Handout 9: Some additional notes.

1. The chain rule.

The proof in the book for the chain rule is probably better than the one I gave in class. Recall that a composition of continuous functions $G(f(x))$ is continuous since $x_n \to c \Rightarrow f(x_n) \to f(c) \Rightarrow G(f(x_n)) \to G(f(c))$.

Theorem (Chain rule) Suppose that $f(x)$ is differentiable at $x = c$ and $g(y)$ is differentiable at $y = f(c)$. Then $(g \circ f)(x) = g(f(x))$ is differentiable at $x = c$, and $(g \circ f)'(c) = g'(f(c))f'(c)$.

Proof: We wish to use the formula

$$\frac{g(f(x)) - g(f(c))}{x - c} = \frac{g(f(x)) - g(f(c))}{f(x) - f(c)} \frac{f(x) - f(c)}{x - c}.$$

but this makes sense only if we know $f(x) - f(c) \neq 0$. That need not be the case even if x is very close to c. We handle this as follows.

The function

$$G_0(y) = \frac{g(y) - g(f(c))}{y - f(c)}$$

is not defined for $y = f(c)$, but it converges as $y \to f(c)$ since

$$\lim_{y \to f(c)} G_0(y) = g'(f(c))$$

We extend $G_0(y)$ to a function G defined at $f(c)$ by letting $G(f(c)) = g'(f(c))$. It is evident that G is continuous at $f(c)$. We have that

$$g(y) - g(f(c)) = G_0(y)(y - f(c))$$

for $y \neq c$, hence

$$g(y) - g(f(c)) = G(y)(y - f(c))$$

even when $y = c$ (since both sides are zero when $y = f(c)$). We conclude that

$$\frac{g(f(x)) - g(f(c))}{x - c} = G(f(x)) \frac{f(x) - f(c)}{x - c}.$$

Since f is differentiable at c, it is continuous at c, hence $G(f(x))$ is continuous at c. We conclude that

$$\lim_{x \to c} \frac{g(f(x)) - g(f(c))}{x - c} = G(f(c))f'(c) = g'(f(c)).$$

2. The Schwarz inequality and the fact that $\|\|_2$ is a norm on $\mathbb{R}^d$.

Recall that if $v = (x_1, \ldots, x_d)$ and $w = (y_1, \ldots, y_d) \in \mathbb{R}^d$, then the dot product $v \cdot w$ is defined by

$$v \cdot w = \sum_{j=1}^{d} v_j w_j,$$

and the corresponding norm $\| \, \|_2$ on $\mathbb{R}^d$ is defined by $\| v\|_2 = (v \cdot v)^{1/2}$

Lemma:For all $v, w \in \mathbb{R}^d$,

$$v \cdot w \leq \frac{\|v\|_2^2 + \|w\|_2^2}{2}.$$

Proof: For any vector $u \in \mathbb{R}^d$, $u \cdot u = \sum u_j^2 \geq 0$, and thus if $u = v - w$,

$$0 \leq (v - w) \cdot (v - w) = v \cdot v + w \cdot w - 2(v \cdot w).$$

Theorem For all $v, w \in \mathbb{R}^d, |v \cdot w| \leq \|v\|_2 \|w\|_2$.

Proof: In general for any $c, d \in \mathbb{R}$, $(cv) \cdot (dw) = (cd)(v \cdot w)$. Thus for any $t > 0$

$$v \cdot w = (t^{1/2} v) \cdot (t^{-1/2} w) \leq \frac{t \|v\|_2^2 + t^{-1} \|w\|_2^2}{2}$$

Let us assume that $\|v\|_2$ and $\|w\|_2 \neq 0$. The result follows by letting $t = \|w\|_2 / \|v\|_2$. If either v or w is zero, both sides of the Schwarz inequality are zero, so it is still true. QED

Finally the only hard part of showing $\| \|_2$ is indeed a norm is the triangle inequality:

$$\|v+w\|_2^2 = (v+w)\cdot(v+w) = \|v\|_2^2 + \|w\|_2^2 + 2v\cdot w \leq \|v\|_2^2 + \|w\|_2^2 + 2\,\|v_2\|_2\,\|w\|_2 = (\|v\|_2 + \|w\|_2)^2,$$

and the desired result follows when one takes the square root of both sides.