2.5 Chain Rule for Multiple Variables

Prof. Tesler

Math 20C
Winter 2018
Review of the chain for functions of one variable

Chain rule

\[ \frac{d}{dx} f(g(x)) = f'(g(x)) \cdot g'(x) \]

Example

\[ \frac{d}{dx} \sin(x^2) = \cos(x^2) \cdot (2x) = 2x \cos(x^2) \]

This is the derivative of the outside function (evaluated at the inside function), times the derivative of the inside function.
Function composition

Composing functions of one variable

- Let \( f(x) = \sin(x) \) and \( g(x) = x^2 \).

- The *composition* of these is the function \( h = f \circ g \):
  \[
  h(x) = f(g(x)) = \sin(x^2)
  \]

- The notation \( f \circ g \) is read as
  “\( f \) composed with \( g \)”
  or “the composition of \( f \) with \( g \).”

Prof. Tesler

2.5 Chain Rule

Math 20C / Winter 2018
A, B, C are sets. They can have different dimensions, e.g.,

\[ A \subseteq \mathbb{R}^n \quad B \subseteq \mathbb{R}^m \quad C \subseteq \mathbb{R}^p \]

\( f, g, \) and \( h \) are functions. Domains and codomains:

\[ f : B \rightarrow C \]
\[ g : A \rightarrow B \]
\[ h : A \rightarrow C \]
Function composition: Multiple variables

\[ f : \mathbb{R}^2 \to \mathbb{R} \quad \text{with} \quad f(x, y) = x^2 + y^2 \]

\[ \vec{r} : \mathbb{R} \to \mathbb{R}^2 \quad \text{with} \quad \vec{r}(t) = \langle x(t), y(t) \rangle = \langle 2t + 1, 3t - 1 \rangle \]

\[ f \circ \vec{r} : \mathbb{R} \to \mathbb{R} \quad \text{with} \quad (f \circ \vec{r})(t) = f(\vec{r}(t)) = f(2t + 1, 3t - 1) = (2t + 1)^2 + (3t - 1) = 4t^2 + 7t \]

**Derivative of \( f(\vec{r}(t)) \)**

- **Notations:** \( \frac{d}{dt} f(\vec{r}(t)) = \frac{d}{dt} (f \circ \vec{r})(t) = (f \circ \vec{r})'(t) \)
- **Example:** \( (f \circ \vec{r})'(t) = 8t + 7 \)
  \( (f \circ \vec{r})'(10) = 8 \cdot 10 + 7 = 87 \)
A mountain has altitude $z = f(x, y)$ above point $(x, y)$.

Plot a hiking trail $(x(t), y(t))$ on the contour map. This gives altitude $z(t) = f(x(t), y(t))$, and 3D trail $(x(t), y(t), z(t))$.

What is the hiker’s vertical speed, $dz/dt$?
What is \( \frac{dz}{dt} = \) vertical speed of hiker?

Let \( \Delta t = \) very small change in time.

The change in altitude is

\[
\Delta z = z(t + \Delta t) - z(t) \\
\approx f_x(x, y) \Delta x + f_y(x, y) \Delta y 
\]

Using the linear approximation.
What is \( \frac{dz}{dt} = \) vertical speed of hiker?

- Let \( \Delta t = \) very small change in time.
- The change in altitude is
  \[
  \Delta z = z(t + \Delta t) - z(t) 
  \approx f_x(x, y) \Delta x + f_y(x, y) \Delta y
  \]
  Using the linear approximation
- The vertical speed is approximately
  \[
  \frac{\Delta z}{\Delta t} \approx f_x(x, y) \frac{\Delta x}{\Delta t} + f_y(x, y) \frac{\Delta y}{\Delta t}
  \]
- The instantaneous vertical speed is the limit of this as \( \Delta t \to 0 \):
  \[
  \frac{dz}{dt} = f_x(x, y) \frac{dx}{dt} + f_y(x, y) \frac{dy}{dt}
  \]
Chain rule for paths
Our book: “First special case of chain rule”

Let $z = f(x, y)$, where $x$ and $y$ are functions of $t$.
So $z(t) = f(x(t), y(t))$. Then

$$\frac{dz}{dt} = f_x(x, y) \frac{dx}{dt} + f_y(x, y) \frac{dy}{dt}$$

or

$$\frac{dz}{dt} = \frac{\partial f}{\partial x} \frac{dx}{dt} + \frac{\partial f}{\partial y} \frac{dy}{dt}$$

Vector version

- Let $z = f(x, y)$ and $\vec{r}(t) = \langle x(t), y(t) \rangle$.
- $z(t) = f(x(t), y(t))$ becomes $z(t) = f(\vec{r}(t))$.
- The chain rule becomes

$$\frac{d}{dt} f(\vec{r}(t)) \approx \nabla f \cdot \vec{r}'(t)$$

where $\nabla f = \langle \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \rangle$ and $\vec{r}'(t) = \langle \frac{dx}{dt}, \frac{dy}{dt} \rangle$. 
Chain rule example

Let \( z = f(x, y) = x^2 + y \)

where \( x = 2t + 1 \) and \( y = 3t - 1 \)

Compute \( \frac{dz}{dt} \).

**First method: Substitution / Function composition**

- Explicitly compute \( z \) as a function of \( t \).
  Plug \( x \) and \( y \) into \( z \), in terms of \( t \):

  \[
  z = x^2 + y = (2t + 1)^2 + (3t - 1)
  = 4t^2 + 4t + 1 + 3t - 1
  = 4t^2 + 7t
  
  Then compute \( \frac{dz}{dt} \):

  \[
  \frac{dz}{dt} = 8t + 7
  
  Prof. Tesler

2.5 Chain Rule
Math 20C / Winter 2018 10 / 39
Chain rule example

Let \( z = f(x, y) = x^2 + y \)

where \( x = 2t + 1 \) and \( y = 3t - 1 \). Compute \( dz / dt \).

Second method: Chain rule

- Chain rule formula:

\[
\frac{dz}{dt} = \frac{\partial z}{\partial x} \frac{dx}{dt} + \frac{\partial z}{\partial y} \frac{dy}{dt}
\]

\[
= 2x \cdot 2 + 1 \cdot 3 = 4x + 3
\]

- Plug in \( x, y \) in terms of \( t \):

\[
= 4(2t + 1) + 3 = 8t + 4 + 3 = 8t + 7
\]

- This agrees with the first method.
Chain rule example

Let \( z = f(x, y) = x^2 + y \)
where \( x = 2t + 1 \) and \( y = 3t - 1 \). Compute \( dz/dt \).

Vector version

- Convert from components \( x(t), y(t) \) to position vector function \( \vec{r}(t) \).
  \[
  \vec{r}(t) = \langle x(t), y(t) \rangle = \langle 2t + 1, 3t - 1 \rangle
  \]

- Compute the derivative \( dz/dt = (f \circ \vec{r})'(t) \):
  \[
  \frac{dz}{dt} = \nabla f \cdot \vec{r}'(t) = \langle 2x, 1 \rangle \cdot \langle 2, 3 \rangle = 4x + 3 = \cdots = 8t + 7 \text{ as before.}
  \]
Tree diagram of chain rule (not in our book)

\[ z = f(x, y) \] where \( x \) and \( y \) are functions of \( t \), gives \( z = h(t) = f(x(t), y(t)) \)

\[ \frac{\partial z}{\partial x} \quad \frac{\partial z}{\partial y} \]

\[ x \quad \frac{dx}{dt} \quad y \quad \frac{dy}{dt} \quad t \]

\[ z = f(x, y) \] depends on two variables. Use partial derivatives.

\[ x \text{ and } y \text{ each depend on one variable, } t. \] Use ordinary derivative.

To compute \( \frac{dz}{dt} \):
- There are two paths from \( z \) at the top to \( t \)'s at the bottom.
- Along each path, multiply the derivatives.
- Add the products over all paths.

\[ \frac{dz}{dt} = \frac{\partial z}{\partial x} \frac{dx}{dt} + \frac{\partial z}{\partial y} \frac{dy}{dt} \]
Tree diagram of chain rule

\( z = f(x, y) \), \( x = g_1(u, v) \), \( y = g_2(u, v) \), gives \( z = h(u, v) = f(g_1(u, v), g_2(u, v)) \)

\[
\begin{align*}
\frac{\partial z}{\partial x} & \quad \frac{\partial z}{\partial y} \\
\frac{\partial x}{\partial u} & \quad \frac{\partial x}{\partial v} \\
\frac{\partial y}{\partial u} & \quad \frac{\partial y}{\partial v}
\end{align*}
\]

\( z = f(x, y) \) depends on two variables. Use partial derivatives.

\( x \) and \( y \) each depend on two variables. Use partial derivatives.

To compute \( \frac{\partial z}{\partial u} \):
- Highlight the paths from the \( z \) at the top to the \( u \)'s at the bottom.
- Along each path, multiply the derivatives.
- Add the products over all paths.

\[
\frac{\partial z}{\partial u} = \frac{\partial z}{\partial x} \frac{\partial x}{\partial u} + \frac{\partial z}{\partial y} \frac{\partial y}{\partial u}
\]
Tree diagram of chain rule

\[ z = f(x, y), \quad x = g_1(u, v), \quad y = g_2(u, v), \] gives \[ z = h(u, v) = f(g_1(u, v), g_2(u, v)). \]

\[
\begin{array}{c}
\frac{\partial z}{\partial x} \\
\frac{\partial z}{\partial y}
\end{array}
\]

\[
\begin{array}{c}
x \\
y
\end{array}
\]

\[
\begin{array}{ll}
\frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\
\frac{\partial y}{\partial u} & \frac{\partial y}{\partial v}
\end{array}
\]

\[ u \quad v \quad u \quad v \]

\[ z = f(x, y) \text{ depends on two variables.} \]
\[ \text{Use partial derivatives.} \]

\[ x \text{ and } y \text{ each depend on two variables.} \]
\[ \text{Use partial derivatives.} \]

To compute \( \frac{\partial z}{\partial v} \):

- Highlight the paths from the \( z \) at the top to the \( v \)'s at the bottom.
- Along each path, multiply the derivatives.
- Add the products over all paths.

\[
\frac{\partial z}{\partial v} = \frac{\partial z}{\partial x} \frac{\partial x}{\partial v} + \frac{\partial z}{\partial y} \frac{\partial y}{\partial v}
\]
Example: Chain rule to convert to polar coordinates

Let \( z = f(x, y) = x^2y \)

where \( x = r \cos(\theta) \) and \( y = r \sin(\theta) \)

**Compute \( \frac{\partial z}{\partial r} \) and \( \frac{\partial z}{\partial \theta} \) using the chain rule**

\[
\frac{\partial z}{\partial r} = \frac{\partial z}{\partial x} \frac{\partial x}{\partial r} + \frac{\partial z}{\partial y} \frac{\partial y}{\partial r} \\
= 2xy(\cos \theta) + x^2(\sin \theta) \\
= 2(r \cos \theta)(r \sin \theta)(\cos \theta) + (r \cos \theta)^2(\sin \theta) \\
= 3r^2 \cos^2 \theta \sin \theta
\]

\[
\frac{\partial z}{\partial \theta} = \frac{\partial z}{\partial x} \frac{\partial x}{\partial \theta} + \frac{\partial z}{\partial y} \frac{\partial y}{\partial \theta} \\
= 2xy(-r \sin \theta) + x^2(r \cos \theta) \\
= 2(r \cos \theta)(r \sin \theta)(-r \sin \theta) + (r \cos \theta)^2(r \cos \theta) \\
= -2r^3 \cos \theta \sin^2 \theta + r^3 \cos^3 \theta
\]
Example: Chain rule to convert to polar coordinates

Let  \( z = f(x, y) = x^2 y \)

where  \( x = r \cos(\theta) \)  and  \( y = r \sin(\theta) \)

Use substitution to confirm it

\[
z = x^2 y = (r \cos \theta)^2 (r \sin \theta) = r^3 \cos^2 \theta \sin \theta
\]

\[
\frac{\partial z}{\partial r} = 3r^2 \cos^2 \theta \sin \theta
\]

\[
\frac{\partial z}{\partial \theta} = r^3 (-2 \cos \theta \sin^2 \theta + \cos^3 \theta)
\]
A balloon is approximately an ellipsoid, with radii $a, b, c$:

$\frac{(x - x_0)^2}{a^2} + \frac{(y - y_0)^2}{b^2} + \frac{(z - z_0)^2}{c^2} = 1$

Radii $a(t), b(t), c(t)$ at time $t$ vary as balloon is inflated/deflated.

Volume $V(t) = \frac{4\pi}{3} a(t) b(t) c(t)$.

Instead of formulas for $a(t), b(t), c(t)$, we have experimental measurements. At time $t = 2$ sec:

$a = 4$ in \quad \frac{da}{dt} = -.5$ in/sec

$b = c = 3$ in \quad \frac{db}{dt} = \frac{dc}{dt} = -.9$ in/sec

What is $\frac{dV}{dt}$ at $t = 2$?
Example: Related rates using measurements

Volume $V(t) = \frac{4\pi}{3} a(t) b(t) c(t)$, and at time $t = 2$:

- $a = 4$ in $\quad \frac{da}{dt} = -0.5$ in/sec
- $b = c = 3$ in $\quad \frac{db}{dt} = \frac{dc}{dt} = -0.9$ in/sec

Without formulas for $a(t), b(t), c(t)$, we can’t compute $V(t)$ as a function and differentiate it to get $V'(t)$ as a function.

But we can still evaluate $V(2) = \frac{4\pi}{3} (4)(3)(3) = 48\pi$ and $V'(2)$:

\[
\frac{dV}{dt} = \frac{\partial V}{\partial a} \frac{da}{dt} + \frac{\partial V}{\partial b} \frac{db}{dt} + \frac{\partial V}{\partial c} \frac{dc}{dt}
\]

\[
= \frac{4\pi}{3} \left( bc \frac{da}{dt} + ac \frac{db}{dt} + ab \frac{dc}{dt} \right)
\]

At time $t=2$:

\[
= \frac{4\pi}{3} \left( (3)(3)(-0.5) + (4)(3)(-0.9) + (4)(3)(-0.9) \right)
\]

\[
= \frac{4\pi}{3} \cdot 26.1 \approx -109.33 \text{ in}^3/\text{sec}
\]
Matrices

- A matrix is a square or rectangular table of numbers.
- An \( m \times n \) matrix has \( m \) rows and \( n \) columns. This is read “\( m \) by \( n \)”.
- This matrix is \( 2 \times 3 \) ("two by three"): 
  \[
  \begin{bmatrix}
  1 & 2 & 3 \\
  4 & 5 & 6 \\
  \end{bmatrix}
  \]
- You may have seen matrices in High School Algebra. Matrices will be covered in detail in Linear Algebra (Math 18).
Matrix multiplication

Let $A$ be $p \times q$ and $B$ be $q \times r$.

The product $AB = C$ is a certain $p \times r$ matrix of dot products:

$$C_{i,j} = \text{entry in } i^{\text{th}} \text{ row, } j^{\text{th}} \text{ column of } C$$

$$= \text{dot product } (i^{\text{th}} \text{ row of } A) \cdot (j^{\text{th}} \text{ column of } B)$$

The number of columns in $A$ must equal the number of rows in $B$ (namely $q$) in order to be able to compute the dot products.
Matrix multiplication

\[
\begin{bmatrix}
1 & 2 & 3 \\
4 & 5 & 6
\end{bmatrix}
\begin{bmatrix}
5 & -2 & 3 & 2 \\
0 & 1 & 1 & -1 \\
-1 & 6 & 4 & 3
\end{bmatrix}
= \begin{bmatrix}
\vdots & \vdots & \vdots & \vdots 
\end{bmatrix}
\]

\[C_{1,1} = 1(5) + 2(0) + 3(-1) = 5 + 0 - 3 = 2\]
Matrix multiplication

\[
\begin{bmatrix}
1 & 2 & 3 \\
4 & 5 & 6
\end{bmatrix}
\begin{bmatrix}
5 & \textcolor{lime}{-2} & 3 \\
0 & 1 & 2 \\
-1 & -1 & 3
\end{bmatrix}
= \begin{bmatrix}
2 & \textcolor{lime}{18} & \cdot & \cdot & \cdot
\end{bmatrix}
\]

\[
C_{1,2} = 1(-2) + 2(1) + 3(6) = -2 + 2 + 18 = 18
\]
Matrix multiplication

\[
\begin{pmatrix}
1 & 2 & 3 \\
4 & 5 & 6
\end{pmatrix}
\begin{pmatrix}
5 & -2 & 3 & 2 \\
0 & 1 & 1 & -1 \\
-1 & 6 & 4 & 3
\end{pmatrix}
= \begin{pmatrix}
\cdot & \cdot & \cdot & \cdot \\
2 & 18 & 17 & \cdot
\end{pmatrix}
\]

\[C_{1,3} = 1(3) + 2(1) + 3(4) = 3 + 2 + 12 = 17\]
Matrix multiplication

\[
\begin{bmatrix}
1 & 2 & 3 \\
4 & 5 & 6 \\
\end{bmatrix}
\begin{bmatrix}
5 & -2 & 3 \\
0 & 1 & 1 \\
-1 & 6 & 4 \\
\end{bmatrix}
= 
\begin{bmatrix}
2 & 18 & 17 & 9 \\
\end{bmatrix}
\]

\[C_{1,4} = 1(2) + 2(-1) + 3(3) = 2 - 2 + 9 = 9\]
Matrix multiplication

\[
\begin{bmatrix}
1 & 2 & 3 \\
4 & 5 & 6 \\
-1 & &
\end{bmatrix}
\begin{bmatrix}
5 & -2 & 3 & 2 \\
0 & 1 & 1 & -1 \\
6 & 4 & 3 &
\end{bmatrix}
= \begin{bmatrix}
2 & 18 & 17 & 9 \\
14 & & &
\end{bmatrix}
\]

\[C_{2,1} = 4(5) + 5(0) + 6(-1) = 20 + 0 - 6 = 14\]
Matrix multiplication

\[
\begin{bmatrix}
1 & 2 & 3 \\
4 & 5 & 6
\end{bmatrix}
\begin{bmatrix}
5 & -2 & 3 & 2 \\
0 & 1 & 1 & -1 \\
-1 & 6 & 4 & 3
\end{bmatrix}
= 
\begin{bmatrix}
2 & 18 & 17 & 9 \\
14 & 33 & \cdot & \cdot
\end{bmatrix}
\]

\( C_{2,2} = 4(-2) + 5(1) + 6(6) = -8 + 5 + 36 = 33 \)
Matrix multiplication

\[
\begin{bmatrix}
1 & 2 & 3 \\
4 & 5 & 6
\end{bmatrix}
\begin{bmatrix}
5 & -2 & 3 \\
0 & 1 & -1 \\
-1 & 6 & 3
\end{bmatrix}
= 
\begin{bmatrix}
2 & 18 & 17 & 9 \\
14 & 33 & 41 & .
\end{bmatrix}
\]

\[C_{2,3} = 4(3) + 5(1) + 6(4) = 12 + 5 + 24 = 41\]
Matrix multiplication

\[
\begin{bmatrix}
1 & 2 & 3 \\
4 & 5 & 6
\end{bmatrix}
\begin{bmatrix}
5 & -2 & 3 \\
0 & 1 & 1 \\
-1 & 6 & 4
\end{bmatrix}
= 
\begin{bmatrix}
2 & 18 & 17 & 9 \\
14 & 33 & 41 & 21
\end{bmatrix}
\]

\[C_{2,4} = 4(2) + 5(-1) + 6(3) = 8 - 5 + 18 = 21\]
Chain rule using matrices

Our earlier example

Let \( z = f(x, y) = x^2 y \)
where \( x = r \cos(\theta) \) and \( y = r \sin(\theta) \)

becomes

Let \( z = f(x, y) = x^2 y \)
where \( (x, y) = g(r, \theta) = (r \cos(\theta), r \sin(\theta)) \)

and set \( h = f \circ g \)

\[
h(r, \theta) = f(g(r, \theta)) = f(r \cos(\theta), r \sin(\theta))
= (r \cos(\theta))^2(r \sin(\theta)) = r^3 \cos^2(\theta) \sin(\theta)
\]
Chain rule using matrices

Let \( z = f(x, y) = x^2 y \)

where \( (x, y) = g(r, \theta) = (r \cos(\theta), r \sin(\theta)) \)

and set \( h = f \circ g \)

\[ h(r, \theta) = f(g(r, \theta)) = \cdots = r^3 \cos^2(\theta) \sin(\theta) \]

\[
\frac{\partial h}{\partial r} = \frac{\partial f}{\partial x} \frac{\partial x}{\partial r} + \frac{\partial f}{\partial y} \frac{\partial y}{\partial r} \quad \frac{\partial h}{\partial \theta} = \frac{\partial f}{\partial x} \frac{\partial x}{\partial \theta} + \frac{\partial f}{\partial y} \frac{\partial y}{\partial \theta}
\]
Chain rule using matrices

Let \( z = f(x, y) = x^2 y \)

where \( (x, y) = g(r, \theta) = (r \cos(\theta), r \sin(\theta)) \)

and set \( h = f \circ g \)

\[ h(r, \theta) = f(g(r, \theta)) = \cdots = r^3 \cos^2(\theta) \sin(\theta) \]

\[
\begin{bmatrix}
\frac{\partial h}{\partial r} & \frac{\partial h}{\partial \theta}
\end{bmatrix} = \begin{bmatrix}
\frac{\partial f}{\partial x} \frac{\partial x}{\partial r} + \frac{\partial f}{\partial y} \frac{\partial y}{\partial r} & \frac{\partial f}{\partial x} \frac{\partial x}{\partial \theta} + \frac{\partial f}{\partial y} \frac{\partial y}{\partial \theta}
\end{bmatrix}
\]

\[
= \begin{bmatrix}
\frac{\partial f}{\partial x} & \frac{\partial f}{\partial y}
\end{bmatrix} \begin{bmatrix}
\frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta}
\end{bmatrix}
\begin{bmatrix}
\frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta}
\end{bmatrix}
\]

\[ Dh(r, \theta) = \left( Df \text{ at } (x, y) = g(r, \theta) \right) \left( Dg(r, \theta) \right) \]

\[ = D(\text{outside function}) D(\text{inside function}) \]
Let \( z = f(x, y) = x^2 y \)

where \( (x, y) = g(r, \theta) = (r \cos(\theta), r \sin(\theta)) \)

and set \( h = f \circ g \)

\[
h(r, \theta) = f(g(r, \theta)) = \cdots = r^3 \cos^2(\theta) \sin(\theta)
\]

\[
\begin{bmatrix}
\frac{\partial h}{\partial r} & \frac{\partial h}{\partial \theta}
\end{bmatrix}
= \begin{bmatrix}
\frac{\partial f}{\partial x} & \frac{\partial f}{\partial y}
\end{bmatrix}
\begin{bmatrix}
\frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\
\frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta}
\end{bmatrix}
= \begin{bmatrix}
2xy & x^2
\end{bmatrix}
\begin{bmatrix}
\cos(\theta) & -r \sin(\theta) \\
\sin(\theta) & r \cos(\theta)
\end{bmatrix}
= \begin{bmatrix}
2xy \cos(\theta) + x^2 \sin(\theta) & -2xy r \sin(\theta) + x^2 r \cos(\theta)
\end{bmatrix}
= \begin{bmatrix}
3r^2 \cos^2(\theta) \sin(\theta) & -2r^3 \cos(\theta) \sin^2(\theta) + r^3 \cos^3(\theta)
\end{bmatrix}
\]
Chain rule using matrices

Let $g : \mathbb{R}^a \to \mathbb{R}^b$  
$\vec{y} = g(\vec{x}) = g(x_1, \ldots, x_a)$  
$= (g_1(x_1, \ldots, x_a), \ldots, g_b(x_1, \ldots, x_a))$

$f : \mathbb{R}^b \to \mathbb{R}^c$  
$\vec{z} = f(\vec{y}) = f(y_1, \ldots, y_b)$  
$= (f_1(y_1, \ldots, y_b), \ldots, f_c(y_1, \ldots, y_b))$

Set $h = f \circ g$:  
$h : \mathbb{R}^a \to \mathbb{R}^c$  
$\vec{z} = h(\vec{x}) = f(g(\vec{x}))$  
$= (h_1(x_1, \ldots, x_a), \ldots, h_c(x_1, \ldots, x_a))$

The chain rule is expressed as a product of derivative matrices:

$$Dh(\vec{x}) = \left( Df(\vec{y}) \text{ at } \vec{y} = g(\vec{x}) \right) \left( Dg(\vec{x}) \right)$$

Size: $c \times a$ $c \times b$ $b \times a$

$D(\text{outside function}) D(\text{inside})$
Derivatives of sums, products, and quotients

Single variable

For single variable functions $f : \mathbb{R} \to \mathbb{R}$, $g : \mathbb{R} \to \mathbb{R}$, and constant $c$:

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

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

$$\frac{d}{dx} (f(x)g(x)) = \left( \frac{d}{dx} f(x) \right) g(x) + f(x) \left( \frac{d}{dx} g(x) \right)$$

$$\frac{d}{dx} \left( \frac{f(x)}{g(x)} \right) = \frac{g(x) \left( \frac{d}{dx} f(x) \right) - f(x) \left( \frac{d}{dx} g(x) \right)}{g(x)^2}$$
Derivatives of sums, products, and quotients

Gradient

For multivariable functions \( f : \mathbb{R}^n \to \mathbb{R} \), \( g : \mathbb{R}^n \to \mathbb{R} \), and constant \( c \):

\[
\nabla (c f(\vec{x})) = c \nabla f(\vec{x})
\]

\[
\nabla \left( f(\vec{x}) + g(\vec{x}) \right) = \nabla \left( f(\vec{x}) \right) + \nabla \left( g(\vec{x}) \right)
\]

\[
\nabla \left( f(\vec{x}) g(\vec{x}) \right) = (\nabla f(\vec{x})) g(\vec{x}) + f(\vec{x}) \left( \nabla g(\vec{x}) \right)
\]

\[
\nabla \left( \frac{f(\vec{x})}{g(\vec{x})} \right) = \frac{g(\vec{x}) \left( \nabla f(\vec{x}) \right) - f(\vec{x}) \left( \nabla g(\vec{x}) \right)}{g(\vec{x})^2}
\]
Derivatives of sums, products, and quotients

Gradient examples

Example

With $\nabla f(x, y) = \langle f_x, f_y \rangle$, we apply the single variable rules for $\frac{\partial}{\partial x}$ in the 1st component and $\frac{\partial}{\partial y}$ in the 2nd component:

$$\nabla (2x^2 y + 3e^x) = 2 \nabla (x^2 y) + 3 \nabla (e^x) = \langle 4xy + 3e^x, 2x^2 \rangle$$

$$\nabla (e^{xy} \cos(x^2)) = \langle y e^{xy}, x e^{xy} \rangle \cos(x^2) + e^{xy} \langle -2x \sin(x^2), 0 \rangle$$

$$= \langle (y \cos(x^2) - 2x \sin(x^2)) e^{xy}, x \cos(x^2) e^{xy} \rangle$$
Derivatives of sums, products, and quotients

Derivative matrix

For \( f : \mathbb{R}^n \to \mathbb{R}^m \), \( g : \mathbb{R}^n \to \mathbb{R}^m \), and constant \( c \)

\[
D(cf(\vec{x})) = c \ Df(\vec{x})
\]

\[
D\left(f(\vec{x}) + g(\vec{x})\right) = Df(\vec{x}) + Dg(\vec{x})
\]

For \( f : \mathbb{R}^n \to \mathbb{R} \), \( g : \mathbb{R}^n \to \mathbb{R} \)

- For multiplying or dividing scalar-valued functions of vectors:

\[
D\left(f(\vec{x})g(\vec{x})\right) = (Df(\vec{x})) \ g(\vec{x}) + f(\vec{x}) \ (Dg(\vec{x}))
\]

\[
D \left(\frac{f(\vec{x})}{g(\vec{x})}\right) = \frac{g(\vec{x}) \ (Df(\vec{x})) - f(\vec{x}) \ (Dg(\vec{x}))}{g(\vec{x})^2}
\]

- This case is identical to the gradient on the previous slides:
  Since \( f, g \) are scalar-valued, \( Df = \nabla f \) and \( Dg = \nabla g \) are just different notations for the same thing.
Derivatives of sums, products, and quotients

Derivative matrix: example

\[
D \left( x^2y + 3e^x, xy^3 + 3e^y \right) = D \left( x^2y, xy^3 \right) + 3D \left( e^x, e^y \right)
\]

\[
= \begin{bmatrix}
2xy & x^2 \\
y^3 & 3xy^2
\end{bmatrix} + 3 \begin{bmatrix}
e^x & 0 \\
0 & e^y
\end{bmatrix}
\]

\[
= \begin{bmatrix}
2xy + 3e^x & x^2 \\
y^3 & 3xy^2 + 3e^y
\end{bmatrix}
\]