

# Analisi Matematica II.

## Differenziabilità

Claudio Saccon<sup>1</sup>

<sup>1</sup>Dipartimento di Matematica, Via F. Buonarroti 1/C, 56127 PISA  
email: [claudio.saccon@unipi.it](mailto:claudio.saccon@unipi.it)  
sito web: <http://pagine.dm.unipi.it/csblog1>  
orario di ricevimento: Venerdì mattina alle 9.30

# Derivabilità per funzioni di più variabili

In tutto quanto segue consideriamo una funzione  $f : A \rightarrow \mathbb{R}$  o  $\mathbf{f} : A \rightarrow \mathbb{R}^M$  dove  $A$  è un **aperto** di  $\mathbb{R}^N$  e  $\mathbf{x}_0 \in A$  è un punto..

Il fatto di prendere  $A$  aperto permette di potere “avvicinare  $\mathbf{x}_0$  in  $A$ , da qualunque direzione”.

## derivate direzionali

Se  $\vec{v}$  è un vettore di  $\mathbb{R}^N$  indichiamo:

$$\frac{df}{d\vec{v}}(\mathbf{x}_0) = f'(\mathbf{x}_0)(\vec{v}) := \lim_{h \rightarrow 0} \frac{f(\mathbf{x}_0 + h\vec{v}) - f(\mathbf{x}_0)}{h}$$

ammesso che questo limite esista:  $f'(\mathbf{x}_0)(\vec{v})$  si chiama la *derivata (direzionale) di  $f$  nella direzione  $\vec{v}$  (o lungo  $\vec{v}$ )*.  $f'(\mathbf{x}_0)(\vec{v})$  si può anche vedere come la derivata della funzione  $\varphi(t) := f(\mathbf{x}_0 + t\vec{v})$  in  $t = 0$ .

Nel caso scalare ( $f : A \rightarrow \mathbb{R}$ )  $f'(\mathbf{x}_0)(\vec{v}) \in \mathbb{R}$  mentre nel caso di funzioni a valori vettoriali ( $\mathbf{f} : A \rightarrow \mathbb{R}^M$ )  $\mathbf{f}'(\mathbf{x}_0)(\vec{v})$  è un vettore di  $\mathbb{R}^M$ .

# Derivabilità per funzioni di più variabili

Consideriamo  $i$  intero tra 1 ed  $N$ .

## derivate parziali

Chiamiamo *derivata parziale  $i$ -esima* (o rispetto a  $x_i$ ) il limite (se esiste):

$$\frac{\partial f}{\partial x_i}(\mathbf{x}_0) := \lim_{\xi \rightarrow x_{0,i}} \frac{f(x_{0,1}, \dots, \xi, \dots, x_{0,N}) - f(x_{0,1}, \dots, x_{0,i}, \dots, x_{0,N})}{\xi - x_{0,i}}$$

cioè la derivata in  $\xi_{0,i} = x_{0,i}$  della funzione  $\xi \mapsto f(x_{0,1}, \dots, \xi, \dots, x_{0,N})$  ottenuta “congelando” le componenti di  $\mathbf{x}$  diverse dalla  $i$ -esima.

Le derivate parziali sono un caso particolare delle derivate direzionali.

Se  $\hat{e}_1, \dots, \hat{e}_N$  sono i vettori della base canonica, si ha:

$$\frac{\partial f}{\partial x_i}(\mathbf{x}_0) = f'(\mathbf{x}_0)(\hat{e}_i).$$

# Derivabilità per funzioni di più variabili

## derivate parziali $\nRightarrow$ continuità

Il difetto delle derivate direzionali (e dunque delle derivate parziali) è che la loro esistenza (da sola) non basta a garantire la continuità della funzione.

**Controesempio** La funzione (già incontrata)  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  definita da:

$$f(x, y) := \begin{cases} \frac{xy}{x^2 + y^2} & \text{se } (x, y) \neq (0, 0), \\ 0 & \text{se } (x, y) = (0, 0), \end{cases}$$

non è continua nell'origine pur avendo derivate parziali (nulle) nell'origine:

$$f(x, 0) = 0 \Rightarrow \frac{\partial f}{\partial x}(0, 0) = 0, \quad f(0, y) = 0 \Rightarrow \frac{\partial f}{\partial y}(0, 0) = 0,$$

ma come già visto  $f$  non è continua in  $(0, 0)$ .

# Differenziabilità per funzioni di più variabili

## differenziale

Supponiamo che  $L : \mathbb{R}^N \rightarrow \mathbb{R}$  sia un'applicazione lineare. Diremo che  $L$  è un *differenziale* per  $f$  in  $\mathbf{x}_0$  se:

$$\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} \frac{f(\mathbf{x}) - f(\mathbf{x}_0) - L(\mathbf{x} - \mathbf{x}_0)}{\|\mathbf{x} - \mathbf{x}_0\|} = 0. \quad (*)$$

Diremo che  $f$  è differenziabile in  $\mathbf{x}_0$  se esiste un differenziale per  $f$  in  $\mathbf{x}_0$

La differenziabilità implica la derivabilità direzionale.

## Theorem

Se  $L$  è un differenziale per  $f$  in  $\mathbf{x}_0$  allora  $f$  ammette derivata direzionale lungo un qualunque vettore  $\vec{v}$  e:

$$f'(\mathbf{x}_0)(\vec{v}) = L(\vec{v}).$$

# Differenziabilità per funzioni di più variabili

*Dimostrazione* Sia  $L$  un differenziale e mettiamo  $\mathbf{x} = \mathbf{x}_0 + t\vec{v}$  in  $(\star)$ :

$$\lim_{t \rightarrow 0} \frac{f(\mathbf{x}_0 + t\vec{v}) - f(\mathbf{x}_0) - tL(\vec{v})}{t} = 0$$

(limite della restrizione). Dunque:

$$\lim_{t \rightarrow 0} \frac{f(\mathbf{x}_0 + t\vec{v}) - f(\mathbf{x}_0)}{t} = L(\vec{v}).$$

Dal teorema segue che ci può essere al più **un** differenziale, perchè se  $L$  esiste, allora  $L\vec{v}$  è univocamente determinato per ogni vettore  $\vec{v}$ .

Dunque se  $f$  è differenziabile in  $\mathbf{x}_0$  chiameremo (il) *differenziale* di  $f$  in  $\mathbf{x}_0$  l'unica applicazione lineare  $L : \mathbb{R}^N \rightarrow \mathbb{R}$  per cui vale  $(\star)$ , e lo indicheremo con  $\boxed{df(\mathbf{x}_0)}$ . Sempre per il teorema si ha:

$$df(\mathbf{x}_0)(\vec{v}) = \frac{\partial f}{\partial x_1}(\mathbf{x}_0)v_1 + \cdots + \frac{\partial f}{\partial x_N}(\mathbf{x}_0)v_N.$$

## Theorem

*Se  $f$  è differenziabile in  $\mathbf{x}_0$ , allora è continua in  $\mathbf{x}_0$ .*

**Dimostrazione** Dalla definizione di differenziale si ottiene:

$$\left| \frac{f(\mathbf{x}) - f(\mathbf{x}_0) - L(\mathbf{x} - \mathbf{x}_0)}{\|\mathbf{x} - \mathbf{x}_0\|} \right| \leq 1 \quad \text{per } \mathbf{x} \in D(\mathbf{x}_0, \rho)$$

pur di prendere  $\rho > 0$  sufficientemente piccolo. Allora:

$$|f(\mathbf{x}) - f(\mathbf{x}_0) - L(\mathbf{x} - \mathbf{x}_0)| \leq \|\mathbf{x} - \mathbf{x}_0\| \quad \text{per } \mathbf{x} \in D(\mathbf{x}_0, \rho)$$

e quindi:

$$|f(\mathbf{x}) - f(\mathbf{x}_0)| \leq (\|L\| + 1)\|\mathbf{x} - \mathbf{x}_0\| \quad \text{per } \mathbf{x} \in D(\mathbf{x}_0, \rho)$$

che implica  $f(\mathbf{x}) \rightarrow f(\mathbf{x}_0)$  per  $\mathbf{x} \rightarrow \mathbf{x}_0$ .

# Differenziabilità per funzioni di più variabili

**Esempio** Sia

$$f(x, y) := \begin{cases} \frac{xy^2}{x^2 + y^2} & \text{se } (x, y) \neq (0, 0), \\ 0 & \text{se } (x, y) = (0, 0), \end{cases}$$

Si può vedere che  $f$  è continua. Peraltro se  $\vec{v} = (v_x, v_y)$ , allora

$$f(0 + tv_x, 0 + tv_y) = \frac{tv_x v_y^2}{v_x^2 + v_y^2} = \text{costante in } t \Rightarrow f'(0, 0)(\vec{v}) = \frac{v_x v_y^2}{v_x^2 + v_y^2}.$$

Questo mostra che  $f$  **non è differenziabile** in  $\mathbf{0} = (0, 0)$ , perché  $\vec{v} \mapsto f'(\mathbf{0})(\vec{v})$  non è lineare.

Si potrebbe far vedere che l'ipotesi che  $f'(\mathbf{x}_0)(\vec{v})$  esista per ogni  $\vec{v}$  e sia lineare in  $\vec{v}$  non basta a garantire la differenziabilità di  $f$  in  $\mathbf{x}_0$ .

# Differenziabilità per funzioni di più variabili

L'esistenza del differenziale in  $\mathbf{x}_0$  si può esprimere:

$$f(\mathbf{x}) = f(\mathbf{x}_0) + df(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0) + o(\|\mathbf{x} - \mathbf{x}_0\|)$$

che si interpreta dicendo che “localmente”  $f(\mathbf{x})$  si comporta come:

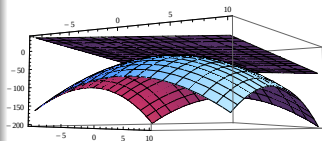
$$r(\mathbf{x}) := f(\mathbf{x}_0) + df(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0)$$

## piano tangente

L'applicazione affine  $r(\mathbf{x})$  ha come grafico il *piano tangente* al grafico di  $f(\mathbf{x})$ . Si dice che:

$$z = f(\mathbf{x}_0) + df(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0)$$

è l'equazione del piano tangente al grafico di  $f$  in  $(\mathbf{x}_0, f(\mathbf{x}_0))$ .



## Theorem (Teorema del differenziale totale)

*Se le derivate parziali esistono in tutto un intorno di  $\mathbf{x}_0$  e sono continue in  $\mathbf{x}_0$ , allora  $f$  è differenziabile in  $\mathbf{x}_0$ .*

*Dimostrazione* Facciamo la dimostrazione nel caso  $N = 2$ . Per brevità chiamo  $P_0$  il punto (invece di  $\mathbf{x}_0$ ) e indico con  $(x_0, y_0)$  le sue coordinate. Analogamente chiamo  $P$ , di coordinate  $(x, y)$  il generico punto di  $\mathbb{R}^2$ .

Poniamo anche  $L(v_x, v_y) := \frac{\partial f}{\partial x}(P_0)v_x + \frac{\partial f}{\partial y}(P_0)v_y$ . Devo dimostrare:

$$\lim_{P \rightarrow P_0} \frac{f(P) - f(P_0) - L(P - P_0)}{\|P - P_0\|} = 0 \quad \text{cioè:}$$

$$\lim_{(x,y) \rightarrow (x_0,y_0)} \frac{f(x,y) - f(x_0,y_0) - \frac{\partial f}{\partial x}(x_0,y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0,y_0)(y - y_0)}{\sqrt{(x - x_0)^2 + (y - y_0)^2}} = 0$$

# Differenziale totale

Chiamiamo  $\Delta(x, y)$  il numeratore. Allora, applicando il teorema di Lagrange rispetto a  $y$  ( $x$  fissata):

$$\begin{aligned}\Delta(x, y) &= \Delta(x, y_0) + \frac{\partial \Delta}{\partial y}(x, \eta(x, y))(y - y_0) = \\ &\quad \Delta(x, y_0) + \left( \frac{\partial f}{\partial y}(x, \eta(x, y)) - \frac{\partial f}{\partial y}(x_0, y_0) \right) (y - y_0)\end{aligned}$$

per un opportuno punto  $\eta(x, y)$  compreso tra  $y$  e  $y_0$ . Analogamente:

$$\begin{aligned}\Delta(x, y_0) &= \Delta(x_0, y_0) + \frac{\partial \Delta}{\partial x}(\xi(x, y_0), y_0)(x - x_0) = \\ &\quad \underbrace{\Delta(x_0, y_0)}_{=0} + \left( \frac{\partial f}{\partial x}(\xi(x, y_0), y_0) - \frac{\partial f}{\partial x}(x_0, y_0) \right) (x - x_0)\end{aligned}$$

per un opportuno  $\xi(x, y_0)$  compreso tra  $x$  e  $x_0$ . Mettendo tutto insieme:

# Differenziale totale

$$\Delta(x, y) = \left( \frac{\partial f}{\partial x}(\xi(x, y_0), y_0) - \frac{\partial f}{\partial x}(x_0, y_0) \right) (x - x_0) + \\ \left( \frac{\partial f}{\partial y}(x, \eta(x, y)) - \frac{\partial f}{\partial y}(x_0, y_0) \right) (y - y_0),$$

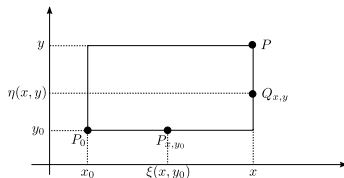
da cui, usando la disuguaglianza di Schwartz:

$$|\Delta(x, y)| \leq \left( \left( \frac{\partial f}{\partial x}(\xi(x, y_0), y_0) - \frac{\partial f}{\partial x}(x_0, y_0) \right)^2 + \right. \\ \left. \left( \frac{\partial f}{\partial y}(x, \eta(x, y)) - \frac{\partial f}{\partial y}(x_0, y_0) \right)^2 \right)^{1/2} \|P - P_0\|$$

e quindi, se  $P_{x,y_0} := (\xi(x, y_0), y_0)$  e  $Q_{x,y} := (x, \eta(x, y))$ , si ha:

$$\left| \frac{\Delta(P)}{\|P - P_0\|} \right| \leq \left\| \left( \frac{\partial f}{\partial x}(P_{x,y_0}) - \frac{\partial f}{\partial x}(P_0), \frac{\partial f}{\partial y}(Q_{x,y}) - \frac{\partial f}{\partial y}(P_0) \right) \right\|$$

# Differenziale totale



Facendo tendere  $(x, y)$  a  $(x_0, y_0)$  si ha che  $Q_{x,y} \rightarrow P_0$  e  $P_{x,y_0} \rightarrow P_0$  (vedi figura) e quindi, per la continuità delle derivate parziali, si ha:

$$\frac{\partial f}{\partial x}(P_{x,y_0}) \rightarrow \frac{\partial f}{\partial x}(P_0), \quad \frac{\partial f}{\partial y}(Q_{x,y}) \rightarrow \frac{\partial f}{\partial y}(P_0).$$

Ne segue la tesi:

$$\left| \frac{\Delta(P)}{\|P - P_0\|} \right| \rightarrow 0.$$

# Differenziale totale: esempio

**Esempio** Prendiamo la funzione:

$$f(x, y) := \begin{cases} \frac{xy^3}{x^2+y^2} & \text{se } (x, y) \neq (0, 0), \\ 0 & \text{se } (x, y) = (0, 0). \end{cases}$$

Allora  $f$  è differenziabile in  $(0, 0)$  e  $df(0, 0) = 0$  (cioè  $df(0, 0)(v_x, v_y) = 0$  per ogni  $v_x, v_y$ ).

**Primo modo** (Usando la definizione di differenziale) mostrare che:

$$\lim_{(x,y) \rightarrow (0,0)} \frac{f(x, y)}{\|(x, y)\|} = 0$$

**Secondo modo** (Usando il teorema del differenziale totale) mostrare che:

$$\lim_{(x,y) \rightarrow (0,0)} \frac{\partial f}{\partial x}(x, y) = 0, \quad \lim_{(x,y) \rightarrow (0,0)} \frac{\partial f}{\partial y}(x, y) = 0.$$

# Gradiente

Abbiamo visto che  $f$  può avere derivate parziali senza essere differenziabile. Però se  $df(\mathbf{x}_0)$  esiste, è completamente determinata dalle derivate parziali.

Ricordiamo che  $df(\mathbf{x}_0)$  è un'applicazione lineare da  $\mathbb{R}^N$  in  $\mathbb{R}$  tale che:

$$df(\mathbf{x}_0)(\vec{v}) = f'(\mathbf{x}_0)(\vec{v}) \quad \forall \vec{v} \in \mathbb{R}^N.$$

Ma allora esiste una matrice  $1 \times N$   $A = (a_1, \dots, a_N)$  tale che  $df(\mathbf{x}_0)\vec{v} = A\vec{v}$  per ogni  $\vec{v}$ . Prendendo  $\vec{v} = \hat{e}_i$  (versore  $i$ -esimo):

$$A\hat{e}_i = f'(\mathbf{x}_0)(\hat{e}_i) \Leftrightarrow a_i = \frac{\partial f}{\partial x_i}(\mathbf{x}_0) \quad i = 1, \dots, N.$$

Chiamiamo *gradiente* di  $f$  in  $\mathbf{x}_0$  il vettore  $\nabla f(\mathbf{x}_0) := A^T$ . Allora:

$$df(\mathbf{x}_0)(\vec{v}) = \nabla f(\mathbf{x}_0) \cdot \vec{v} \quad \forall \vec{v}.$$

(rappresentazione delle forme lineari)

$$\nabla f(\mathbf{x}_0) = \begin{pmatrix} \frac{\partial f}{\partial x_1}(\mathbf{x}_0) \\ \vdots \\ \frac{\partial f}{\partial x_N}(\mathbf{x}_0) \end{pmatrix}$$

# Proprietà del gradiente

Dunque, se  $f$  è scalare ed è differenziabile:

$$f'(\mathbf{x}_0)(\vec{v}) = \nabla f(\mathbf{x}_0) \cdot \vec{v} \quad \forall \vec{v} \in \mathbb{R}^N.$$

Limitiamoci ai vettori  $\vec{v}$  di norma 1:  $\|\vec{v}\| = 1$ . Allora (usando la disuguaglianza di Schwartz) la formula sopra ci dice che la derivata direzionale  $f'(\mathbf{x}_0)(\vec{v})$  è **massima** se  $\vec{v} = \frac{\nabla f(\mathbf{x}_0)}{\|\nabla f(\mathbf{x}_0)\|}$ , in sostanza la direzione di  $\nabla f(\mathbf{x}_0)$  è la “direzione di massima crescita per  $f$ ”.

La formula dice anche che  $\|\nabla f(\mathbf{x}_0)\|$  è il **valore massimo** delle derivate direzionali rispetto a tutti i versori.

## Caso a valori vettoriali

Se  $\mathbf{f} : A \rightarrow \mathbb{R}^M$  la definizione di differenziabilità è la stessa: se

$$\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} \frac{\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{x}_0) - L(\mathbf{x} - \mathbf{x}_0)}{\|\mathbf{x} - \mathbf{x}_0\|} = 0. \quad (\star)$$

con  $L : \mathbb{R}^N \rightarrow \mathbb{R}^M$  lineare, diciamo che  $\mathbf{f}$  è differenziabile in  $\mathbf{x}_0$  e indichiamo l'applicazione  $L$  con  $d\mathbf{f}(\mathbf{x}_0)$ . Come prima:

$$d\mathbf{f}(\mathbf{x}_0)(\vec{v}) = \mathbf{f}'(\mathbf{x}_0)(\vec{v}) = \frac{d\mathbf{f}}{d\vec{v}}(\mathbf{x}_0) \quad \forall \vec{v} \in \mathbb{R}^N$$

Ora in arrivo c'è  $\mathbb{R}^M$  invece di  $\mathbb{R}$ . Quindi  $(\star)$  significa che ognuna delle  $M$  componenti dell'espressione scritta deve tendere a zero. Analogamente nell'ultima relazione le derivate direzionali  $\mathbf{f}'(\mathbf{x}_0)(\vec{v})$  sono dei vettori di  $\mathbb{R}^M$ .

## Caso a valori vettoriali

$\mathbf{f}$  è differenziabile in  $\mathbf{x}_0$  se e solo se ogni componente  $f_j$  di  $\mathbf{f}$  lo è; inoltre:

$$d\mathbf{f}(\mathbf{x}_0)(\vec{v}) = \begin{pmatrix} df_1(\mathbf{x}_0)(\vec{v}) \\ \vdots \\ df_M(\mathbf{x}_0)(\vec{v}) \end{pmatrix}$$

### Matrice Jacobiana

Chiamiamo *matrice Jacobiana* di  $\mathbf{f}$  in  $\mathbf{x}_0$  la matrice

$$J_{\mathbf{f}}(\mathbf{x}_0) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\mathbf{x}_0) & \frac{\partial f_1}{\partial x_2}(\mathbf{x}_0) & \cdots & \frac{\partial f_1}{\partial x_N}(\mathbf{x}_0) \\ \vdots & \vdots & & \vdots \\ \frac{\partial f_M}{\partial x_1}(\mathbf{x}_0) & \frac{\partial f_M}{\partial x_2}(\mathbf{x}_0) & \cdots & \frac{\partial f_M}{\partial x_N}(\mathbf{x}_0) \end{pmatrix}$$

$J_{\mathbf{f}}(\mathbf{x}_0)\vec{v} = d\mathbf{f}(\mathbf{x}_0)(\vec{v})$  per ogni  $\vec{v} \in \mathbb{R}^N$  (rappresentazione matriciale di  $d\mathbf{f}$ ).

# Teoremi di calcolo

- Se  $\mathbf{f}, \mathbf{g}$  sono differenziabili in  $\mathbf{x}_0$  e  $\lambda, \mu \in \mathbb{R}$  anche  $\lambda\mathbf{f} + \mu\mathbf{g}$  lo è e:

$$d(\lambda\mathbf{f} + \mu\mathbf{g})(\mathbf{x}_0) = \lambda d\mathbf{f}(\mathbf{x}_0) + \mu d\mathbf{g}(\mathbf{x}_0);$$

in termini di matrici Jacobiane:

$$J_{\lambda\mathbf{f} + \mu\mathbf{g}}(\mathbf{x}_0) = \lambda J_{\mathbf{f}}(\mathbf{x}_0) + \mu J_{\mathbf{g}}(\mathbf{x}_0).$$

- Se  $\mathbf{f}, g$  sono differenziabili in  $\mathbf{x}_0$  ( $g$  è scalare!) anche  $g\mathbf{f}$  lo è e:

$$d(g\mathbf{f})(\mathbf{x}_0)(\vec{v}) = dg(\mathbf{x}_0)(\vec{v})\mathbf{f}(\mathbf{x}_0) + g(\mathbf{x}_0)d\mathbf{f}(\mathbf{x}_0)(\vec{v}) \quad \forall \vec{v} \in \mathbb{R}^N;$$

in termini di matrici Jacobiane:

$$J_{g\mathbf{f}}(\mathbf{x}_0) = \nabla g(\mathbf{x}_0) \times \mathbf{f}(\mathbf{x}_0) + g(\mathbf{x}_0)J_{\mathbf{f}}(\mathbf{x}_0)$$

(se  $\vec{v} = (v_i)_{i=1, \dots, N_1}$ ,  $\vec{w} = (w_j)_{j=1, \dots, N_2}$  allora  $\vec{v} \times \vec{w} = (v_i w_j)_{i,j}$ ).

# Teoremi di calcolo

- Se  $\mathbf{f} : A \rightarrow B \subset \mathbb{R}^M$  è differenziabile in  $\mathbf{x}_0 \in A \subset \mathbb{R}^N$ , se  $\mathbf{g} : B \rightarrow \mathbb{R}^K$  è differenziabile in  $\mathbf{y}_0 := \mathbf{f}(\mathbf{x}_0)$ , allora  $\mathbf{g} \circ \mathbf{f} : A \rightarrow \mathbb{R}^K$  è differenziabile in  $\mathbf{x}_0$  e si ha:

$$d(\mathbf{g} \circ \mathbf{f})(\mathbf{x}_0)(\vec{v}) = d\mathbf{g}(\mathbf{y}_0)(d\mathbf{f}(\mathbf{x}_0)(\vec{v})) \quad \forall \vec{v} \in \mathbb{R}^N;$$

in termini di matrici Jacobiane questo equivale a:

$$J_{\mathbf{g} \circ \mathbf{f}}(\mathbf{x}_0) = J_{\mathbf{g}}(\mathbf{y}_0) J_{\mathbf{f}}(\mathbf{x}_0).$$

- Se  $\mathbf{f}, \mathbf{g}$  sono differenziabili in  $\mathbf{x}_0$  anche  $\mathbf{f} \cdot \mathbf{g}$  lo è e:

$$d(\mathbf{f} \cdot \mathbf{g})(\mathbf{x}_0)(\vec{v}) = d\mathbf{f}(\mathbf{x}_0)(\vec{v}) \cdot \mathbf{g}(\mathbf{x}_0) + \mathbf{f}(\mathbf{x}_0) \cdot d\mathbf{g}(\mathbf{x}_0)(\vec{v}) \quad \forall \vec{v} \in \mathbb{R}^N;$$

in termini di matrici Jacobiane questo equivale a:

$$\nabla(\mathbf{f} \cdot \mathbf{g})(\mathbf{x}_0) = J_{\mathbf{f}}(\mathbf{x}_0)^T \mathbf{g}(\mathbf{x}_0) + J_{\mathbf{g}}(\mathbf{x}_0)^T \mathbf{f}(\mathbf{x}_0).$$

# Teoremi di calcolo

**Dimostrazione** della regola di composizione.

Per definizione di differenziale:

$$\mathbf{f}(\mathbf{x}) = \mathbf{f}(\mathbf{x}_0) + d\mathbf{f}(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0) + o(\|\mathbf{x} - \mathbf{x}_0\|_N)$$

$$\mathbf{g}(\mathbf{y}) = \mathbf{g}(\mathbf{y}_0) + d\mathbf{g}(\mathbf{y}_0)(\mathbf{y} - \mathbf{y}_0) + o(\|\mathbf{y} - \mathbf{y}_0\|_M)$$

(al solito  $o(h)$  indica un infinitesimo di ordine superiore ad  $h$ , cioè una quantità  $k$  tale che  $k/h \rightarrow 0$  – analogamente  $O(h)$  indica una  $k$  tale che  $k/h$  ha limite finito). Allora:

$$\begin{aligned} \mathbf{g}(\mathbf{f}(\mathbf{x})) &= \mathbf{g}(\mathbf{f}(\mathbf{x}_0)) + d\mathbf{g}(\mathbf{y}_0)(d\mathbf{f}(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0)) + \\ &\quad \underbrace{d\mathbf{g}(\mathbf{y}_0)(o(\|\mathbf{x} - \mathbf{x}_0\|_N))}_{=o(\|\mathbf{x} - \mathbf{x}_0\|_N)} + \underbrace{o(d\mathbf{f}(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0) + o(\|\mathbf{x} - \mathbf{x}_0\|_N))}_{=O(\|\mathbf{x} - \mathbf{x}_0\|_N)} = \\ &\quad \mathbf{g}(\mathbf{f}(\mathbf{x}_0)) + d\mathbf{g}(\mathbf{y}_0)(d\mathbf{f}(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0)) + o(\|\mathbf{x} - \mathbf{x}_0\|_N) \end{aligned}$$

(per le proprietà degli infinitesimi). Ne segue la tesi.

## derivate parziali della funzione composta

La regola di composizione, letta sulle componenti, corrisponde a:

$$\frac{\partial g_k(f_1(x_1, \dots, x_N), \dots, f_M(x_1, \dots, x_N))}{\partial x_i}(\mathbf{x}_0) = \sum_{j=1}^M \frac{\partial g_k}{\partial y_j}(\mathbf{y}_0) \frac{\partial f_j}{\partial x_i}(\mathbf{x}_0)$$

per  $i = 1, \dots, N$ ,  $k = 1, \dots, K$ , dove ricordiamo che  $\mathbf{y}_0 = \mathbf{f}(\mathbf{x}_0)$ .

Infatti la sommatoria scritta sopra è proprio l'elemento  $(k, i)$  della matrice prodotto tra la matrice  $J_{\mathbf{g}}(\mathbf{y}_0) = \left( \frac{\partial \mathbf{g}_k}{\partial y_j}(\mathbf{y}_0) \right)_{\substack{k=1, \dots, K \\ j=1, \dots, M}}$  e la matrice

$$J_{\mathbf{f}}(\mathbf{x}_0) = \left( \frac{\partial \mathbf{f}_j}{\partial x_i}(\mathbf{x}_0) \right)_{\substack{j=1, \dots, M \\ i=1, \dots, N}}.$$

# Curve di livello

Data  $f : A \rightarrow \mathbb{R}$  e  $c \in \mathbb{R}$  chiamo *insieme di livello* per  $f$  relativo a  $c$  l'insieme:

$$\{f = c\} := \{\mathbf{x} \in A : f(\mathbf{x}) = c\}.$$

Se  $\gamma$  ha supporto in  $\{f = c\}$ , cioè se  $f(\gamma(t)) = c$  per ogni  $t$ , dico che  $\gamma$  è una *curva di livello*  $c$  per  $f$ .

Se  $\gamma$  è una curva di livello regolare per  $f$ , allora in ogni punto  $P = \gamma(t)$  il gradiente  $\nabla f(P)$  è ortogonale alla direzione tangente  $\gamma'(P)$ .

Infatti da  $f(\gamma(t)) = c$  si ottiene  $\frac{d}{dt}f(\gamma(t)) = 0$  per ogni  $t$ . Usando la regola sulla derivazione della funzione composta:

$$0 = \frac{d}{dt}f(\gamma(t)) = \sum_{i=1}^M \frac{\partial f}{\partial x_i}(\gamma(t)) \gamma'_i(t) = \nabla f(\gamma(t)) \cdot \gamma'(t).$$

# Derivate seconde

## Derivate direzionali seconde

Siano  $\vec{v}_1$  e  $\vec{v}_2$  due vettori di  $\mathbb{R}^N$ . Poniamo:

$$\frac{d^2 f}{d\vec{v}_1 d\vec{v}_2}(\mathbf{x}_0) = f''(\mathbf{x}_0)(\vec{v}_1, \vec{v}_2) := g'(\mathbf{x}_0)(\vec{v}_1) \quad \text{dove } g(\mathbf{x}) =: f'(\mathbf{x})(\vec{v}_2)$$

cioè la derivata direzione lungo  $\vec{v}_1$  della derivata direzionale lungo  $\vec{v}_2$  (supponendo che tutto esista). Convenzione:  $f''(\mathbf{x}_0)(\vec{v}^2) := f''(\mathbf{x}_0)(\vec{v}, \vec{v})$ .

## Derivate parziali seconde

Chiamiamo derivata parziale seconda rispetto a  $x_i$  e  $x_j$ :

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x}_0) := \frac{\partial}{\partial x_i} \frac{\partial}{\partial x_j} f(\mathbf{x}_0) = f''(\mathbf{x}_0)(\hat{e}_i, \hat{e}_j),$$

A priori  $f''(\mathbf{x}_0)(\vec{v}_1, \vec{v}_2) \neq f''(\mathbf{x}_0)(\vec{v}_2, \vec{v}_1)$ .

## Derivate seconde

**Controesempio** Prendiamo la funzione  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  definita da:

$$f(x, y) := \begin{cases} \frac{xy(x^2 - y^2)}{x^2 + y^2} & \text{se } (x, y) \neq (0, 0). \\ 0 & \text{se } (x, y) = (0, 0). \end{cases}$$

Si vede che  $f$  ha derivate parziali prime continue in  $(0, 0)$ , punto in cui entrambe queste derivate fanno zero. Con qualche calcolo:

$$\frac{\partial f}{\partial x}(x, y) = \frac{y(x^4 + 4x^2y^2 - y^4)}{(x^2 + y^2)^2}, \quad \frac{\partial f}{\partial y}(x, y) = \frac{x(x^4 - 4x^2y^2 - y^4)}{(x^2 + y^2)^2}.$$

$$\frac{\partial}{\partial x} \frac{\partial}{\partial y} f(0, 0) = \lim_{x \rightarrow 0} \frac{1}{x} \left( \frac{\partial f}{\partial y}(x, 0) - \frac{\partial f}{\partial y}(0, 0) \right) = \lim_{x \rightarrow 0} \frac{x^4}{x^4} = 1,$$

$$\frac{\partial}{\partial y} \frac{\partial}{\partial x} f(0, 0) = \lim_{y \rightarrow 0} \frac{1}{y} \left( \frac{\partial f}{\partial x}(0, y) - \frac{\partial f}{\partial x}(0, 0) \right) = \lim_{y \rightarrow 0} \frac{-y^4}{y^4} = -1.$$

Quindi cambiando l'ordine di derivazione il risultato cambia.

## Theorem (Theorema di Schwartz)

*Supponiamo che le derivate parziali seconde esistano in un intorno di  $\mathbf{x}_0$  e siano continue in  $\mathbf{x}_0$ . Allora  $\frac{\partial^2}{\partial x_i \partial x_j} f(\mathbf{x}_0) = \frac{\partial^2}{\partial x_j \partial x_i} f(\mathbf{x}_0)$ .*

**Dimostrazione (con  $N = 2$ )** Consideriamo la funzione (definita in un intorno di  $(x_0, y_0)$ ):

$$\Delta(x, y) := f(x, y) - f(x_0, y) - f(x, y_0) + f(x_0, y_0).$$

Se poniamo anche  $\Phi_x(y) := f(x, y) - f(x_0, y)$  (dove  $x$  è considerato un parametro) possiamo scrivere:

$$\Delta(x, y) = \Phi_x(y) - \Phi_x(y_0) = \Phi'_x(\eta)(y - y_0)$$

per un opportuno  $\eta = \eta(x, y)$  tra  $y_0$  e  $y$  (uso Lagrange). Calcolando  $\Phi'_x$ :

$$\Delta(x, y) = \left( \frac{\partial f}{\partial y}(x, \eta) - \frac{\partial f}{\partial y}(x_0, \eta) \right) (y - y_0) =$$
$$\frac{\partial}{\partial x} \frac{\partial f}{\partial y}(\xi, \eta)(x - x_0)(y - y_0)$$

per un opportuno  $\xi = \xi(x, y)$  tra  $x_0$  e  $x$  (riapplico Lagrange a  $x \mapsto \frac{\partial f}{\partial y}(x, \eta)$ ). Questo procedimento si può ripetere invertendo  $x$  e  $y$  (cominciando con lo scrivere  $\Delta(x, y) = \Psi_y(x) - \Psi_y(x_0)$  dove  $\Psi_y(x) = f(x, y) - f(x, y_0)$  e facendo gli stessi calcoli). Si trova:

$$\Delta(x, y) = \frac{\partial}{\partial y} \frac{\partial f}{\partial x}(\xi', \eta')(y - y_0)(x - x_0).$$

con  $\xi' = \xi'(x, y)$  compreso tra  $x_0$  e  $x$  e  $\eta' = \eta'(x, y)$  compreso tra  $y_0$  e  $y$ .

## Derivate seconde

Eguagliando e semplificando  $(x - x_0)(y - y_0)$ :

$$\frac{\partial}{\partial x} \frac{\partial f}{\partial y}(\xi, \eta) = \frac{\partial}{\partial y} \frac{\partial f}{\partial x}(\xi', \eta')$$

A questo punto facciamo tendere  $(x, y) \rightarrow (x_0, y_0)$ : notiamo che i punti passano al limite:

$$(\xi(x, y), \eta(x, y)) \rightarrow (x_0, y_0); \quad (\xi'(x, y), \eta'(x, y)) \rightarrow (x_0, y_0)$$

(per come sono stati costruiti) e allora, a causa dell'ipotesi di continuità delle derivate seconde:

$$\frac{\partial}{\partial x} \frac{\partial f}{\partial y}(\xi, \eta) \rightarrow \frac{\partial}{\partial x} \frac{\partial f}{\partial y}(x_0, y_0), \quad \frac{\partial}{\partial y} \frac{\partial f}{\partial x}(\xi', \eta') \rightarrow \frac{\partial}{\partial y} \frac{\partial f}{\partial x}(x_0, y_0).$$

Dato che l'eguaglianza passa al limite:  $\frac{\partial}{\partial x} \frac{\partial f}{\partial y}(x_0, y_0) = \frac{\partial}{\partial y} \frac{\partial f}{\partial x}(x_0, y_0).$

## Hessiano

Chiamiamo matrice Hessiana la matrice  $N \times N$  definita da

$$H_f(\mathbf{x}_0) := \begin{pmatrix} \frac{\partial^2}{\partial x_1 \partial x_1} f(\mathbf{x}_0) & \cdots & \frac{\partial^2}{\partial x_N \partial x_1} f(\mathbf{x}_0) \\ \vdots & & \vdots \\ \frac{\partial^2}{\partial x_1 \partial x_N} f(\mathbf{x}_0) & \cdots & \frac{\partial^2}{\partial x_N \partial x_N} f(\mathbf{x}_0) \end{pmatrix}$$

## simmetria dell'Hessiano

Il teorema di Schwartz di che se  $H_f$  è continuo, allora è **simmetrico**.

## Derivate seconde

Si vede subito che  $H_f$  è lo Jacobiano dell'applicazione  $\mathbf{x} \mapsto \nabla f(\mathbf{x})$ . Se  $H_f$  è continua allora  $\nabla f$  è differenziabile (teor. del differenziale totale). Allora:

$$\begin{aligned} f''(\mathbf{x}_0)(\vec{v}_1, \vec{v}_2) &= \frac{d}{d\vec{v}_1} \frac{df}{d\vec{v}_2}(\mathbf{x}_0) = \frac{d}{d\vec{v}_1} (\nabla f(\mathbf{x}_0) \cdot \vec{v}_2) = \\ &= \left( \frac{d}{d\vec{v}_1} \nabla f(\mathbf{x}_0) \right) \cdot \vec{v}_2 = (J_{\nabla f}(\mathbf{x}_0) \vec{v}_1) \cdot \vec{v}_2 = (H_f(\mathbf{x}_0) \vec{v}_1) \cdot \vec{v}_2 \end{aligned}$$

Dunque la mappa  $(\vec{v}_1, \vec{v}_2) \mapsto f''(\mathbf{x}_0)(\vec{v}_1, \vec{v}_2)$  è **bilineare**. Questa applicazione bilineare viene detta *differenziale secondo* di  $f$  in  $\mathbf{x}_0$  e si indica con  $d^2 f(\mathbf{x}_0) : \mathbb{R}^N \times \mathbb{R}^N \rightarrow \mathbb{R}$ :

$$d^2 f(\mathbf{x}_0)(\vec{v}_1, \vec{v}_2) = \vec{v}_1 \cdot H_f(\mathbf{x}_0) \vec{v}_2 = \vec{v}_1^T H_f(\mathbf{x}_0) \vec{v}_2$$

Si potrebbe introdurre  $d^2 f(\mathbf{x}_0)$  come il *differenziale del differenziale* di  $f$ , senza l'ipotesi che  $H_f$  sia continua – rinunciando a questa generalità.

## Formula di Taylor al secondo ordine

Ricordiamo che  $A$  si dice *convesso* se dati due punti  $x_1, x_2$  in  $A$ , allora il segmento che li congiunge è tutto contenuto in  $A$ :

$$x_1, x_2 \in A \Rightarrow tx_1 + (1 - t)x_2 \in A \quad \forall t \in [0, 1].$$

### Formula di Taylor con resto di Lagrange

Supponiamo che  $A$  sia convesso e che  $f$  sia di classe  $\mathcal{C}^2$ , Siano  $x_0, x \in A$ .

- Esiste un punto  $\xi$  interno al segmento tra  $x_0$  e  $x$  (cioè  $\xi = tx_1 + (1 - t)x_2$  per un  $t$  in  $]0, 1[$ ) tale che:

$$f(x) = f(x_0) + \nabla f(\xi) \cdot (x - x_0)$$

- Esiste un punto  $\xi'$  interno al segmento tra  $x_0$  e  $x$  tale che:

$$f(x) = f(x_0) + \nabla f(x_0) \cdot (x - x_0) + \frac{1}{2}(x - x_0) \cdot H_f(\xi')(x - x_0),$$

## Formula di Taylor al secondo ordine

**Dimostrazione** Sia  $\varphi(t) := f(\mathbf{x}_0 + t(\mathbf{x} - \mathbf{x}_0))$ ;  $\varphi$  è derivabile due volte e  
 $\varphi'(t) = \nabla f(\mathbf{x}_0 + t(\mathbf{x} - \mathbf{x}_0)) \cdot (\mathbf{x} - \mathbf{x}_0)$ ,  $\varphi''(t) = (\mathbf{x} - \mathbf{x}_0) \cdot H_f(\mathbf{x}_0 + t(\mathbf{x} - \mathbf{x}_0))(\mathbf{x} - \mathbf{x}_0)$ .

Applicando il teorema di Lagrange a  $\varphi$  trovo  $\tau$  tra 0 e 1 tale che:

$$\varphi(1) = \varphi(0) + \varphi'(\tau)(1 - 0) \quad \text{e cioè}$$

$$f(\mathbf{x}) = f(\mathbf{x}_0) + \nabla f(\xi) \cdot (\mathbf{x} - \mathbf{x}_0) \quad \text{dove } \xi := \mathbf{x}_0 + \tau(\mathbf{x} - \mathbf{x}_0)$$

Analogamente, usando Taylor secondo Lagrange (al secondo ordine) per  $\varphi$ :

$$\varphi(1) = \varphi(0) + \varphi'(0)(1 - 0) + \frac{1}{2}\varphi''(\tau')(1 - 0)^2$$

per un opportuno  $\tau'$  tra 0 e 1. Sostituendo a  $\varphi$ ,  $\varphi'$  e  $\varphi''$  le rispettive espressioni, e ponendo  $\xi' := \mathbf{x}_0 + \tau'(\mathbf{x} - \mathbf{x}_0)$ :

$$f(\mathbf{x}) = f(\mathbf{x}_0) + \nabla f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0) + \frac{1}{2}(\mathbf{x} - \mathbf{x}_0) \cdot H_f(\xi')(\mathbf{x} - \mathbf{x}_0),$$

# Formula di Taylor al secondo ordine

## Formula di Taylor con resto di Peano

Supponiamo che  $f$  sia di classe  $\mathcal{C}^2$ . Allora:

$$\frac{f(\mathbf{x}) - f(\mathbf{x}_0) - \nabla f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0) - \frac{1}{2}(\mathbf{x} - \mathbf{x}_0) \cdot H_f(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0)}{\|\mathbf{x} - \mathbf{x}_0\|^2} = 0$$

In altri termini:

$$f(\mathbf{x}) = f(\mathbf{x}_0) + \nabla f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0) + \frac{1}{2}(\mathbf{x} - \mathbf{x}_0) \cdot H_f(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0) + R(\mathbf{x}, \mathbf{x}_0)$$

dove:

$$\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} \frac{R(\mathbf{x}, \mathbf{x}_0)}{\|\mathbf{x} - \mathbf{x}_0\|^2} = 0$$

# Formula di Taylor al secondo ordine

**Dimostrazione** Prendiamo  $R > 0$  in modo che  $D(\mathbf{x}_0, R) \subset A$ . Allora  $D(\mathbf{x}_0, R)$  è convesso e per ogni  $\mathbf{x} \in D(\mathbf{x}_0, R)$  possiamo usare Taylor con il resto di Lagrange:

$$\begin{aligned} f(\mathbf{x}) - f(\mathbf{x}_0) - \nabla f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0) - \frac{1}{2}(\mathbf{x} - \mathbf{x}_0) \cdot H_f(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0) = \\ \frac{1}{2}(\mathbf{x} - \mathbf{x}_0) \cdot H_f(\xi)(\mathbf{x} - \mathbf{x}_0) - \frac{1}{2}(\mathbf{x} - \mathbf{x}_0) \cdot H_f(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0) =: R(\mathbf{x}, \mathbf{x}_0) \end{aligned}$$

dove  $\xi = \xi(\mathbf{x}, \mathbf{x}_0)$  è sul segmento tra  $\mathbf{x}_0$  e  $\mathbf{x}$ . Allora:

$$\left| \frac{R(\mathbf{x}, \mathbf{x}_0)}{\|\mathbf{x} - \mathbf{x}_0\|^2} \right| \leq \frac{1}{2} \|H_f(\xi) - H_f(\mathbf{x}_0)\| \xrightarrow{\mathbf{x} \rightarrow \mathbf{x}_0} 0$$

perché  $\xi(\mathbf{x}, \mathbf{x}_0) \rightarrow \mathbf{x}_0$  quando  $\mathbf{x} \rightarrow \mathbf{x}_0$ . Nella formula sopra si è usata la *norma della matrice*  $H_f(\xi) - H_f(\mathbf{x}_0)$  (definizione proveniente dall'algebra lineare) che (si può vedere facilmente) tende a zero se il suo argomento tende a zero (componente per componente).

# Funzioni con gradiente nullo

Supponiamo che  $A$  sia un aperto connesso e  $f : A \rightarrow \mathbb{R}$  una funzione differenziabile con  $\nabla f(\mathbf{x}) = \vec{0}$  Per ogni  $\mathbf{x}$  in  $A$ . Allora  $f$  è costante in  $A$ .

**Dimostrazione** Fissiamo a caso un punto  $\mathbf{x}_0 \in A$ . Se  $\mathbf{x} \in A$  sappiamo che esiste una curva  $\gamma : [a, b] \rightarrow A$  tale che  $\gamma(a) = \mathbf{x}_0$ ,  $\gamma(b) = \mathbf{x}$ . Si può dimostrare (non lo facciamo) che tale  $\gamma$  può essere trovata regolare. Allora:

$$\frac{d}{dt}f(\gamma(t)) = \nabla f(\gamma(t)) \cdot \gamma'(t) = 0 \quad \forall t \in [a, b]$$

$\Rightarrow f(\gamma(t))$  costante. In particolare  $f(\mathbf{x}_0) = f(\gamma(a)) = f(\gamma(b)) = f(\mathbf{x})$ .  
Dato che questo è vero per ogni  $\mathbf{x} \in A$ :

$$f(\mathbf{x}) = f(\mathbf{x}_0) \quad \forall \mathbf{x} \in A,$$

dunque  $f$  è costante in  $A$ .

## Punti di massimo/minimo relativo: condizione necessaria

Sia  $f : A \rightarrow \mathbb{R}$  dove  $A$  è un sottoinsieme (qualunque) di  $\mathbb{R}^N$  e sia  $\mathbf{x}_0 \in A$ .  $\mathbf{x}_0$  si dice *punto di massimo (minimo) relativo per  $f$*  se esiste  $\rho > 0$  tale che:

$$f(\mathbf{x}) \leq f(\mathbf{x}_0) \text{ (} \geq f(\mathbf{x}_0) \text{)} \quad \forall \mathbf{x} \in A \text{ con } \|\mathbf{x} - \mathbf{x}_0\| < \rho.$$

### Teorema di Fermat in più variabili

Supponiamo che  $\mathbf{x}_0$  sia interno ad  $A$  e che sia di massimo o minimo relativo per  $f$ . Supponiamo che  $f$  sia differenziabile in  $\mathbf{x}_0$ . Allora  
 $\nabla f(\mathbf{x}_0) = \vec{0}$ .

Quando  $\nabla f(\mathbf{x}_0) = \vec{0}$  diciamo che  $\mathbf{x}_0$  è un punto stazionario, o un punto critico, per  $f$ . Dunque:

$$\mathbf{x}_0 \text{ pto di max/min. rel.} \Rightarrow \mathbf{x}_0 \text{ stazionario}$$

Già in una variabile si vede che non vale il viceversa.

**Dimostrazione** Prendiamo un vettore qualunque  $\vec{v}$  e consideriamo la solita restrizione di  $f$  alla retta per  $\mathbf{x}_0$  diretta lungo  $\vec{v}$ :

$$\varphi(t) := f(\mathbf{x}_0 + t\vec{v}).$$

Dato che  $\mathbf{x}_0$  è di massimo (minimo) locale per  $f$ , allora  $t = 0$  è di massimo (minimo) locale per  $\varphi$  e quindi  $\varphi'(0) = 0$ . Ma  $\varphi'(0) = \mathbf{f}'(\mathbf{x}_0)(\vec{v})$  e quindi

$$\mathbf{f}'(\mathbf{x}_0)(\vec{v}) = 0 \quad \forall \vec{v} \in \mathbb{R}^N \quad \text{dunque } \nabla f(\mathbf{x}_0) = \vec{0}.$$

## Punti di massimo/minimo relativo: condizione sufficiente

Supponiamo che  $x_0$  sia interno ad  $A$ , che  $f$  sia  $\mathcal{C}^2$  in un intorno di  $x_0$  e che  $\nabla f(x_0) = \vec{0}$  ( $x_0$  sia stazionario). Si ha:

- se  $H_f(x_0)$  è definito positivo, allora  $x_0$  è un punto di minimo relativo per  $f$ ;
- se  $H_f(x)$  è semidefinito positivo in un intorno di  $x_0$ , allora  $x_0$  è un punto di minimo relativo per  $f$ ;
- se  $H_f(x_0)$  è definito negativo, allora  $x_0$  è un punto di massimo relativo per  $f$ ;
- se  $H_f(x)$  è semidefinito negativo in un intorno di  $x_0$ , allora  $x_0$  è un punto di massimo relativo per  $f$ ;
- se  $H_f(x_0)$  è indefinito, allora  $x_0$  non è né di massimo né di minimo locale – se  $H_f(x_0)$  è indefinito diremo che  $x_0$  è un *punto di sella* per  $f$ .

# Punti di massimo/minimo relativo: condizione sufficiente

**Dimostrazione** Prendiamo un disco  $D(\mathbf{x}_0, R)$  che sia contenuto in  $A$  e prendiamo  $\mathbf{x} \in D(\mathbf{x}_0, R)$ . Per Taylor (tenendo conto di  $\nabla f(\mathbf{x}_0) = \vec{0}$ ):

$$f(\mathbf{x}) - f(\mathbf{x}_0) = \frac{1}{2} H_f(\xi(\mathbf{x}, \mathbf{x}_0))(\mathbf{x} - \mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0),$$

dove  $\xi(\mathbf{x}, \mathbf{x}_0)$  è un punto sul segmento tra  $\mathbf{x}_0$  e  $\mathbf{x}$  e dunque  $\xi(\mathbf{x}, \mathbf{x}_0) \rightarrow \mathbf{x}_0$  per  $\mathbf{x} \rightarrow \mathbf{x}_0$ . Facciamo il caso di  $H_f(\mathbf{x}_0)$  definita positiva: allora esiste  $\nu > 0$  tale che tutti gli autovalori di  $H_f(\xi(\mathbf{x}, \mathbf{x}_0))$  sono maggiori di  $\nu$ . Per la continuità delle derivate seconde, se  $R$  è abbastanza piccolo, gli autovalori di  $H_f(\xi(\mathbf{x}, \mathbf{x}_0))$  sono tutti maggiori di  $\nu/2$ , quindi:

$$\forall \mathbf{x} \in D(\mathbf{x}_0, R) \quad H_f(\xi(\mathbf{x}, \mathbf{x}_0))\vec{v} \cdot \vec{v} \geq \frac{\nu}{2} \|\vec{v}\|^2 \quad \forall \vec{v} \in \mathbb{R}^N.$$

Ne segue che  $\mathbf{x}_0$  è di minimo in  $D(\mathbf{x}_0, R)$ . Se invece  $H_f(\mathbf{x})$  è semidefinito positivo in un intorno  $U$  di  $\mathbf{x}_0$ , possiamo prendere  $R > 0$  in modo che  $D(\mathbf{x}_0, R) \subset U$ . Allora  $\xi \in U$  e quindi  $H_f(\xi(\mathbf{x}, \mathbf{x}_0))(\mathbf{x} - \mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0) \geq 0$ . Ne segue di nuovo che  $\mathbf{x}_0$  è di minimo in  $D(\mathbf{x}_0, R)$ .

Per massimo il discorso è analogo.

## Punti di massimo/minimo relativo: condizione sufficiente

Supponiamo ora che  $H_f(\mathbf{x}_0)$  sia indefinita. Allora  $H_f(\mathbf{x}_0)$  ha un autovalore positivo  $\lambda_+$  ed uno negativo  $\lambda_-$ . Questo vuol dire che esistono due autovettori  $\vec{e}_+$  ed  $\vec{e}_-$ , di norma uno tali che:

$$H_f(\mathbf{x}_0)\vec{e}_+ = \lambda_+\vec{e}_+, \quad H_f(\mathbf{x}_0)\vec{e}_- = \lambda_-\vec{e}_-$$

Consideriamo la restrizione di  $f$  a una retta passante per  $\mathbf{x}_0$ :

$$\begin{aligned} \varphi(t) &= f(\mathbf{x}_0 + t\vec{v}) \quad \text{per un } \vec{v} \in \mathbb{R}^N. \quad \text{Allora:} \\ \varphi(0) &= f(\mathbf{x}_0), \quad \varphi'(0) = \nabla f(\mathbf{x}_0)(\vec{v}) = 0, \quad \varphi''(0) = H_f(\mathbf{x}_0)\vec{v} \cdot \vec{v} \end{aligned}$$

Ne segue che nella direzione  $\vec{v} = \vec{e}_+$  la  $\varphi$  ha un minimo locale in  $t = 0$  ( $\varphi'(0) = 0$  e  $\varphi''(0) = \lambda_+ > 0$ ), mentre nella direzione  $\vec{v} = \vec{e}_-$  la  $\varphi$  ha un massimo locale ( $\varphi'(0) = 0$  e  $\varphi''(0) = \lambda_- < 0$ ).

Dunque  $\mathbf{x}_0$  non può essere né di massimo né di minimo per  $f$ .

Cercare i punti critici delle seguenti funzioni  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  e classificarne la natura.

❶  $f(x, y) := 4xy - x^2 - y^4;$

❷  $f(x, y) := e^{x^2+y^2} - 2exy;$

❸  $f(x, y) := y^2 + \frac{16}{\sqrt{1+y^2-x^2}};$

❹  $f(x, y) := e^{x^2+y^2} - 2e^2xy;$

❺  $f(x, y) := 2e^{xy+1} + x^2 + y^2;$

❻  $f(x, y) := \frac{1}{x^2 + y^2} + 2xy;$

❼  $f(x, y) := 3\ln(1 + x^2 + y^2) + 2xy.$

Si tratta di esercizi presi da compiti d'esame, le cui soluzioni sono reperibili sul sito <http://pagine.dm.unipi.it/csblog1>.

Suggerisco inoltre di vedere gli esercizi a pag. 146–147, 157 dell'Adams.

# Funzioni omogenee

Sia  $\alpha > 0$ . Una funzione  $f : \mathbb{R}^N \rightarrow \mathbb{R}$  si dice *omogenea di grado  $\alpha$*  se:

$$f(t\mathbf{x}) = t^\alpha f(\mathbf{x}) \quad \text{per ogni } t > 0 \text{ e ogni } \mathbf{x} \text{ in } \mathbb{R}^N$$

Per esempio  $f(x, y, z) = x^2y - 2xyz$  è omogenea di grado 3.

Se  $f$  è  $\mathcal{C}^1$  ed è omogenea di grado  $\alpha$ , allora

$$\nabla f(\mathbf{x}) \cdot \mathbf{x} = \alpha f(\mathbf{x}) \quad \forall \mathbf{x} \in \mathbb{R}^N.$$

Per vederlo basta derivare rispetto a  $t$  la relazione di omogeneità:

$$\nabla f(t\mathbf{x}) \cdot \mathbf{x} = \alpha t^{\alpha-1} f(\mathbf{x})$$

e mettere  $t = 1$ .

# Funzioni convesse

Sia  $A$  un insieme convesso di  $\mathbb{R}^N$  e sia  $f : A \rightarrow \mathbb{R}$

## Funzione convessa

Si dice che  $f$  è convessa se:

$$f(tx_1 + (1-t)x_2) \leq tf(x_1) + (1-t)f(x_2) \quad \forall x_1, x_2 \in A, \forall t \in [0, 1].$$

Notiamo che la curva

$$\gamma(t) := tx_1 + (1-t)x_2 \quad t \in [0, 1]$$

descrive il segmento tra  $x_1$  e  $x_2$  (che per ipotesi è tutto contenuto in  $A$ ), mentre la curva

$$\gamma_1(t) = (tx_1 + (1-t)x_2, tf(x_1) + (1-t)f(x_2)) \quad t \in [0, 1]$$

descrive il segmento (in  $\mathbb{R}^{N+1}$ ) tra  $(x_1, f(x_1))$  e  $(x_2, f(x_2))$ .

Dunque  $f$  è convessa se e solo se per ogni coppia di punti  $P_1$  e  $P_2$  del grafico di  $f$  il segmento che li congiunge passa sopra il grafico di  $f$ .

# Funzioni convesse

Supponiamo  $A$  aperto convesso. Allora: (per le DIM. vedi le note)

## Convessità e piani tangenti

Sia  $f : A \rightarrow \mathbb{R}$  di classe  $\mathcal{C}^1$ . Allora  $f$  è convessa se e solo se:

$$f(\mathbf{x}) \geq f(\mathbf{x}_0) + \nabla f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0) \quad \forall \mathbf{x}, \mathbf{x}_0 \in A.$$

(cioè se il grafico di  $f$  sta sopra il piano tangente a  $f$  in un qualunque punto  $\mathbf{x}_0$ ).

Se  $f$  è convessa  $\mathcal{C}^1$  e se  $\nabla f(\mathbf{x}_0) = \vec{0}$ , allora  $\mathbf{x}_0$  è un punto di minimo assoluto per  $f$ .

## Convessità e derivate seconde

Sia  $f : A \rightarrow \mathbb{R}$  di classe  $\mathcal{C}^2$ . Allora  $f$  è convessa se e solo se la matrice Hessiana  $H_f(\mathbf{x}_0)$  è semidefinita positiva in ogni  $\mathbf{x}_0$  di  $A$ .

# Regressione lineare

Consideriamo una  $n$ -upla di punti  $(x_1, y_1), \dots, (x_n, y_n)$  di punti in  $\mathbb{R}^2$ . Cerchiamo una retta  $y = mx + q$  il cui grafico “meglio approssimi” questi  $n$  punti. Chiaramente per risolvere questo problema dovremo precisare il senso di questa approssimazione. Con ragionamenti che non approfondiamo si perviene all'idea di scegliere  $m$  e  $q$  in modo che risulti minima la somma dei quadrati delle distanze:

$$F(m, q) := \sum_{i=1}^n (y_i - mx_i - q)^2.$$

Poniamo:

$$\mathbf{x} := (x_1, \dots, x_n), \quad \mathbf{y} := (y_1, \dots, y_n)$$

e per un generico  $(v_1, \dots, v_n)$  indichiamo con  $\bar{v} := \frac{\sum_{i=1}^n v_i}{n}$  la sua media.

# Regressione lineare

Per trovare  $m$  e  $q$  calcoliamo le derivate parziali di  $F$ :

$$\frac{\partial F}{\partial m}(m, q) = -2 \sum_{i=1}^n (y_i - mx_i - q)x_i, \quad \frac{\partial F}{\partial q}(m, q) = -2 \sum_{i=1}^n (y_i - mx_i - q),$$

da cui, eguagliando a zero, si ottiene:

$$\sum_{i=1}^n x_i y_i = m \sum_{i=1}^n x_i^2 + q \sum_{i=1}^n x_i, \quad \sum_{i=1}^n y_i = m \sum_{i=1}^n x_i + nq,$$

cioè:

$$\mathbf{x} \cdot \mathbf{y} = m \|\mathbf{x}\|^2 + qn\bar{x}, \quad \bar{y} = m\bar{x} + q$$

e in definitiva:

$$m = \frac{\mathbf{x} \cdot \mathbf{y} - n\bar{x}\bar{y}}{\|\mathbf{x}\|^2 - n\bar{x}^2} = \frac{\overline{\mathbf{x}\mathbf{y}} - \bar{x}\bar{y}}{\overline{\mathbf{x}\mathbf{x}} - \bar{x}^2}, \quad q = \bar{y} - m\bar{x}$$

## Theorem (Dini)

Siano  $A \subset \mathbb{R}^{N+1}$  aperto,  $G : A \rightarrow \mathbb{R}$  una funzione  $\mathcal{C}^1$ . Poniamo:

$$M := \{(\mathbf{x}, y) \in Q_{\delta, \varepsilon}(\mathbf{x}_0, y_0) : G(\mathbf{x}, y) = 0\}$$

( $M$  è il luogo degli zeri di  $G$ ). Supponiamo che  $(\mathbf{x}_0, y_0) \in A$ , sia tale che:

$$G(\mathbf{x}_0, y_0) = 0, \quad \frac{\partial G}{\partial y}(\mathbf{x}_0, y_0) \neq 0.$$

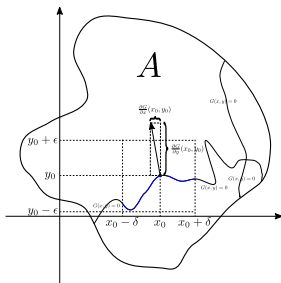
(in particolare  $(\mathbf{x}_0, y_0) \in M$ ). Indichiamo:

$$Q_{\delta, \varepsilon}(\mathbf{x}_0, y_0) := B(\mathbf{x}_0, \delta) \times ]y_0 - \varepsilon, y_0 + \varepsilon[$$

Allora esistono  $\delta > 0, \varepsilon > 0$  ed esiste  $f : B(\mathbf{x}_0, \delta) \rightarrow ]y_0 - \varepsilon, y_0 + \varepsilon[$  tali che:

$$Q_{\delta, \varepsilon}(\mathbf{x}_0, y_0) \subset A, \quad M \cap [Q_{\delta, \varepsilon}(\mathbf{x}_0, y_0) = \{(\mathbf{x}, f(\mathbf{x})) : \mathbf{x} \in B(\mathbf{x}_0, \delta)\}$$

# Funzioni implicite



Inoltre la funzione  $f$  è di classe  $\mathcal{C}^1$  su  $B(\mathbf{x}_0, \delta)$  e:

$$\frac{\partial f}{\partial x_i}(\mathbf{x}) = - \frac{\frac{\partial G}{\partial x_i}(\mathbf{x}, f(\mathbf{x}))}{\frac{\partial G}{\partial y}(\mathbf{x}, f(\mathbf{x}))} \quad \forall \mathbf{x} \in B(\mathbf{x}_0, \delta).$$

# Funzioni implicite

Per la dimostrazione del teorema nel caso  $N = 1$  vedi le note.

Nel teorema il fatto di avere esplicitato rispetto all'ultima variabile di  $G$  (che è stata chiamata  $y$  per distinguerla dalle altre) è arbitrario. In realtà vale:

## Theorem (Dini)

*Sia  $A$  un aperto di  $\mathbb{R}^N$  con  $N \geq 2$  e sia  $G : A \rightarrow \mathbb{R}$  una funzione  $\mathcal{C}^1$ . Se  $\mathbf{x}_0 \in A$  è un punto tale che  $G(\mathbf{x}_0) = 0$  e se esiste  $i$  tra 1 ed  $N$  tale che  $\frac{\partial G}{\partial x_i}(\mathbf{x}_0) \neq 0$ , allora l'insieme  $M := \{\mathbf{x} \in A : G(\mathbf{x}) = 0\}$  è "localmente" ("vicino a  $\mathbf{x}_0$ ") grafico di una funzione*

$$x_i = f(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_N)$$

*Dunque, se  $\nabla G(\mathbf{x}_0) \neq \vec{0}$ , allora "vicino  $\mathbf{x}_0$ " l'insieme  $M$  è grafico di una funzione di  $N - 1$  variabili – si può "esplicitare" una variabile  $x_i$  se  $i$  è tale che la componente  $i$ -esima di  $\nabla G(\mathbf{x}_0) \neq \vec{0}$  non è nulla.*

Nel caso  $N = 2$ , cioè  $G : A(\subset \mathbb{R}^2) \rightarrow \mathbb{R}$ , se

$$M := \{(x, y) \in \mathbb{R}^2 : G(x, y) = 0\}$$

e se

$$\nabla G(x, y) \neq \vec{0} \quad \forall (x, y) \in A,$$

allora  $M$  è localmente descrivibile come supporto di una curva regolare.

Se  $N = 3$ , vedremo che  $M$  è descrivibile come supporto di una superficie (da definire).

## Fatto

Dati  $x_0, y_0, \varepsilon_0 > 0$  e  $\delta_0 > 0$  tali che:

$$G(x_0, y_0) = 0, \quad \frac{\partial G}{\partial y} > 0 \text{ in } Q_{\varepsilon_0, \delta_0}(x_0, y_0) \subset A,$$

esiste  $\delta$  con  $0 < \delta \leq \delta_0$  tale che:

$\forall x \in ]x_0 - \delta, x_0 + \delta[$  esiste uno e uno solo  $y \in ]y_0 - \varepsilon_0, y_0 + \varepsilon_0[$  t.c.

$$M \cap Q_{\varepsilon_0, \delta_0}(x_0, y_0) = \{(x, y)\}.$$

# Funzioni implicite

**Dim.** Fissato  $x \in ]x_0 - \delta, x_0 + \delta_0[$  consideriamo la funzione  $\psi_x(y) := G(x, y)$ , definita per  $y \in ]y_0 - \varepsilon_0, y_0 + \varepsilon_0[$ . Tale funzione è strettamente crescente dato che  $\psi'_x(y) = \frac{\partial G}{\partial y}(x, y) > 0$ . Dato che  $\psi_{x_0}(y_0) = G(x_0, y_0) = 0$  si ha  $G(x_0, y_0 - \varepsilon_0) = \psi_{x_0}(y_0 - \varepsilon_0) < 0$  e  $G(x_0, y_0 + \varepsilon_0) = \psi_{x_0}(y_0 + \varepsilon_0) > 0$ . Applicando la permanenza del segno a  $\varphi_-(x) := G(x, y_0 - \varepsilon_0)$  e a  $\varphi_+(x) := G(x, y_0 + \varepsilon_0)$ , possiamo trovare  $\delta$  in  $]0, \delta_0]$  tale che  $G(x, y_0 - \varepsilon_0) < 0$  e  $G(x, y_0 + \varepsilon_0) > 0$  per tutte le  $x$  in  $]x_0 - \delta, x_0 + \delta[$ . Dunque se  $x_0 - \delta < x < x_0 + \delta$  si ha:

$$\psi_x(y_0 - \varepsilon_0) < 0, \quad \psi_x(y_0 + \varepsilon_0) > 0.$$

Dato che  $\psi_x$  è strettamente crescente esiste uno e uno solo  $y$  in  $]y_0 - \varepsilon_0, y_0 + \varepsilon_0[$  per cui  $G(x, y) = \psi_x(y) = 0$ .

## Domini regolari

Sia  $A \subset \mathbb{R}^N$  aperto. Dico che  $D \subset A$  è un *dominio regolare chiuso in A* se:

$$D = \{\mathbf{x} \in A : G(\mathbf{x}) \leq 0\}$$

dove  $G : A \rightarrow \mathbb{R}$  è una funzione  $\mathcal{C}^1$  tale che:

$$\nabla G(\mathbf{x}) \neq \vec{0} \quad \forall \mathbf{x} \in A \text{ tale che } G(\mathbf{x}) = 0.$$

Se  $A = \mathbb{R}^N$  dico semplicemente che  $D$  è un dominio regolare chiuso. Notiamo che, nel caso  $A = \mathbb{R}^N$ :

$$\bar{D} = D, \quad \text{int}(D) = \{G(\mathbf{x}) < 0\}, \quad \partial D = \{G(\mathbf{x}) = 0\}$$

( $D$  è effettivamente chiuso). Chiamo poi *dominio regolare aperto* la parte interna di un dominio regolare chiuso, cioè l'insieme  $\{\mathbf{x} \in A : G(\mathbf{x}) < 0\}$

# Massimi e minimi vincolati

Consideriamo  $A$  aperto di  $\mathbb{R}^N$ , una funzione  $G : A \rightarrow \mathbb{R}$  di classe  $\mathcal{C}^1$  t.c.:

$$\nabla G(\mathbf{x}) \neq \vec{0} \quad \text{per ogni } \mathbf{x} \in A \text{ tale che } G(\mathbf{x}) = 0$$

e poniamo:

$$\Omega := \{\mathbf{x} \in A : G(\mathbf{x}) \leq 0\}, \quad \Omega_0 := \{\mathbf{x} \in A : G(\mathbf{x}) = 0\}.$$

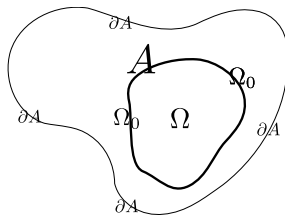
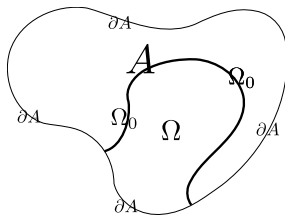
Come già osservato, essendo  $\Omega_0$  una superficie di livello per  $G$ , per ogni  $\mathbf{x} \in \Omega_0$  il vettore  $\nabla G(\mathbf{x})$  è diretto **ortogonalmente** a  $\Omega_0$  (nel senso che è ortogonale a qualunque curva che giace in  $\Omega_0$ ).

Considereremo anche una funzione  $f : A \rightarrow \mathbb{R}$ : ci interessa **trovare i punti di estremo (relativo o assoluto) di  $f$  su  $\Omega$  o su  $\Omega_0$ .**

# Massimi e minimi vincolati

È chiaro che:

$$\text{int}(\Omega) = \{x \in A : G(x) < 0\}, \quad \Omega_0 = \partial\Omega \cap A \subset \partial\Omega.$$



Notiamo anche che  $\Omega$  è chiuso se e solo se  $\Omega_0$  è chiuso se e solo se  $\Omega_0 = \partial\Omega$  (cioè  $\partial\Omega \subset A$ ). Se  $\Omega$  è limitato e chiuso, allora (per Weierstrass) esistono sicuramente un punto di massimo e un punto di minimo per  $f$  su  $\Omega$ .

## Theorem (Estremi vincolati a $\partial\Omega \cap A$ )

*Nelle ipotesi scritte prima, se  $\mathbf{x}_0 \in \Omega_0$  (cioè  $G(\mathbf{x}_0) = 0$ ) è un punto di massimo o minimo relativo per  $f$  su  $\Omega_0$ , allora esiste  $\lambda \in \mathbb{R}$  tale che:*

$$\nabla f(\mathbf{x}_0) = \lambda \nabla G(\mathbf{x}_0).$$

Se un punto  $\mathbf{x}_0 \in \Omega_0$  verifica  $\nabla f(\mathbf{x}_0) = \lambda \nabla G(\mathbf{x}_0)$  per un opportuno  $\lambda \in \mathbb{R}$  diremo che  $\mathbf{x}_0$  è un **punto stazionario (o critico) vincolato** a  $\Omega_0$ .

Nei punti critici vincolati a  $\Omega_0$  il gradiente di  $f$  è perpendicolare a  $\Omega_0$ , dunque la sua “componente tangente” a  $\Omega_0$  è nulla.

Il numero  $\lambda$  è detto *moltiplicatore di Lagrange*. Se  $\lambda = 0$   $\mathbf{x}_0$  è un punto stazionario “libero” per  $f$ .

# Massimi e minimi vincolati

**Dimostrazione del Teorema** Supponiamo per esempio  $\frac{\partial G}{\partial x_N}(\mathbf{x}_0) \neq 0$ .

Se  $\mathbf{x}$  di  $\mathbb{R}^N$  scriviamo  $\mathbf{x} = (\mathbf{x}', x_N)$  con  $\mathbf{x}' \in \mathbb{R}^{N-1}$  e  $x_N \in \mathbb{R}$ . Indichiamo anche con  $B'(\mathbf{x}', r')$  il disco in  $\mathbb{R}^{N-1}$ , di centro  $\mathbf{x}'$  e raggio  $r' > 0$ .

Per Dini esistono  $\delta > 0$ ,  $\varepsilon > 0$  e  $\varphi : B'(\mathbf{x}'_0, \delta) \rightarrow \mathbb{R}$  tali che:

$$\Omega_0 \cap B'(\mathbf{x}'_0, \delta) = \{(\mathbf{x}', \varphi(\mathbf{x}')) : \mathbf{x}' \in B'(\mathbf{x}'_0, \delta)\}.$$

Definiamo  $\Phi : B'(\mathbf{x}'_0, \delta) \rightarrow \mathbb{R}^N$  ponendo  $\Phi(\mathbf{x}') := (\mathbf{x}', \varphi(\mathbf{x}'))$  – notiamo che  $\Phi(\mathbf{x}'_0) = \mathbf{x}_0$  e che  $\Phi(\mathbf{x}') \in \Omega_0$  per ogni  $\mathbf{x}' \in U'$ .

Dato che  $\mathbf{x}_0$  è punto di estremo relativo per  $f$  su  $\Omega_0$ ,  $\mathbf{x}'_0$  è di estremo relativo per  $\mathbf{x}' \mapsto f(\mathbf{x}', \varphi(\mathbf{x}'))$ . Per esempio, se  $\mathbf{x}_0$  è di minimo per  $f$  su  $\partial\Omega$ :

$$f(\Phi(\mathbf{x}'_0)) = f(\mathbf{x}'_0, \varphi(\mathbf{x}'_0)) = f(\mathbf{x}_0) \leq f(\mathbf{x}', \varphi(\mathbf{x}')) = f(\Phi(\mathbf{x}'))$$

per ogni  $\mathbf{x}' \in B'(\mathbf{x}'_0, \delta)$  e dunque  $\mathbf{x}'_0$  è di minimo per  $f \circ \Phi$ . Se il minimo è locale si fa allo stesso modo restringendosi su un opportuno intorno – analogo ragionamento per i massimi.

# Massimi e minimi vincolati

Dato che  $\mathbf{x}'_0$  è ovviamente interno a  $B'(\mathbf{x}'_0, \delta)$  possiamo usare Fermat:

$$\frac{\partial f \circ \Phi}{\partial x_i}(\mathbf{x}'_0) = 0 \quad i = 1, \dots, N-1.$$

Facendo i calcoli (e usando ancora il teorema del Dini):

$$\begin{aligned} 0 &= \frac{\partial f(x_1, \dots, x_i, \dots, x_{N-1}, \varphi(x_1, \dots, x_{N-1}))}{\partial x_i}(\mathbf{x}'_0) = \\ &\quad \frac{\partial f}{\partial x_i}(\mathbf{x}'_0, \varphi(\mathbf{x}'_0)) + \frac{\partial f}{\partial x_N}(\mathbf{x}'_0, \varphi(\mathbf{x}'_0)) \frac{\partial \varphi}{\partial x_i}(\mathbf{x}'_0) = \\ &\quad \frac{\partial f}{\partial x_i}(\mathbf{x}_0) - \frac{\partial f}{\partial x_N}(\mathbf{x}_0) \frac{\frac{\partial G}{\partial x_i}(\mathbf{x}_0)}{\frac{\partial G}{\partial x_N}(\mathbf{x}_0)} \quad i = 1, \dots, N-1. \end{aligned}$$

# Massimi e minimi vincolati

Queste eguaglianze si possono scrivere:

$$\frac{\partial f}{\partial x_1}(\mathbf{x}_0) = \lambda \frac{\partial G}{\partial x_1}(\mathbf{x}_0), \dots, \frac{\partial f}{\partial x_{N-1}}(\mathbf{x}_0) = \lambda \frac{\partial G}{\partial x_{N-1}}(\mathbf{x}_0),$$

dove:

$$\lambda := \frac{\frac{\partial f}{\partial x_N}(\mathbf{x}_0)}{\frac{\partial G}{\partial x_N}(\mathbf{x}_0)}$$

ed è chiaro dalla definizione di  $\lambda$  che vale anche:

$$\frac{\partial f}{\partial x_N}(\mathbf{x}_0) = \lambda \frac{\partial G}{\partial x_N}(\mathbf{x}_0),$$

di modo che  $\nabla f(\mathbf{x}_0) = \lambda \nabla G(\mathbf{x}_0)$ , cioè la tesi.

# Massimi e minimi vincolati

Dalla dimostrazione si vede che l'ipotesi  $\nabla G(\mathbf{x}) \neq \vec{0}$  serve solo in  $\mathbf{x} = \mathbf{x}_0$ . Dunque se  $\mathbf{x}_0$  è punto di estremo relativo per  $f$  su  $\Omega_0 = \{G = 0\}$ , se  $f$  è differenziabile in  $\mathbf{x}_0$ , se  $G$  è differenziabile in  $\mathbf{x}_0$  e  $\nabla G(\mathbf{x}_0) \neq \vec{0}$ , allora  $\mathbf{x}_0$  è un punto critico vincolato per  $f$  su  $\partial\Omega$ .

## Lagrangiana

Si vede facilmente che  $\mathbf{x}_0$  appartiene a  $\Omega_0$  ed è punto critico vincolato per  $f$  su  $\Omega_0$  se e solo se esiste  $\lambda_0 \in \mathbb{R}$  tale che il punto  $(\mathbf{x}_0, \lambda_0)$  (che si trova in  $A \times \mathbb{R} \subset \mathbb{R}^{N+1}$ ) è critico per la *Lagrangiana*:

$$L(\mathbf{x}, \lambda) := f(\mathbf{x}) - \lambda G(\mathbf{x})$$

( $L$  va da  $\mathbb{R}^{N+1}$  in  $\mathbb{R}$ ). In effetti:

$$\nabla L(\mathbf{x}, \lambda) = \begin{pmatrix} \nabla f(\mathbf{x}) - \lambda \nabla G(\mathbf{x}) \\ -G(\mathbf{x}) \end{pmatrix} \quad \text{dunque} \quad \nabla L(\mathbf{x}, \lambda) = \vec{0} \Leftrightarrow \begin{cases} \nabla f(\mathbf{x}) = \lambda \nabla G(\mathbf{x}) \\ G(\mathbf{x}) = 0 \end{cases}$$

## Estremi per $f$ su $\Omega$

Supponiamo valide tutte le ipotesi scritte prima. Sia  $\mathbf{x}_0$  un punto di  $\Omega$ . Se  $\mathbf{x}_0$  è un estremo relativo per  $f$  su  $\Omega$  allora vale una delle due alternative:

- ①  $G(\mathbf{x}_0) < 0$  e  $\nabla f(\mathbf{x}_0) = \vec{0}$ ;
- ②  $G(\mathbf{x}_0) = 0$  ed esiste  $\lambda \in \mathbb{R}$  tale che  $\nabla f(\mathbf{x}_0) = \lambda \nabla G(\mathbf{x}_0)$ .

In sostanza  $\mathbf{x}_0$  può essere interno a  $\Omega$ , e in quel caso è stazionario libero, oppure sulla frontiera di  $\Omega$ , e in quel caso è stazionario vincolato.

## Lagrangiana – condizioni di Kuhn-Tucker

Sia  $\mathbf{x}_0$  in  $A$ . Allora  $\mathbf{x}_0$  è in  $\Omega$  e verifica una delle due alternative scritte sopra se e solo se esistono  $\lambda_0, u_0 \in \mathbb{R}$  e tali che il punto  $(\mathbf{x}_0, \lambda_0, u_0)$  (che è in  $A \times \mathbb{R} \times \mathbb{R}$ ) è stazionario per la *Lagrangiana*:

$$L(\mathbf{x}, \lambda, u) := f(\mathbf{x}) - \lambda(G(\mathbf{x}) - u^2).$$

# Massimi e minimi vincolati

In effetti si ha:

$$\nabla L(\mathbf{x}, \lambda, u) = \begin{pmatrix} \nabla f(\mathbf{x}) - \lambda \nabla G(\mathbf{x}) \\ -G(\mathbf{x}) + u^2 \\ 2\lambda u \end{pmatrix}$$

per cui  $(\mathbf{x}_0, \lambda_0, u_0)$  è stazionario per  $L$  se e solo se:

$$\begin{cases} \nabla f(\mathbf{x}) = \lambda \nabla G(\mathbf{x}) \\ G(\mathbf{x}) = -u^2 \\ \lambda u = 0 \end{cases}$$

La terza condizione equivale a  $\lambda = 0$  oppure  $u = 0$ . Se  $\lambda = 0$  allora  $\nabla f(\mathbf{x}) = 0$  (per la prima) e  $G(\mathbf{x}) \leq 0$  (per la seconda) – in questo caso  $\mathbf{x}_0$  è stazionario libero. Se invece  $u = 0$  si ricava  $G(\mathbf{x}) = 0$  (dalla seconda), cioè  $\mathbf{x} \in \Omega_0$ , ed è stazionario vincolato per  $f$  su  $\Omega_0$  (per la prima).

Le relazioni scritte sopra, a destra della graffa, sono dette *condizioni di Kuhn-Tucker*.

# Funzioni implicite – caso generale

Considereremo ora una funzione  $\mathbf{G}$  definita in un aperto di  $\mathbb{R}^{N+M}$  a valori in  $\mathbb{R}^M$ . Se indichiamo i punti di  $\mathbb{R}^{N+M}$  come  $(\mathbf{x}, \mathbf{y})$ , dove  $\mathbf{x} = (x_1, \dots, x_N) \in \mathbb{R}^N$  e  $\mathbf{y} = (y_1, \dots, y_M) \in \mathbb{R}^M$  allora:

$$\mathbf{G}(\mathbf{x}, \mathbf{y}) = \begin{pmatrix} G_1(x_1, \dots, x_N, y_1, \dots, y_M) \\ \vdots \\ G_M(x_1, \dots, x_N, y_1, \dots, y_M) \end{pmatrix}$$

che ha come matrice Jacobiana:

$$J_{\mathbf{G}}(\mathbf{x}, \mathbf{y}) = \begin{pmatrix} \frac{\partial G_1}{\partial x_1}(\mathbf{x}, \mathbf{y}) & \cdots & \frac{\partial G_1}{\partial x_N}(\mathbf{x}, \mathbf{y}) & \frac{\partial G_1}{\partial y_1}(\mathbf{x}, \mathbf{y}) & \cdots & \frac{\partial G_1}{\partial y_M}(\mathbf{x}, \mathbf{y}) \\ \vdots & & \vdots & \vdots & & \vdots \\ \frac{\partial G_M}{\partial x_1}(\mathbf{x}, \mathbf{y}) & \cdots & \frac{\partial G_M}{\partial x_N}(\mathbf{x}, \mathbf{y}) & \frac{\partial G_M}{\partial y_1}(\mathbf{x}, \mathbf{y}) & \cdots & \frac{\partial G_M}{\partial y_M}(\mathbf{x}, \mathbf{y}) \end{pmatrix}$$

# Funzioni implicite – caso generale

## Notazione

Introduciamo gli *Jacobiani parziali* rispetto ai gruppi di variabili  $\mathbf{x}$  e  $\mathbf{y}$ :

$$\frac{\partial \mathbf{G}}{\partial \mathbf{x}}(\mathbf{x}, \mathbf{y}) = \frac{\partial (G_1, \dots, G_M)}{\partial (x_1, \dots, x_N)}(\mathbf{x}, \mathbf{y}) := \begin{pmatrix} \frac{\partial G_1}{\partial x_1}(\mathbf{x}, \mathbf{y}) & \cdots & \frac{\partial G_1}{\partial x_N}(\mathbf{x}, \mathbf{y}) \\ \vdots & & \vdots \\ \frac{\partial G_M}{\partial x_1}(\mathbf{x}, \mathbf{y}) & \cdots & \frac{\partial G_M}{\partial x_N}(\mathbf{x}, \mathbf{y}) \end{pmatrix}$$

$$\frac{\partial \mathbf{G}}{\partial \mathbf{y}}(\mathbf{x}, \mathbf{y}) = \frac{\partial (G_1, \dots, G_M)}{\partial (y_1, \dots, y_M)}(\mathbf{x}, \mathbf{y}) := \begin{pmatrix} \frac{\partial G_1}{\partial y_1}(\mathbf{x}, \mathbf{y}) & \cdots & \frac{\partial G_1}{\partial y_M}(\mathbf{x}, \mathbf{y}) \\ \vdots & & \vdots \\ \frac{\partial G_M}{\partial y_1}(\mathbf{x}, \mathbf{y}) & \cdots & \frac{\partial G_M}{\partial y_M}(\mathbf{x}, \mathbf{y}) \end{pmatrix}$$

Dunque  $\frac{\partial \mathbf{G}}{\partial \mathbf{x}}(\mathbf{x}, \mathbf{y})$  è una matrice  $M \times N$ ,  $\frac{\partial \mathbf{G}}{\partial \mathbf{y}}(\mathbf{x}, \mathbf{y})$  è una matrice  $M \times M$  (quadrata!) e  $J_{\mathbf{G}}(\mathbf{x}, \mathbf{y}) = \left( \frac{\partial \mathbf{G}}{\partial \mathbf{x}}(\mathbf{x}, \mathbf{y}), \frac{\partial \mathbf{G}}{\partial \mathbf{y}}(\mathbf{x}, \mathbf{y}) \right)$ .

## Theorem (delle funzioni implicite)

Siano  $A \subset \mathbb{R}^{N+M}$  aperto,  $\mathbf{G} : A \rightarrow \mathbb{R}^M$  una funzione  $\mathcal{C}^1$  e  $(\mathbf{x}_0, \mathbf{y}_0) \in A$  t.c.:

$$\mathbf{G}(\mathbf{x}_0, \mathbf{y}_0) = 0, \quad \det \left( \frac{\partial \mathbf{G}}{\partial \mathbf{y}}(\mathbf{x}_0, \mathbf{y}_0) \right) \neq 0$$

(notiamo che  $\frac{\partial \mathbf{G}}{\partial \mathbf{y}}(\mathbf{x}_0, \mathbf{y}_0)$  è invertibile avendo determinante non nullo).

Allora esistono  $\delta > 0, \varepsilon > 0$  ed esiste  $\mathbf{f} : B_N(\mathbf{x}_0, \delta) \rightarrow B_M(\mathbf{y}_0, \varepsilon)$  tali che:

$$Q_{\delta, \varepsilon}(\mathbf{x}_0, \mathbf{y}_0) := B_N(\mathbf{x}_0, \delta) \times B_M(\mathbf{y}_0, \varepsilon) \subset A, \\ \{(\mathbf{x}, \mathbf{y}) \in Q_{\delta, \varepsilon}(\mathbf{x}_0, \mathbf{y}_0) : \mathbf{G}(\mathbf{x}, \mathbf{y}) = 0\} = \{(\mathbf{x}, \mathbf{f}(\mathbf{x})) : \mathbf{x} \in B(\mathbf{x}_0, \delta)\}$$

## Funzioni implicite – caso generale

Inoltre la funzione  $\mathbf{f}$  è di classe  $\mathcal{C}^1$  su  $B(\mathbf{x}_0, \delta)$  e:

$$J_{\mathbf{f}}(\mathbf{x}) \left( = \frac{\partial \mathbf{f}}{\partial \mathbf{x}}(\mathbf{x}) \right) = - \left( \frac{\partial \mathbf{G}}{\partial \mathbf{y}}(\mathbf{x}, \mathbf{f}(\mathbf{x})) \right)^{-1} \left( \frac{\partial \mathbf{G}}{\partial \mathbf{x}}(\mathbf{x}, \mathbf{f}(\mathbf{x})) \right) \quad \forall \mathbf{x} \in B(\mathbf{x}_0, \delta).$$

Una conseguenza è il seguente teorema.

### Theorem (invertibilità locale)

*Supponiamo che  $\mathbf{f} : A \rightarrow \mathbb{R}^N$  sia un'applicazione  $\mathcal{C}^1$  definita su un aperto  $A \subset \mathbb{R}^N$ . Sia  $\mathbf{x}_0 \in A$  tale che  $\det(J_{\mathbf{f}}(\mathbf{x}_0)) \neq 0$  (cioè  $J_{\mathbf{f}}(\mathbf{x}_0)$  invertibile). Allora esiste  $\rho > 0$  tale che  $\mathbf{f} : B(\mathbf{x}_0, \rho) \rightarrow \mathbf{f}(B(\mathbf{x}_0, \rho))$  risulta invertibile, l'inversa  $\mathbf{f}^{-1} : \mathbf{f}(B(\mathbf{x}_0, \rho)) \rightarrow B(\mathbf{x}_0, \rho)$  è  $\mathcal{C}^1$  e*

$$J_{\mathbf{f}^{-1}}(\mathbf{y}) = (J_{\mathbf{f}}(\mathbf{x}))^{-1} \quad \text{dove } \mathbf{f}(\mathbf{x}) = \mathbf{y} \quad \forall \mathbf{y} \in \mathbf{f}(B(\mathbf{x}_0, \rho))$$

**Dim.** Basta applicare il teorema a  $\mathbf{G} : \mathbb{R}^{N+N} \rightarrow \mathbb{R}^N$  definita da  $\mathbf{G}(\mathbf{x}, \mathbf{y}) := \mathbf{y} - \mathbf{f}(\mathbf{x})$  (invertendo i ruoli di  $\mathbf{x}$  e  $\mathbf{y}$ ).

# Moltiplicatori di Lagrange – caso generale

## Theorem (Caso di $M$ vincoli di eguaglianza)

Supponiamo ora  $A \subset \mathbb{R}^N$  aperto, siano date  $M$  funzioni di classe  $\mathcal{C}^1$   $G_1, \dots, G_M : A \rightarrow \mathbb{R}$  e poniamo

$$V := \{\mathbf{x} \in A : G_1(\mathbf{x}) = \dots = G_M(\mathbf{x}) = 0\}.$$

Indichiamo  $\mathbf{G}(\mathbf{x}) := (G_1(\mathbf{x}), \dots, G_M(\mathbf{x}))$  ( $\mathbf{G} : A \rightarrow \mathbb{R}^M$ ). Supponiamo che:

$$\forall \mathbf{x}_0 \in V \text{ la matrice jacobiana } \frac{\partial \mathbf{G}}{\partial \mathbf{x}}(\mathbf{x}_0) = J_{\mathbf{G}}(\mathbf{x}_0) \text{ ha rango } M. \quad (\star\star)$$

Se  $f : A \rightarrow \mathbb{R}$  e  $\mathbf{x}_0 \in V$  è punto estrema per  $f$  su  $V$ , allora esistono  $M$  moltiplicatori  $\lambda_1, \dots, \lambda_M$  in  $\mathbb{R}$  tali che:

$$\nabla f(\mathbf{x}_0) = \lambda_1 \nabla G_1(\mathbf{x}_0) + \dots + \lambda_M \nabla G_M(\mathbf{x}_0).$$

La  $(\star\star)$  equivale a dire che  $\nabla G_1(\mathbf{x}_0), \dots, \nabla G_M(\mathbf{x}_0)$  sono linearmente indipendenti. Ne segue in particolare che  $M \leq N$ .

# Moltiplicatori di Lagrange – caso generale

## Theorem (Caso generale)

Supponiamo ora  $A \subset \mathbb{R}^N$  aperto, siano date  $M_1 + M_2$  funzioni di classe  $\mathcal{C}^1$   $G_1, \dots, G_{M_1}, H_1, \dots, H_{M_2} : A \rightarrow \mathbb{R}$  e poniamo:

$$V := \{\mathbf{x} \in A : G_1(\mathbf{x}) = \dots = G_{M_1}(\mathbf{x}) = 0, H_1(\mathbf{x}) \leq 0, \dots, H_{M_2}(\mathbf{x}) \leq 0\}.$$

Supponiamo che per ogni  $\mathbf{x}_0$  in  $V$ :

se  $r \leq M_2$  e  $1 \leq i_1 < i_2 < \dots < i_r \leq M_2$  sono  $r$  indici tali che  $H_{i_1}(\mathbf{x}_0) = H_{i_2}(\mathbf{x}_0) = \dots = H_{i_r}(\mathbf{x}_0) = 0$ , allora: (★★)

$$\frac{\partial G_1 \dots \partial G_{M_1} \partial H_{i_1} \dots \partial H_{i_r}}{\partial x_1 \dots \partial x_N}(\mathbf{x}_0) \text{ ha rango } M_1 + r$$

# Moltiplicatori di Lagrange – caso generale

In queste ipotesi, se  $f : A \rightarrow \mathbb{R}$  e  $\mathbf{x}_0 \in V$  è punto estemale per  $f$  su  $V$ , allora esistono  $M_1 + M_2$  moltiplicatori  $\lambda_1, \dots, \lambda_{M_1}, \mu_1, \dots, \mu_{M_2}$  in  $\mathbb{R}$  tali che:

$$\begin{aligned}\nabla f(\mathbf{x}_0) = & \lambda_1 \nabla G_1(\mathbf{x}_0) + \dots + \lambda_{M_1} \nabla G_{M_1}(\mathbf{x}_0) + \\ & + \mu_1 \nabla H_1(\mathbf{x}_0) + \dots + \mu_{M_2} \nabla H_{M_2}(\mathbf{x}_0).\end{aligned}$$

e inoltre:

$$\mu_i = 0 \text{ per gli indici } i \text{ tali che } H_i(\mathbf{x}_0) < 0$$

(i moltiplicatori compaiono solo per i vincoli in cui si realizza l'eguaglianza).

Se  $\mathbf{x}_0$  verifica la relazione sopra, per opportuni  $\lambda_1, \dots, \lambda_{M_1}$  e  $\mu_1, \dots, \mu_{M_2}$  in  $\mathbb{R}$ , diciamo che  $\mathbf{x}_0$  è un punto critico vincolato per  $f$  su  $V$ .

## Lagrangiana

Come nei casi precedenti possiamo affermare che la caratterizzazione seguente. Sono equivalenti;

- $\mathbf{x}_0$  in  $V$  è punto critico su  $V$  per  $f$ ;
- esistono  $\boldsymbol{\lambda}_0 = (\lambda_{01}, \dots, \lambda_{0M_1}) \in \mathbb{R}^{M_1}$ ,  $\boldsymbol{\mu}_0 = (\mu_{01}, \dots, \mu_{0M_2}) \in \mathbb{R}^{M_2}$  e  $\mathbf{u}_0 = (u_{01}, \dots, u_{0M_2}) \in \mathbb{R}^{M_2}$  tali che  $(\mathbf{x}_0, \boldsymbol{\lambda}_0, \boldsymbol{\mu}_0, \mathbf{u}_0)$  è critico per la Lagrangiana:

$$L(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \mathbf{u}) := f(\mathbf{x}) - \sum_{i=1}^{M_1} \lambda_i G_i(\mathbf{x}) - \sum_{i=1}^{M_2} \mu_i (H_i(\mathbf{x}) + u_i^2).$$

# Domini con spigoli

Estendiamo la nozione di dominio nel modo seguente: diremo che un insieme  $D$  contenuto in  $\mathbb{R}^N$  è un *dominio chiuso “con spigoli”* se

$$D = \left\{ \mathbf{x} \in \mathbb{R}^N : G_1(\mathbf{x}) \leq 0, \dots, G_k(\mathbf{x}) \leq 0 \right\}$$

dove  $G_1, \dots, G_k : \mathbb{R}^N \rightarrow \mathbb{R}$  sono  $k$  funzioni  $\mathcal{C}^1$  tali che:

se  $G_{i_1}(\mathbf{x}_0) = \dots = G_{i_r}(\mathbf{x}_0) = 0$  allora  $\frac{\partial G_{i_1} \cdots \partial G_{i_r}}{\partial \mathbf{x}}(\mathbf{x}_0)$  ha rango  $r$ .

(l'ipotesi  $(\star\star)$ ) di prima senza vincoli di eguaglianza). In sostanza un tale insieme è intersezione di  $k$  domini regolari “i cui bordi si intersecano trasversalmente”. Chiaramente  $D$  è chiuso, mentre la sua parte interna è

$$\text{int}(D) = \left\{ \mathbf{x} \in \mathbb{R}^N : G_1(\mathbf{x}) < 0, \dots, G_k(\mathbf{x}) < 0 \right\}$$

che verrà detto *dominio aperto “con spigoli”*.

# Domini con spigoli

Per esempio il cubo  $Q := \{(x, y, z) : |x| \leq 1, |y| \leq 1, |z| \leq 1\}$  è un dominio con spigoli ottenuto mediante le sei funzioni:

$$\begin{aligned} G_1(x, y, z) &= x - 1, \quad G_2(x, y, z) = x + 1, \quad G_3(x, y, z) = y - 1, \\ G_4(x, y, z) &= y + 1, \quad G_5(x, y, z) = z - 1, \quad G_6(x, y, z) = z + 1. \end{aligned}$$

Per verificarlo si mostra che:

- 1 per ogni  $i = 1, \dots, 6$   $\nabla G_i(x, y, z) \neq \vec{0}$  in ogni punto  $(x, y, z)$  tale che  $G_i(x, y, z) = 0$  (in realtà per ogni  $(x, y, z)$ ): **regolarità delle facce**;
- 2 per ogni coppia di indici  $i, j$  tra 1 e 6 se  $G_i(x, y, z) = G_j(x, y, z) = 0$  allora  $\nabla G_i(x, y, z)$  e  $\nabla G_j(x, y, z)$  sono linearmente indipendenti: **regolarità degli spigoli**;
- 3 per ogni terna di indici  $i, j, k$  tra 1 e 6 se  $G_i(x, y, z) = G_j(x, y, z) = G_k(x, y, z) = 0$  allora  $\nabla G_i(x, y, z)$ ,  $\nabla G_j(x, y, z)$  e  $\nabla G_k(x, y, z)$  sono linearmente indipendenti: **regolarità dei vertici**;