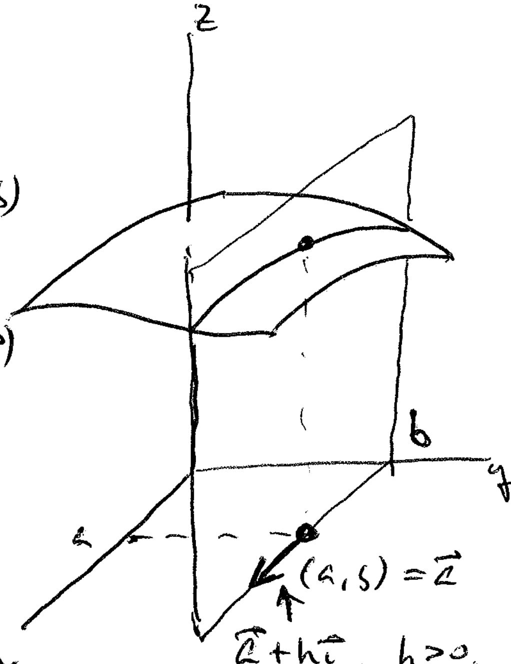
Section 2.6.

Now, in the same way that we compute the partial derivative of a function as the derivative of a slice of a function in a ~~the~~ coordinate direction,



$$\frac{\partial f}{\partial x}(a,b) = \lim_{h\to 0} \frac{f(a+h,b)-f(a,b)}{h}$$

$$= \lim_{h\to 0} \frac{f(\vec{a}+h\vec{i})-f(\vec{a})}{h}.$$

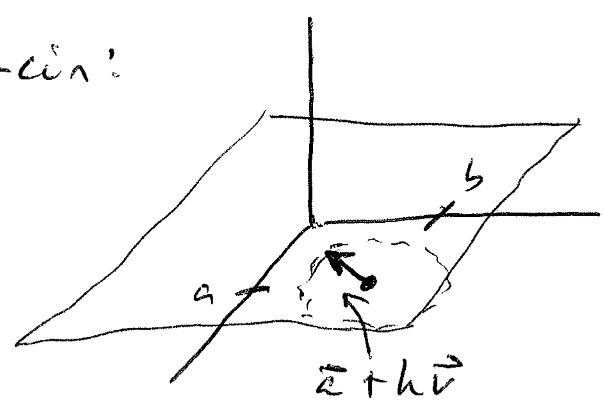
(here, $(a+h,b) = (a,b)+(h,0)$

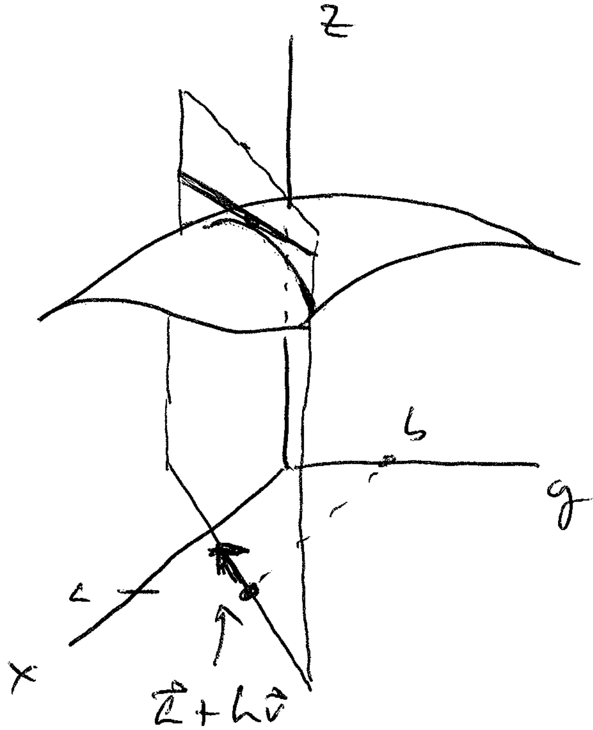
and in vector notation

$$\begin{bmatrix} a+h \\ b \end{bmatrix} = \begin{bmatrix} a \\ b \end{bmatrix} + \begin{bmatrix} h \\ 0 \end{bmatrix} = \begin{bmatrix} a \\ b \end{bmatrix} + h\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \vec{a} + h\vec{i}.$$

One can take a derivative of f at (a,b) in <u>any direction</u> in the domain:

$$\lim_{h\to 0} \frac{f(\vec{a}+h\vec{v})-f(\vec{a})}{h}.$$



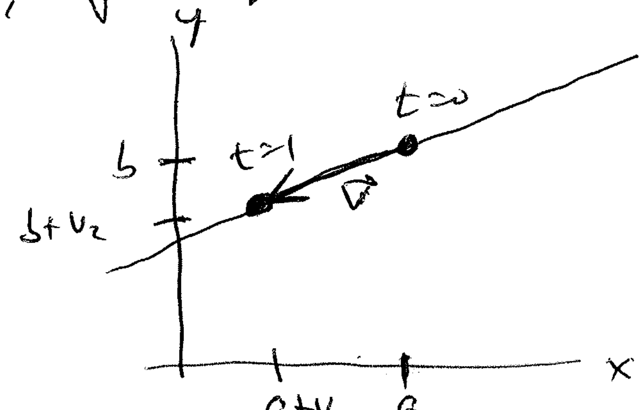


This is the derivative of f at (a,b) in the direction of $\vec v$, or the directional derivative of f at (a,b) wrt $\vec v$:

$$D_{\vec v} f(\vec a) = \lim_{h\to 0} \frac{f(\vec a + h\vec v) - f(\vec a)}{h}.$$

How does this work? Let $f: X \subset \mathbb{R}^2 \to \mathbb{R}$ be diff at $(a,b) \in X$. Compose f with the affine function $\vec g: \mathbb{R} \to \mathbb{R}^2$, $\vec g(t) = \vec a + t\vec v$ where $\vec a = \begin{bmatrix} a \\ b \end{bmatrix}$. $\vec v = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$ is some vector in $\mathbb{R}^2$.

Here $\vec g(t)$ parameterizes a line in $\mathbb{R}^2$ where @ $t=0$, $\vec g(0) = \begin{bmatrix} a \\ b \end{bmatrix}$, $\vec g(1) = \begin{bmatrix} a \\ b \end{bmatrix} + \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$



$\vec g(t)$ is diff and $\vec g'(t) = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \vec v$

$\vec g'(0) = \vec v$.

Now let $F(t) = f(\vec{g}(t)) = f \circ \vec{g}\,(t)$

$= f(\vec{a} + t\vec{v})$ like in our def of directional derivative.

Here F, as the composition of 2 diff. functions, is diff. and

$$F'(0) = \lim_{t \to 0} \frac{F(t) - F(0)}{t - 0}$$

$$= \lim_{t \to 0} \frac{f(\vec{a} + t\vec{v}) - f(\vec{a})}{t}$$

Hence

$$D_{\vec{v}} f(\vec{a}) = \frac{d}{dt}\Big|_{t=0} f(\vec{a} + t\vec{v}) = Df(\vec{g}(0))\,\vec{g}\,'(0)$$

$$= Df(\vec{a})\,\vec{v}$$

Def. Let $f : X \subset \mathbb{R}^n \to \mathbb{R}$ be C^1. Then the gradient function of f is the function

$$\nabla f : X \subset \mathbb{R}^n \to \mathbb{R}^n, \quad \nabla f(\vec{x}) = \begin{bmatrix} f_{x_1}(\vec{x}) \\ \vdots \\ f_{x_n}(\vec{x}) \end{bmatrix}.$$

and the gradient vector of f at $\vec{a}$ is

$$\nabla f(\vec{a}) = \begin{bmatrix} f_{x_1}(\vec{a}) \\ \vdots \\ f_{x_n}(\vec{a}) \end{bmatrix} \in \mathbb{R}^n$$

Notes: ① The gradient carries the same information as the derivative matrix of f, but is a vector of functions:

$$Df(\bar{x}) = (\nabla f)^T \qquad (T \text{ is transpose}).$$

② Only defined for real-valued functions.

With this, we can write

$$D_{\vec{v}} f(\vec{x}) = \underbrace{Df(\vec{x})\,\vec{v}}_{\text{matrix mult}} = \underbrace{\nabla f(\vec{x}) \cdot \vec{v}}_{\text{dot product}}$$

Warning: The choice of $\vec{v}$ is really a choice of direction! Thus, it is important that $\|\vec{v}\| = 1$ for our choice.

Exercise: Show for $\vec{w} = k\vec{v}$, that $D_{\vec{w}} f(\vec{x}) = k D_{\vec{v}} f(\vec{x})$

The directional derivative specifies ~~the~~ how f is changing in the direction of $\vec{v}$ in X.

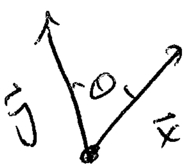
What does this mean?

You are standing in $\mathbb{R}^n$ at a point $\vec{a}$ where a real valued f is defined and differentiable.

How is f changing in any direction? For any

$$v \in \mathbb{R}^n \ni \|\vec{v}\| = 1, \quad D_{\vec{v}} f(\vec{a}) = \nabla f(\vec{a}) \cdot \vec{v}$$

Recall $\vec{x} \cdot \vec{y} = \|\vec{x}\| \, \|\vec{y}\| \cos\theta$ where θ - angle between $\vec{x}$ and $\vec{y}$



(even in $\mathbb{R}^n$, any 2 vectors (not co-linear) span a plane, so there is a well defined angle between them).

Here $D_{\vec{v}} f(\vec{a}) = \nabla f(\vec{a}) \cdot \vec{v} = \|\nabla f(\vec{a})\| \cos\theta \quad (\|\vec{v}\| = 1)$

Hence $-\|\nabla f(\vec{a})\| \le D_{\vec{v}} f(\vec{a}) \le \|\nabla f(\vec{a})\|$

and will achieve a maximum when $\theta = 0$

a minimum when $\theta = \pi$

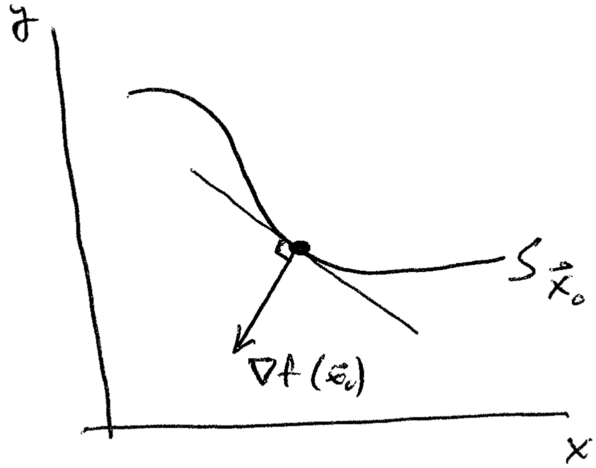
and is 0 when $\theta = \frac{\pi}{2}$ (perpendicular).

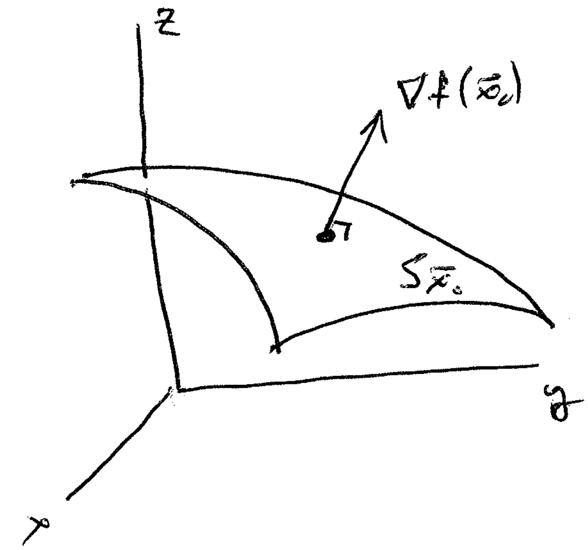
Beautiful interpretation

Thm Let $\mathbb{X} \subset \mathbb{R}^n$ be open and $f: \mathbb{X} \to \mathbb{R}$ a C^1-function. For $\vec{x}_0 \in \mathbb{X}$, let

$$S_{\vec{x}_0} = \{ \vec{x} \in \mathbb{X} \mid f(\vec{x}) = f(\vec{x}_0) = c \}$$

Then $\nabla f(\vec{x}_0) \perp S$.





Another way to say this is that, any vector $\vec{v}$ tangent to $S_{\vec{x}_0}$ will be perpendicular to $\nabla f(\vec{x}_0)$

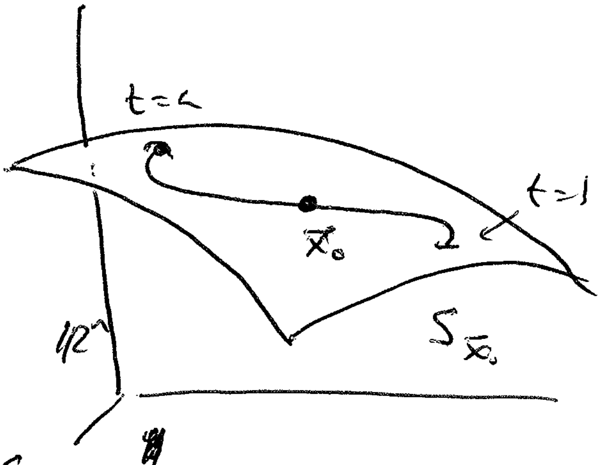
proof Let $\vec{c}: (a,b) \to \mathbb{R}^n$ be a C^1-parameterized curve, $\vec{c}(t) = \begin{bmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{bmatrix}$ such that ① $\vec{c}(t_0) = \vec{x}_0$ for some $t_0 \in (a,b)$, and ② $\vec{c}((a,b)) \subset S_{\vec{x}_0}$.

Then $(f \circ \vec{c}) : (a,1) \to \mathbb{R}$

where $f \in C^1$, and

$f(\vec{c}(t)) =$

$f(x_1(t), x_2(t), \dots, x_n(t)) = c$



Given this last implicit function of t is C^1, we diff.

$$\underbrace{\frac{d}{dt} f(x_1(t), \dots, x_n(t))} = \frac{d}{dt} c = 0$$

$$Df(\vec{c}(t))\, \vec{c}\,'(t) = 0$$

and at $t = t_0$, $Df(\vec{x}_0)\vec{v} = 0$ where $\vec{v} = \vec{c}\,'(t_0)$,

a vector tangent to $\vec{c}$ and hence tangent to $S_{\vec{x}_0}$.

$$Df(\vec{x}_0)\vec{v} = \boxed{\nabla f(\vec{x}_0) \cdot \vec{v} = 0}$$ ■

Def. For any $(n-1)$-dimensional hypersurface in $\mathbb{R}^n$ defined by ~~[illegible]~~ as the c-level set of a C^1 function $f : X \subset \mathbb{R}^n \to \mathbb{R}$,

$$S_c = \{ \vec{x} \in X \mid f(\vec{x}) = c \}$$

The tangent space to S at $\vec{a} \in S$ is the space of all vectors perpendicular to $\nabla f(\vec{a}_0)$; defined by $h(\vec{x}) = \nabla f(\vec{a}_0) \cdot (\vec{x} - \vec{a}_0) = 0$, or

$$h(\vec{x}) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\vec{a}_0)(x_i - a_i) = 0.$$

Note: Compare this to the tangent space of $\text{graph}(f) \subset \mathbb{R}^3$ where $f: \mathbb{R}^2 \to \mathbb{R}$ and $\text{graph}(f)$ is defined by the equation in $\mathbb{R}^3$ $z = f(x,y)$.