DIFFERENTIABILITY

Linear maps and matrices (Recollection)

<u>Def</u>: A map $\varphi : \mathbb{R}^n \to \mathbb{R}^k$ is linear if
- $\forall u, v \in \mathbb{R}^n$, $\varphi(u+v) = \varphi(u) + \varphi(v)$
- $\forall u \in \mathbb{R}^n, \forall \lambda \in \mathbb{R}$, $\varphi(\lambda u) = \lambda \varphi(u)$

<u>Notation</u>: $e_i^n = \underbrace{(0, \dots, 0, \underset{\text{ith component}}{1}, 0, \dots, 0)}_{n \text{ components}}$

So that $(e_i^n)_{i=1 \dots n}$ is the standard basis of $\mathbb{R}^n$

<u>Remark</u>: A linear map $\varphi : \mathbb{R}^n \to \mathbb{R}^k$ is entirely determined by the values $\varphi(e_i^n)$:

Let $\underset{(u_1, \dots, u_n)}{u} \in \mathbb{R}^n$, then $u = \sum_{i=1}^{n} u_i e_i^n$ and $\varphi(u) = \sum_{i=1}^{n} u_i \varphi(e_i^n)$

<u>Def</u>: We denote by $(a_{ij})_{i=1 \dots k}$ the components of $\varphi(e_j^n)$, where $\varphi : \mathbb{R}^n \to \mathbb{R}^k$ is linear

<u>ie</u>: $\varphi(e_j^n) = \sum_{i=1}^{k} a_{ij} e_i^k$

The matrix of φ (in the standard bases) is

$$\mathrm{Mat}(\varphi) := \underbrace{\begin{array}{c} \begin{array}{cccc} \varphi(e_1^n) & \varphi(e_2^n) & & \varphi(e_n^n) \\ \downarrow & \downarrow & & \downarrow \end{array} \\ \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kn} \end{pmatrix} \end{array}}_{n \text{ columns}} \left.\vphantom{\begin{pmatrix} a \\ a \\ a \\ a \end{pmatrix}}\right\} k \text{ rows} \in M_{k,n}(\mathbb{R})$$

<u>Remark</u>: φ is entirely determined by the above matrix.

Def: $M = \begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{k1} & \cdots & a_{km} \end{pmatrix} \in M_{km}(\mathbb{R})$, $N = \begin{pmatrix} b_{11} & \cdots & b_{1m} \\ \vdots & & \vdots \\ b_{k1} & \cdots & b_{km} \end{pmatrix} \in M_{km}(\mathbb{R})$

then $M + N := \begin{pmatrix} c_{11} & \cdots & c_{1m} \\ \vdots & \ddots & \vdots \\ c_{k1} & \cdots & c_{km} \end{pmatrix} \in M_{km}(\mathbb{R})$

where $c_{ij} = a_{ij} + b_{ij}$

Def: $M = \begin{pmatrix} a_{11} & \cdots & a_{1l} \\ \vdots & \ddots & \vdots \\ a_{k1} & \cdots & a_{kl} \end{pmatrix} \in M_{kl}(\mathbb{R})$, $N = \begin{pmatrix} b_{11} & \cdots & b_{1m} \\ \vdots & \ddots & \\ b_{l1} & \cdots & b_{lm} \end{pmatrix} \in M_{lm}(\mathbb{R})$

$$MN := \begin{pmatrix} c_{11} & \cdots & c_{1m} \\ \vdots & \ddots & \vdots \\ c_{k1} & \cdots & c_{km} \end{pmatrix} \in M_{km}(\mathbb{R})$$

where $c_{ij} = \sum_{x=1}^{l} a_{ix} b_{xj}$

! $M_{kl} \times M_{lm} \longrightarrow M_{km}$

the number of columns of the first matrix must be equal to the number of lines of the second matrix.

Prop: Let $\varphi : \mathbb{R}^m \to \mathbb{R}^l$ and $\psi : \mathbb{R}^l \to \mathbb{R}^k$ be two linear maps then the matrix of the linear map $\psi \circ \varphi : \mathbb{R}^m \to \mathbb{R}^k$ is

$$\mathrm{Mat}(\psi \circ \varphi) = \mathrm{Mat}(\psi)\, \mathrm{Mat}(\varphi)$$

Prop: Let $\varphi : \mathbb{R}^m \to \mathbb{R}^k$ linear, $v = \sum_{i=1}^{m} v_i e_i^m \in \mathbb{R}^m$

then $\varphi(v) = \sum_{i=1}^{k} c_i e_i^k$ where

$$\begin{pmatrix} c_1 \\ \vdots \\ c_k \end{pmatrix} = \underbrace{\mathrm{Mat}(\varphi)}_{\in M_{km}(\mathbb{R})} \underbrace{\begin{pmatrix} v_1 \\ \vdots \\ v_m \end{pmatrix}}_{\in M_{m1}(\mathbb{R})} \in M_{k1}(\mathbb{R})$$

DIFFERENTIABILITY

Vector-valued functions

Def: $\mathcal{U} \subset \mathbb{R}^m$ open set, $f: \mathcal{U} \longrightarrow \mathbb{R}^k$, $x_0 \in \mathcal{U}$.

We say that f is differentiable at x_0 if there exists a linear map $d_{x_0}f: \mathbb{R}^m \longrightarrow \mathbb{R}^k$ s.t.

$$\lim_{h \to 0} \frac{f(x_0+h) - f(x_0) - d_{x_0}f(h)}{\|h\|} = 0$$

or equivalently $f(x_0+h) = f(x_0) + d_{x_0}f(h) + E(h)$

with $\displaystyle\lim_{h \to 0} \frac{E(h)}{\|h\|} = 0$

$d_{x_0}f: \mathbb{R}^m \longrightarrow \mathbb{R}^k$ is called the differential or total derivative of f at x_0

Def: $\mathcal{U} \subset \mathbb{R}^m$ open set, $f = (f_1, \dots, f_k): \mathcal{U} \longrightarrow \mathbb{R}^k$, $x_0 \in \mathcal{U}$

Assume that all the partial derivatives $\left(\frac{\partial f_i}{\partial x_j}(x_0)\right)_{\substack{i=1,\dots,k \\ j=1,\dots,m}}$ exist. The we define the Jacobian matrix of f at x_0 by

$$\underbrace{Df(x_0)}_{} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x_0) & \cdots & \frac{\partial f_1}{\partial x_m}(x_0) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_k}{\partial x_1}(x_0) & \cdots & \frac{\partial f_k}{\partial x_m}(x_0) \end{pmatrix}$$

Notation used in the textbook.
The notation $J_f(x_0)$ is very common

Theorem: $\mathcal{U} \subset \mathbb{R}^m$ open, $f = (f_1, \dots, f_k) : \mathcal{U} \longrightarrow \mathbb{R}^k$, $x_0 \in \mathcal{U}$.

f is differentiable at $x_0 \Leftrightarrow \forall i$, f_i is differentiable at x_0.

Moreover, if the above holds,

$$d_{x_0} f = (d_{x_0} f_1, \dots, d_{x_0} f_k) : \mathbb{R}^m \longrightarrow \mathbb{R}^k$$

Δ Notice that componentwise, if $\ell : \mathbb{R}^m \to \mathbb{R}^k$ is linear,

$$f(x_0 + h) = f(x_0) + \ell(h) + \varepsilon(h)$$

becomes

$$\begin{pmatrix} f_1(x_0+h) \\ \vdots \\ f_k(x+h) \end{pmatrix} = \begin{pmatrix} f_1(x_0) \\ \vdots \\ f_k(x_0) \end{pmatrix} + \begin{pmatrix} \ell_1(h) \\ \vdots \\ \ell_k(h) \end{pmatrix} + \begin{pmatrix} \varepsilon_1(h) \\ \vdots \\ \varepsilon_k(h) \end{pmatrix}$$

and $\frac{1}{\|h\|} \varepsilon(h) = \begin{pmatrix} \frac{\varepsilon_1(h)}{\|h\|} \\ \vdots \\ \frac{\varepsilon_k(h)}{\|h\|} \end{pmatrix}$

$\square$

Hence, it is enough to understand well the real-valued case:

Theorem: $\mathcal{U} \subset \mathbb{R}^m$ open, $f = (f_1, \dots, f_k) : \mathcal{U} \longrightarrow \mathbb{R}^k$, $x_0 \in \mathcal{U}$

If f is differentiable at x_0 then all the directional derivatives $\partial_v f_i(x_0)$ exist ($i = 1, \dots, k$, $v \in \mathbb{R}^m$)

And

$$\mathrm{Mat}(d_{x_0} f) = D(f)(x_0) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x_0) & \cdots & \frac{\partial f_1}{\partial x_m}(x_0) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_k}{\partial x_1}(x_0) & \cdots & \frac{\partial f_k}{\partial x_m}(x_0) \end{pmatrix}$$

Δ We apply the result from the real-valued case to $f_1, \dots, f_k$ and then use the previous theorem to get that

$$d_{x_0} f(e_i^m) = \left(d_{x_0} f_1(e_i^m), \dots, d_{x_0} f_k(e_i^m)\right)$$

$$= \left(\frac{\partial f_1}{\partial x_i}(x_0), \dots, \frac{\partial f_k}{\partial x_i}(x_0)\right)$$

Hence

$$\begin{array}{cccccc} & d_{x_0} f(e_1^m) & d_{x_0} f(e_2^m) & & d_{x_0} f(e_m^m) \\ & \downarrow & \downarrow & & \downarrow \\ \mathrm{Mat}(d_{x_0} f) = & \multicolumn{4}{c}{\begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x_0) & \frac{\partial f_1}{\partial x_2}(x_0) & \cdots & \frac{\partial f_1}{\partial x_m}(x_0) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_k}{\partial x_1}(x_0) & \frac{\partial f_k}{\partial x_2}(x_0) & \cdots & \frac{\partial f_k}{\partial x_m}(x_0) \end{pmatrix}} \end{array}$$

$\square$

<u>Theorem</u>: $U \subset \mathbb{R}^m$ open, $f = (f_1, \dots, f_k) : U \to \mathbb{R}^k$, $x_0 \in U$

If all the partial derivatives $\frac{\partial f_i}{\partial x_j}$ exist on U and are continuous at x_0 then f is differentiable at x_0.

Δ We apply the theorem from the real-valued case to $f_1, \dots, f_k$ $\square$

Theorem: If f is differentiable at x_0 then f is continuous at x_0

Proof: apply the result from the real-valued case to the components. QED

Ex: $f = (f_1, \dots, f_k) : (a,b) \longrightarrow \mathbb{R}^k$

f is differentiable at $t_0 \in (a,b)$ iff $f_1'(t_0), \dots, f_k'(t_0)$ exist and then

$$Df(t_0) = \begin{pmatrix} f_1'(t_0) \\ \vdots \\ f_k'(t_0) \end{pmatrix}$$

In this case it is common to use the notation $f'(t_0)$ instead of $Df(t_0)$.

Comment: if f is differentiable at t_0 and $f'(t_0) \neq \vec{0}$ then:

1. $f'(t_0)$ is tangent to the parametrized curve at $f(t_0)$
2. $\theta(h) = f(t_0) + h f'(t_0)$, $h \in \mathbb{R}$, parametrizes the tangent line ~~at~~ of the parametrized curve at $f(t_0)$

See examples 2 and 3 in Section 2.2.

Ex: $U \subset \mathbb{R}^m$ open, $f : U \longrightarrow \mathbb{R}$, $x_0 \in U$

If f is differentiable at x_0 then

$$Df(x_0) = \left(\frac{\partial f}{\partial x_1}(x_0) \dots \frac{\partial f}{\partial x_m}(x_0) \right)$$

Notice that $Df(x_0) \cdot \begin{pmatrix} h_1 \\ \vdots \\ h_m \end{pmatrix} = \sum_{i=1}^{m} \frac{\partial f}{\partial x_i}(x_0) \cdot h = \nabla f(x_0) \cdot h$

We recover the gradient from the previous section.

But be careful:

$Df(x_0)$ is the $1 \times m$ matrix of a linear map

$\nabla f(x_0)$ is a vector of $\mathbb{R}^m$