2.5 Chain Rule for Multiple Variables

Prof. Tesler

Math 20C
Fall 2018

Review of the chain for functions of one variable

Chain rule

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

Example

$$\frac{d}{dx} \sin(x^2) = \cos(x^2) \cdot (2x) = \boxed{\mathbf{2\,x\cos(x^2)}}$$

- 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 $\quad f(x) = \text{sin}(x) \qquad g(x) = x^2$
- The *composition* of these is the function $h = f \circ g$:
$$h(x) = f(g(x)) = \text{sin}(x^2)$$
- The notation $f \circ g$ is read as
"f composed with g"
or "the composition of f with g."

Function composition: Diagram

$$
\begin{tikzcd}
A \arrow[r, "g"', "\text{inside function}" below] \arrow[rr, bend left, "h = f \circ g", "\text{composition}"'] & B \arrow[r, "f", "\text{outside function}" below] & C
\end{tikzcd}
$$

- A, B, C are sets. They can have different dimensions, e.g.,

$$A \subseteq \mathbb{R}^n \qquad B \subseteq \mathbb{R}^m \qquad C \subseteq \mathbb{R}^p$$

- f, g, and h are functions. Domains and codomains:

$$
\begin{aligned}
f &: B \to C \\
g &: A \to B \\
h &: A \to C
\end{aligned}
$$

Function composition: Multiple variables

$$f : \mathbb{R}^2 \to \mathbb{R} \qquad f(x,y) = x^2 + y$$

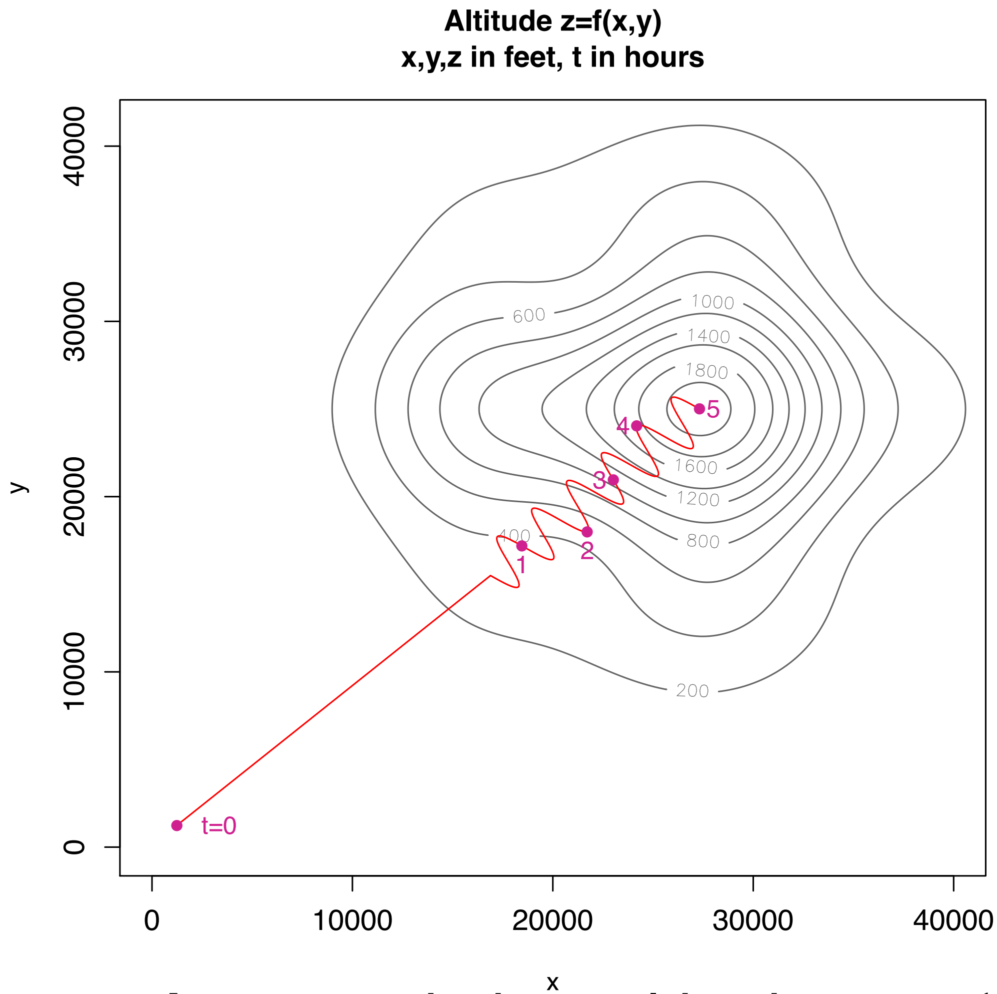
$$\vec{r} : \mathbb{R} \to \mathbb{R}^2 \qquad \begin{aligned} \vec{r}(t) &= \langle x(t), y(t) \rangle \\ &= \langle 2t+1, 3t-1 \rangle \end{aligned}$$

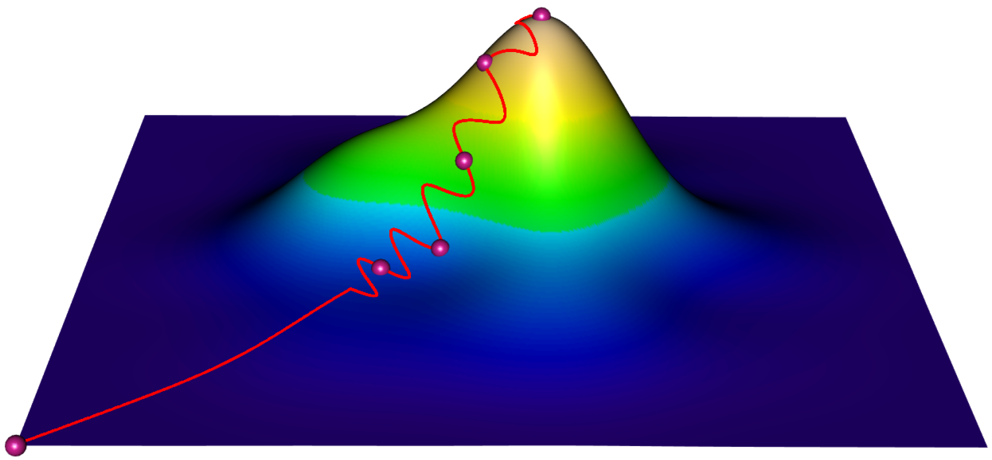
$$f \circ \vec{r} : \mathbb{R} \to \mathbb{R} \qquad \begin{aligned} (f \circ \vec{r})(t) &= f(\vec{r}(t)) \\ &= f(2t+1, 3t-1) \\ &= (2t+1)^2 + (3t-1) \\ &= 4t^2 + 7t \end{aligned}$$

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:** $(f \circ \vec{r})'(t) = 8t + 7$
  $(f \circ \vec{r})'(10) = 8 \cdot 10 + 7 = 87$

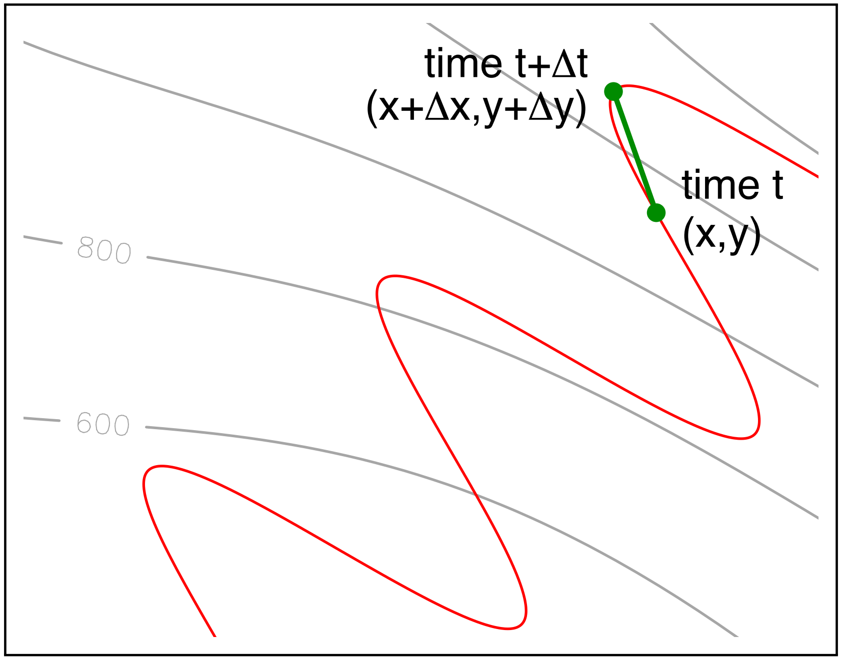
Hiking trail





- 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 dz/dt = vertical speed of hiker?



- Let Δt = very small change in time.
- The change in altitude is

$$\begin{aligned} \Delta z &= z(t+\Delta t) - z(t) \\ &\approx f_x(x,y)\Delta x \;+\; f_y(x,y)\Delta y \qquad \text{Using the linear approximation} \end{aligned}$$

What is dz/dt = vertical speed of hiker?

- Let Δt = very small change in time.
- The change in altitude is

$$\begin{aligned} \Delta z &= z(t+\Delta t) - z(t) \\ &\approx f_x(x,y)\Delta x \ + \ f_y(x,y)\Delta y \qquad \text{Using the linear approximation} \end{aligned}$$

- 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 $\Delta t \to 0$:

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

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

$$\boxed{\frac{dz}{dt} = f_x(x,y)\frac{dx}{dt} + f_y(x,y)\frac{dy}{dt}} \quad \text{or} \quad \boxed{\frac{dz}{dt} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt}}$$

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

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

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

$$\text{Let } z = f(x,y) = x^2 + y$$
$$\text{where } x = 2t + 1 \quad \text{and} \quad y = 3t - 1$$

Compute 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:

$$\begin{aligned} z = x^2 + y &= (2t+1)^2 + (3t-1) \\ &= 4t^2 + 4t + 1 + 3t - 1 \\ &= 4t^2 + 7t \end{aligned}$$

- Then compute dz/dt:

$$\frac{dz}{dt} = 8t + 7$$

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:

$$= 4(2t + 1) + 3 = 8t + 4 + 3 = \boxed{\mathbf{8t + 7}}$$

- 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 $z = h(t) = f(x(t), y(t))$

$$\begin{tikzcd}
 & z \arrow[dl, dash, "\frac{\partial z}{\partial x}"'] \arrow[dr, dash, "\frac{\partial z}{\partial y}"] & \\
x \arrow[d, dash, "\frac{dx}{dt}"'] & & y \arrow[d, dash, "\frac{dy}{dt}"] \\
t & & t
\end{tikzcd}$$

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

x and y each depend on one variable, t.
Use ordinary derivative.

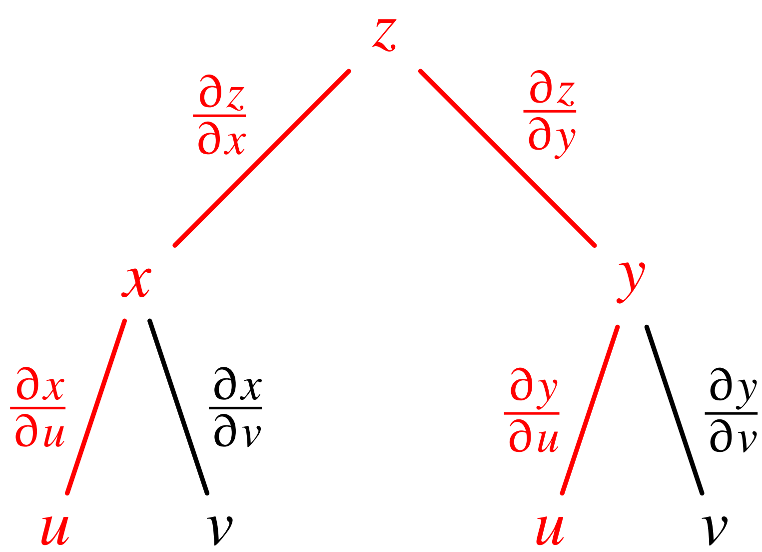
To compute $\frac{dz}{dt}$:

- 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 $z = h(u, v) = f(g_1(u, v), g_2(u, v))$



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

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

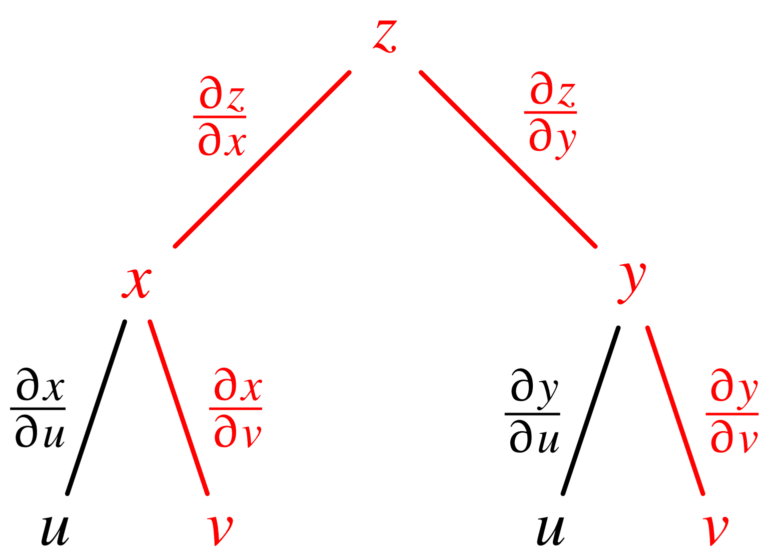
To compute $\frac{\partial z}{\partial u}$:

- 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

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



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

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

To compute $\frac{\partial z}{\partial v}$:

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

Compute $\partial z/\partial r$ and $\partial z/\partial\theta$ using the chain rule

$$
\begin{aligned}
\frac{\partial z}{\partial r} &= \frac{\partial z}{\partial x}\frac{\partial x}{\partial r} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial r} \\
&= 2xy(\cos\theta) + x^2(\sin\theta) \\
&= 2(r\cos\theta)(r\sin\theta)(\cos\theta) + (r\cos\theta)^2(\sin\theta) \\
&= 3r^2\cos^2\theta\sin\theta
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial z}{\partial\theta} &= \frac{\partial z}{\partial x}\frac{\partial x}{\partial\theta} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial\theta} \\
&= 2xy(-r\sin\theta) + x^2(r\cos\theta) \\
&= 2(r\cos\theta)(r\sin\theta)(-r\sin\theta) + (r\cos\theta)^2(r\cos\theta) \\
&= -2r^3\cos\theta\sin^2\theta + r^3\cos^3\theta
\end{aligned}
$$

Example: Chain rule to convert to polar coordinates

$$\text{Let } z = f(x, y) = x^2 y$$
$$\text{where } x = r\cos(\theta) \quad \text{and} \quad y = r\sin(\theta)$$

Use substitution to confirm it

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

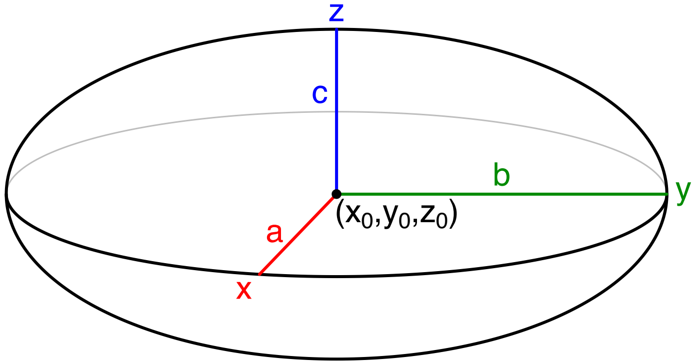
$$\frac{\partial z}{\partial r} = 3r^2 \cos^2\theta \sin\theta$$

$$\frac{\partial z}{\partial \theta} = r^3(-2\cos\theta \sin^2\theta + \cos^3\theta)$$

Example: Related rates using measurements

- 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:

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

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

- 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 $t = 2$:

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

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

- 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 $V'(2)$:

$$\begin{aligned}
\frac{dV}{dt} &= \frac{\partial V}{\partial a}\frac{da}{dt} + \frac{\partial V}{\partial b}\frac{db}{dt} + \frac{\partial V}{\partial c}\frac{dc}{dt} \\
&= \frac{4\pi}{3}\left(bc\frac{da}{dt} + ac\frac{db}{dt} + ab\frac{dc}{dt}\right) \\
\text{At time } t{=}2\text{:} \quad &= \frac{4\pi}{3}\Big((3)(3)(-.5) + (4)(3)(-.9) + (4)(3)(-.9)\Big) \\
&= \frac{4\pi}{3}(-26.1) \approx -109.33 \text{ in}^3/\text{sec}
\end{aligned}$$

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 2×3 ("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

$$\begin{array}{cccc} A & B & = & C \\ \underbrace{\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}}_{2\times 3} & \underbrace{\begin{bmatrix} 5 & -2 & 3 & 2 \\ 0 & 1 & 1 & -1 \\ -1 & 6 & 4 & 3 \end{bmatrix}}_{3\times 4} & = & \underbrace{\begin{bmatrix} \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{bmatrix}}_{2\times 4} \end{array}$$

- 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:

$$\begin{aligned} C_{i,j} &= \text{entry in } i^{\text{th}} \text{ row, } j^{\text{th}} \text{ column of } C \\ &= \text{dot product } (i^{\text{th}} \text{ row of } A) \cdot (j^{\text{th}} \text{ column of } B) \end{aligned}$$

- 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

$$\begin{bmatrix} \boxed{1 \quad 2 \quad 3} \\ 4 \quad 5 \quad 6 \end{bmatrix} \begin{bmatrix} 5 & -2 & 3 & 2 \\ 0 & 1 & 1 & -1 \\ -1 & 6 & 4 & 3 \end{bmatrix} = \begin{bmatrix} 2 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{bmatrix}$$

$$C_{1,1} = 1(5) + 2(0) + 3(-1) = 5 + 0 - 3 = 2$$

Matrix multiplication

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \begin{bmatrix} 5 & -2 & 3 & 2 \\ 0 & 1 & 1 & -1 \\ -1 & 6 & 4 & 3 \end{bmatrix} = \begin{bmatrix} 2 & 18 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{bmatrix}$$

$$C_{1,2} = 1(-2) + 2(1) + 3(6) = -2 + 2 + 18 = 18$$

Matrix multiplication

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \begin{bmatrix} 5 & -2 & 3 & 2 \\ 0 & 1 & 1 & -1 \\ -1 & 6 & 4 & 3 \end{bmatrix} = \begin{bmatrix} 2 & 18 & 17 & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{bmatrix}$$

$$C_{1,3} = 1(3) + 2(1) + 3(4) = 3 + 2 + 12 = 17$$

Matrix multiplication

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

$$C_{1,4} = 1(2) + 2(-1) + 3(3) = 2 - 2 + 9 = 9$$

Matrix multiplication

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

$$C_{2,1} = 4(5) + 5(0) + 6(-1) = 20 + 0 - 6 = 14$$

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

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

$$C_{2,3} = 4(3) + 5(1) + 6(4) = 12 + 5 + 24 = 41$$

Matrix multiplication

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

$$C_{2,4} = 4(2) + 5(-1) + 6(3) = 8 - 5 + 18 = 21$$

Chain rule using matrices

Our earlier example

$$\text{Let } z = f(x,y) = x^2 y$$
$$\text{where } x = r\cos(\theta) \quad \text{and} \quad y = r\sin(\theta)$$

becomes

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

$$\text{and set } h = f \circ g$$

$$\begin{aligned} h(r,\theta) &= f(g(r,\theta)) = f(r\cos(\theta), r\sin(\theta)) \\ &= (r\cos(\theta))^2 (r\sin(\theta)) = r^3\cos^2(\theta)\sin(\theta) \end{aligned}$$

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 $h = f \circ g$

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

$$\frac{\partial h}{\partial r} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial r} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial r} \qquad \frac{\partial h}{\partial \theta} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial \theta} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial \theta}$$

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 $h = f \circ g$

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

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

$$\begin{aligned}
Dh(r,\theta) &= \Big(Df \text{ at } (x,y) = g(r,\theta)\Big)\Big(Dg(r,\theta)\Big) \\
&= D(\text{outside function})\, D(\text{inside function})
\end{aligned}$$

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 $h = f \circ g$

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

$$\begin{aligned}
\begin{bmatrix} \frac{\partial h}{\partial r} & \frac{\partial h}{\partial \theta} \end{bmatrix} &= \begin{bmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \end{bmatrix} \begin{bmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{bmatrix} \\
&= \begin{bmatrix} 2xy & x^2 \end{bmatrix} \begin{bmatrix} \cos(\theta) & -r\sin(\theta) \\ \sin(\theta) & r\cos(\theta) \end{bmatrix} \\
&= \begin{bmatrix} 2\,x\,y\cos(\theta) + x^2\sin(\theta) & -2\,x\,y\,r\sin(\theta) + x^2\,r\cos(\theta) \end{bmatrix} \\
&= \begin{bmatrix} 3r^2\cos^2(\theta)\sin(\theta) & -2r^3\cos(\theta)\sin^2(\theta) + r^3\cos^3(\theta) \end{bmatrix}
\end{aligned}$$

Chain rule using matrices

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

$$= (g_1(x_1, \ldots, x_a), \ldots, g_b(x_1, \ldots, x_a))$$

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

$$= (f_1(y_1, \ldots, y_b), \ldots, f_c(y_1, \ldots, y_b))$$

Set $h = f \circ g$:

$h : \mathbb{R}^a \to \mathbb{R}^c \qquad \vec{z} = h(\vec{x}\,) = f(g(\vec{x}\,))$

$$= (h_1(x_1, \ldots, x_a), \ldots, h_c(x_1, \ldots, x_a))$$

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

$$Dh(\vec{x}) \quad = \quad \Big(Df(\vec{y}\,) \text{ at } \vec{y}=g(\vec{x}\,)\Big) \ \Big(Dg(\vec{x}\,)\Big)$$

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

D(outside function) D(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:

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

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

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

$$\frac{d}{dx}\left(\frac{f(x)}{g(x)}\right) = \frac{g(x)\left(\frac{d}{dx} f(x)\right) - f(x)\left(\frac{d}{dx} g(x)\right)}{g(x)^2}$$

Derivatives of sums, products, and quotients
Gradient

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

$$\nabla(cf(\vec{x})) = c\nabla f(\vec{x})$$

$$\nabla\Big(f(\vec{x}) + g(\vec{x})\Big) = \nabla\Big(f(\vec{x})\Big) + \nabla\Big(g(\vec{x})\Big)$$

$$\nabla\Big(f(\vec{x})g(\vec{x})\Big) = (\nabla f(\vec{x}))\, g(\vec{x}) + f(\vec{x})\,(\nabla g(\vec{x}))$$

$$\nabla\left(\frac{f(\vec{x})}{g(\vec{x})}\right) = \frac{g(\vec{x})\,(\nabla f(\vec{x})) - f(\vec{x})\,(\nabla g(\vec{x}))}{g(\vec{x})^2}$$

Derivatives of sums, products, and quotients

Gradient examples

Example

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

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

$$\begin{aligned}
\nabla(e^{xy}\cos(x^2)) &= (\nabla(e^{xy}))\cos(x^2) + e^{xy}\nabla(\cos(x^2)) \\
&= \langle y\,e^{xy}, x\,e^{xy}\rangle \cos(x^2) + e^{xy}\left\langle -2x\sin(x^2), 0 \right\rangle \\
&= \left\langle (y\cos(x^2) - 2x\sin(x^2))e^{xy}, x\cos(x^2)e^{xy} \right\rangle
\end{aligned}$$

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

$$D(c\,f(\vec{x})) = c\,Df(\vec{x})$$

$$D\Big(f(\vec{x}) + g(\vec{x})\Big) = Df(\vec{x}) + Dg(\vec{x})$$

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

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

$$D\Big(f(\vec{x})g(\vec{x})\Big) = (Df(\vec{x}))\,g(\vec{x}) + f(\vec{x})\,(Dg(\vec{x}))$$

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

- 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

$$\begin{aligned} D\left\langle x^2y+3e^x, xy^3+3e^y\right\rangle &= D\left\langle x^2y, xy^3\right\rangle + 3D\left\langle e^x, e^y\right\rangle \\ &= \begin{bmatrix} 2xy & x^2 \\ y^3 & 3xy^2 \end{bmatrix} + 3\begin{bmatrix} e^x & 0 \\ 0 & e^y \end{bmatrix} \\ &= \begin{bmatrix} 2xy+3e^x & x^2 \\ y^3 & 3xy^2+3e^y \end{bmatrix} \end{aligned}$$