

The Derivative Mapping on \mathbb{R}^n

305a

These notes attempt an improvement of my notes 306 → 316. This version connects the general derivative mapping (Frechet Derivative) to the "directional derivative". I start with a few motivations from calculus I then give the general def². The other version builds up to the general case by looking at $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ and $f: \mathbb{R}^3 \rightarrow \mathbb{R}$ to begin. (read 306-311 for more)

~~H~~ (examples given in other notes)

CALCULUS I : LINEARIZATION \Leftrightarrow DERIVATIVE

Consider that the tangent line approximates the function near the base point $(a, f(a))$

$$f(a+h) \approx f(a) + m h = L_f(a)(h)$$

We insist this approximation hold to first order,

$$\lim_{h \rightarrow 0} \left[\frac{f(a+h) - L_f(a)(h)}{h} \right] = 0$$

this says $L_f(a)(h)$ is the best linear approximation to $f(x)$ near $x=a$

In other words,

$$\lim_{h \rightarrow 0} \left(\frac{f(a+h) - f(a) - mh}{h} \right) = 0$$

$$\lim_{h \rightarrow 0} \left(\frac{f(a+h) - f(a)}{h} \right) = \lim_{h \rightarrow 0} \left(\frac{mh}{h} \right) = \lim_{h \rightarrow 0} (m) = m.$$

We find that we could interchange the derivative $f'(a)$ with the linearization $L_f(a)$.

$$L_f(a)(h) = f(a) + mh$$

is best linear approximation
to $y = f(x)$ at $x=a$

$$\Leftrightarrow \frac{df}{dx}(a) = \lim_{h \rightarrow 0} \left[\frac{f(a+h) - f(a)}{h} \right]$$

These comments help motivate the def² to follow.

We found that in principle the derivative is contained within the best linear approximation for some function. It turns out that this gives us a way to implicitly define the derivative for $f: U \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$.

Defⁿ / Let $U \subseteq \mathbb{R}^n$ be open and $f: U \rightarrow \mathbb{R}^m$. Let $\vec{a} \in U$ then f is differentiable at \vec{a} if $Df(\vec{a}): \mathbb{R}^n \rightarrow \mathbb{R}^m$ is linear with the function $L_f(\vec{a}): \mathbb{R}^n \rightarrow \mathbb{R}^m$ defined by

$$L_f(\vec{a})(\vec{h}) = f(\vec{a}) + Df(\vec{a})(\vec{h})$$

being the best linear approximation to f at \vec{a} , meaning:

$$\lim_{\vec{h} \rightarrow 0} \frac{\|f(\vec{a} + \vec{h}) - f(\vec{a}) - Df(\vec{a})(\vec{h})\|_m}{\|\vec{h}\|_n} = 0$$

$$\text{Here } \| \langle x_1, x_2, \dots, x_m \rangle \|_m = \sqrt{x_1^2 + x_2^2 + \dots + x_m^2} = \sqrt{\vec{x} \cdot \vec{x}}$$

$$\text{and } \|\vec{y}\|_n = \sqrt{\vec{y} \cdot \vec{y}}$$

This definition is quite general. Let me list what it encompasses in various cases:

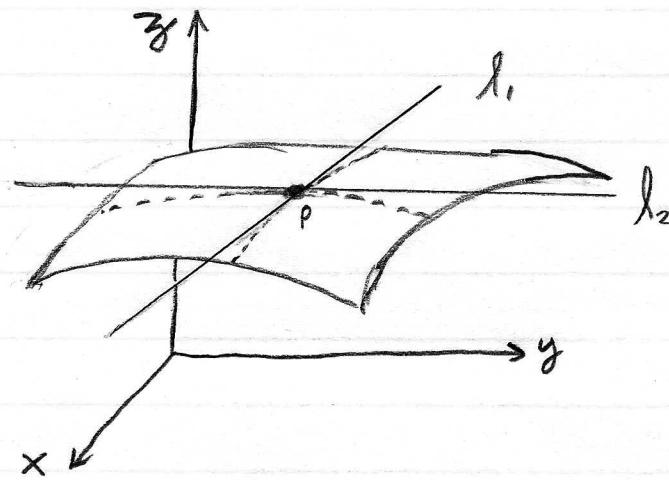
- i.) $m = n = 1$: calculus I : $\frac{df}{dx}$
- ii.) $m = 3, n = 1$: derivative of space curve : $\frac{d\vec{r}}{dt}$
- iii) $m = 2, n = 2$: directional derivative and $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \nabla f$
- iv.) $m = 3, n = 3$: directional derivative, $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}, \nabla f$
- v.) $m = n = N$: directional derivative, $\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \dots, \frac{\partial f}{\partial x_n}, \nabla f$

all of these derivatives are hidden away inside this general Frechet derivative. We already saw i.) as a motivation. I'll return to show how iii) falls out later.

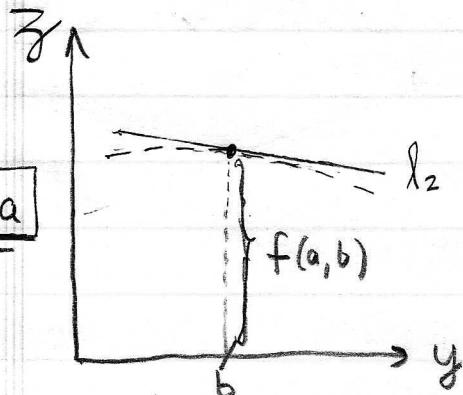
Geometry of $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ and Tangent Plane

(305c)

Together they give the tangent plane once we pick a point. Let's verify that via the picture on pg. 292 plus some analytic geometry.

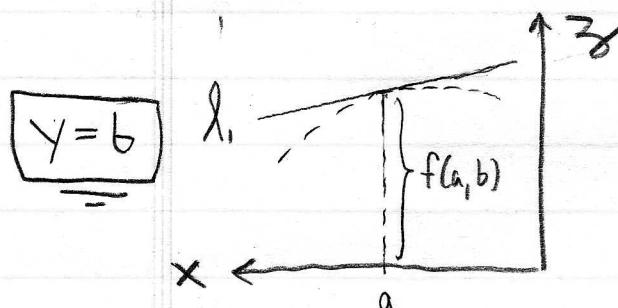


Look at the cross-sections containing l_1 & l_2



$$z = f(a, b) + \frac{\partial f}{\partial y} \Big|_{(a, b)} (y - b)$$

l_2 has slope $\frac{\partial f}{\partial y}(a, b)$ in the $x = a$ plane.



$$z = f(a, b) + \frac{\partial f}{\partial x} \Big|_{(a, b)} (x - a)$$

l_1 has slope $\frac{\partial f}{\partial x}(a, b)$ in the $y = b$ plane

We can parametrize l_1 and l_2 via $\vec{r}_1(y)$ and $\vec{r}_2(x)$

$$\vec{r}_1(y) = \langle a, y, f(a, b) + f_y(a, b)(y - b) \rangle$$

$$\vec{r}_2(x) = \langle x, b, f(a, b) + f_x(a, b)(x - a) \rangle$$

Continuing: Goal: connect $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ to our eq² for the tangent plane at $(a, b, f(a, b))$. We have two lines in the tangent plane.

305d

$$\vec{r}_1(y) = \langle a, 0, f(a, b) - bf_y(a, b) \rangle + y \langle 0, 1, f_y(a, b) \rangle$$

$$\vec{r}_2(x) = \langle 0, b, f(a, b) - af_x(a, b) \rangle + x \langle 1, 0, f_x(a, b) \rangle$$

Identify direction vectors for \vec{r}_1 and \vec{r}_2 are $\vec{v}_1 = \langle 0, 1, f_y(a, b) \rangle$ and $\vec{v}_2 = \langle 1, 0, f_x(a, b) \rangle$ respectively. So we can take $\vec{v}_1 \times \vec{v}_2$ for a normal vector to the plane containing \vec{r}_1 , \vec{r}_2 ,

$$\vec{v}_1 \times \vec{v}_2 = \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ 0 & 1 & f_y \\ 1 & 0 & f_x \end{vmatrix} = \langle f_x, f_y, -1 \rangle$$

Thus the eq² to the tangent plane is

$$f_x(a, b)(x-a) + f_y(a, b)(y-b) - z = 0$$

This is the same as the graph of the linearization $L_f(a, b)(x, y) = f(a, b) + f_x(a, b)(x-a) + f_y(a, b)(y-b)$ which is $z = L_f(a, b)(x, y)$.

Partial Derivatives and Tangent Plane

305e

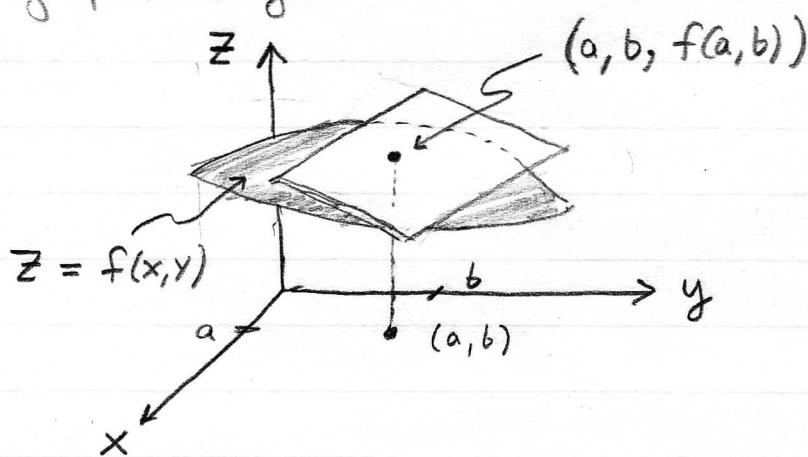
Consider $U \subset \mathbb{R}^2$ and $f: U \rightarrow \mathbb{R}$ differentiable at (a, b) . It turns out that this implies $f_x(a, b)$ and $f_y(a, b)$ exist. Moreover, we can write

$$f(x, y) \cong f(a, b) + \underbrace{[f_x(a, b)](x-a) + [f_y(a, b)](y-b)}_{(\nabla f(a, b)) \cdot (x-a, y-b)}$$

Or we can use the gradient of f

$$\nabla f \equiv \left\langle \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right\rangle$$

To write $Df(a, b)(\vec{h}) = (\nabla f(a, b)) \cdot \vec{h}$. Notation aside, notice what this linearization of $f(x, y)$ means graphically.



The tangent plane is the graph of the linearization.

$$\begin{aligned} z &= f(a, b) + [(\nabla f)(a, b)] \cdot \langle x-a, y-b \rangle \\ &= f(a, b) + \left\langle \frac{\partial f}{\partial x}(a, b), \frac{\partial f}{\partial y}(a, b) \right\rangle \cdot \langle x-a, y-b \rangle \\ &= f(a, b) + \underbrace{\left(\frac{\partial f}{\partial x}(a, b)(x-a) + \frac{\partial f}{\partial y}(a, b)(y-b) \right)}_{\text{linearization}} \end{aligned}$$

It quantifies how f changes in all directions about the point (a, b) .

DIRECTIONAL DERIVATIVE

305f

The notes on (314) and §15.6 could be confusing since the notation $\hat{u} = \langle a, b \rangle$ is used. I have used (a, b) for the point at which the derivative is considered. To avoid confusion let's use (x_0, y_0) for the point now.

Defⁿ/ Let $f: U \subseteq \mathbb{R}^2 \rightarrow \mathbb{R}$. We define the directional derivative of f in the direction $\hat{u} = \langle a, b \rangle$ at the point (x_0, y_0) to be

$$D_{\hat{u}} f(x_0, y_0) \equiv \lim_{h \rightarrow 0} \left(\frac{f(x_0 + ah, y_0 + bh) - f(x_0, y_0)}{h} \right).$$

This gives the change of f in the \hat{u} -direction at the point (x_0, y_0) . Stewart has a nice proof on pg. 948 that shows

$$D_{\hat{u}} f(x_0, y_0) = f_x(x_0, y_0) a + f_y(x_0, y_0) b$$

Which is better written $D_{\hat{u}} f(x_0, y_0) = (\nabla f(x_0, y_0)) \cdot \hat{u}$.

- Likewise, for $f: U \subseteq \mathbb{R}^3 \rightarrow \mathbb{R}$ we can define $D_{\hat{u}} f(p)$ for $p = (x_0, y_0, z_0)$ and $\hat{u} = \langle a, b, c \rangle$,

$$D_{\hat{u}} f(p) = (\nabla f(p)) \cdot \hat{u}$$

What Good is This?

- We can quantify the rate of change for a function of several variables in some given direction
 - $\text{grad}(f) = \nabla f$ points in max. rate of change while $-\nabla f$ points in min. rate of change.
Generally $D_{\hat{u}} f(p) = |\nabla f(p)| \cos \theta$
-

DERIVATIVE \Rightarrow DIRECTIONAL DERIVATIVE

305g

Consider $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ Differentiable at $(x_0, y_0) = \vec{x}_0$

This means the following limit exists, and $Df(\vec{x}_0)$ is a linear mapping from $\mathbb{R}^2 \rightarrow \mathbb{R}^2$,

$$\lim_{\vec{h} \rightarrow 0} \frac{\|f(\vec{x}_0 + \vec{h}) - f(\vec{x}_0) - Df(\vec{x}_0)(\vec{h})\|_1}{\|\vec{h}\|_2} = 0$$

Whis is equivalent to the following,

$$\lim_{\vec{h} \rightarrow \vec{0}} \frac{f(\vec{x}_0 + \vec{h}) - f(\vec{x}_0)}{\|\vec{h}\|_2} = \lim_{\vec{h} \rightarrow 0} \left(\frac{1}{\|\vec{h}\|_2} Df(\vec{x}_0)(\vec{h}) \right)$$

Notice by linearity: $\frac{1}{\|\vec{h}\|_2} Df(\vec{x}_0)(\vec{h}) = Df(\vec{x}_0)\left(\frac{1}{\|\vec{h}\|_2} \vec{h}\right)$
 Therefore,

$$\lim_{\vec{h} \rightarrow 0} \left(\frac{f(\vec{x}_0 + \vec{h}) - f(\vec{x}_0)}{\|\vec{h}\|_2} \right) = \lim_{\vec{h} \rightarrow 0} \left(Df(\vec{x}_0)\left(\frac{\vec{h}}{\|\vec{h}\|_2}\right) \right)$$

The limit $\vec{h} \rightarrow 0$ means that for all possible paths going to $(0,0)$ in \mathbb{R}^2 the limit holds. In particular we can choose a unit vector \hat{u} and take the paths approaching $(0,0)$ along the \hat{u} -direction. That is let $\vec{h} = h\hat{u}$ and let $h \rightarrow 0$.

$$\lim_{h \rightarrow 0} \left(\frac{f(\vec{x}_0 + h\hat{u}) - f(\vec{x}_0)}{\|h\hat{u}\|_2} \right) = \lim_{h \rightarrow 0} \left(Df(\vec{x}_0)\left(\frac{h\hat{u}}{\|h\hat{u}\|_2}\right) \right)$$

Since $\|h\hat{u}\|_2 = h\|\hat{u}\|_2 = h$ we obtain

$\lim_{h \rightarrow 0} (Df(\vec{x}_0)(\hat{u}))$ on the RHS. Therefore, using $\hat{u} = \langle a, b \rangle$,

$$Df(\vec{x}_0)(\hat{u}) = \lim_{h \rightarrow 0} \left(\frac{f(x_0 + ah, y_0 + bh) - f(x_0, y_0)}{h} \right)$$