

Section 2.6.

I

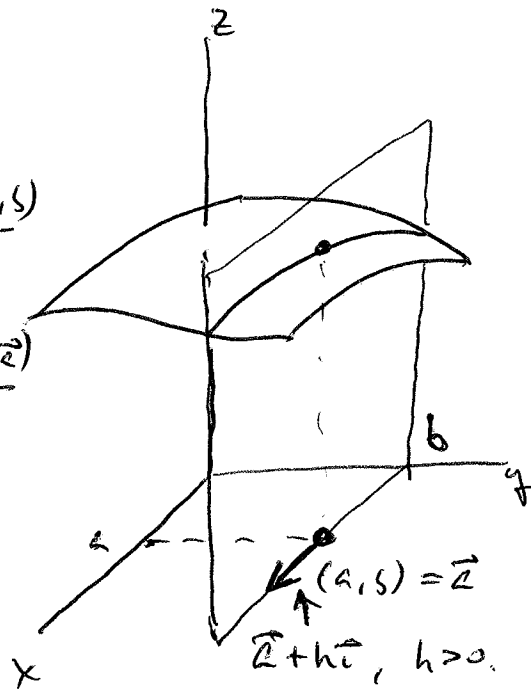
Now, in the same way that we compute the partial derivative of a function as the derivative of a slice of a function in a coordinate direction,

$$\begin{aligned}\frac{\partial f}{\partial x}(a,b) &= \lim_{h \rightarrow 0} \frac{f(a+h, b) - f(a, b)}{h} \\ &= \lim_{h \rightarrow 0} \frac{f(\vec{a} + h\vec{e}_1) - f(\vec{a})}{h}\end{aligned}$$

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

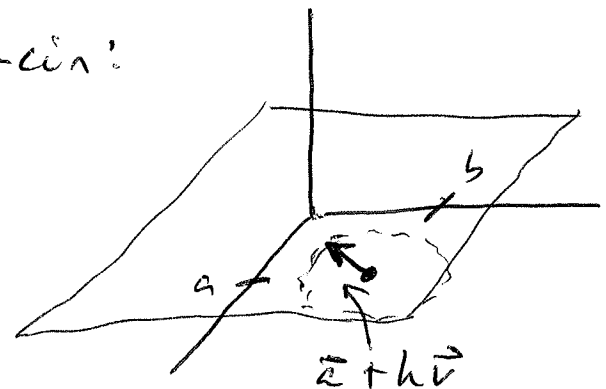
and in vector notation

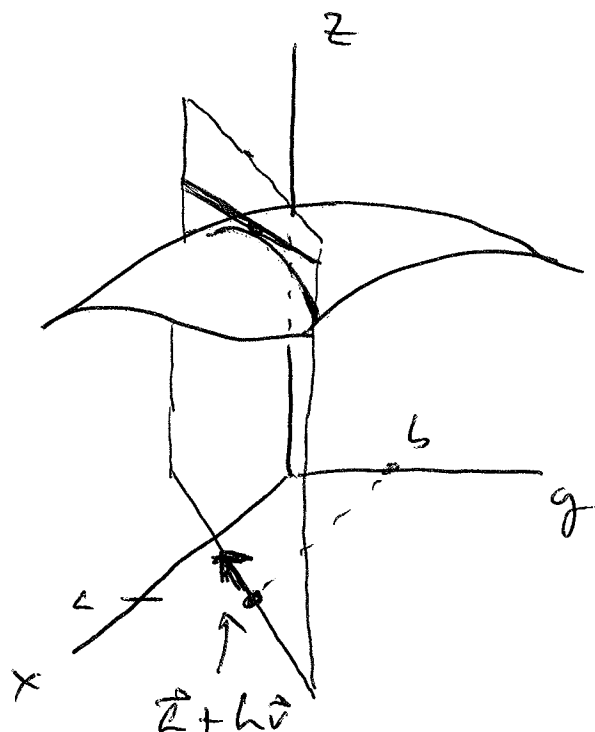
$$\begin{bmatrix} a+h \\ b \end{bmatrix} = \begin{bmatrix} a \\ b \end{bmatrix} + \begin{bmatrix} h \\ 0 \end{bmatrix} = \begin{bmatrix} a \\ b \end{bmatrix} + h \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \vec{a} + h\vec{e}_1.$$



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

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





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 $\vec{v}$:

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

How does this work? Let $f: X \subset \mathbb{R}^2 \rightarrow \mathbb{R}$.

be diff at $(a, b) \in X$. Compute f with

the affine function $\vec{g}: \mathbb{R} \rightarrow \mathbb{R}^2$, $\vec{g}(t) = \vec{a} + t\vec{v}$

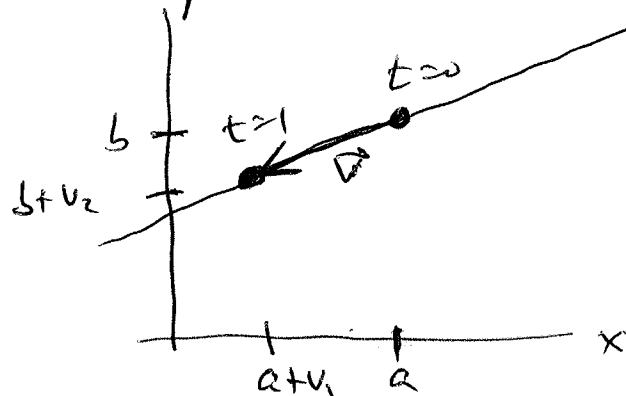
where $\vec{a} = \begin{bmatrix} a \\ b \end{bmatrix}$, $\vec{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$ is some vector in $\mathbb{R}^2$

Here $\vec{g}(t)$ parameterizes a line in $\mathbb{R}^2$ where

$$\text{at } t=0, \vec{g}(0) = \begin{bmatrix} a \\ b \end{bmatrix}, \quad \vec{g}(1) = \begin{bmatrix} a \\ b \end{bmatrix} + \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$$

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

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



Now let $F(t) = f(\bar{g}(t)) = f \circ \bar{g}(t)$
 $= f(\bar{a} + t\bar{v})$ (this is our def of directional derivative.

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

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

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

Hence $D_{\bar{v}} f(\bar{a}) = \frac{d}{dt} \bigg|_{t=0} f(\bar{a} + t\bar{v}) = Df(\bar{g}(0)) \bar{g}'(0)$
 $= Df(\bar{a}) \bar{v}$

~~or $D_{\bar{v}} f(\bar{a}) = \nabla f(\bar{a}) \cdot \bar{v}$~~

Def. Let $f: \mathbb{R} \subset \mathbb{R}^n \rightarrow \mathbb{R}$ be C^1 . Then the

gradient function of f is the function

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

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

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

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

$$Df(\vec{x}) = (\nabla f)^T \quad (\text{T is transpose}),$$

② Only defined for real-valued functions.

With this, we can write

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

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 $D_{\vec{w}} f(\vec{z}) = k D_{\vec{v}} f(\vec{z})$

The directional derivative specifies how f is changing in the direction of $\vec{v}$ in $\mathbb{R}^n$.

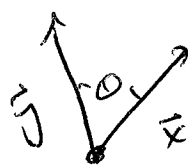
What does this mean?

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

How is f changing in any direction? For any

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

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

Hence $-\|\nabla f(\vec{z})\| \leq D_{\vec{v}} f(\vec{z}) \leq \|\nabla f(\vec{z})\|$

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

a minimum when $\theta = \pi$

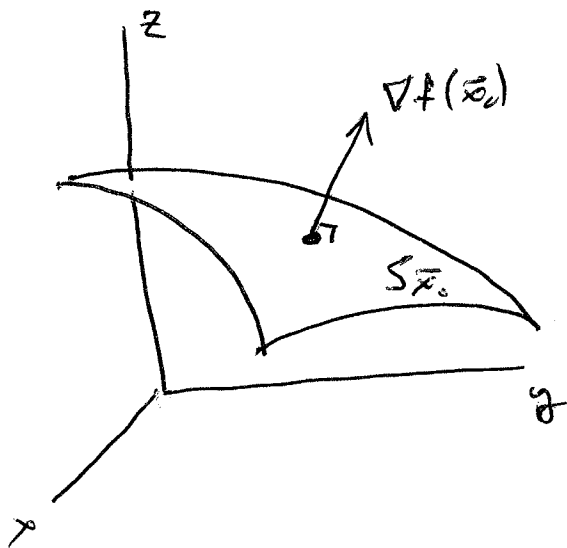
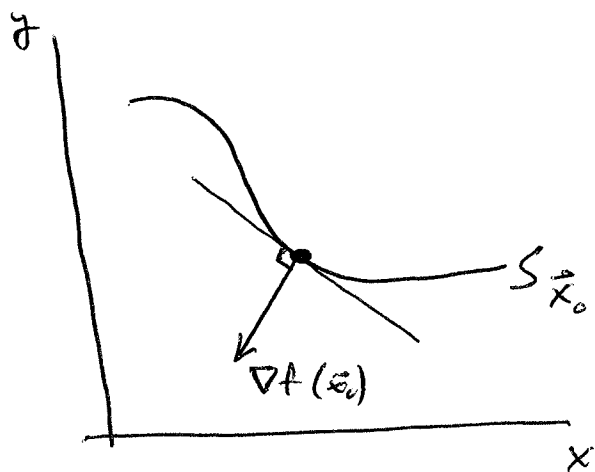
and is 0 when $\theta = \frac{\pi}{2}$ (perpendicular).

Beautiful interpretation

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

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

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 $\nabla f(\vec{x}_0)$

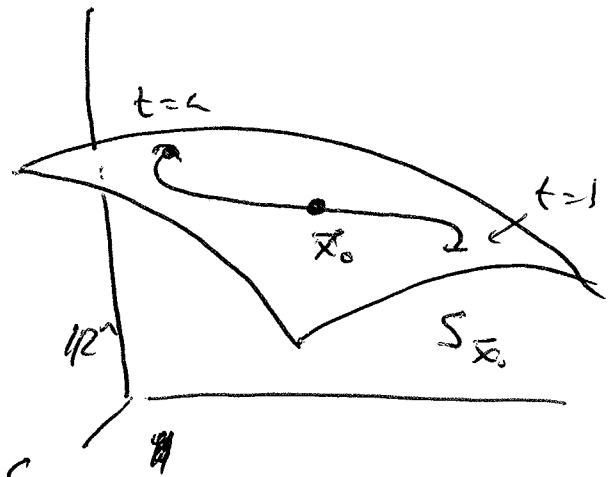
proof Let $\vec{c}: (a,b) \rightarrow \mathbb{R}^n$ be a C^1 -parameterized curve, $\vec{c}(t) = \begin{bmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{bmatrix}$ such that (1) $\vec{c}(t_0) = \vec{x}_0$ for some $t_0 \in (a,b)$, and (2) $\vec{c}((a,b)) \subset S_{\vec{x}_0}$.

Then $(f \circ \vec{c}) : (a, b) \rightarrow \mathbb{R}$.

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 $t \in C^1$, we

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

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

and at $t=t_0$, $DA(\vec{x}_0) \vec{v} = 0$ where $\vec{v} = \vec{c}'(t_0)$,
a vector tangent to $\vec{c}$ and hence tangent to $S_{\vec{x}_0}$.

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

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

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

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

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

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