## 1.2 The total differential

in general not a vectorfield

**Definition:** Let $D \subset \mathbb{R}^n$ open, $\mathbf{x}^0 \in D$ and $\mathbf{f} : D \to \mathbb{R}^m$. The function $\mathbf{f}(\mathbf{x})$ is called differentiable in $\mathbf{x}^0$ (or totally differentiable in $\mathbf{x}_0$), if there exists a linear map

$$\mathbf{l}(\mathbf{x}, \mathbf{x}^0) := \mathbf{A} \cdot (\mathbf{x} - \mathbf{x}^0)$$

with a matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$ which satisfies the following approximation property

$$\mathbf{f}(\mathbf{x}) = \mathbf{f}(\mathbf{x}^0) + \mathbf{A} \cdot (\mathbf{x} - \mathbf{x}^0) + \mathbf{o}(\|\mathbf{x} - \mathbf{x}^0\|)$$

i.e.

$$\lim_{\mathbf{x} \to \mathbf{x}^0} \frac{\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{x}^0) - \mathbf{A} \cdot (\mathbf{x} - \mathbf{x}^0)}{\|\mathbf{x} - \mathbf{x}^0\|} = 0.$$

# The total differential and the Jacobian matrix.

**Notation:** We call the linear map $\mathbf{l}$ the differential or the total differential of $\mathbf{f}(\mathbf{x})$ at the point $\mathbf{x}^0$. We denote $\mathbf{l}$ by $\mathbf{df}(\mathbf{x}^0)$.

The related matrix $\mathbf{A}$ is called Jacobi–matrix of $\mathbf{f}(\mathbf{x})$ at the point $\mathbf{x}^0$ and is denoted by $\mathbf{J\,f}(\mathbf{x}^0)$ (or $\mathbf{Df}(\mathbf{x}^0)$ or $\mathbf{f}'(\mathbf{x}^0)$).

**Remark:** For $m = n = 1$ we obtain the well known relation

$$f(x) = f(x_0) + \underbrace{f'(x_0)}_{A}(x - x_0) + o(|x - x_0|)$$

$(1 \times 1)$

for the derivative $f'(x_0)$ at the point $x_0$.

**Remark:** In case of a scalar function ($m = 1$) the matrix $\mathbf{A} = \mathbf{a}$ is a row vextor and $\mathbf{a}(\mathbf{x} - \mathbf{x}^0)$ a scalar product $\langle \mathbf{a}^T, \mathbf{x} - \mathbf{x}^0 \rangle$.

$n=2$
$m=1$

$$f(x_1,x_2) = f(x_1^0,x_2^0) + \frac{\partial f}{\partial x_1}(x_1^0,x_2^0)(x_1-x_1^0) + \frac{\partial f}{\partial x_2}(x_1^0,x_2^0)(x_2-x_2^0) + \ldots = f(x_1^0,x_2^0) + \underbrace{\overbrace{\left(\frac{\partial f}{\partial x_1}(x_1^0,x_2^0), \frac{\partial f}{\partial x_2}(x_1^0,x_2^0)\right)}^{A}}_{1\times 2} \cdot \begin{pmatrix} x_1 - x_1^0 \\ x_2 - x_2^0 \end{pmatrix} + \ldots$$

# Total and partial differentiability.

**Theorem:** Let $\mathbf{f} : D \to \mathbb{R}^m$, $\mathbf{x}^0 \in D \subset \mathbb{R}^n$, $D$ open.

a) If $\mathbf{f}(\mathbf{x})$ is differentiable in $\mathbf{x}^0$, then $\mathbf{f}(\mathbf{x})$ is continuous in $\mathbf{x}^0$.

b) If $\mathbf{f}(\mathbf{x})$ is differentiable in $\mathbf{x}^0$, then the (total) differential and thus the Jacobi–matrix are uniquely determined and we have

$m \times n$

$$\mathbf{J}\,\mathbf{f}(\mathbf{x}^0) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\mathbf{x}^0) & \cdots & \frac{\partial f_1}{\partial x_n}(\mathbf{x}^0) \\ \vdots & & \vdots \\ \underbrace{\frac{\partial f_m}{\partial x_1}(\mathbf{x}^0)}_{\frac{\partial f}{\partial x_n}(x^0)} & \cdots & \underbrace{\frac{\partial f_m}{\partial x_n}(\mathbf{x}^0)}_{\frac{\partial f}{\partial x_n}(x^0)} \end{pmatrix} = \begin{pmatrix} Df_1(\mathbf{x}^0) \\ \vdots \\ Df_m(\mathbf{x}^0) \end{pmatrix}$$

c) If $\mathbf{f}(\mathbf{x})$ is a $C^1$–function on $D$, then $\mathbf{f}(\mathbf{x})$ is differentiable on $D$.

# Proof of a).

If $\mathbf{f}$ is differentiable in $\mathbf{x}^0$, then by definition

$$\lim_{\mathbf{x}\to\mathbf{x}^0} \frac{\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{x}^0) - \mathbf{A}\cdot(\mathbf{x}-\mathbf{x}^0)}{\|\mathbf{x}-\mathbf{x}^0\|} = 0$$

Thus we conclude

$$\lim_{\mathbf{x}\to\mathbf{x}^0} \underbrace{\|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{x}^0) - \mathbf{A}\cdot(\mathbf{x}-\mathbf{x}^0)\|} = 0$$

and we obtain

$$\|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{x}^0)\| \leq \overbrace{\|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{x}^0) - \mathbf{A}\cdot(\mathbf{x}-\mathbf{x}^0)\|}^{} + \underbrace{\|\mathbf{A}\cdot(\mathbf{x}-\mathbf{x}^0)\|}_{\substack{\|A\|\|x-x^0\| \\ \to 0}}$$

$\to 0$ $\quad \to \quad 0 \qquad \text{as } \mathbf{x}\to\mathbf{x}^0$

Therefore the function $\mathbf{f}$ is continuous at $\mathbf{x}^0$.

# Proof of b).

Let $\mathbf{x} = \mathbf{x}^0 + t\mathbf{e}_i$, $|t| < \varepsilon$, $i \in \{1, \ldots, n\}$. Since $\mathbf{f}$ in differentiable at $\mathbf{x}^0$, we have

$$\lim_{\mathbf{x} \to \mathbf{x}^0} \frac{\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{x}^0) - \mathbf{A} \cdot (\mathbf{x} - \mathbf{x}^0)}{\|\mathbf{x} - \mathbf{x}^0\|_\infty} = 0$$

We write $x - x^0 = te_i$ $\quad \|x - x^0\|_\infty = |t|$

$$\underbrace{\frac{\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{x}^0) - \mathbf{A} \cdot (\mathbf{x} - \mathbf{x}^0)}{\|\mathbf{x} - \mathbf{x}^0\|_\infty}}_{\to 0} = \frac{t\left(\mathbf{f}(\mathbf{x}^0 + t\mathbf{e}_i) - \mathbf{f}(\mathbf{x}^0)\right)}{|t|\,t} - \frac{t\mathbf{A}\mathbf{e}_i}{|t|}$$

$$= \frac{t}{|t|} \cdot \left( \frac{\mathbf{f}(\mathbf{x}^0 + t\mathbf{e}_i) - \mathbf{f}(\mathbf{x}^0)}{t} - \mathbf{A}\mathbf{e}_i \right)$$

$$\to \quad 0 \qquad \text{as } t \to 0$$

Thus

$$\overbrace{\lim_{t \to 0} \frac{\mathbf{f}(\mathbf{x}^0 + t\mathbf{e}_i) - \mathbf{f}(\mathbf{x}^0)}{t}}^{\frac{\partial f}{\partial x_i}} = \overbrace{\mathbf{A}\mathbf{e}_i}^{i \text{ column of } A} \qquad i = 1, \ldots, n$$

# Examples.

- Consider the scalar function $f(x_1, x_2) = x_1 e^{2x_2}$. Then the Jacobian is given by: $n=2$ $m=1$

$$1\times 2 \qquad Jf(x_1, x_2) = Df(x_1, x_2) = e^{2x_2}(1, 2x_1) = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}\right)$$

- Consider the function $\mathbf{f} : \mathbb{R}^3 \to \mathbb{R}^2$ defined by

$$\mathbf{f}(x_1, x_2, x_3) = \begin{pmatrix} x_1 x_2 x_3 \\ \sin(x_1 + 2x_2 + 3x_3) \end{pmatrix}$$

The Jacobian is given by

$m=2$ $n=3$
$2\times 3$

$$\mathbf{Jf}(x_1, x_2, x_3) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \frac{\partial f_1}{\partial x_3} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \frac{\partial f_2}{\partial x_3} \end{pmatrix} = \begin{pmatrix} x_2 x_3 & x_1 x_3 & x_1 x_2 \\ \cos(s) & 2\cos(s) & 3\cos(s) \end{pmatrix}$$

with $s = x_1 + 2x_2 + 3x_3$.

## Further examples.

- Let $\mathbf{f}(\mathbf{x}) = \mathbf{A}\mathbf{x}$, $\mathbf{A} \in \mathbb{R}^{m\times n}$ and $\mathbf{x} \in \mathbb{R}^n$. Then $f(x) = Ax = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \dots \\ \\ a_{m1}x_1 + \dots \end{pmatrix}$

$$\mathbf{J}\mathbf{f}(\mathbf{x}) = \mathbf{A} \qquad \text{for all } \mathbf{x} \in \mathbb{R}^n$$

- Let $f(\mathbf{x}) = \mathbf{x}^T\mathbf{A}\mathbf{x} = \langle \mathbf{x}, \mathbf{A}\mathbf{x}\rangle$, $\mathbf{A} \in \mathbb{R}^{n\times n}$ and $\mathbf{x} \in \mathbb{R}^n$. $f: \mathbb{R}^n \to \mathbb{R}$ $Jf$ $1\times n$
  Then we have

$\frac{\partial x}{\partial x_i} = e_i$

$$\underbrace{\frac{\partial f}{\partial x_i}}_{\text{ith colom. of } Jf} = \langle \mathbf{e}_i, \mathbf{A}\mathbf{x}\rangle + \langle \mathbf{x}, \mathbf{A}\mathbf{e}_i\rangle$$

$e_i^T Ax = (e_i^T Ax)^T = x^T A^T e_i$

$$= \mathbf{e}_i^T\mathbf{A}\mathbf{x} + \mathbf{x}^T\mathbf{A}\mathbf{e}_i$$

$$= \underbrace{\mathbf{x}^T(\mathbf{A}^T + \mathbf{A})\mathbf{e}_i}_{\text{ith colom of } x^T(A^T+A)}$$

We conclude

$$\mathbf{J}f(\mathbf{x}) = \text{grad} f(\mathbf{x}) = \mathbf{x}^T(\mathbf{A}^T + \mathbf{A})$$

# Rules for the differentiation.

**Theorem:**

a) **Linearität:** LET $\mathbf{f}, \mathbf{g} : D \to \mathbb{R}^m$ be differentiable in $\mathbf{x}^0 \in D$, $D$ open. Then $\alpha\,\mathbf{f}(\mathbf{x}^0) + \beta\,\mathbf{g}(\mathbf{x}^0)$, and $\alpha, \beta \in \mathbb{R}$ are differentiable in $\mathbf{x}^0$ and we have

$$\begin{aligned} \mathbf{d}(\alpha\mathbf{f} + \beta\mathbf{g})(\mathbf{x}^0) &= \alpha\,\mathbf{df}(\mathbf{x}^0) + \beta\,\mathbf{dg}(\mathbf{x}^0) \\ \mathbf{J}(\alpha\mathbf{f} + \beta\mathbf{g})(\mathbf{x}^0) &= \alpha\,\mathbf{Jf}(\mathbf{x}^0) + \beta\,\mathbf{Jg}(\mathbf{x}^0) \end{aligned}$$

b) **Chain rule:** Let $\mathbf{f} : D \to \mathbb{R}^m$ be differentiable in $\mathbf{x}^0 \in D$, $D$ open. Let $\mathbf{g} : E \to \mathbb{R}^k$ be differentiable in $\mathbf{y}^0 = f(\mathbf{x}^0) \in E \subset \mathbb{R}^m$, $E$ open. Then $g \circ f$ is differentiable in $\mathbf{x}^0$.

For the differentials it holds

$$\mathbf{d}(\mathbf{g} \circ \mathbf{f})(\mathbf{x}^0) = \mathbf{dg}(\mathbf{y}^0) \circ \mathbf{df}(\mathbf{x}^0)$$

and analoglously for the Jacobian matrix

$$\underbrace{\mathbf{J}(\mathbf{g} \circ \mathbf{f})(\mathbf{x}^0)}_{k \times n} = \underbrace{\mathbf{Jg}(\mathbf{y}^0)}_{k \times m} \cdot \underbrace{\mathbf{Jf}(\mathbf{x}^0)}_{m \times n}$$

# Examples for the chain rule.

Let $I \subset \mathbb{R}$ be an intervall. Let $\mathbf{h} : I \to \mathbb{R}^n$ be a curve, differentiable in $t_0 \in I$ with values in $D \subset \mathbb{R}^n$, $D$ open. Let $f : D \to \mathbb{R}$ be a scalar function, differentiable in $\mathbf{x}^0 = \mathbf{h}(t_0)$.

$h : \mathbb{R}^1 \to \mathbb{R}^n$

$f : \mathbb{R}^n \to \mathbb{R}^1$ $\quad f \circ h : \mathbb{R}^1 \to \mathbb{R}^1$

Then the composition

$$(f \circ \mathbf{h})(t) = f(h_1(t), \ldots, h_n(t))$$

is differentiable in $t_0$ and we have for the derivative:

$$\begin{aligned}
(f \circ \mathbf{h})'(t_0) &= \underbrace{\mathbf{J}f(\mathbf{h}(t_0))}_{1 \times n} \cdot \underbrace{\mathbf{Jh}(t_0)}_{n \times 1} \\
&= \operatorname{grad} f(\mathbf{h}(t_0)) \cdot \mathbf{h}'(t_0) \\
&= \sum_{k=1}^{n} \frac{\partial f}{\partial x_k}(\mathbf{h}(t_0)) \cdot h_k'(t_0)
\end{aligned}$$