Math 3220 # The Chain Rule **March 2019**

We all know that the chain rule is the main work horse for computing derivatives. Other rules, like linearity and the product rule, are very striaght forward and rather easy to prove. The chain rule is a bit more slippery. First, we will give an "approximate proof".

Suppose that g is differentiable at $x = a$ and f is differentiable at $x = g(a)$. This means that f and g can be well approximated (near $x = g(a)$ and $x = a$ respectively) by their *linearizations* (whose graphs are tangent lines). Specifically, we have that...

$$g(x) \approx g(a) + g'(a)(x-a) \qquad \text{and} \qquad f(X) \approx f(g(a)) + f'(g(a))(X - g(a))$$

Now we want to compose f and g (let $X = g(x)$ and then substitute in g's linearization):

$$\begin{aligned}(f\circ g)(x) = f(g(x)) &\approx f(g(a)) + f'(g(a))(g(x)-g(a)) \approx f(g(a)) + f'(g(a))\big((g(a)+g'(a)(x-a)) - g(a)\big) \\ &\approx f(g(a)) + f'(g(a))g'(a)(x-a)\end{aligned}$$

Therefore, the derivative of the composition *should* be: $(f\circ g)'(a) = f'(g(a))g'(a)$. Let's now prove this carefully/formally. First, we recall a lemma:

Lemma Given, $g'(a) \neq 0$, there is some $\delta > 0$ such that for all $0 < |x-a| < \delta$, $g(x) \neq g(a)$.

proof: Let $\epsilon = |g'(a)|$ (this is positive since $g'(a) \neq 0$). There exists some $\delta > 0$ such that for all $0 < |x-a| < \delta$ we have $\left|\dfrac{g(x)-g(a)}{x-a} - g'(a)\right| < \epsilon = |g'(a)|$. Thus, $g'(a) - |g'(a)| < \dfrac{g(x)-g(a)}{x-a} < g'(a) + |g'(a)|$. If $g'(a) > 0$, then $g'(a) - |g'(a)| = g'(a) - g'(a) = 0$. If $g'(a) < 0$, then $g'(a) + |g'(a)| = g'(a) - g'(a) = 0$. So either way, we get that $\dfrac{g(x)-g(a)}{x-a} \neq 0$. Therefore, $g(x) - g(a) \neq 0$. We have shown that $g(x) \neq g(a)$ for all $0 < |x-a| < \delta$. ■

Theorem [Chain Rule]: Let g be differentiable at $x = a$ and let f be differentiable at $x = g(a)$. Then $f \circ g$ is differentiable at $x = a$. Moreover, $(f\circ g)'(a) = f'(g(a))g'(a)$.

proof: We consider 2 cases: $g'(a) \neq 0$ and $g'(a) = 0$.
Case 1: Suppose that $g'(a) \neq 0$. By our lemma, there is some $\delta > 0$ such that $g(x) \neq g(a)$ for all $x \in I = (a-\delta, a+\delta)$ and $x \neq a$. We restrict the domain of g to this interval I.

$$(f\circ g)'(a) = \lim_{x\to a} \frac{(f\circ g)(x) - (f\circ g)(a)}{x-a} = \lim_{x\to a} \frac{f(g(x)) - f(g(a))}{x-a} \cdot \frac{g(x)-g(a)}{g(x)-g(a)}$$

[*Note:* Since the limit is computed with x's such that $x \neq a$ but near a, we can assume $x \in I$ and $x \neq a$ so that $g(x) - g(a) \neq 0$. Thus we are not dividing by zero!]

$$= \lim_{x\to a} \frac{f(g(x)) - f(g(a))}{g(x)-g(a)} \cdot \frac{g(x)-g(a)}{x-a} = \lim_{x\to a} \frac{f(g(x)) - f(g(a))}{g(x)-g(a)} \cdot \lim_{x\to a} \frac{g(x)-g(a)}{x-a} = f'(g(a))g'(a)$$

In the above calculation, the limits exist because of our assumption that f and g are differentiable at $x = g(a)$ and $x = a$ respectively as well as the fact that the limit of the product is the product of the limits (given limits exist).

Case 2: Suppose that $g'(a) = 0$. This means that $f'(g(a))g'(a) = f'(g(a)) \cdot 0 = 0$, so our goal is to show that $(f\circ g)'(a) = 0$.

Since f is differentiable at $X = g(a)$, the limit $\displaystyle\lim_{X\to g(a)} \frac{f(X) - f(g(a))}{X - g(a)}$ exists. In particular, let $\epsilon = 1$. Then there exists $\delta' > 0$ such that $\left|\dfrac{f(X)-f(g(a))}{X-g(a)} - f'(g(a))\right| < 1$ whenever $0 < |X - g(a)| < \delta'$. Thus letting $M = |f'(g(a))| + 1$, we have $\left|\dfrac{f(X)-f(g(a))}{X-g(a)}\right| < M$ whenever $0 < |f(X) - f(g(a))| < \delta'$. Therefore, $|f(X) - f(g(a))| \leq M|x - g(a)|$ whenever $|X - g(a)| < \delta'$. Note that $M > 0$.

Now we embark on showing our derivative is 0 using the limit definition: Let $\epsilon > 0$, so $\epsilon/M > 0$. Now since $g'(a) = 0$, there exists some $\delta'' > 0$ such that $\left|\dfrac{g(x)-g(a)}{x-a} - 0\right| < \epsilon/M$ whenever $0 < |x-a| < \delta''$. Therefore, $|g(x) - g(a)| \leq (\epsilon/M)|x-a|$ whenever $|x-a| < \delta''$.

Next, since g is differentiable at $x = a$, it is also continuous at $x = a$. This means that there is some $\delta''' > 0$ such that whenever $|x-a| < \delta'''$ we have $|g(x) - g(a)| < \delta'$ (here δ' is our "ϵ").

Now let $\delta = \min\{\delta'', \delta'''\}$ and suppose that $|x-a| < \delta$. This implies that $|x-a| < \delta'''$ so that $|g(x) - g(a)| < \delta'$. This in turn implies that $|f(g(x)) - f(g(a))| \leq M|g(x) - g(a)|$. But $|g(x) - g(a)| < (\epsilon/M)|x-a|$ since $|x-a| < \delta''$. Therefore, $|f(g(x)) - f(g(a))| \leq M(\epsilon/M)|x-a| = \epsilon|x-a|$.

Finally, consider $0 < |x-a| < \delta$. Then $x - a \neq 0$ and $|f(g(x)) - f(g(a))| \leq \epsilon|x-a|$ so that $\left|\dfrac{f(g(x)) - f(g(a))}{x-a}\right| \leq \epsilon$.

Therefore, $\displaystyle\lim_{x\to a} \frac{f(g(x)) - f(g(a))}{x-a} = 0$ (i.e., $(f\circ g)'(a) = 0$) as we wanted to show. ■