Directional Derivatives. The partial derivatives $\partial_j f$ give information about how $f(\mathbf{x})$ varies as $\mathbf{x}$ moves along lines parallel to the coordinate axes. Sometimes we wish to study the variation of f along oblique lines instead. Thus, given a unit vector $\mathbf{u}$ and a base point $\mathbf{a}$, we consider the line passing through $\mathbf{a}$ in the direction $\mathbf{u}$, which can be represented parametrically by $\mathbf{g}(t) = \mathbf{a} + t\mathbf{u}$. The **directional derivative** of f at $\mathbf{a}$ in the direction $\mathbf{u}$ is defined to be

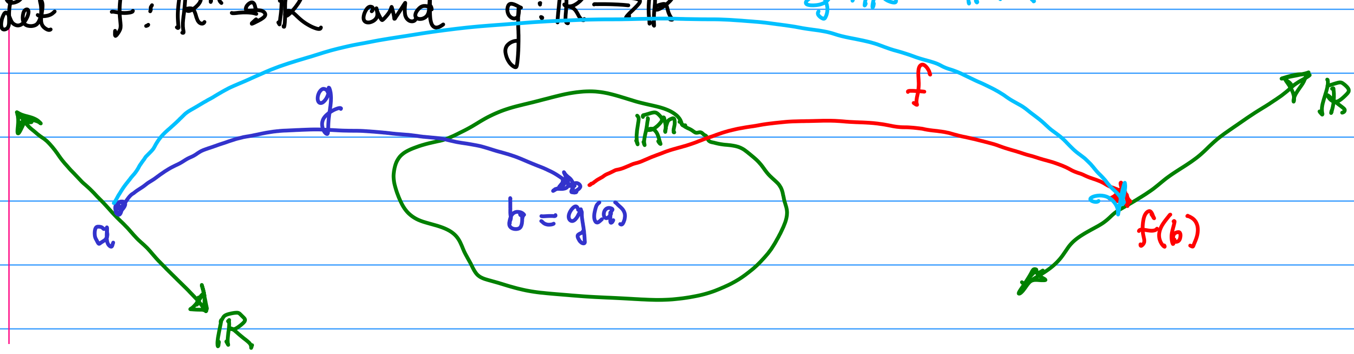
$$\partial_{\mathbf{u}} f(\mathbf{a}) = \frac{d}{dt} f(\mathbf{a} + t\mathbf{u})\Big|_{t=0} = \lim_{t \to 0} \frac{f(\mathbf{a} + t\mathbf{u}) - f(\mathbf{a})}{t},$$

provided that the limit exists. For example, if $\mathbf{u}$ is the unit vector in the positive

Theorem $\partial_u f(a) = \nabla f(a) \cdot u$

2.23 Theorem. *If f is differentiable at $\mathbf{a}$, then the directional derivatives of f at $\mathbf{a}$ all exist, and they are given by*

Let $f: \mathbb{R}^n \to \mathbb{R}$ and $g: \mathbb{R} \to \mathbb{R}^n$ $\quad f \circ g: \mathbb{R} \to \mathbb{R}$.



2.26 Theorem (Chain Rule I). *Suppose that $\mathbf{g}(t)$ is differentiable at $t = a$, $f(\mathbf{x})$ is differentiable at $\mathbf{x} = \mathbf{b}$, and $\mathbf{b} = \mathbf{g}(a)$. Then the composite function $\varphi(t) = f(\mathbf{g}(t))$ is differentiable at $t = a$, and its derivative is given by*

$$\varphi'(a) = \nabla f(\mathbf{b}) \cdot \mathbf{g}'(a),$$

Chain rule in 1D.. Idea..

$$\lim_{h\to 0} \frac{f\circ g(x+h) - f\circ g(x)}{h} = \lim_{h \to 0} \underbrace{\frac{f\circ g(x+h) - f\circ g(x)}{g(x+h) - g(x)}}_{f'(g(x))} \lim_{h\to 0} \underbrace{\frac{g(x+h)-g(x)}{h}}_{g'(x)}$$

difficulty is that this denominator might be zero

Let $f: \mathbb{R}^n \to \mathbb{R}$ and $g: \mathbb{R} \to \mathbb{R}^n$

2.26 Theorem (Chain Rule I). *Suppose that* $\mathbf{g}(t)$ *is differentiable at* $t = a$, $f(\mathbf{x})$ *is differentiable at* $\mathbf{x} = \mathbf{b}$, *and* $\mathbf{b} = \mathbf{g}(a)$. *Then the composite function* $\varphi(t) = f(\mathbf{g}(t))$ *is differentiable at* $t = a$, *and its derivative is given by*

$$\varphi'(a) = \nabla f(\mathbf{b}) \cdot \mathbf{g}'(a),$$

The point is $E_1(0)$ actually makes sense?

Proof: f differentiable at b means (and g diff. at a means)

$$f(b+h) = f(b) + \nabla f(b)\cdot h + E_1(h) \quad \text{where} \quad \frac{E_1(h)}{|h|} \to 0 \text{ as } h \to 0.$$

$$g(a+u) = g(a) + g'(a)u + E_2(u) \quad \text{where} \quad \frac{|E_2(u)|}{u} \to 0 \text{ as } u \to 0.$$

$g(a+u) - g(a) = g'(a)\cdot u + E_2(u)$

want to show $\varphi(t) = f\circ g(t)$ is differentiable at a.

$$\varphi(a+u) = \varphi(a) + \boxed{\varphi'(a)}u + \boxed{E_3(u)} \quad \text{where} \quad \frac{E_3(u)}{u} \to 0 \text{ as } u \to 0.$$

need to show this

$$\varphi(a+u) = f(g(a+u)) = f(b+h) = f(b) + \nabla f(b)\cdot h + E_1(h)$$

note $g(a+u) = b+h = g(a)+h$ so $h = g(a+u) - g(a)$

$$= f(b) + \nabla f(b)\cdot \big(g(a+u) - g(a)\big) + E_1(h)$$

$$= f(b) + \nabla f(b)\cdot \big(g'(a)u + E_2(u)\big) + E_1(h)$$

$$= f(b) + \boxed{\nabla f(b)\cdot g'(a)}u + \boxed{\nabla f(b)\cdot E_2(u) + E_1(h)}$$

Claim setting $\varphi'(a) = \nabla f(b)\cdot g'(a)$ and $E_3(u) = \nabla f(b)\cdot E_2(u) + E_1(h)$ leads to being able to show $\frac{E_3(u)}{u} \to 0$ as $u \to 0$.

$$\left|\frac{\nabla f(b)\cdot E_2(u)}{u}\right| \le \frac{|\nabla f(b)|\cdot|E_2(u)|}{|u|} = \underbrace{|\nabla f(b)|}_{\text{const}}\underbrace{\left|\frac{E_2(u)}{u}\right|}_{\text{goes to zero}} \to 0 \text{ as } u \to 0$$

Now consider

recall $h = g(a+u) - g(a)$

$$\left|\frac{E_1(h)}{u}\right| = \left|\frac{E_1(g(a+u)-g(a))}{u}\right|$$

Idea from 310

$$= \left|\frac{E_1(g(a+u)-g(a))}{\underbrace{g(a+u)-g(a)}_{\text{might be zero...}}}\right|\left|\frac{g(a+u)-g(a)}{u}\right|$$

From text...

$$|\mathbf{h}| = |u\mathbf{g}'(a) + \mathbf{E}_2(u)| \leq \big(|\mathbf{g}'(a)| + 1\big)|u|.$$

instead of solving for h, let's bound it ...

Thus,

$$|h| = |g(a+u)-g(a)| = |g'(a)u + E_2(u)| \leq |g'(a)||u| + |E_2(u)|$$

By hypothesis $\frac{|E_2(u)|}{u} \to 0$ as $u \to 0$.

$\forall \varepsilon > 0 \ \exists \delta > 0$ s.t $|u| < \delta$ implies $\left|\frac{E_2(u)}{u}\right| < \varepsilon$.

Choose $\varepsilon = 1$ $\exists \delta > 0$ s.t $|u| < \delta$ implies $\left|\frac{E_2(u)}{u}\right| < 1$.

$$\left|\frac{E_2(u)}{u}\right| < 1 \quad \text{means} \quad |E_2(u)| < |u|$$

It follows

$$|h| \leq |g'(a)||u| + |u| = (|g'(a)| + 1)|u|$$

this inequality means $h \to 0$ as $u \to 0$

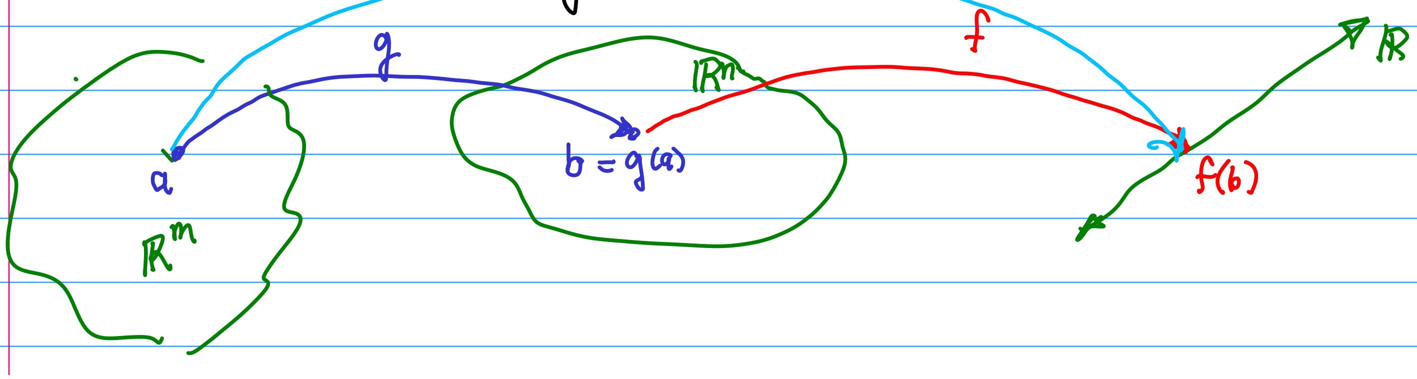
Now... provided $h \neq 0$ then...

$$\left|\frac{E_1(h)}{u}\right| = \underbrace{\frac{|E_1(h)|}{|h|}\frac{|h|}{|u|}}_{\text{here we are ignoring the exact dep. of } h \text{ on } u.} \leq \frac{|E_1(h)|}{|h|}\underbrace{(|g'(a)|+1)}_{\text{constant..}}$$

only using this ineq.

$$\lim_{u \to 0} \left| \frac{E_1(h)}{u} \right| = \lim_{h \to 0} \frac{|E_1(h)|}{|h|} \left(|g'(a)| + 1 \right) = 0 \qquad \square$$

Let $f : \mathbb{R}^n \to \mathbb{R}$ and $g : \mathbb{R}^m \to \mathbb{R}^n$ ~~$g : \mathbb{R} \to \mathbb{R}$.~~



2.29 Theorem (Chain Rule II). *Suppose that $g_1, \ldots, g_n$ are functions of $\mathbf{t} = (t_1, \ldots, t_m)$ and f is a function of $\mathbf{x} = (x_1, \ldots, x_n)$. Let $\mathbf{b} = \mathbf{g}(\mathbf{a})$ and $\varphi = f \circ \mathbf{g}$. If $g_1, \ldots, g_n$ are differentiable at $\mathbf{a}$ (resp. of class C^1 near $\mathbf{a}$) and f is differentiable at $\mathbf{b}$ (resp. of class C^1 near $\mathbf{b}$), then φ is differentiable at $\mathbf{a}$ (resp. of class C^1 near $\mathbf{a}$), and its partial derivatives are given by*

$$\frac{\partial \varphi}{\partial t_k} = \frac{\partial f}{\partial x_1}\frac{\partial x_1}{\partial t_k} + \cdots + \frac{\partial f}{\partial x_n}\frac{\partial x_n}{\partial t_k}, \tag{2.30}$$

where the derivatives $\partial f / \partial x_j$ are evaluated at $\mathbf{b}$ and the derivatives $\partial \varphi / \partial t_k$ and $\partial x_j / \partial t_k = \partial g_j / \partial t_k$ are evaluated at $\mathbf{a}$.

Proof is the same ….