Longo: Math 20C - Winter 2017
Lecture Notes

Date:

Section:

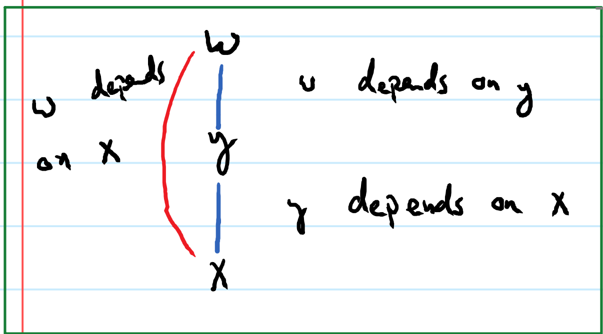
- §2.5 (part 2)

Topics Covered:

- The chain rule for multivariable functions

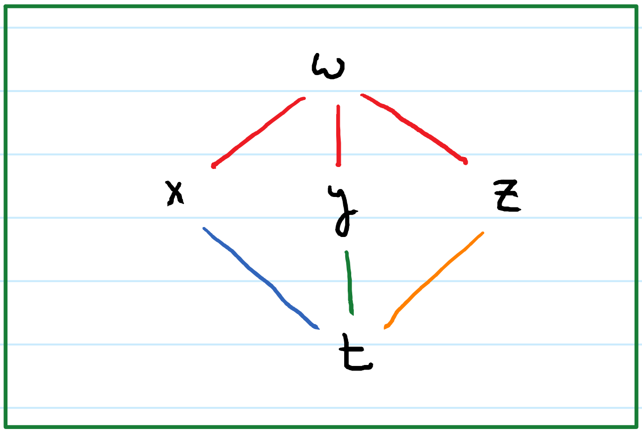
Multivariable Chain Rule:

Let $f, g: \mathbb{R} \to \mathbb{R}$ be diff'ble fcns. Then $f \circ g: \mathbb{R} \to \mathbb{R}$ is also diff'ble and $(f \circ g)'(x) = f'(g(x)) \cdot g'(x)$. In Leibniz notation, if $y = f(x)$ and $w = g(y)$, we see w is a fcn of x and $\frac{dw}{dx} = \left(\frac{dw}{dy}\right) \cdot \left(\frac{dy}{dx}\right)$.



For multivariable fcns, the situation is more complicated (as expected). Let's start with a simple example.

Suppose w is a function of 3 variables: x, y, z. Suppose x, y, z all depend on a single variable, t.



Then w is a function of t, so it makes sense to ask how w changes with respect to a change in t. It turns out

$$\frac{dw}{dt} = \left(\frac{\partial w}{\partial x}\right)\left(\frac{dx}{dt}\right) + \left(\frac{\partial w}{\partial y}\right) \cdot \left(\frac{dy}{dt}\right) + \left(\frac{\partial w}{\partial x}\right)\left(\frac{dx}{dt}\right)$$

We will discuss where this formula comes from shortly, but for now think of it as the sum of the contributions in change coming from the changes in each of the variables.

Example: A bee flies around the room. Suppose at time $t > 0$, the position of the bee is given by

$\vec{c}(t) = (t^2+1,\ t^3+2t,\ \frac{1}{t})$. Suppose the temperature, T, in Fahrenheit at any point in the room is given by

$$T(x,y,z) = 10e^{-x-2y-4z} + 80.$$

① After 2 seconds, what is the bee's temperature.
② How is the bee's temperature changing at 2 seconds?

Sol: ① At the 2 second mark, the bee is at the point $\vec{c}(2) = (2^2+1,\ 2^3+2(2),\ \frac{1}{2})$. The bee's temperature is

$$T(\vec{c}(2)) = 10\,e^{-5-24-2} + 80 = \boxed{\frac{10}{e^{-31}} + 80}$$

② The change in the bee's temperature at time $t=2$ is given by:

$$\left(\frac{dT}{dt}\right)\Big|_{t=2} = \left(\frac{\partial T}{\partial x}\right)\Big|_{\vec{c}(2)} \left(\frac{dx}{dt}\right)\Big|_{t=2} + \left(\frac{\partial T}{\partial y}\right)\Big|_{\vec{c}(2)} \left(\frac{dy}{dt}\right)\Big|_{t=2} + \left(\frac{\partial T}{\partial z}\right)\Big|_{\vec{c}(2)} \left(\frac{dz}{dt}\right)\Big|_{t=2}.$$

Let's compute: $\left(\frac{\partial T}{\partial x}\right)\Big|_{\vec{c}(2)} = \left(-10e^{-x-2y-4z}\right)\Big|_{\vec{c}(2)} = \boxed{-10\,e^{-31}}$

$$\left(\frac{\partial T}{\partial x}\right)\Big|_{\vec{c}(2)} = \left(-20e^{-x-2y-4z}\right)\Big|_{\vec{c}(2)} = \boxed{-20\,e^{-31}}$$

$$\left(\frac{\partial T}{\partial z}\right)\Big|_{\vec{c}(2)} = \left(-40e^{-x-2y-4z}\right)\Big|_{\vec{c}(2)} = \boxed{-40\,e^{-31}}$$

$$\left(\frac{dx}{dt}\right)\Big|_{t=2} = (2t)\Big|_{t=2} = \boxed{4} \qquad \left(\frac{dz}{dt}\right)\Big|_{t=2} = \left(\frac{-1}{t^2}\right)\Big|_{t=2} = \boxed{\frac{-1}{4}}$$

$$\left(\frac{dy}{dt}\right)\Big|_{t=2} = (3t^2)\Big|_{t=2} = \boxed{12}$$

All together, we get

$$\left(\frac{dT}{dt}\right)\Big|_{t=2} = (-10e^{-31})(4) + (-20e^{-31})(12) + (-40e^{-31})\left(-\frac{1}{4}\right) = \boxed{-240e^{-31}} \ {}^\circ F/s$$

Where did this formula come from?

Notice we could rewrite $\frac{dw}{dt} = \left(\frac{\partial w}{\partial x}\right)\left(\frac{dx}{dt}\right) + \left(\frac{\partial w}{\partial y}\right)\left(\frac{dy}{dt}\right) + \left(\frac{\partial w}{\partial z}\right)\left(\frac{dz}{dt}\right)$ as

$$\nabla w(\vec{c}(t)) \bullet \vec{c}\,'(t) \quad \text{where} \quad \vec{c}(t) = (x(t), y(t), z(t))$$

Recall that ∇w and $\vec{c}\,'(t)$ are just special notations for the total derivatives $[Dw]$ and $[D\vec{c}]$.

So we could further rewrite the above equation as

$$[D(w \circ \vec{c})](t) = [Dw](\vec{c}(t))\,[D\vec{c}](t)$$

This is a 1x3 matrix

This is 3x1 matrix

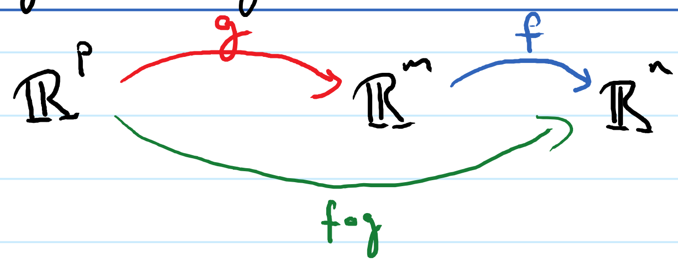
matrix mult. makes sense here

The (total) derivative of $w \circ \vec{c}$ at t is the total derivative of w, evaluated at $\vec{c}(t)$ times the (total) derivative of $\vec{c}$ evaluated at t. This is essentially the same as in single variable, except now you multiply matrices!

More generally:

Theorem: (Chain Rule): Suppose $f: \mathbb{R}^n \to \mathbb{R}^m$, $g: \mathbb{R}^p \to \mathbb{R}^{n'}$ are both diff'ble. Then

$(f \circ g)(\vec{x}) = f(g(\vec{x}))$ is a fcn from $\mathbb{R}^p$ to $\mathbb{R}^n$



If $\vec{x}$ is a p-dim. vector, $g(\vec{x})$ is m-dim. so it makes sense to plug it into f. $f(g(\vec{x}))$ is an n-dim. vector.

and
$$[D(f \circ g)](\vec{x}) = [Df](g(\vec{x}))\,[Dg](\vec{x}).$$

$[D(f \circ g)](\vec{x})$: $n \times p$ matrix

$[Df](g(\vec{x}))$: $n \times m$ matrix

$[Dg](\vec{x})$: $m \times p$ matrix

matrix mult. makes sense

Example: Let $g:\mathbb{R}^2 \to \mathbb{R}^2$, $g(r,\theta) = (r\cos\theta, r\sin\theta)$

$f:\mathbb{R}^2 \to \mathbb{R}^3$, $f(x,y) = (y^2, x+y, x-y^2)$

Compute: ① $[Dg](2,\frac{\pi}{2})$

② $[Df](g(2,\frac{\pi}{2}))$

③ $[D(f \circ g)](2,\frac{\pi}{2})$.

① $$[Dg] = \begin{bmatrix} \frac{\partial g_1}{\partial r} & \frac{\partial g_1}{\partial \theta} \\ \frac{\partial g_2}{\partial r} & \frac{\partial g_2}{\partial \theta} \end{bmatrix} = \begin{bmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{bmatrix}$$

$$[Dg](2,\tfrac{\pi}{2}) = \begin{bmatrix} \cos(\frac{\pi}{2}) & -2\sin(\frac{\pi}{2}) \\ \sin(\frac{\pi}{2}) & 2\cos(\frac{\pi}{2}) \end{bmatrix} = \begin{bmatrix} 0 & -2 \\ 1 & 0 \end{bmatrix}$$

② $$[Df] = \begin{bmatrix} 0 & 2y \\ 2x & 1 \\ 1 & -2y \end{bmatrix}$$

Notice: $g(2, \frac{\pi}{2}) = (2\cos(\frac{\pi}{2}), 2\sin(\frac{\pi}{2}))$
$= (0, 2)$

$$[Df](g(2, \tfrac{\pi}{2})) = [Df](0,2) = \begin{bmatrix} 0 & 4 \\ 0 & 1 \\ 1 & -4 \end{bmatrix}$$

③ $[D(f \circ g)](2, \frac{\pi}{2}) = [Df](g(2,\frac{\pi}{2}))\,[Dg](2, \frac{\pi}{2})$

$$= \begin{bmatrix} 0 & 4 \\ 0 & 1 \\ 1 & -4 \end{bmatrix} \begin{bmatrix} 0 & -2 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0+(4)(1) & (0)(-2)+(4)(0) \\ (0)(0)+(1)(1) & (0)(-2)+(1)(0) \\ (1)(0)+(-4)(1) & (1)(-2)+(-4)(0) \end{bmatrix}$$

$$= \begin{bmatrix} 4 & 0 \\ 1 & 0 \\ -4 & -2 \end{bmatrix}$$

Remark: If we write $(f \circ g)(r,\theta) = ((f \circ g)_1(r,\theta), (f \circ g)_2(r,\theta), (f \circ g)_3(r,\theta))$.
Then since

$$[D(f \circ g)](2, \pi/2) = \begin{bmatrix} 4 & 0 \\ 1 & 0 \\ -4 & -2 \end{bmatrix}$$

we have, for example:

$$\left(\frac{\partial (f \circ g)_1}{\partial r}\right)(2, \tfrac{\pi}{2}) = 4$$

$$\left(\frac{\partial (f \circ g)_3}{\partial \theta}\right)(2, \tfrac{\pi}{2}) = -2$$

I.e., this matrix carries all the information about all the partial derivatives of $f \circ g$.

Shortcut diagram for partial derivatives / implicit differentiation / tangent planes to surfaces.