seen that y is a function of x near $(3,3)$, so $\underbrace{y = h(x)}$, with $h(x)$ unclear from the context.
implicit function.

- x and y are related by $\underbrace{x^3 + y^3 - 6xy = 0}_{F(x,y)} \leftarrow (*)$
- Treating y as a function of x , apply $\frac{d}{dx}$ to $(*)$:

(This is O.K. for the point $(3,3)$ on the curve.)

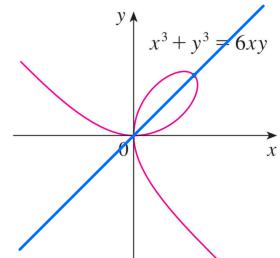
$$\begin{aligned} 3x^2 + 3y^2 \cdot \frac{dy}{dx} - 6\left(x \frac{dy}{dx} + 1 \cdot y\right) &= 0 \\ \Rightarrow (3y^2 - 6x) \frac{dy}{dx} &= 6y - 3x^2 \\ \Rightarrow \frac{dy}{dx} &= \frac{2y - x^2}{y^2 - 2x} \\ \Rightarrow \text{slope at } (3,3) &= \frac{dy}{dx} \Big|_{(x,y)=(3,3)} = \frac{6-9}{9-6} = -1 \end{aligned}$$

Note: In order for the approach above to work, y has to be a function of x at the required point "locally" (i.e. in a small scale), and it cannot have a vertical tangent. e.g., it does not work at $(x,y) = (0,0)$, since no matter how close you "zoom" in at $(0,0)$, the curve is not the graph of a function $y = h(x)$.

Def The normal line to a curve at (x_0, y_0) is the line perpendicular to the tangent line to the curve at (x_0, y_0) \downarrow " \perp "

Remark: For $y = h(x)$, the slope of normal line at (x_0, y_0) is $-\frac{1}{h'(x_0)}$, provided that $h'(x_0)$ exists and $\neq 0$.

e.g. For $x^3 + y^3 - 6xy = 0$, at point $(3, 3)$,
slope of normal line is $-\frac{1}{(-1)} = 1$.



The strategy for finding higher-order derivatives for an implicit function is similar.

e.g. 3 Given the curve $2x^3 - 3y^2 = 8$, find $\frac{d^2}{dx^2}y$.

Sol. $\frac{d}{dx}(2x^3 - 3y^2) = \frac{d}{dx}8$

$$= \frac{2x^2 - 6y \cdot 2y'}{y^3} \quad \text{for } y \neq 0.$$

(i.e., except $(\sqrt[3]{4}, 0)$)

Linearization

Which could be complicated

Given a general function $f(x)$ and a fixed x_0 , it may be computationally expensive to compute $f(x_0)$, e.g., $\sqrt[3]{27.135} = ?$.

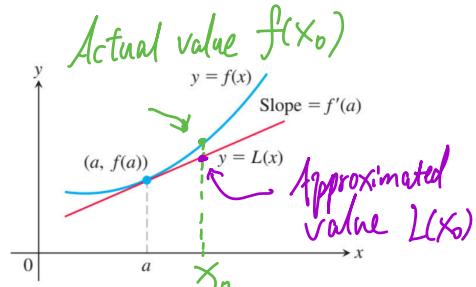
Q: What could be a way to approximate $f(x_0)$ while costing a lot less computational resources?

A: Use tangent lines!

(Linear functions are among the simplest functions!)

- If f is differentiable at a and x_0 is "near a ", then

$$\frac{f(x_0) - f(a)}{x_0 - a} \approx f'(a)$$



$$\Rightarrow f(x_0) \approx f(a) + f'(a)(x_0 - a) \quad L(x_0) = L_a(x_0)$$

DEFINITIONS If f is differentiable at $x = a$, then the approximating function

$$L(x) = f(a) + f'(a)(x - a) \quad \leftarrow \text{Can also write } L_a(x)$$

is the **linearization** of f at a . The approximation

$$f(x) \approx L(x)$$

of f by L is the **standard linear approximation** of f at a . The point $x = a$ is the **center** of the approximation.

e.g.4 Approximate the numerical value of $\sqrt[3]{27.4}$ using standard linear approximation with center 27.

$$\text{Sol: } f(x) = x^{\frac{1}{3}}, \quad f'(x) = \frac{1}{3}x^{-\frac{2}{3}} = \frac{1}{3x^{\frac{2}{3}}}$$

$$\Rightarrow f'(27) = \frac{1}{3 \cdot 9} = \frac{1}{27}$$

$$\Rightarrow L_{27}(x) = f(27) + f'(27)(x-27) = 3 + \frac{1}{27}(x-27)$$

$$= \underbrace{\frac{1}{27}x}_{\text{More friendly than the function } f(x) = \sqrt[3]{x}} + 2.$$

$$\sqrt[3]{27.4} \approx \frac{1}{27}(27.4) + 2 = 3.0148148\dots$$

For reference, $\sqrt[3]{27.4} = 3.01474225\dots$, error < 0.0001 .

e.g.5 Approximate $(1+\varepsilon)^k$ for small ε , ← Could be negative.
using standard linear approximation.

Sol: Set $f(x) = x^k$, consider $L(x)$, with center $x=1$.

$$f'(x) = kx^{k-1} \Rightarrow f'(1) = k$$

$$\Rightarrow L(x) = L_1(x) = f(1) + f'(1)(x-1) = 1 + k(x-1).$$

For $x=1+\varepsilon$,

$$(1+\varepsilon)^k = f(1+\varepsilon) \approx L(1+\varepsilon) = 1 + k\varepsilon.$$

In particular,

$$\sqrt{1.01} = 1.01^{\frac{1}{2}} \approx 1 + \frac{1}{2}(0.01) \approx 1.005$$

(Exact: 1.00498756...)

$$\sqrt[3]{1+5x^4} = \underbrace{(1+5x^4)^{\frac{1}{3}}}_{\varepsilon} \approx 1 + \frac{1}{3}(5x^4) = 1 + \frac{5}{3}x^4,$$

when x is "small".

Differentials

Up to this point, Leibniz's notation $\frac{dy}{dx}$ is just a fixed symbol — it is not technically a ratio.

Here, we define dy and dx separately in a way that is consistent with Leibniz's notation.

DEFINITION Let $y = f(x)$ be a differentiable function. The **differential** dx is an independent variable. The **differential** dy is

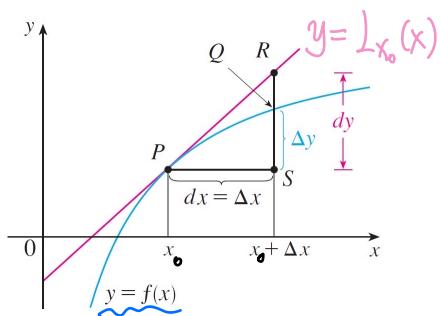
$$dy = f'(x) dx.$$

- dx and dy are both called **differentials**.
- dx is **independent**.
- dy , or technically $dy(x, dx)$, depends on both x and dx .
- For $dx \neq 0$, we have $f'(x) = \frac{dy}{dx}$, consistent with previous notation.

E.g. 6 For $y = f(x) = x^5 + 37x$, when $x=1$ and $dx=0.2$,

$$dy = f'(x) dx = (5x^4 + 37)|_{x=1} (0.2) = 8.4.$$

Geometric meaning



1. $dx = \Delta x$ = independent variable
= change in x -value from $x=x_0$;
2. Δy = Exact change in y -value (for f);
3. dy = Approximated change in y -value (using L_{x_0});
4. Both Δy and dy depend on both x_0 and Δx .

Remarks

1. Now we can treat $\frac{dy}{dx}$ as fractions:

e.g. $y = f(x) \Rightarrow \frac{dy}{dx} = f'(x) \Rightarrow dy = f'(x) dx$

2. Differentials can be used to approximate changes:

e.g. If the radius of a circle increases from 10m to 10.022m, then its change of area can be approximated by

$$dA = A'(10) dr = 20\pi(0.022) = 0.44\pi (m^2).$$

3. If $y = f(x)$, we can write df to mean dy .

Error of Standard Linear Approximation

Q: Error := $E(x_0, \Delta x) := f(x_0 + \Delta x) - L(x_0 + \Delta x) = ?$

$$\begin{aligned}
 f(x_0 + \Delta x) - L(x_0 + \Delta x) &= f(x_0 + \Delta x) - f(x_0) - f'(x_0)\Delta x \\
 &= \left[\frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x} - f'(x_0) \right] \Delta x \\
 &\quad \underbrace{\varepsilon := \frac{\Delta y}{\Delta x} - \frac{dy}{dx} \Big|_{x=x_0}}_{\text{secant vs tangent.}}
 \end{aligned}$$

$$\begin{aligned}
 \text{Note that } \lim_{\Delta x \rightarrow 0} \varepsilon &= \lim_{\Delta x \rightarrow 0} \left[\frac{\Delta y}{\Delta x} - f'(x_0) \right] = \left(\lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x} \right) - f'(x_0) \\
 &= f'(x_0) - f'(x_0) = 0.
 \end{aligned}$$

This means that

$$\lim_{\Delta x \rightarrow 0} \frac{f(x_0 + \Delta x) - L(x_0 + \Delta x)}{\Delta x} = \lim_{\Delta x \rightarrow 0} \varepsilon = 0 ,$$

So the error $E(x_0, \Delta x) := f(x_0 + \Delta x) - L(x_0 + \Delta x)$ does not just go to zero, but it does so faster than Δx ! (Since $\frac{E(x_0, \Delta x)}{\Delta x} \rightarrow 0$ as $\Delta x \rightarrow 0$.)

Let us make a note that

$$E(x_0, \Delta x) := f(x_0 + \Delta x) - L(x_0 + \Delta x) = \varepsilon \Delta x .$$

Since

$$\underbrace{f(x_0 + \Delta x) - L(x_0 + \Delta x)}_{\varepsilon \Delta x} = f(x_0 + \Delta x) - f(x_0) - f'(x_0) \Delta x \\ = \Delta y - f'(x_0) \Delta x, \quad \begin{matrix} \text{this is} \\ \Delta y - dy \end{matrix}$$

We also have the following formula for the exact change Δy .

Fix $x = x_0$. If f is differentiable at x_0 , then

$$\Delta y = f'(x_0) \Delta x + \varepsilon \Delta x$$

for some ε with $\varepsilon \rightarrow 0$ as $\Delta x \rightarrow 0$.