Introduction to Mathematics and Modeling

lecture 5
The chain rule and optimization

UNIVERSITY OF TWENTE.

academic year : 18-19
lecture : 5
build : December 10, 2018
slides : 31
This week

1. Section 3.6: the chain rule
2. Section 3.8: derivatives of logarithms (only pages 176–181)
3. Section 4.1: extreme values
The **composition of** \( f \) and \( g \) is the function that maps \( x \) to \( f(g(x)) \).

The composition is denoted as \( f \circ g \), and is pronounced as “\( f \) after \( g \).”

**Example:** let \( f(x) = x^2 \) and let \( g(x) = x + 1 \), then

\[
\begin{align*}
    f \circ g(x) &= f(g(x)) = f(x + 1) = (x + 1)^2 = x^2 + 2x + 1 \\
    g \circ f(x) &= g(f(x)) = g(x^2) = x^2 + 1.
\end{align*}
\]

⚠️ In general \( f \circ g \) and \( g \circ f \) are **not** identical.
Let \( f(x) = ax + b \) and \( g(x) = \sin(x) \) and define \( h = f \circ g \), then

\[
h(x) = f \circ g(x) = f(g(x)) = a \sin(x) + b
\]

Using the sum rule and constant multiple rule we know that

\[
h'(x) = a \cos(x)
\]

Now let \( h = g \circ f \) then

\[
h(x) = g(f(x)) = \sin(ax + b)
\]

The sum- and constant multiple rule cannot be applied
Consider the special case \( h(x) = \sin(3x) \). The graph of \( h \) is obtained by scaling \( \sin x \) in horizontal direction.

- The slopes of all tangents are scaled too!

- By scaling back \( \sin(3x) \) in vertical direction, this effect is cancelled out:

\[
\frac{d}{dx} \left( \frac{1}{3} \sin(3x) \right) = \cos(3x),
\]

in other words:

\[
\frac{d}{dx} \sin(3x) = 3 \cos(3x).
\]
Composition with a linear function

We see that

\[ f(x) = \sin(ax) \quad \Rightarrow \quad f'(x) = a \cos(ax) \]

By shifting a graph horizontally, the slopes must shift accordingly:

\[ f(x) = \sin(ax + b) \quad \Rightarrow \quad f'(x) = a \cos(ax + b) \]

Chain rule, simple version

Let \( f \) be a differentiable function. Then for any constant \( a \) and \( b \) the following holds:

\[ \frac{d}{dx} \left( f(ax + b) \right) = af'(ax + b). \]

Warning: \( \frac{d}{dx} (f(ax + b)) \) is the derivative of the composition \( f(ax + b) \), but \( f'(ax + b) \) is the composition of \( f' \) and \( y = ax + b \).
Examples

- The derivative of \( \sin(2x) \) is \( 2 \cos(2x) \).

- Define \( y = \sqrt{5 - 3x} \), then

\[
\frac{dy}{dx} = -\frac{3}{2\sqrt{5 - 3x}}
\]

since \( \frac{d}{dx}(\sqrt{x}) = \frac{1}{2\sqrt{x}} \).

Also: write \( 5 - 3x = (-3)x + 5 \), hence \( a = -3 \) and \( b = 5 \).

\[
\frac{d}{dx} \left( \frac{1}{2e^x} \right) =
\]
See lecture 4: if we define \( f(x) = a^x \), then

\[
    f'(x) = k_a a^x
\]

where

\[
    k_a = \lim_{h \to 0} \frac{a^h - 1}{h} = f'(0).
\]

With the simple version of the chain rule we can prove:

\[
    k_a = \ln a
\]
Chain rule

Let $f$ and $g$ be differentiable functions, then

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

- In words: multiply the composition of the derivative of $f$ with $g$ by the derivative of $g$.

- Work inward:
  - differentiate the ‘outer function’ $f$, but keep the ‘inner function’ $g$ intact;
  - then multiply with the derivative of the ‘inner function’ $g$. 

Example

Find the derivative of \( h(x) = (3x^2 + 1)^2 \).

- The function \( h \) is equal to \( h = f \circ g \), where
  \[
  f(x) = x^2 \quad \text{and} \quad g(x) = 3x^2 + 1.
  \]

- Apply the chain rule:
  \[
  h'(x) = \]
Anonymous functions

- If a function is named \( f \), the derivative is denoted as \( f' \).
- If \( y \) is an anonymous function of \( x \), the derivative is denoted as \( \frac{dy}{dx} \).

If \( y \) is a function of \( x \) and \( z \) is a function of \( y \), then \( z \) is (by composition) a function of \( x \). In this case the chain rule is

\[
\frac{dz}{dx} = \frac{dz}{dy} \frac{dy}{dx}.
\]

⚠️ Note that \( \frac{dz}{dy} \) is expressed in terms of \( y \), hence afterwards you should replace all occurrences of \( y \) with the corresponding expression in \( x \).
Example

Let $y = 3x^2 + 1$ and $z = y^2$, find $\frac{dz}{dx}$.

Apply the chain rule (anonymous variant):

$$\frac{dz}{dx} = \frac{dz}{dy} \cdot \frac{dy}{dx}$$
Example

Find the derivative of \( f(x) = \frac{1}{\sqrt{x^2 + 1}} \).

- Avoid using the quotient rule by writing
  \[ f(x) = (x^2 + 1)^{-1/2}. \]

- Apply the chain rule:
  \[ f'(x) = \]
Example

Calculate the derivative of \( f(x) = \sqrt{\frac{1 - x^2}{1 + x^2}} \).

Combine the chain rule with the quotient rule:

\[
 f'(x) = 
\]
■ The **logarithmic function with base** $a$ is the inverse of the base-$a$ exponential function:

$$y = a^x \iff x = \log_a y$$

■ The **natural logarithm** is the logarithm with base $e$:

$$\ln x = \log_e x \quad \text{where} \quad e \approx 2.71828\ldots$$

■ Examples:

$$\log_2 8 = 3 \quad \text{because} \quad 2^3 = 8$$

$$\log_{10} 100 = 2 \quad \text{because} \quad 10^2 = 100$$

$$\ln e\sqrt{e} = \frac{3}{2} \quad \text{because} \quad e^{\frac{3}{2}} = e\sqrt{e}$$
\[ \log_a 1 = 0 \quad \text{and} \quad \log_a a = 1 \]

\[ \log_a xy = \log_a x + \log_a y \]

\[ \log_a \frac{x}{y} = \log_a x - \log_a y \]

\[ \log_a \frac{1}{y} = -\log_a y \]

\[ \log_a x^p = p \log_a x \]

\[ \log_a x = \frac{\log_b x}{\log_b a}, \quad \text{in particular} \quad \log_a x = \frac{\ln x}{\ln a} \]

\[ a^x = b^{x \log_b a}, \quad \text{in particular} \quad a^x = e^{x \ln a} \]
The derivative of $\ln(x)$

- Note that $e^x$ and $\ln(x)$ are each others inverse:

$$e^{\ln(x)} = x.$$

- Now take derivatives on both sides and apply the chain rule to the left-hand side:

$$e^{\ln(x)} \ln'(x) = 1,$$

$$x \ln'(x) = 1,$$

$$\ln'(x) = \frac{1}{x}.$$

⚠️ This holds for $x > 0$. 

\[\text{Graph: } e^x \text{ and } \ln(x)\]
The derivative of \( \log_a(x) \) is

\[
\frac{1}{x \ln(a)}.
\]

From the change-of-base formula for logarithms follows

\[
\log_a(x) = \frac{\ln x}{\ln a}.
\]

Apply the constant-multiple rule:

\[
f'(x) =
\]
Example

Find the derivative of $f(x) = \ln(x^2 + 3)$.

Apply the chain rule:

$$f'(x) =$$
Consider a function $f : D \rightarrow \mathbb{R}$.

- $f$ has an **absolute maximum** in $c$ if
  $$f(x) \leq f(c) \quad \text{for all } x \in D$$
- $f$ has an **absolute minimum** in $c$ if
  $$f(x) \geq f(c) \quad \text{for all } x \in D$$

⚠ Extreme values do not necessarily have to exist!

?? If they exist, how do we find them?
Extreme values of a function are domain dependent

On $D = (-\infty, \infty)$ the function $f(x) = x^2$ has

- an absolute minimum in $x = 0$;
- no absolute maximum.

The range of $f$ is $[0, \infty)$. 
On $D = [0, 2]$ the function $f(x) = x^2$ has

- an absolute minimum in $x = 0$;
- an absolute maximum in $x = 2$.

The range of $f$ is $[0, 4]$. 
On $D = (0, 2]$ the function $f(x) = x^2$ has

- no absolute minimum;
- an absolute maximum in $x = 2$.

The range of $f$ is $(0, 4]$. 
On $D = (0, 2)$ the function $f(x) = x^2$ has

- no absolute minimum;
- no absolute maximum.

The range of $f$ is $(0, 4)$. 
Extreme values of a function are domain dependent

**Extreme Value Theorem**

Theorem 1, page 223

A **continuous function** on a **finite closed interval** attains both an **absolute maximum** and an **absolute minimum**.

⚠️ The theorem tells us that extreme values do exist, but *not* where to find them!
Definition

Consider a function $f : D \rightarrow \mathbb{R}$.

- $f$ has a **local maximum** in $c$ if
  \[ f(x) \leq f(c) \quad \text{for all } x \text{ in an environment of } c \]

- $f$ has an **local minimum** in $c$ if
  \[ f(x) \geq f(c) \quad \text{for all } x \text{ in an environment of } c \]
Critical points

**First Derivative Theorem**

If \( f \) is differentiable at \( c \) and \( f \) attains a local maximum or local minimum at \( c \) then \( f'(c) = 0 \).

**Definition**

The number \( c \) is a **critical point of** \( f \) if

\[
f'(c) = 0 \quad \text{or} \quad f \text{ is not differentiable at } c.
\]

Critical points are *candidates* for a local maximum or minimum.
The function $F$ is defined on the domain $[a, b]$.

- Critical points: $c_1$, $c_2$, $c_3$, $c_4$ and $c_5$.
- Local extremes: $a$, $c_2$, $c_3$, $c_4$ and $b$.
- Absolute extremes: $a$ and $c_4$. 
Recipe for computing the extreme values of a continuous function

\[ f : [a, b] \to \mathbb{R} \]

1. Find all critical points of \( f \) in \([a, b]\), i.e., solve the equation \( f'(x) = 0 \) and retain all solutions \( x \) in \([a, b]\); then add all points where \( f \) is not differentiable.

2. Evaluate \( f \) at the critical points and at the end points \( x = a \) and \( x = b \).

3. Take the largest and smallest values found in step 2: these are the absolute maximum and minimum of \( f \) on the interval \([a, b]\).
Example

Find extremes for \( f(x) = 10x(2 - \ln x) \) on \([1, e^2]\).

1. Find the critical points:

2. Evaluate \( f \) at the critical points and at the endpoints:
\[
\begin{align*}
 f(1) &= \\
 f(e^2) &=
\end{align*}
\]

3. Take the largest and smallest values of step 2:
   - The absolute maximum is 
   - The absolute minimum is
Example

Find extremes for $f(x) = xe^{-x}$ on $[-1, 1]$.

1. Find the critical points:

2. Evaluate $f$ at the critical points and at the endpoints:

$$f(-1) =$$
$$f(1) =$$

3. Take the largest and smallest values of step 2:
   - The absolute maximum is 
   - The absolute minimum is
**Example**

*Find extremes for* \( f(x) = 3x^2 - 2x^3 \) *on* \([-\frac{1}{2}, 2]\).

1. Find the critical points:

2. Evaluate \( f \) at the critical points and at the endpoints:

   \[ f\left(-\frac{1}{2}\right) = \quad f(2) = \]

3. Take the largest and smallest values of step 2:
   - The absolute maximum is __________
   - The absolute minimum is __________