PARTIAL DERIVATIVES AND THE CHAIN RULE - M. Hopkins

1. Consider the physics formula $a = \frac{F}{m}$, i.e. the acceleration of an object is directly proportional to the applied force F and inversely proportional to its mass m.

To answer a question like "how does the acceleration change as we vary the applied force?", we can compute the partial derivative $\frac{\partial a}{\partial F}$, which is the usual derivative of a with respect to F, given that we treat the other quantities (i.e. mass m) as constants.

$$\frac{\partial a}{\partial F} = \frac{1}{m}$$

2. This is fine, but what about a similar formula describing the internal air pressure of a balloon as a function of altitude and temperature:

$$p = \frac{t}{a^2}$$

To answer a question like "how does the pressure change as we change the altitude?", we could compute $\frac{\partial p}{\partial a}$...

PARTIAL DERIVATIVES AND THE CHAIN RULE – M. Hopkins

③ ... but what if the temperature also changes when we change the altitude? e.g. say that $t = -a$.

Then there are two partial derivatives that might make sense:

$\left(\frac{\partial p}{\partial a}\right)_t$ denotes the derivative of p wrt a if we keep t fixed

$\frac{\partial p}{\partial a}$ denotes the derivative of p wrt a if we allow t to change as a function of t

---

④ This can all get rather confusing. It helps to first specify a causal network over the variables, e.g.:

$$\begin{tikzcd} F \arrow[dr] & & m \arrow[dl] \\ & a & \end{tikzcd} \qquad \text{or} \qquad \begin{tikzcd} a \arrow[r] \arrow[d] & t \arrow[dl] \\ p & \end{tikzcd}$$

Here, the edges have a causal interpretation, specifically a node is a function of its parents. Also, we assume that if I change a variable, then the values of its descendants are updated, but the values of its nondescendant remain the same.

For instance, if I change the altitude (in the network above), then the temperature changes. But if I change the temperature, the altitude stays the same.

⑤ This interpretation makes it easier to talk about partial derivatives. In the network:

$$\begin{tikzcd} a \arrow[r] \arrow[d] & t \arrow[dl] \\ p & \end{tikzcd}$$

changing a causes p to change in two different ways, each corresponding to a path from a to p:

$$\begin{tikzcd} a \arrow[r] \arrow[d, very thick] & t \arrow[dl] \\ p & \end{tikzcd} \quad \text{and} \quad \begin{tikzcd} a \arrow[r, very thick] \arrow[d] & t \arrow[dl, very thick] \\ p & \end{tikzcd}$$

Since we want to know about the total effect of changing a on the value of p, we define $\frac{\partial a}{\partial p}$ as the derivative of p wrt a, keeping all nondescendants (here, nothing) fixed.

(6) To hammer home the value of the network, consider the following equations:

$$x = u$$
$$y = u + v$$
$$z = x + y$$

Now let's compute $\frac{\partial z}{\partial x}$ and $\frac{\partial z}{\partial u}$.

$$\frac{\partial z}{\partial x} = \frac{\partial (x+y)}{\partial x} = 1 + 0 = 1$$

$$\frac{\partial z}{\partial u} = \frac{\partial (x+y)}{\partial u} = \frac{\partial x}{\partial u} + \frac{\partial y}{\partial u} = \frac{\partial u}{\partial u} + \frac{\partial (u+v)}{\partial u} = 1 + 1 + 0 =$$

It's possibly confusing that $\frac{\partial z}{\partial x} \neq \frac{\partial z}{\partial u}$, given $x = u$.

It only makes sense under a causal (assymetric) interpretation of $x = u$ as $x := u$:

$$\begin{tikzcd}
u \arrow[r] \arrow[dr] & x \arrow[dr] & \\
& & z \\
v \arrow[r] & y \arrow[ur] &
\end{tikzcd}$$

Under this interpretation, u affects z through two different causal paths, while x affects z through only one, explaining the difference.

(7) In a causal network like

$$\begin{tikzcd} u \arrow[r] \arrow[dr] & x \arrow[dr] & \\ & & z \\ v \arrow[r] & y \arrow[ur] & \end{tikzcd} \qquad \begin{aligned} x &= u \\ y &= u+v \\ z &= x+y \end{aligned}$$

we only directly express the value of a variable in terms of its parents. The Chain Rule of Partial Derivatives makes it easier to compute derivatives like $\frac{\partial z}{\partial u}$, where we don't have a direct expression of z in terms of u.

(8) Specifically the Chain Rule says the following. Suppose S is a set of variables that separate variable b from its ancester a in the causal network. Then: $$\frac{\partial b}{\partial a} = \sum_{s \in S} \frac{\partial b}{\partial s} \frac{\partial s}{\partial a}$$

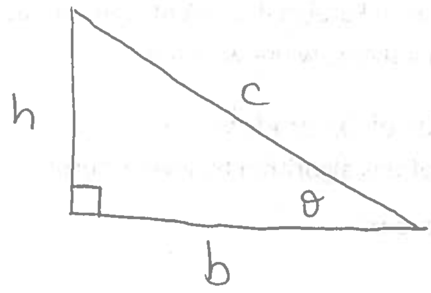
(9) For instance, variables x and y separate z from u in the above network.

$$\begin{tikzcd} u \arrow[r] \arrow[dr] & x \arrow[dr] & \\ & & z \\ v \arrow[r] & y \arrow[ur] & \end{tikzcd}$$

From the Chain Rule: $$\frac{\partial z}{\partial u} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial u} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial u} = 1 \cdot 1 + 1 \cdot 1 = 2$$

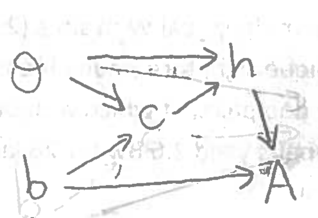
PARTIAL DERIVATIVES AND THE CHAIN RULE – M. Hopkins

(10) Let's exercise our newfound partial derivative skills to solve the following problem. Consider the following right triangle:



How does the area of the right triangle change as we modify θ?

(11) First we set up a causal network that models the problem:



$$c = \frac{b}{\cos\theta} \qquad h = c\sin\theta$$

$$A = \frac{1}{2}bh$$

By making c and h descendants of θ and b, we ensure the network can only model right triangles.

(for a right triangle, $\cos\theta = \frac{b}{c} \Rightarrow c = \frac{b}{\cos\theta}$

and $\sin\theta = \frac{h}{c} \Rightarrow h = c\sin\theta$)

The area A is simply the usual formula: half base times height for a triangle

(12) Now let's try to compute $\dfrac{\partial A}{\partial \theta}$.

Because h separates A from θ, we can apply the Chain Rule:

$$\frac{\partial A}{\partial \theta} = \frac{\partial A}{\partial h}\frac{\partial h}{\partial \theta}$$

$$= \frac{b}{2} \cdot \frac{\partial h}{\partial \theta} \longleftarrow \text{what's this?}$$

(13) Next we need to compute $\dfrac{\partial h}{\partial \theta}$. Unfortunately there are no separating sets that separate h from θ. We can however fix this by introducing a copy θ' of θ into the network:

$$\begin{tikzcd}
\theta \arrow[r] \arrow[dr] & \theta' \arrow[r] & h \arrow[dd] \\
 & c \arrow[ur] & \\
b \arrow[ur] \arrow[rr] & & A
\end{tikzcd}$$

$\theta' = \theta \qquad h = c \sin \theta'$

$c = \dfrac{b}{\cos \theta} \qquad A = \dfrac{1}{2} bh$

Now the set $\{\theta', c\}$ separates h from θ, so by the Chain Rule:

$$\frac{\partial h}{\partial \theta} = \frac{\partial h}{\partial \theta'} \cdot \frac{\partial \theta'}{\partial \theta} + \frac{\partial h}{\partial c} \cdot \frac{\partial c}{\partial \theta}$$

$$= c \cos \theta' \cdot 1 + \sin \theta' \cdot \left(\frac{b \tan \theta}{\cos \theta}\right)$$

$$= c \cos \theta + b \tan \theta \left(\frac{\sin \theta}{\cos \theta}\right)$$

$$= b + b \tan^2 \theta \qquad \left[\text{b/c } c = \frac{b}{\cos \theta} \text{ (above) and } \frac{\sin \theta}{\cos \theta} = \tan \theta \text{ (trig ident)} \right]$$

(14) Putting it all together, we get:

$$\frac{\partial A}{\partial \theta} = \frac{b}{2} \cdot \frac{\partial h}{\partial \theta}$$

$$= \frac{b}{2}\left(b + b\tan^2\theta\right)$$

$$= \frac{b^2}{2}\left(1 + \tan^2\theta\right)$$

$$= \frac{b^2 \sec^2\theta}{2}$$ [b/c $\sec^2\theta = 1 + \tan^2\theta$ is a trig ident]