Pre-class Warm-up!!!

1. $\dfrac{d}{dx} \sin(2x^2) =$

   a. $\cos(4x)$

   b. $2x^2 \sin(2x^2)$

   c. $4x \cos(2x^2)$ ✓

   d. $2x^2 \cos(4x)$

   e. None of the above

2. To do question 1 we use:

   a. a calculator

   b. Leibniz' rule

   c. the chain rule ✓

   d. guess work

   e. None of the above

2.5 The chain rule and other properties of the derivative.

Mostly this section is about the chain rule, and the most important thing is what it looks like, why you might expect it to be this way, and how to do the HW questions

$$y = y(u), \quad u = u(x)$$

$$\left.\frac{dy}{dx}\right|_a = \left.\frac{dy}{du}\right|_{u(a)} \cdot \left.\frac{du}{dx}\right|_a$$

We could write $y(u) = f(u)$
$$u(x) = g(x)$$

Review of the 1-variable case.

Informally the chain rule says
dy/dx = (dy/du)(du/dx)
when y is a function of u and u a function of x.

Really we want $\left.\frac{dy}{dx}\right|_{x=a} = \frac{dy}{dx}(a)$

E.g. u = x^2 + x, y = 2u^2
- y = 2(x^2+x)^2 = 2x^4 + 4x^3 + 2x^2
- dy/dx = 8x^3 + 12x^2 + 4x

- dy/du = 4u   du/dx = 2x+1
- dy/dx = (4u) (2x+1) = 4(x^2+x) (2x+1)
  = 8x^3 + 12x^2 + 4x

← multiplication

Or: (f ° g)′ (a) = f′(g(a)) g′(a)

Special case of the chain rule

Let $c : R \to R^3$ and $f : R^3 \to R$ be

$$R \xrightarrow{\;c\;} R^3 \xrightarrow{\;f\;} R$$

$c(t) = (t^2, 2t, \sin t)$, $f(x,y,z) = 2x + yz^2$

Find df/dt

Question: was it right to use d rather than ∂ just now?

a. Yes ✓
b. No

$$\frac{df}{dt} = 2 \cdot 2t + z^2 \cdot 2 + 2yz \cos t$$

$$= 4t + 2\sin^2 t + 4t \sin t \cos t \; .$$

Would it be correct to write ∂ in $\frac{\partial f}{\partial t}$?

Yes.

The chain rule says

$$\frac{df}{dt}\Big|_{t=a} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial t}\Big|_{c(a)}\Big|_{a} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial t}\Big|_{c(a)}\Big|_{a} + \frac{\partial f}{\partial z}\frac{\partial z}{\partial t}\Big|_{c(a)}\Big|_{a} \; .$$

Notice: the derivative matrices of f and c are

$$Df = \left[\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z} \right], \quad Dc = \begin{bmatrix} \frac{\partial c_1}{\partial t} \\ \frac{\partial c_2}{\partial t} \\ \frac{\partial c_3}{\partial t} \end{bmatrix}$$

The chain rule is

$$D(f \circ c)(a) = Df(c(a)) \cdot Dc(a)$$

and $$\frac{df}{dt} = Df \cdot Dc$$

matrix multiplication .

The chain rule

Theorem 11. Let $g : \mathbb{R}^n \to \mathbb{R}^m$ and $f : \mathbb{R}^m \to \mathbb{R}^p$ be differentiable. Then $f \circ g$ is differentiable and

$D(f \circ g)(a) = Df(g(a)) \circ Dg(a)$ — matrix multiplication = composition of functions

$$\mathbb{R}^n \xrightarrow{\ g\ } \mathbb{R}^m \xrightarrow{\ f\ } \mathbb{R}^p$$

Idea of proof: we go to the definition of the derivative and show that

$$\lim_{x \to a} \frac{\| (f \circ g)(x) - (f \circ g)(a) - Df(g(a)) \circ Dg(a)\,(x-a) \|}{\| x - a \|}$$

$$\simeq 0$$

$$\frac{\| fg(x) - fg(a) - Df(g(a)) \circ Dg(a)\,(x-a) \|}{\| x \to a \|}$$

$$= \| fg(x) - fg(a) - Df(g(a))\,(g(x)-g(a))$$

$$+ Df(g(a))\,(g(x) - g(a) - Dg(a)(x-a) \|$$

$$\overline{\qquad\qquad \| x-a \| \qquad\qquad}$$

$$\leq \| fg(x) - fg(a) - Df(g(a)\,(g(x)-g(a)) \|$$

$$+ \quad etc$$

Other properties of the derivative

They may all be proved by going to the definition of the derivative and showing that the given expression for the derivative satisfies the definition.

Theorem 10 (i) and (ii)
Let $f, g : R^n \to R^m$ be differentiable and let c, d be scalars.
Then $cf + dg$ is differentiable and
$$D(cf + dg)(a) = c\, Df(a) + d\, Dg(a)$$

D is linear.

Example: $D\left(2t^2 + 3\sin t\right) = 2\,D(t^2) + 3\,D(\sin t)$

Parts (iii) and (if) of Theorem 10 do not have much use because they apply to ~only~ differentiable functions $f, g : R^n \to R$

(iii) The product gf is differentiable and
$$D(fg)(a) = g(a)\, Df(a) + Dg(a)\, f(a)$$

(iv) If g is never zero then f/g is differentiable and
$$D(f/g)(a) = g(a)\, Df(a) - f(a)\, Dg(a) / [g(a)]^2$$

Exercise 1. If $f : R^n \to R$ is differentiable, prove that $x \to f^2(x) + 2f(x)$ is differentiable and find its derivative in terms of Df.

Most questions get you to calculate $D(f \circ g)$ in particular cases, and they can mostly be done by first computing $f \circ g$ and then finding the derivative without the chain rule.

Example
f(u,v) = (u-v, 2v^2)

$f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$

g(x,y,z) = (xy+z, x^2 + y^2 + z^2)

$g : \mathbb{R}^3 \to \mathbb{R}^2$

Find $D(f \circ g) \, (1,0,1)$

Example (like Ex 33):

Let c = c(t) : R -> R^2 have c(1) = (3,4) and
c′(1) = (1,2).
Let f = f(x,y) : R^2-> R have ∂f/∂x (3,4) = -1.
Suppose that d/dt (f ° c) = 5t.
Find ∂f/∂y (3,4).

$$R \xrightarrow{\ c\ } R^2 \xrightarrow{\ f\ } R$$

Consider

$$\frac{df}{dt}(1) = \frac{\partial f}{\partial x}(c(1)) \frac{\partial x}{\partial t}(1) + \frac{\partial f}{\partial y} c(1) \frac{\partial y}{\partial t}(1)$$

$$= \underbrace{\frac{\partial f}{\partial x}(3,4)}_{-1} \cdot 1 + \frac{\partial f}{\partial y}(3,4) \cdot 2 = 5 \cdot 1$$

$$5 = -1 + 2\frac{\partial f}{\partial y}(3,4), \qquad \frac{\partial f}{\partial y}(3,4) = 3.$$

$$\mathbb{R}^n \xrightarrow{f} \mathbb{R}^m \xrightarrow{g} \mathbb{R}^p$$

$Df \rightsquigarrow$ best linear approxn to f

$Dg \rightsquigarrow$ ‾ ‾ ‾ g

$D(g \circ f) \longrightarrow$ $g \circ f$

To get the best approximation to $g \circ f$ we compose the best approximation for f with the best approximation for g.