Derivazione di Funzioni Composte
La regola della catena

$*$

Dimostrazione e conseguenze

Raul Paolo Serapioni

Analisi Matematica A – Secondo modulo
Corso di Laurea in Matematica
Università di Trento

Lezione 22

Example

Se $g : \mathbb{R} \to \mathbb{R}$, $h : \mathbb{R} \to \mathbb{R}$ sono derivabili allora

$$g \circ f : \mathbb{R} \to \mathbb{R}$$

è differenziabile e

$$\frac{d}{dx}(g \circ h)(x) = g'(h(x))h'(x).$$

Example

Se $f : \mathbb{R}^2 \to \mathbb{R}$ e $h : \mathbb{R} \to \mathbb{R}$ sono differenziabili allora

$$h \circ f : \mathbb{R}^2 \to \mathbb{R}$$

è differenziabile e

$$\nabla(h \circ f)(\mathbf{x}) = h'(f(\mathbf{x}))\nabla f(\mathbf{x})$$

Example

Se $f : \mathbb{R}^2 \to \mathbb{R}$, $\mathbf{r} : \mathbb{R} \to \mathbb{R}^2$ sono differenziabili allora

$$f \circ \mathbf{r} : \mathbb{R} \to \mathbb{R}$$

è differenziabile e

$$\begin{aligned}
\frac{d}{dt}(f \circ \mathbf{r})(t) &= \frac{d}{dt} f(r_1(t), r_2(t)) \\
&= J_{f\circ\mathbf{r}}(t) = J_f(\mathbf{r}(t)) \cdot J_{\mathbf{r}}(t) \qquad \text{regola della catena} \\
&= \langle \nabla f(\mathbf{r}(t)), \mathbf{r}'(t) \rangle \\
&= (\partial_1 f(\mathbf{r}(t)), \partial_2 f(\mathbf{r}(t))) \cdot \begin{pmatrix} r_1'(t) \\ r_2'(t) \end{pmatrix} \\
&= (\partial_1 f)(\mathbf{r}(t)) \cdot r_1'(t) + (\partial_2 f)(\mathbf{r}(t)) \cdot r_2'(t).
\end{aligned}$$

Example (Derivazione di funzioni composte: esempio)

Se $f : \mathbb{R}^2 \to \mathbb{R}$ e $\mathbf{r} : \mathbb{R} \to \mathbb{R}^2$ sono differenziabili, allora

$$(\mathbf{r} \circ f)(\mathbf{x}) := \begin{pmatrix} r_1(f(\mathbf{x})) \\ r_2(f(\mathbf{x})) \end{pmatrix} : \mathbb{R}^2 \to \mathbb{R}^2$$

è differenziabile perché entrambe le componenti lo sono e

$$\begin{aligned} J_{\mathbf{r}\circ f}(\mathbf{x}) &= \begin{bmatrix} \partial_1(r_1 \circ f)(\mathbf{x}) & \partial_2(r_1 \circ f)(\mathbf{x}) \\ \partial_1(r_2 \circ f)(\mathbf{x}) & \partial_2(r_2 \circ f)(\mathbf{x}) \end{bmatrix} \\ &= \begin{bmatrix} r_1'(f(\mathbf{x}))\partial_1 f(\mathbf{x}) & r_1'(f(\mathbf{x}))\partial_2 f(\mathbf{x}) \\ r_2'(f(\mathbf{x}))\partial_1 f(\mathbf{x}) & r_2'(f(\mathbf{x}))\partial_2 f(\mathbf{x}) \end{bmatrix} \\ &= \begin{pmatrix} r_1'(f(\mathbf{x})) \\ r_2'(f(\mathbf{x})) \end{pmatrix} \cdot (\partial_1 f(\mathbf{x}), \partial_2 f(\mathbf{x})) = J_{\mathbf{r}}(f(\mathbf{x})) \cdot J_f((\mathbf{x}) \end{aligned}$$

Teorema: la regola della catena

Siano $\mathbf{F} : \mathbb{R}^n \to \mathbb{R}^p$, $\mathbf{G} : \mathbb{R}^p \to \mathbb{R}^m$ tali che

- sia definita $\mathbf{G} \circ \mathbf{F} : \mathbb{R}^n \to \mathbb{R}^m$,
- $\mathbf{F}$ sia differenziabile in $\mathbf{x}_0 \in \mathbb{R}^n$,
- $\mathbf{G}$ sia differenziabile in $\mathbf{y}_0 := \mathbf{F}(\mathbf{x}_0) \in \mathbb{R}^p$

allora

$$\mathbf{G} \circ \mathbf{F} \quad \text{è differenziabile in } \mathbf{x}_0 \in \mathbb{R}^n$$

e

$$J_{\mathbf{G}\circ\mathbf{F}}(\mathbf{x}_0) = J_{\mathbf{G}}(\mathbf{F}(\mathbf{x}_0)) \cdot J_{\mathbf{F}}(\mathbf{x}_0).$$

Traccia di prova:
Per le ipotesi di differenziabilità di **F** e di **G**

$$\mathbf{F}(\mathbf{x}_0+\mathbf{h}) = \mathbf{F}(\mathbf{x}_0) + J_{\mathbf{F}}(\mathbf{x}_0)\cdot\mathbf{h} + \mathbf{o}_{\mathbb{R}^p}(\|\mathbf{h}\|) \quad \text{per } \mathbf{h}\to\mathbf{0} \text{ in } \mathbb{R}^n;$$
$$\mathbf{G}(\mathbf{y}_0+\mathbf{k}) = \mathbf{G}(\mathbf{y}_0) + J_{\mathbf{G}}(\mathbf{y}_0)\cdot\mathbf{k} + \mathbf{o}_{\mathbb{R}^m}(\|\mathbf{k}\|) \quad \text{per } \mathbf{k}\to\mathbf{0} \text{ in } \mathbb{R}^p.$$

$$\begin{aligned}
(\mathbf{G}\circ\mathbf{F})(\mathbf{x}_0+\mathbf{h}) &= \mathbf{G}(\mathbf{F}(\mathbf{x}_0+\mathbf{h}))\\
&= \mathbf{G}\left(\mathbf{F}(\mathbf{x}_0) + J_{\mathbf{F}}(\mathbf{x}_0)\cdot\mathbf{h} + \mathbf{o}_{\mathbb{R}^p}(\|\mathbf{h}\|)\right)\\
&= \mathbf{G}\left(\mathbf{F}(\mathbf{x}_0)+\mathbf{k}\right)\\
&\qquad\qquad \text{abbiamo denotato } \mathbf{k} := J_{\mathbf{F}}(\mathbf{x}_0)\cdot\mathbf{h} + \mathbf{o}_{\mathbb{R}^p}(\|\mathbf{h}\|)\\
&= \mathbf{G}(\mathbf{F}(\mathbf{x}_0)) + J_{\mathbf{G}}(\mathbf{F}(\mathbf{x}_0))\cdot\mathbf{k} + \mathbf{o}_{\mathbb{R}^m}(\|\mathbf{k}\|)\\
&= \mathbf{G}(\mathbf{F}(\mathbf{x}_0)) + J_{\mathbf{G}}(\mathbf{F}(\mathbf{x}_0))\cdot J_{\mathbf{F}}(\mathbf{x}_0)\mathbf{h} +\\
&\qquad + J_{\mathbf{G}}(\mathbf{F}(\mathbf{x}_0))\cdot\mathbf{o}_{\mathbb{R}^p}(\|\mathbf{h}\|) + \mathbf{o}_{\mathbb{R}^m}(\|\mathbf{k}\|)
\end{aligned}$$

La prova si conclude se dimostriamo che

$$J_{\mathbf{G}}(\mathbf{F}(\mathbf{x}_0)) \cdot \mathbf{o}_{\mathbb{R}^p}(\|\mathbf{h}\|) + \mathbf{o}_{\mathbb{R}^m}(\|\mathbf{k}\|) = \mathbf{o}_{\mathbb{R}^m}(\|\mathbf{h}\|) \qquad \text{per } \mathbf{h} \to \mathbf{0}.$$

Infatti, in questo caso otterremmo che, per $\mathbf{h} \to 0$

$$(\mathbf{G} \circ \mathbf{F})(\mathbf{x}_0 + \mathbf{h}) = \mathbf{G}(\mathbf{F}(\mathbf{x}_0)) + J_{\mathbf{G}}(\mathbf{F}(\mathbf{x}_0)) \cdot J_{\mathbf{F}}(\mathbf{x}_0)\mathbf{h} + \mathbf{o}_{\mathbb{R}^m}(\mathbf{h}).$$

Questo dice che

- $\mathbf{G} \circ \mathbf{F}$ è differenziabile in $\mathbf{x}_0$
- per l'unicità del differenziale,

$$J_{\mathbf{G}\circ\mathbf{F}}(\mathbf{x}_0) = J_{\mathbf{G}}(\mathbf{F}(\mathbf{x}_0)) \cdot J_{\mathbf{F}}(\mathbf{x}_0).$$

Dimostriamo quindi che

$$J_{\mathbf{G}}(\mathbf{F}(\mathbf{x}_0)) \cdot \mathbf{o}_{\mathbb{R}^p}(\|\mathbf{h}\|) + \mathbf{o}_{\mathbb{R}^m}(\|\mathbf{k}\|) = \mathbf{o}_{\mathbb{R}^m}(\|\mathbf{h}\|) \qquad \text{per } \mathbf{h} \to \mathbf{0}$$

o, esplicitando $\mathbf{k}$, dimostriamo che per $\mathbf{h} \to \mathbf{0}$,

$$J_{\mathbf{G}}(\mathbf{F}(\mathbf{x}_0)) \cdot \mathbf{o}_{\mathbb{R}^p}(\|\mathbf{h}\|) + \mathbf{o}_{\mathbb{R}^m}\left(\|J_{\mathbf{F}}(\mathbf{x}_0) \cdot \mathbf{h} + \mathbf{o}_{\mathbb{R}^p}(\|\mathbf{h}\|)\|\right) = \mathbf{o}_{\mathbb{R}^m}(\|\mathbf{h}\|)$$

Consideriamo separatamente i due addendi e dimostriamo che per $\mathbf{h} \to \mathbf{0}$,

$$J_{\mathbf{G}}(\mathbf{F}(\mathbf{x}_0)) \cdot \mathbf{o}_{\mathbb{R}^p}(\|\mathbf{h}\|) = \mathbf{o}_{\mathbb{R}^m}(\|\mathbf{h}\|).$$

$$\mathbf{o}_{\mathbb{R}^m}(\|J_{\mathbf{F}}(\mathbf{x}_0) \cdot \mathbf{h} + \mathbf{o}_{\mathbb{R}^p}(\|\mathbf{h}\|)\|) = \mathbf{o}_{\mathbb{R}^m}(\|\mathbf{h}\|).$$

$J_{\mathbf{G}}(\mathbf{F}(\mathbf{x}_0))$ è una matrice $(m \times p)$ e quindi esiste una costante $c_1 > 0$ tale che per ogni $\mathbf{v} \in R^p$ vale

$$\|J_{\mathbf{G}}(\mathbf{F}(\mathbf{x}_0)) \cdot \mathbf{v}\| \leq c_1 \|\mathbf{v}\|$$

Quindi

$$\lim_{\mathbf{h}\to\mathbf{0}} \frac{\|J_{\mathbf{G}}(\mathbf{F}(\mathbf{x}_0)) \cdot \mathbf{o}_{\mathbb{R}^p}(\|\mathbf{h}\|)\|}{\|\mathbf{h}\|} \leq \lim_{\mathbf{h}\to\mathbf{0}} \frac{c_1 \|\mathbf{o}_{\mathbb{R}^p}(\|\mathbf{h}\|)\|}{\|\mathbf{h}\|} = 0,$$

e abbiamo provato che

$$J_{\mathbf{G}}(\mathbf{F}(\mathbf{x}_0)) \cdot \mathbf{o}_{\mathbb{R}^p}(\|\mathbf{h}\|) = \mathbf{o}_{\mathbb{R}^m}(\|\mathbf{h}\|).$$

Esiste $c_2 > 0$ tale che $\|J_{\mathbf{F}}(\mathbf{x}_0)\cdot\mathbf{h}\| \le c_2\|\mathbf{h}\|$ per ogni $\mathbf{h}\in R^n$; esiste $\delta > 0$ tale che $\|\mathbf{o}_{\mathbb{R}^p}(\|\mathbf{h}\|)\| \le \|\mathbf{h}\|$ per $\|\mathbf{h}\| < \delta$.
Quindi

$$\|J_{\mathbf{F}}(\mathbf{x}_0)\cdot\mathbf{h} + \mathbf{o}_{\mathbb{R}^p}(\|\mathbf{h}\|)\| \le (c_2+1)\|\mathbf{h}\| \qquad \text{per } \|\mathbf{h}\| < \delta.$$

Infine: $$\lim_{\mathbf{h}\to\mathbf{0}} \frac{\|\mathbf{o}_{\mathbb{R}^m}(\|J_{\mathbf{F}}(\mathbf{x}_0)\cdot\mathbf{h} + \mathbf{o}_{\mathbb{R}^p}(\|\mathbf{h}\|)\|)\|}{\|\mathbf{h}\|}$$

$$= \lim_{\mathbf{h}\to\mathbf{0}} \frac{\|\mathbf{o}_{\mathbb{R}^m}(\|J_{\mathbf{F}}(\mathbf{x}_0)\cdot\mathbf{h} + \mathbf{o}_{\mathbb{R}^p}(\|\mathbf{h}\|)\|)\|}{\|J_{\mathbf{F}}(\mathbf{x}_0)\cdot\mathbf{h} + \mathbf{o}_{\mathbb{R}^p}(\|\mathbf{h}\|)\|} \frac{\|J_{\mathbf{F}}(\mathbf{x}_0)\cdot\mathbf{h} + \mathbf{o}_{\mathbb{R}^p}(\|\mathbf{h}\|)\|}{\|\mathbf{h}\|}$$

$$\le \lim_{\mathbf{h}\to\mathbf{0}} \frac{\|\mathbf{o}_{\mathbb{R}^m}(\|J_{\mathbf{F}}(\mathbf{x}_0)\cdot\mathbf{h} + \mathbf{o}_{\mathbb{R}^p}(\|\mathbf{h}\|)\|)\|}{\|J_{\mathbf{F}}(\mathbf{x}_0)\cdot\mathbf{h} + \mathbf{o}_{\mathbb{R}^p}(\|\mathbf{h}\|)\|} (c_2+1) = 0$$

Quindi anche

$$\mathbf{o}_{\mathbb{R}^m}(\|J_{\mathbf{F}}(\mathbf{x}_0)\cdot\mathbf{h} + \mathbf{o}_{\mathbb{R}^p}(\|\mathbf{h}\|)\|) = \mathbf{o}_{\mathbb{R}^m}(\|\mathbf{h}\|).$$

Example

- A sia un aperto di $\mathbb{R}^2$, $f : A \to \mathbb{R}$ e $f \in C^1(A)$;
- $\mathbf{r} : (t_0 - \varepsilon, t_0 + \varepsilon) \to A$ sia una curva C^1;
- $\mathbf{r}$ parametrizzi un tratto di una linea di livello di f, cioè

$$f(\mathbf{r}(t)) = \text{costante, per ogni } t \in (t_0 - \varepsilon, t_0 + \varepsilon).$$

Allora, $f \circ \mathbf{r} : (t_0 - \varepsilon, t_0 + \varepsilon) \to \mathbb{R}$ è una funzione costante per $t \in (t_0 - \varepsilon, t_0 + \varepsilon)$, quindi

$$\begin{aligned} 0 = \frac{d}{dt} f(\mathbf{r}(t)) &= \partial_1 f(\mathbf{r}(t))\, r_1'(t) + \partial_2 f(\mathbf{r}(t))\, r_2'(t) \\ &= \langle (\nabla f)(\mathbf{r}(t)), \mathbf{r}'(t) \rangle \end{aligned}$$

otteniamo che

$$\langle (\nabla f)(\mathbf{r}(t)), \mathbf{r}'(t) \rangle = 0.$$

Corollario

Se

- $f \in C^1(A)$ e $\nabla f \neq 0$,
- una linea di livello di f si può parametrizzare con una curva $\mathbf{r}$ di classe C^1 e con $\mathbf{r}'(t) \neq 0$,

allora, nei punti della linea di livello,

la linea di livello è ortogonale a ∇f.