Analisi Matematica II, Anno Accademico 2021-2022.
Ingegneria

Vincenzo M. Tortorelli

FOGLIO DI TEORIA n. 10
DERIVATYE PARZIALI, DIREZIONALI
E DIFFERENZIABILITÀ

1 Derivate parizali e direzionali

1.1 Derivate parziali - si dice che una funzione di due variabili reali $(x, y) \mapsto f(x, y) \in \mathbf{R}^m$ ammette derivata parziale rispetto alla prima variabile nel punto (a, b) se la funzione di una variabile $t \mapsto f(t, b)\mathbf{R}^m$ è derivabile in a.
Analogamente rispetto alla seconda variabile considerando la derivabilità in b di $t \mapsto f(a, t)$.
- Per più di due variabili si estende la definizione come segue:
si dice che $f(x_1, \ldots, x_M) \in \mathbf{R}^m$ ammette derivata parziale rispetto all' $i^a - esima$ variabile nel punto $p = (p_1, \ldots, p_M)$ se la funzione di una variabile $t \mapsto f(p_1 \ldots, p_i + t, \ldots p_M) = f(p + te_i) \in \mathbf{R}^m$ è derivabile in $t = 0$.
Equivalentemente $f(p_1 \ldots, t, \ldots p_M)$ è derivabile in $t = p_i$: $\quad \exists \lim_{t\to 0} \dfrac{f(p + te_i) - f(p)}{t}$.
Per tali eventuali valori, detti **derivate parziali**, si usano le notazioni $\dfrac{\partial f}{\partial x_i}(p)$, $\partial_{x_i} f(p)$, $\partial_i f(p)$.

Osservazione 1: per calcolare la derivata parziale di una funzione $f(x, y, z)$ rispetto alla seconda variabile in $(1, 2, 3)$ si calcola la funzione in $(1, y, 3)$ poi la derivata di $y \mapsto f(1, y, 3)$ per $y = 2$.

Esempio 1: - Per $f(x, y, z, u, v) = 3xyzuv + \left(\log \dfrac{2\cos\left(\log \frac{x^2+z^2}{u^2+v^2}\right)}{y^2 + z^2 + u^2 + v^2}\right) \cdot \text{artan} \log\left(\cos \dfrac{y^2 - z^2}{u^2 + v^2}\right)$

si ha: $\dfrac{\partial f}{\partial x}(1, 1, 1, 1, 1,) = (f(t, 1, 1, 1, 1))'_{|t=1} = \left(3t + \left(\log \dfrac{2\cos\left(\log \frac{t^2+1}{2}\right)}{4}\right) \cdot \text{artan} \log\left(\cos \dfrac{0}{2}\right)\right)'_{|t=1} =$

$= (3t + 0)'_{|t=1} = 3.$

Osservazione 2: Una *funzione vettoriale* a valori in $\mathbf{R}^m$ ammette derivate parziali se e solo se ciò accade per ognuna delle sue componenti.

Funzione derivata parziale. Se in ogni $x \in D \subseteq \text{Dom}f$ vi è $\dfrac{\partial f}{\partial x_i}(x)$ la funzione che ad x associa la derivata parziale in x si dice *funzione derivata parziale* i^a di f e si indica con $\dfrac{\partial f}{\partial x_i}$.

Proposizione 1. Sia Ω aperto connesso di $\mathbf{R}^M$. Allora una funzione che abbia le derivate parziali nulle in ogni punto di Ω è costante su Ω.
Dimostrazione 1: per la proposizione 5 in FT6.5 due punti di Ω si collegano con una spezzata con *lati paralleli agli assi*. Ma la funzione ristretta a questi lati ha derivata nulla per ipotesi quindi è costante su ognuno dei lati. In particolare ha lo stesso valore nei due arbitrari punti di Ω.
Dimostrazione 2: si usa direttamente l'argomento che dimostra anche la citata proposizione 5 FT6.5. Dato $p \in \Omega$ si mostra che $A = \{x \in \Omega : f(p) = f(x)\}$ ed $\Omega \setminus A$ sono aperti.
- A è aperto: se $y \in A \subseteq \Omega$ poichè Ω è aperto vi è $r > 0$ per cui $B = B(y, r) \subseteq \Omega$, per la diseguaglianza $|y| \leq \sqrt{M} \max\{|y_1|, \ldots, |y_M|\}$ l'ipercubo $Q = y + \left(-\frac{r}{\sqrt{M}}; \frac{r}{\sqrt{M}}\right)^M \subset B \subseteq \Omega$. Poichè $\max\{|y_1|, \ldots, |y_M|\} \leq |y|$ si ha $Q \supset B(y, \frac{r}{\sqrt{M}})$ è un intorno di y, e tutti i suoi punti sono raggiungibili da spezzate con lati paralleli agli assi a partire da y: pertanto $f(x) = f(y) = f(p)$ per $x \in Q$, cioè $Q \subseteq A$.
- $\Omega \setminus A$ è aperto: se $y \in \Omega \setminus A$ come sopra si considera un ipercubo $Q \subseteq \Omega$ centrato in y: quindi $f(x) = f(y) \neq f(p)$ per $x \in Q$, per cui $Q \subseteq \Omega \setminus A$.
Poichè Ω è connesso e $A \neq \emptyset$ deve esser $\Omega \setminus A = \emptyset$.

Derivate parziali successive. Per semplicità sia p interno a $\mathrm{Dom} f$.
Le derivate parziali successive in p, nel caso esistano, sono date induttivamente rispetto all'*ordine di derivazione* $k \in \mathbf{N}$, usando le derivate parziali prime:

$$\frac{\partial^0 f}{\partial x_i^0}(p) = f(p),\ \frac{\partial^{k+1} f}{\partial x_{i_{k+1}} \ldots \partial x_{i_1}}(p) = \frac{\partial}{\partial x_{i_{k+1}}}\left(\frac{\partial^k f}{\partial x_{i_k} \ldots \partial x_{i_1}}\right)(p) = \left(\frac{\partial^k f}{\partial x_{i_k} \ldots \partial x_{i_1}}(p + te_{i_{k+1}})\right)'(0).$$

- Tali derivate successive, se esistono, si indicano anche con $\partial_{i_k} \ldots \partial_{i_1} f(p)$, $\partial_{i_k \ldots i_1} f(p)$.
- Se $i_1 = \cdots = i_k = i$ la derivata $\dfrac{\partial^k f}{\partial x_i \ldots \partial x_i}(p)$ si indica anche con $\dfrac{\partial^k f}{\partial x_i^k}(p)$ o con $\partial_i^k f(p)$.

Osservazione 3: - affinchè la definizione abbia senso le derivate precedenti devono *esistere su tutto un segmento per p parallelo alla direzione dell'ultima derivata.*
- Per calcolarle è utile fissare le variabili rispetto a cui non vengono fatte le derivate successive.
Osservazione 4: come per le derivate parziali prime si considereranno le funzioni derivate parziali di ordine k.

Osservazione 5: in generale *non* si può *scambiare l'ordine di derivazione.*

Esempio 2: $f(x, y) = \begin{cases} xy\dfrac{x^2 - y^2}{x^2 + y^2}, & (x, y) \neq (0, 0) \\ 0, & x = y = 0 \end{cases}$:

$$\begin{cases} \dfrac{\partial f}{\partial x}(0, 0) = \lim\limits_{x \to 0} \dfrac{f(x, 0) - f(0, 0)}{x} = 0 \\ \dfrac{\partial f}{\partial y}(0, 0) = 0 \\ \dfrac{\partial f}{\partial x}(0, y) = \lim\limits_{x \to 0} \dfrac{f(x, y) - f(0, y)}{x} = \lim\limits_{x \to 0} y\dfrac{x^2 - y^2}{x^2 + y^2} = -y \\ \dfrac{\partial f}{\partial y}(x, 0) = \lim\limits_{y \to 0} \dfrac{f(x, y) - f(x, 0)}{y} = \lim\limits_{x \to 0} x\dfrac{x^2 - y^2}{x^2 + y^2} = x \end{cases}, \quad \begin{cases} \dfrac{\partial^2 f}{\partial y \partial x}(0, 0) = -1 \\ \dfrac{\partial^2 f}{\partial x \partial y}(0, 0) = 1 \end{cases}.$$

Vi sono diversi criteri che garantiscono lo scambio dell'ordine di derivazione. Una prima versione è la seguente (con $e_1, \ldots e_M$ si intendono vettori della base canonica di $\mathbf{R}^M$):

Teorema 1 di Schwarz, prima versione: siano date la funzione $f : \mathrm{D}omf \subseteq \mathbf{R}^M \to \mathbf{R}$, $1 \leq i < j \leq M$, il punto $p \in Domf$, il piano affine $V = \{x \in \mathbf{R}^M : x_h = p_h \quad h \neq i,\ j\}$ $(x \in V \Leftrightarrow x = (p_1, \ldots, p_{i-1}, x_i, \ldots, x_j, p_{j+1} \ldots) = p + (x_i - p_i)e_i + (x_j - p_j)e_j)$.

Se in un intorno U di p si ha $U \cap V \subseteq Domf$, cioè p sta nella parte interna relativa a V di $V \cap Domf$, ed inoltre per tale intorno:

1)$\exists \dfrac{\partial f}{\partial x_i}(x),\ \dfrac{\partial f}{\partial x_j}(x), \quad x \in U \cap V$

2)$\exists \dfrac{\partial^2 f}{\partial x_i \partial x_j}(x), \quad x \in U \cap V$; allora $\exists \dfrac{\partial^2 f}{\partial x_j \partial x_i}(p) = \dfrac{\partial^2 f}{\partial x_i \partial x_j}(p)$.

3)$\dfrac{\partial^2 f}{\partial x_i \partial x_j}$ ristretta a $U \cap V$ è continua in p;

Dimostrazione: - si restringe la funzione f al piano affine $p + xe_i + ye_j$, $x,\ y \in \mathbf{R}^2$, per cui $x_i = p_i + x$, $x_j = p_j + y$.

- - In questo modo ci si riduce a dimostrare il teorema per una funzione $g(x, y) = f(p + xe_i + ye_j) = f(p_1, \ldots, p_{i-1}, p_i + x, \ldots, p_j + y, \ldots, p_M)$, di due variabili, e per le sue derivate successive nell'origine $(0, 0)$, infatti:

$\dfrac{\partial f}{\partial x_i}(p + xe_i + ye_j) = \dfrac{\partial g}{\partial x}(x, y)$, $\dfrac{\partial f}{\partial x_j}(p + xe_i + ye_j) = \dfrac{\partial g}{\partial y}(x, y)$,

$\dfrac{\partial^2 f}{\partial x_i \partial x_j}(p + xe_i + ye_j) = \dfrac{\partial^2 g}{\partial x \partial y}(x, y)$, $\dfrac{\partial^2 f}{\partial x_j \partial x_i}(p + xe_i + ye_j) = \dfrac{\partial^2 g}{\partial y \partial x}(x, y)$.

- Basta quindi dimostrare che se esistono $\partial_x g$, $\partial_y g$, $\partial_x \partial_y g$ in un intorno di $(0, 0)$ e la derivata seconda è continua in $(0, 0)$, allora $\exists \partial_y \partial_x g(0, 0) = \partial_x \partial_y g(0, 0)$.

Va studiato il limite di $R(y) = \dfrac{\partial_x g(0, y) - \partial_x g(0, 0)}{y}$ per $y \to 0$,

che se esiste per definzione è $\partial_y \partial_x g(0, 0)$.

- - Per ricondursi alle ipotesi si osserva che tale rapporto incrementale è limite di un rapporto incrementale di rapporti incrementali (a posteriori uniforme localmente in y)

$$RR(x, y) = \frac{\frac{g(x,y)-g(0,y)}{x} - \frac{g(x,0)-g(0,0)}{x}}{y} \xrightarrow[x \to 0]{} R(y),$$

questo è utile perchè con gli incrementi discreti *si scambiano i ruoli* di y ed x associando in modo diverso gli addendi al numeratore:

$$\frac{\frac{g(x,y)-g(0,y)}{x} - \frac{g(x,0)-g(0,0)}{x}}{y} = \frac{g(x, y) - g(0, y) - g(x, 0) + g(0, 0)}{xy} = \frac{\frac{g(x,y)-g(0,y)-g(x,0)+g(0,0)}{y}}{x}.$$

- - - Fissato x si applica il teorema di Lagrange alle funzioni $\phi(y) = g(x, y) - g(x, 0) - g(0, y) + g(0, 0)$, nulle per $y = 0$: $\exists \theta = \theta(x, y)$ compreso tra y e 0

$$RR(x, y) = \frac{\partial_y g(x, \theta(x, y)) - \partial_y g(0, \theta(x, y))}{x} = \frac{\partial_y g(x, \theta) - \partial_y g(0, \theta)}{x}.$$

- - - Riapplicando Lagrange alle funzioni $\psi(z) = \partial_y g(z, \theta(x, y))$, si ottiene quindi per qualche $\eta = \eta(x, \theta(x, y))$ compreso tra x e 0

$$RR(x, y) = \partial_x \partial_y g(\eta, \theta).$$

- - Per continuità di $\partial_x \partial_y g$ in $(0, 0)$ essendo $\eta^2 + \theta^2 \leq x^2 + y^2$, fissato $\varepsilon > 0$ vi è r per cui

$$\sup_{x^2+y^2 \leq 2r^2} |RR(x, y) - \partial_x \partial_y g(0, 0)| \leq \varepsilon$$

- - Concludendo:

$$\text{dato } \varepsilon \text{ vi è } r>0 \text{ per cui se } x^2+y^2 \le 2r^2 \text{ si ha}$$

$$|R(y)-\partial_x\partial_y g(0,0)| \le |R(y)-RR(x,y)| + |RR(x,y)-\partial_x\partial_y g(0,0)| \le |R(y)-RR(x,y)|+\varepsilon,$$

facendo tendere x a 0 si ottiene in particolare che

$$\text{per } |y|\le r \text{ è } |R(y)-\partial_x\partial_y g(0,0)| \le \varepsilon,$$

cioè $\exists \lim\limits_{y\to 0} R(y) = \lim\limits_{y\to 0} \dfrac{\partial_x g(0,y)-\partial_x g(0,0)}{y} = \partial_x\partial_y g(0,0).$

Corollario: se p è interno a Domf ed in un intorno U di p esistono $\dfrac{\partial f}{\partial x_i}(x),\ \dfrac{\partial f}{\partial x_j}(x),\ \dfrac{\partial^2 f}{\partial x_i \partial x_j}(x)$, $x\in U$, e la derivata seconda è continua in p allora esiste anche $\dfrac{\partial^2 f}{\partial x_j \partial x_i}(p) = \dfrac{\partial^2 f}{\partial x_i \partial x_j}(p)$.

Corollario : se una funzione ha tutte le derivate parziali, sino all'ordine $k \in \mathbf{N}$, *continue* allora per tali derivate si può scambiare l'ordine di derivazione.

Multindici: - Un vettore $\mathbf{p}=(p_1,\dots,p_M)\in \mathbf{N}^M$ a componenti intere non negative, si dice *M-multi-indice*

La dimensione M si dice *lunghezza* del multi-indice.

Si dice *peso* o *norma* (è in effetti una norma): $\quad |\mathbf{p}| = \sum\limits_{i=1}^{M} p_i$.

Dati due multindici di egual lunghezza si scrive $\mathbf{q}\le\mathbf{p}$ per $q_1\le p_1,\ \dots, q_M \le p_M$.

Si definiscono: $\mathbf{p}! = p_1!\cdot\ldots\cdot p_M!\,,\quad \dbinom{\mathbf{p}}{\mathbf{q}} = \dbinom{p_1}{q_1}\cdot\ldots\cdot\dbinom{p_M}{q_M},\quad \mathbf{q}\le\mathbf{p},$

e per M variabili $\mathbf{y}=(y_1,\dots,y_M)$ si pone $\mathbf{y}^{\mathbf{p}} =: y_1^{p_1}\cdot\ldots y_M^{p_M}$.

Notazione con multindici per le derivate parziali: $\dfrac{\partial^{|\mathbf{p}|}}{\partial x_1 \cdots_{p_1\ volte} \ldots \partial x_M \cdots_{p_M\ volte}} =$

$$= \frac{\partial^{|\mathbf{p}|}}{\partial x_1^{p_1}\ldots\partial x_M^{p_M}} = \frac{\partial^{|\mathbf{p}|}}{\partial \mathbf{x}^{\mathbf{p}}} = \partial_1^{p_1}\partial_2^{p_2}\ldots\partial_M^{p_M} = \partial^{\mathbf{p}}$$

Spazi C^k:
i-sia $A\subseteq \mathbf{R}^M$ aperto, $k\in\mathbf{N}$ o $k=\infty$. L'insieme delle funzioni che hanno le tutte le derivate parziali in A sino all'ordine k, rispettivamente di ogni ordine, si indica con $C^k(A)$.

- Tale insieme è uno *spazio vettoriale e un algebra.*

- Le derivazioni parziali sono *operatori lineari* da C^{k+1} a C^k.

ii- Dato $B\subseteq\mathbf{R}^M$ chiuso, con $C^k(B)$ (in senso forte) si intende l'insieme delle f definite su B per cui vi è $\Omega\supseteq B$ aperto, e $g\in C^k(\Omega)$ con $g_{|B}\equiv f$.

1.2 Derivate direzionali. - Una funzione $f: A\subseteq B\to C$, B e C spazi normati, si dice che ammette derivata direzionale nella direzione $v\in B$, $v\neq 0$, in un punto p del suo dominio A e di accumulazione, se la funzione di una variabile $t\mapsto f(p+tv)$ è derivabile in $t=0$.

- In altre parole: la composizione della funzione con la retta parametrica affine per p e di direzione v, $t\mapsto p+tv$, risulta derivabile per $t=0$.

- Altrimenti: il grafico di f intersecato il piano "verticale" per la retta di direzione v e passante per p è il grafico di una funzione derivabile di una variabile in p.

Tale *derivata direzionale* $(f(p+tv))'_{|t=0}$ si denota con $\dfrac{\partial f}{\partial v}(p)$, $\partial_v f(p)$.

Osservazione 6: - Ammettere derivata parziale rispetto alla variabile *i-esima* è la stessa cosa che avere derivata direzionale nella direzione dell'*i-esimo* asse coordinato: $\frac{\partial f}{\partial x_i}(p) = \frac{\partial f}{\partial e_i}(p)$.

Osservazione 7: - Sia f reale e continua con derivata nella direzione v in p, e si consideri il cammino sul grafico dato da $\gamma(t) = (p + tv, f(p + tv))$.
Si ha che esiste $\gamma'(0) = (v, \partial_v f(p))$ il vettore velocità tangente al sostegno del cammino in $(p, f(p)) = \gamma(0)$.
Quindi la derivata direzionale dà la pendenza, ortogonalmente rispetto al dominio $p + \mathbf{R} \cdot v$ con unità di misura $|v|_M$, della retta tangente a tale curva.
- Quindi se vi è derivata nella direzione v in p il vettore $(v, \partial_v f(p))$ intuitivamente dovrebbe essere tangente al grafico di f nel punto $(p, f(p))$ essendo ivi tangente ad un cammino, $t \mapsto (p + t, f(p + tv))$, il cui sostegno sta sul grafico stesso.

Osservazione 8: Se $\rho \neq 0$ si ha $\frac{\partial f}{\partial \rho v}(x) = \rho \frac{\partial f}{\partial v}(x)$: ovvero la pendenza cambia in modo omogeneo se si cambia unità di misura sull'asse di definizione della funzione.

Osservazione 9: Se $u,\ v, u + v \neq \vec{0}$, pur esistendo le tre derivate direzionali, può essere
$$\frac{\partial f}{\partial (u+v)} \neq \frac{\partial f}{\partial u} + \frac{\partial f}{\partial v}.$$

Esempio 3: - sia $f(x, y)$ un'arbitraria funzione continua che valga 0 sulle bisettrici ($y = x$, $y = -x$), e valga x sul primo asse ($y = 0$): *e.g.* $f(x, y) = \begin{cases} x\dfrac{x^2 - y^2}{x^2 + y^2}, & (x, y) \neq (0, 0) \\ 0, & x = y = 0 \end{cases}$.

In $(0, 0)$ tali funzioni hanno derivate direzionali rispetto ai vettori dati $u = (1, 1)$, $v = (1, -1)$, $u + v = (2, 0)$ rispettivamente eguali a 0, 0, 2: $0 + 0 \neq 2$.
- In *coordinte polari* r, θ sul dominio, si scrive $f(x, y) = \begin{cases} r \cdot \cos\theta \cdot \cos 2\theta, & r \neq 0 \\ 0, & r = 0 \end{cases}$.

In generale si stanno considerando funzioni continue che fissato θ siano lineari in x:
$f(x, y) = r \cdot \cos\theta \cdot \varphi(\theta) = x \cdot \varphi(\theta)$, $r > 0$, $f(0, 0) = 0$, con
- - φ funzione 2π-periodica continua per ottenere la continuità di f,
- - e in più per rispettare le condizioni imposte sulle tre rette $\mathbf{R} \cdot (1, 1)$, $\mathbf{R} \cdot (1, -1)$, $\mathbf{R} \cdot (1, 0)$: $0 = \varphi(\frac{\pi}{4}) = \varphi(\frac{3}{4}\pi) = \varphi(\frac{5}{4}\pi) = \varphi(\frac{7}{4}\pi)$, e $\varphi(0) = 1$.

Pensando quindi alle coordinate polari è agevole costruire funzioni con comportamenti simili: ristrette a rette passanti per l'origine siano lineari in una variabile con derivate direzionali non additive:

Esempio 4: $f(x, y) = \begin{cases} \dfrac{2x^2 y}{x^2 + y^2}, & (x, y) \neq (0, 0) \\ 0, & x = y = 0 \end{cases} = \begin{cases} x \cdot \sin 2\theta, & r > 0 \\ 0, & r = 0 \end{cases}$

Esempio 5: $f(x, y) = \begin{cases} (\alpha x + \beta y)\dfrac{x^2}{x^2 + y^2}, & (x, y) \neq (0, 0) \\ 0 & x^2 + y^2 = 0 \end{cases} = \begin{cases} (\alpha x + \beta y)\dfrac{1}{1 + (\tan\theta)^2}, & r > 0 \\ 0, & r = 0 \end{cases}$

Osservazione 10: - interpretazione grafica degli esempi del tipo precedente:
si considerano in $\mathbf{R}^3$ rette passanti per l'asse "verticale" delle z, che non siano né "verticali", né giacenti sullo stesso piano "verticale". Si ottiene, escluso l'asse verticale, un grafico "rigato" di funzione di due variabili definita sul piano delle $x,\ y$, privato dell'origine, cioè

$$f(x,y) = f(r\cos\theta, r\sin\theta) = r\cdot\varphi(\theta) + \psi(\theta)$$

$$\text{con } \psi(\theta) = \psi(\theta+\pi),\ \varphi(\theta) = -\varphi(\theta+\pi)^{\cdot}$$

Se vi sono rette che passano per lo stesso punto $(0,0,\overline{\alpha})$ dell'asse verticale, la funzione prolungata con $\overline{f}(0,0) = \overline{\alpha}$, avrà, lungo i vettori unitari delle proiezioni ortogonali, sul piano $x,\ y$, di tali rette, derivate direzionali in $(0,0)$: le pendenze delle rette rispetto al piano.
Vista la pressochè totale arbitrarietà della scelta di tali pendenze non è detto che sussista tra loro alcuna relazione.
Osservazione 11: quindi vi possono essere grafici di funzioni reali di M variabili (che dovrebbero avere intuitivamente dimensione M) che in un punto hanno più di M vettori tangenti a cammini con sostegno sul grafico *linearmente indipendenti*. Ovvero l'insieme di questi vettori tangenti a cammini sul grafico può non essere un piano affine M-dimensionale in $\mathbf{R}^{M+1}$ (ambiente del grafico). Perciò l'esistenza delle derivate in ogni direzione in un punto non comporterebbe l'esistenza di un "piano tangente": in qualsiasi senso si voglia intendere la nozione di "piano tangente ad un grafico" di una funzione di M variabili reali, questa dovrebbe definire un sottospazio affine M-dimensionale in $\mathbf{R}^{M+1}$.

Derivate direzionali iterate: per semplicità sia p interno al dominio di f. Dati $(v^i)_{i\in\mathbf{N}}$ vettori non nulli, le derivate direzionali iterate rispetto ad essi, se esistono, sono date induttivamente sull'ordine di derivazione:

$$\frac{\partial^0 f}{\partial(v^i)^0}(p) = f(p),\ \frac{\partial^{k+1} f}{\partial v^{k+1}\dots\partial v^1}(p) = \frac{\partial}{\partial v^{k+1}}\left(\frac{\partial^k f}{\partial v^k\dots\partial v^1}\right)(p) = \left(\frac{\partial^k f}{\partial v^k\dots\partial v^1}(p+tv^{k+1})\right)'(0).$$

- Tali derivate successive si indicano anche con $\partial_{v^k}\dots\partial_{v^1}f(p)$, $\partial_{v^k\dots v^1}f(p)$.
- Se $v^1 = \dots = v^k = v$ la derivata $\dfrac{\partial^k f}{\partial v\dots\partial v}(p)$ si indica anche con $\dfrac{\partial^k f}{\partial(v)^k}(p)$ o con $\partial_v^k f(p)$.

Scambio dell'ordine di derivazione: vale un criterio analogo di scambio dell'ordine di derivazione per le derivate direzionali iterate con *direzioni fisse*. Per esempio per due derivate in direzioni indipendenti u, v, fissato il punto p in cui si deriva, ci si restringe al piano per tale punto generato dalle due direzioni. Soddisfando la funzione $g(s,t) = f(p+su+tv)$ le ipotesi del teorema di Schwarz, o dei suoi corollari, in $(0,0)$ rispetto alle variabili s, t si ottiene l'eguaglianza delle due derivate seconde direzionali miste di f in p.

Funzione derivata direzionale rispetto a direzioni variabili: dati $f : A\subseteq\mathbf{R}^M\to\mathbf{R}^m$ e $v : A\to\mathbf{R}^M$ (campo di vettori) quando definita la funzione

$$x\in A\mapsto\frac{\partial f}{\partial v(x)}(x)$$

si indica comunque con

$$\frac{\partial f}{\partial v}.$$

2 Differenziabilità: approssimazione lineare e tangenza a grafici

PROBLEMATICA:
a - individuare una famiglia di sottoinsiemi di $\mathbf{R}^N$, oltre ai piani affini M dimensionali, $M \leq N$, per cui sia sensato dire che hanno dimensione M;
b - per un tale insieme individuare una nozione di piano M-dimensionale tangente in un suo assegnato punto p in modo che la *giacitura* del tangente venga a coincidere

con i vettori velocità in p di cammini passanti p e sostegni contenuti nell'insieme in questione.

Una risposta elementare, tra le varie possibili, basata sull'analisi delle funzioni di più variabili, alla prima richiesta è la seguente:

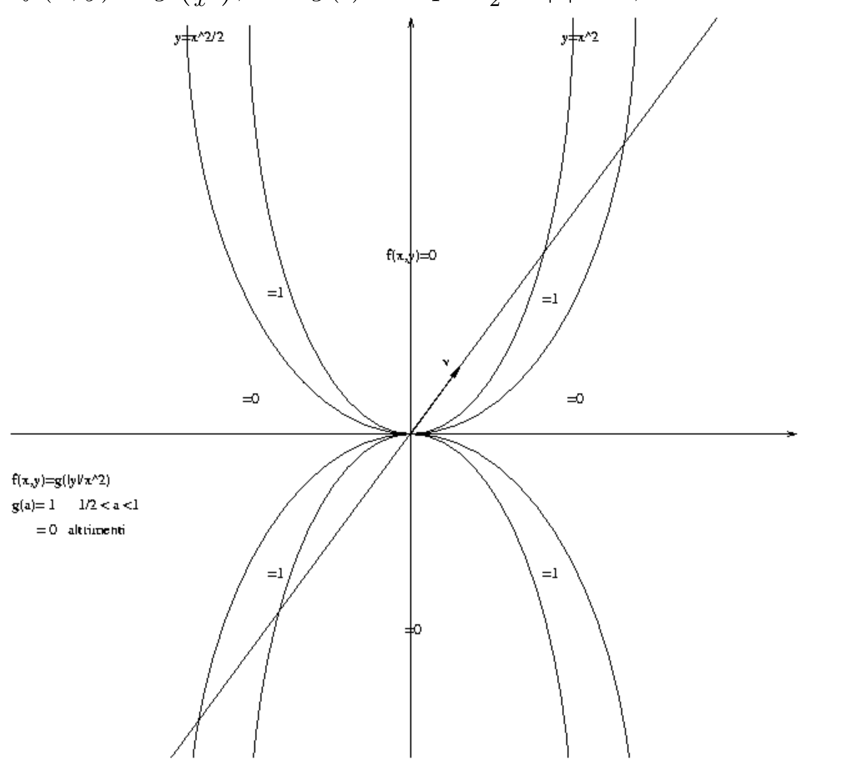
2.1 Grafici locali:

RISPOSTA: a - i sottoinsiemi V di $\mathbf{R}^N$ che possono essere considerati M-dimensionali sono quelli per cui per ogni $p \in V$ vi è un intorno $U(p)$ per cui $V \cap U$ è grafico di una funzione continua di M tra le coordinate di $\mathbf{R}^N$, a valori in $\mathbf{R}^{N-M}$, identificato con il sottospazio vettoriale di $\mathbf{R}^N$ delle rimanenti $N - M$ coordinate.
- In breve si dice che *localmente* il sottoinsieme *è un grafico* di una funzione di M variabili.
- Cioè si identifica la dimensione con il numero "dei gradi di libertà" necessari per descrivere l'insieme: cioè le M coordinate "indipendenti" della funzione.

L'esistenza di tutte le derivate direzionali in un punto $p = (p_1, \ldots, p_M)$ per una funzione f di M variabili non garantisce che i vettori tangenti $(v, \partial_v f(p))$ ai cammini $t \mapsto (p+tv, f(p+tv)) \in$ Graf$f$, e quindi intuitivamente al grafico, formino un sottospazio affine M dimensionale (esmpi 3, 4 , 5). Altri esempi:

Esempio 6: - $f(x,y) = \begin{cases} 1, & \dfrac{x^2}{2} < |y| < x^2 \\ 0, & \text{altrimenti} \end{cases}$ ovvero

$f(x,y) = g\left(\frac{y}{x^2}\right)$, con $g(t) = 1$ per $\frac{1}{2} < |t| < 1$, altrimenti nulla [cfr. figura].



La funzione ha tutte le derivate direzionali in $(0,0)$ poichè, per convessità di $y = x^2$, qualsiasi direzione $(a,b) \neq (0,0)$ dall'origine individua un segmento centrato in $(0,0)$ ove f è nulla. Pertanto la restrizione di f a queste rette $f(ta,tb)$, essendo costante intorno all'origine, ha derivata nulla per $t = 0$.
Si noti che la funzione f risulta discontinua in $(0,0)$.

Esempio 7: - moltiplicando per x la funzione f si ottiene una funzione ϕ continua in $(0,0)$. Infatti f è limitata (assume due valori) ed $x \to 0$ se $x^2 + y^2 \to 0$.
- La ϕ ha tutte le derivate direzionali nulle in $(0,0)$ essendo nulla, per convessità di $y = x^2$, sui segmenti per $(0,0)$.
Per la richiesta fatta in b), se ci fosse un *piano tangente bidimensionale*, conterrebbe le velocità $(a,b,0)$ date da $(at, bt, \phi(ta,tb))'_{|t=0}$: dovrebbe essere il piano "orizzontale" per l'origine definito da $z = 0$.
- Restringendosi invece alle parabole di apertura $\alpha \in (\frac{1}{2}; 1)$, o meglio considerando le composizioni $\phi(t, \alpha t^2) = t$, e considerando le curve sollevate sul grafico $(t, \alpha t^2, t)$ passanti per $(0,0,0)$, si ottengono, per $t = 0$, le velocità $(1,0,1)$, che non giacciono sul candidato piano tangente.
• Osservazione 12: per tali funzioni vale anche $\partial_u \phi(0,0) + \partial_{\rho v} \phi(0,0) = \partial_{u+\rho v} \phi(0,0) = 0$.
Ovvero non solo esistono tutte le derivate direzionali ma sono in relazione di linearità.

Esempio: 8 - con la stessa tecnica, moltiplicando per x una funzione che è costante su ogni curva di una "stella" di curve centrata in $(0,0)$, si ottengono altri svariati esempi di funzioni F con tutte le derivate direzionali in $(0,0)$, ma con vettori tangenti ai cammini sul grafico sollevati di tali curve, che non descrivono un piano bidimensionale.
- Appunto considerando ancora come famiglia di cammini le parable per l'origine le funzioni del tipo $F(x,y) = x \cdot G\left(\frac{y}{x^2}\right)$, al variare di G continua di una variabile, si hanno diversi esempi di funzioni continue in ogni punto con tutte le derivate direzionali in $(0,0)$ ma con velocità, in $(0,0,F(0,0)) = (0,0,0)$, di cammini sul grafico non complanari:
cfr. FT6-6 controesempio 16 finale con $G(t) = (\sin(2\pi t))^2,\ 1 < 2t < 2$, $G(t) = 0$ altrimenti;
FT6-1 esempio 3 con $G(t) = \frac{t}{1+t^2}$.

2.2 Differenziabilità, approssimazione lineare:

RISPOSTA: b- Il concetto *sufficiente* per avere un grafico con piano tangente, che si possa studiare con gli strumenti del calcolo differenziale (le derivate parziali), è quello più impegnativo di *approssimazione lineare* della funzione:

Differenziabilità in punctis: $m + M = N$. Dati

$$f : A \subseteq \mathbf{R}^M \to \mathbf{R}^m,\ f(x) = f(x_1, \ldots, x_m) = (f_1(x_1, \ldots, x_M), \ldots f_m(x_1, \ldots, x_M))$$

$$p = (p_1, \ldots p_m) \quad \textit{interno ad } A$$

La funzione f si dice *differenziabile* in p se

vi è una funzione lineare $L_p : \mathbf{R}^M \to \mathbf{R}^m$ (l'approssimante lineare) per cui

$$f(x) = f(p) + L_p(x - p) + \varepsilon$$

$$\text{con } \lim_{r \to 0} \sup_{|x-p|_M \leq r} \frac{|\varepsilon|_m}{|x-p|_M} = 0$$

$$\text{cioè } \varepsilon(x,p) = o(|x-p|),\ |x-p| \to 0$$

- Nel caso si usa la notazione $L_p = D_p f$: tale applicazione lineare si dice *differenziale* di f in p.
- Si ha per $v \in \mathbf{R}^M$: $D_p f v = (D_p f_1 v, \ldots, D_p f_m v)$.

Osservazione 13: - per funzioni di una variabile $M=1$, le funzioni lineari da $\mathbf{R}$ in $\mathbf{R}^m$ sono le moltiplicazioni per un fissato vettore $a\in\mathbf{R}^m$, $L(t)=a\cdot t$, la differenziabilità è la derivabilità:
$$\exists f'(p)=a \iff \exists\lim_{x\to p}\frac{f(x)-f(p)}{x-p}=a \iff f(x)=f(p)+a\cdot(x-p)+o(x-p)$$
Se esiste, il differenziale per funzioni di una sola variabile è la funzione lineare $u\mapsto f'(p)\cdot u$.

- Nel caso $M=2$, $m=1$, $p=(x_0,y_0)$, le funzioni lineari da $\mathbf{R}^2$ ad $\mathbf{R}$ sono i prodotti scalari per un fissato vettore $(a,b)\in\mathbf{R}^2$, $L(x,y)=ax+by$, affinchè la funzione $f(x,y)$ sia differenziabile in (x_0,y_0) devono esistere due numeri a_p, b_p per cui:
$$\lim_{(x,y)\to(x_0,y_0)}\frac{f(x,y)-f(x_0,y_0)-a(x-x_0)-b(y-y_0)}{\sqrt{(x-x_0)^2+(y-y_0)^2}}=0$$
D_pf è quindi la funzione lineare $(u,v)\mapsto au+bv=\langle(a,b)\cdot(u,v)\rangle=(a,b)\begin{pmatrix}x\\y\end{pmatrix}$.

2.2.1 Matrice Jacobiana: se esiste D_pf, la matrice $m\times M$ associata, nelle basi canoniche, a tale applicazione lineare si dice matrice **Jacobiana** di$f$ in p. Si indica con J_pf o con $Jf(p)$
Osservazione 14: si può esprimere in modo equivalente la differenziabilità in p: la funzione f, è *differenziabile* in p se e solo se vi è una matrice $Jf(p)=J=(J_i^j)_{\substack{1\le i\le m\\1\le j\le M}}$, $m\times M$, per cui
$$f(x)=f(p)+J\ {}^t(x-p)+\varepsilon$$
$$\text{con}\ \lim_{r\to0}\sup_{|x-p|_M\le r}\frac{|\varepsilon|_m}{|x-p|_M}=0.$$
$$i.e.\quad(\varepsilon=o(|x-p|),\ |x-p|\to0)$$

Osservazione 15: $$Jf(p)=(J_i^j)_{\substack{1\le i\le m\\1\le j\le M}}=(J^1\dots J^M)=\begin{pmatrix}J_1\\ \vdots\\ J_m\end{pmatrix}=\begin{pmatrix}Jf_1(p)\\ \vdots\\ Jf_m(p)\end{pmatrix}.$$
Le *righe* $J_i=Jf_i(p)$, $1\le i\le m$, rappresentano i differenziali delle funzioni componenti D_pf_i:
$$\begin{cases}f_1(x)=f_1(p)+\sum_{j=1}^M J_1^j(x_j-p_j)+o_1(|x-p|_M)\\ f_2(x)=f_2(p)+\sum_{j=1}^M J_2^j(x_j-p_j)+o_2(|x-p|_M)\\ \vdots\\ f_m(x)=f_m(p)+\sum_{j=1}^M J_m^j(x_j-p_j)+o_m(|x-p|_M)\end{cases},\quad\begin{cases}f_1(p+v)=f_1(p)+\sum_{j=1}^M J_1^jv_j+o_1(|v|_M)\\ f_2(p+v)=f_2(p)+\sum_{j=1}^M J_2^jv_j+o_2(|v|_M)\\ \vdots\\ f_m(p+v)=f_m(p)+\sum_{j=1}^M J_m^jv_j+o_m(|v|_M)\end{cases}$$
2.2.2 Gradiente: - si dice matrice **gradiente** in p di f, ivi differenziabile, la matrice $M\times m$ *trasposta* di $Jf(p)$: $(G_j^i)_{\substack{1\le j\le M\\1\le i\le m}}=(J_1\dots J_m)=\begin{pmatrix}J^1\\ \vdots\\ J^M\end{pmatrix}$. Si denota con $\nabla f(p)$ o grad$f(p)$.

Osservazione 16: Le sue *colonne* sono i gradienti $M\times1$ delle funzioni componenti
$$\nabla f(p)=\big(\ \nabla f_1(p)\dots\nabla f_m(p)\ \big).$$
Importanti ed immediate conseguenze della definzione sono:

Teorema 2: 1) $f = (f_1, \dots, f_m) : A \subseteq \mathbf{R}^M \to \mathbf{R}^m$ è differenziabile in p interno ad A se e solo se lo sono le funzioni componenti: $D_p f = (D_p f_1, \dots, D_p f_m)$.
2) Le funzioni differenziabili in p sono uno *spazio vettoriale* e D_p è *lineare* a valori operatori lineari: $$D_p(rf + g) = rD_p f + D_p g.$$
3) - Le funzioni c costanti sono differenziabili in ogni punto p: $D_p c \equiv \mathbf{0}_{m\times M}$.
- Le funzioni lineari L sono differenziabili in ogni punto p:
il loro differenziale è costantemente uguale alla funzione stessa, $D_p L = L$.
4) Se f è differenziabile in p allora f è *continua* in p.
5) Se f è differenziabile in p il differenziale è *unico*:
- esistono tutte le derivate parziali di f in p e sono i coefficienti della matrice Jacobiana:

$$\frac{\partial f_i}{\partial x_j}(p) = D_p f_i[e_j] = Jf_i(p)\begin{pmatrix} 0 \\ \vdots \\ 1_{j^o\text{posto}} \\ \vdots \\ 0 \end{pmatrix} = (Jf(p))_i^j = \langle e_j \cdot \nabla f_i(p)\rangle = (\nabla f_i(p))_j = (\nabla f(p))_j^i, \quad \text{cioè}$$

$$Jf(p) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(p)\dots & \frac{\partial f_1}{\partial x_j}(p)\dots & \frac{\partial f_1}{\partial x_M}(p) \\ & \vdots & \\ \frac{\partial f_i}{\partial x_1}(p)\dots & \frac{\partial f_i}{\partial x_j}(p)\dots & \frac{\partial f_i}{\partial x_M}(p) \\ & \vdots & \\ \frac{\partial f_m}{\partial x_1}(p)\dots & \frac{\partial f_m}{\partial x_j}(p)\dots & \frac{\partial f_m}{\partial x_M}(p) \end{pmatrix} = \left(\frac{\partial f}{\partial x_1}(p)\dots\dots\frac{\partial f}{\partial x_M}(p)\right), \quad \nabla f(p) = \begin{pmatrix} \frac{\partial\, {}^t f}{\partial x_1}(p) \\ \vdots \\ \frac{\partial\, {}^t f}{\partial x_M}(p) \end{pmatrix}$$

- esistono le derivate direzionali in p per $v = (v_1, \dots, v_M) \neq 0_{\mathbf{R}^M}$: $$\frac{\partial f}{\partial v}(p) = \begin{pmatrix} \frac{\partial f_1}{\partial v}(p) \\ \vdots \\ \frac{\partial f_m}{\partial v}(p) \end{pmatrix} =$$

$$= D_p f(p)[v] = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(p)\dots & \frac{\partial f_1}{\partial x_j}(p)\dots & \frac{\partial f_1}{\partial x_M}(p) \\ & \vdots & \\ \frac{\partial f_i}{\partial x_1}(p)\dots & \frac{\partial f_i}{\partial x_j}(p)\dots & \frac{\partial f_i}{\partial x_M}(p) \\ & \vdots & \\ \frac{\partial f_m}{\partial x_1}(p)\dots & \frac{\partial f_m}{\partial x_j}(p)\dots & \frac{\partial f_m}{\partial x_M}(p) \end{pmatrix} \begin{pmatrix} v_1 \\ \vdots \\ \vdots \\ \vdots \\ v_M \end{pmatrix} = v_1 \frac{\partial f}{\partial x_1}(p) + \dots + v_M \frac{\partial f}{\partial x_M}(p) =$$

$$= \left(v_1 \frac{\partial}{\partial x_1} + \dots + v_M \frac{\partial}{\partial x_M}\right) f(p) = (v_1, \dots, v_M)\nabla f(p) = \begin{pmatrix} \langle v \cdot \nabla f_1(p)\rangle \\ \vdots \\ \langle v \cdot \nabla f_m(p)\rangle \end{pmatrix} =: \langle v \cdot \nabla\rangle f(p)$$

$v,\ w,\ v + \rho w \in \mathbf{R}^M$ sono *non nulli*, $\rho \in \mathbf{R}$, si ha $\dfrac{\partial f}{\partial(v + \rho w)}(p) = \dfrac{\partial f}{\partial v}(p) + \rho \dfrac{\partial f}{\partial w}(p)$.

2.3 Piano tangente ad un grafico

Piano tangente ad un grafico: $f : A \subseteq \mathbf{R}^M \to \mathbf{R}^m$, p interno ad A.
Se f è differenziabile in p si dice *piano tangente* al <u>grafico</u> di f in $(p, f(p))$, denotato con $T_{(p,f(p))}$, il piano M-dimensionale affine in $\mathbf{R}^{M+m}$

traslato nel punto $(p, f(p))$ del <u>*grafico del differenziale*</u>.

Giacitura del tangente: - il *grafico del differenziale* è un sottospazio vettoriale M-dimensionale di cui il tangente $T_{(p,f(p))}$ è il traslato. Si indica con T_p, tavolta confondendolo con il tangente.
- Considerando che $Graf D_p f = T_p = \mathrm{Im} \begin{pmatrix} Id_M \\ Jf(p) \end{pmatrix} = \mathrm{Ker}\,(Jf(p)| - Id_m)$ e che

$$\begin{pmatrix} Id_M \\ J_p f \end{pmatrix} = \begin{pmatrix} 1\ 0\ \dots & & \dots\ 0 \\ \vdots & & \vdots \\ 0\ \dots & & \dots\ 0\ 1 \\ \frac{\partial f_1}{\partial x_1}(p) & \dots & \frac{\partial f_1}{\partial x_M}(p) \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1}(p) & \dots & \frac{\partial f_m}{\partial x_M}(p) \end{pmatrix}$$

$$(Jf(p)| - Id_{m\times m}) = D_{(p,f(p))}(f - Id_m) = \left.\begin{pmatrix} \frac{\partial f_1}{\partial x_1}(p), \dots, \frac{\partial f_1}{\partial x_M}(p) & -1\ \dots 0 \\ & \vdots \\ \frac{\partial f_m}{\partial x_1}(p), \dots, \frac{\partial f_m}{\partial x_M}(p) & \underbrace{0 \dots\ -1}_{\mathrm{m}} \end{pmatrix}\right\} m, \quad \text{si ha}$$

- una *base* della giacitura *del tangente* è $\begin{pmatrix} 0 \\ \vdots \\ 1_{j^o \text{posto}} \\ \vdots \\ 0 \\ \dfrac{\partial f}{\partial x_j}(p) \end{pmatrix} = \dfrac{\partial}{\partial x_j} \begin{pmatrix} x \\ f(x) \end{pmatrix}(p), \quad 1 \le j \le M;$

- una *base dell'ortogonale* alla giacitura *del tangente* $\left(\dfrac{\partial f_i}{\partial x_1}(p), \dots, \dfrac{\partial f_i}{\partial x_M}(p), 0 \dots, -1_{i^o \text{ posto}} \dots 0\right)$, $1 \le i \le m$.
- Per $m = 1$ un vettore ortogonale a T_p è $\left(\dfrac{\partial f_i}{\partial x_1}(p), \dots, \dfrac{\partial f_i}{\partial x_M}(p), -1\right) \sim \begin{pmatrix} \nabla f(p) \\ -1 \end{pmatrix}$.

Forma parametrica

- quindi il piano tangente ad un grafico è il piano M-dimensionale affine

immagine dell'applicazione affine $v \mapsto (p, f(p)) + (Id_M, D_p f)v : \mathbf{R}^M \to \mathbf{R}^{M+m}$,

cioè il piano M dimensionale in $\mathbf{R}^{M+m}$ descritto dai parameteri $(v_1, \dots v_M)$ come segue

$$(p, f(p)) + (v, D_p f v) = (p, f(p)) + \left(v, \frac{\partial f}{\partial v}(p)\right), \text{ cioè}$$

$$(p, f(p)) + v_1 \left(e_1^{\mathbf{R}^M}, \frac{\partial f}{\partial x_1}(p)\right) + \cdots + v_M \left(e_M^{\mathbf{R}^M}, \frac{\partial f}{\partial x_M}(p)\right)$$

$$(p, f(p)) + v_1 \left(1, 0 \dots, 0, \frac{\partial f}{\partial x_1}(p)\right) + \cdots + v_M \left(0, \dots, 0, 1, \frac{\partial f}{\partial x_M}(p)\right)$$

- un'altra forma parametrica usando come parametri $x = v + p$ è

$$(p, f(p)) + (x_1 - p_1) \left(1, 0 \dots, 0, \frac{\partial f}{\partial x_1}(p)\right) + \cdots + (x_M - p_M) \left(0, \dots, 0, 1, \frac{\partial f}{\partial x_M}(p)\right), \text{ cioè}$$

$$\left(x_1, \dots, x_M, f(p) + \frac{\partial f}{\partial x_1}(p)(x_1 - p_1) + \cdots + \frac{\partial f}{\partial x_M}(p)(x_M - p_M)\right).$$

Forma cartesiana

quindi il piano tangente ad un grafico è *luogo di zeri* in $\mathbf{R}^{M+m}$ della funzione

$$G(x, z) = f(p) + D_p f[x - p] - z, \; G : \mathbf{R}^{M+m} \to \mathbf{R}^m$$

$$G(x, z) = (Jf(p)|Id_{m\times m}) \begin{pmatrix} x - p \\ f(p) - z \end{pmatrix} = D_{(p, f(p))}(f - Id_m) \begin{bmatrix} x - p \\ f(p) - z \end{bmatrix} = 0_{\mathbf{R}^m} :$$

$$(*) \begin{cases} z_1 = & f_1(p) + \dfrac{\partial f_1}{\partial x_1}(p)(x_1 - p_1) + \cdots + \dfrac{\partial f_1}{\partial x_M}(p)(x_M - p_M) \\ \vdots & \qquad\qquad \vdots \\ z_m = & f_m(p) + \dfrac{\partial f_m}{\partial x_1}(p)(x_1 - p_1) + \cdots + \dfrac{\partial f_m}{\partial x_M}(p)(x_M - p_M) \end{cases}$$

2.4 approssimazione metrica: analogamente a quanto mostrato per le velocità di cammini (cfr. FT5-4 ultimo paragrafo), si deduce che il grafico di una funzione $f : A \subseteq \mathbf{R}^M \to \mathbf{R}^m$ differenziabile in p è *"approssimato al primo ordine"* nel suo punto $(p, f(p))$, dal traslato del grafico del differenziale .

Il senso dell'approssimazione è il seguente:
ricordando la nozione di distanza tra un punto ed un insieme $\operatorname{dist}(q, A) = \inf_{a\in A} d(q, a)$, posto $P = (p, f(p))$ (quindi con T_P si indica il piano tangente in $(p, f(p))$ al grafico di f), $Q = (x, f(x))$ per $x \in \operatorname{Dom} f$, si ha:

$$\lim_{\substack{Q\in \operatorname{Graf} f \\ Q\to P}} \frac{\operatorname{dist}(Q, T_P)}{\operatorname{dist}(Q, P)} = 0.$$

Infatti, per continuità di f in p, $Q \to P$ se e solo se $x \to p$, e

$$\frac{\operatorname{dist}(Q, T_P)}{\operatorname{dist}(Q, P)} \le \frac{\operatorname{dist}(Q, T_P)}{|x - p|_M} \le \frac{|(x, f(x)) - (x, D_p f(x - p) + f(p))|_{M+m}}{|x - p|_M} =$$

$$= \frac{|f(x) - f(p) - D_p f(x - p)|_m}{|x - p|_M} \to 0 \,, \quad x \to p.$$

2.5 Vettori tangenti al grafico

Vettori tangenti. Che *le velocità* in $(p, f(p))$ di cammini *sul grafico* di $f : A \subseteq \mathbf{R}^M \to \mathbf{R}^m$, differenziabile in p, diano *tutto* $T_p f$, si deduce della regola della catena (differenziale di funzioni composte cfr. FT11) nel caso particolare di composizione con cammini (funzioni di una variabile):

Proposizione 2, regola della catena per cammini: se $f : A \subseteq \mathbf{R}^M \to \mathbf{R}^m$ è differenziabile in p e $\gamma(t)$ è un cammino in Dom$f$, derivabile per t=0, con $\gamma(0) = p$, allora $f(\gamma(t))$ è derivabile per t=0 e

$$(f\circ\gamma)'(0) = \gamma_1'(0)\frac{\partial f}{\partial x_1}(0) + \cdots + \gamma_M'(0)\frac{\partial f}{\partial x_M}(0) = Jf(p)[\gamma'(0)] = (\gamma'(0)\cdot\nabla)f(p) = \frac{\partial f}{\partial \gamma'(0)}(p).$$

Dimostrazione: per differenziabilità di f: $f(\gamma(t)) - f(p) = Jf(p)(\gamma(t) - p) + o(|\gamma(t) - p|_M)$
per derivabilità di γ: $\gamma(t) - p = \gamma'(0)t + \vec{o}(t), \qquad$ ove $|\vec{o}(t)|_M = o(t)$.
Sostituendo la seconda nella prima

$$f(\gamma(t)) - f(p) = tJf(p)(\gamma'(0)) + Jf(p)\vec{o}(t) + o(|\gamma'(0)t + \vec{o}(t)|_M)$$

- Per Cauchy-Schwarz se A è una matrice $m \times M$, $v \in \mathbf{R}^m$ si ha $|Av|_m \le |v|_M\sqrt{\sum(A_i^j)^2}$: quindi il penultimo addendo del secondo membro è un $o(t)$.
- Poichè $o(|u+v|_M) \subseteq o(|u|_M + |v|_M)$ e se $|u|_M = o(|v|_M)$ si ha $o(|u|_M + |v|_M) = o(|u|_M)$ anche l'ultimo addendo del secondo membro è $o(t)$. Concludendo

$$f(\gamma(t)) - f(p) = tJf(p)(\gamma'(0)) + o(t), \qquad \text{cioè } \frac{f(\gamma(t)) - f(p)}{t} \underset{t\to 0}{\longrightarrow} Jf(p)(\gamma'(0)) = \frac{\partial f}{\partial \gamma'(0)}(p).$$

Corollario: se f è differenziabile in p il suo differenziale $D_p f$ è l'applicazione lineare

$$\frac{d\gamma}{dt}(0) \mapsto \frac{df\circ\gamma}{dt}(0),$$

per ogni γ cammino in Dom$f$, derivabile per $t = 0$ con $\gamma(0) = p$.

Teorema 3: La giacitura del piano tangente al grafico di f in $(p, f(p))$ sono esattamente i vettori V di $\mathbf{R}^{M+m}$, oltre a quello nullo, del tipo $V = \left(v, \frac{\partial f}{\partial v}(p)\right)$ con $v \in \mathbf{R}^M$ non nullo: cioè del tipo $\Gamma'(0) = \left(\gamma'(0), \frac{\partial f}{\partial \gamma'(0)}(p)\right)$, $\Gamma(t) = (\gamma(t), f(\gamma(t))$, $\gamma(0) = p$, $\gamma'(0) = v \neq 0_{\mathbf{R}^M}$.

Dimostrazione: - se $\Gamma(t) = (\gamma(t), f(\gamma(t)) \in$ Graf$f$, con $\gamma(t)$ cammino derivabile per $t = 0$ e $\gamma(0) = p$, si ha per la regola della catena : $\Gamma'(0) = (\gamma'(0), D_p f\gamma'(0)) = \left(\gamma'(0), \frac{\partial f}{\partial \gamma'(0)}(p)\right) \in T_p$.

- Per quanto visto al punto 5 del teorema 2, tali vettori formano uno spazio vettoriale. In particolare per i cammini che danno le derivate parziali $\gamma_j(t) = p + te_j$, $1 \le j \le M$, si ha ${\Gamma_j}'(0) \sim (e_j, \partial_j f(p))$, che sono, come già osservato, una base di T_p.

2.6 Interpretazione geometrica del gradiente: equazioni del tangente ad un insieme di livello (preimmagine)

Gradiente come vettore nello spazio del dominio: se $f : D \subseteq \mathbf{R}^M \to \mathbf{R}$ è differenziabile in p, allora $\nabla f(p)$ è *ortogonale* in p *al livello per* p, $Z = \{x \in D : \ f(x) = f(p)\}$ nel senso seguente: per ogni cammino $\gamma : I \to Z$, $\gamma(0) = p$, derivabile per $t = 0$ si ha:

$$\gamma_1'(0)\frac{\partial f}{\partial x_1}(p) + \cdots + \gamma_M'(0)\frac{\partial f}{\partial x_M}(p) = \langle \nabla f(p) \cdot \gamma'(0)\rangle = 0.$$

Dimostrazione: per la regola della catena per la composizione con funzioni di una variabile si ha che $t \mapsto f(\gamma(t))$ è derivabile per $t = 0$. Inoltre:

$(f_\circ\gamma)'(0) = \langle \nabla f(p) \cdot \gamma'(0)\rangle$, e con l'ipotesi $\gamma(t) \in Z,\ t \in I$, ovvero $f(\gamma(t)) = 0,\ t \in I$, si ha

$$\gamma_1'(0)\frac{\partial f}{\partial x_1}(p) + \cdots + \gamma_M'(0)\frac{\partial f}{\partial x_M}(p) = (f_\circ\gamma)'(0) = 0.$$

Corollario, equazione del piano tangente ad un insieme di livello:
- sia $f : D \subseteq \mathbf{R}^M \to \mathbf{R}$ è C^1 con $\underline{\nabla f(p) \neq \vec{0}_{\mathbf{R}^M}}$,
se $Z = f^{-1}(\{f(p)\})$, intersecato un intorno del punto p è il grafico di una funzione reale di $(M-1)$-variabili differenziabile,
il suo piano tangente $(M-1)$-dimensionale in p è ben defnito e l'equazione di questo è:

$$\langle (x-p) \cdot \nabla f(p)\rangle = (x_1 - p_1)\frac{\partial f}{\partial x_1}(p) + \ \cdots + (x_M - p_M)\frac{\partial f}{\partial x_M}(p) = 0.$$

- Sia $f : D \subseteq \mathbf{R}^M \to \mathbf{R}^m$, $m < M$, è C^1 con $\underline{\nabla f(p) \text{ di rango massimo } m}$,
se $Z = f^{-1}(\{f(p)\})$ in un intorno di p è grafico di una funzione differenziabile di $(M-m)$-variabili a valori in qualche sottospazio di $\mathbf{R}^M$ di dimensione m e quindi ha piano tangente $M - m$ dimensionale in p, l'equazione di questo è:

$$D_p f[x-p] = \vec{0}_{\mathbf{R}^m} \text{ cioè } \begin{cases} (x_1 - p_1)\dfrac{\partial f_1}{\partial x_1}(p) + \ \cdots + (x_M - p_M)\dfrac{\partial f_1}{\partial x_M}(p) = 0 \\ \vdots \\ (x_1 - p_1)\dfrac{\partial f_m}{\partial x_1}(p) + \ \cdots + (x_M - p_M)\dfrac{\partial f_m}{\partial x_M}(p) = 0 \end{cases}.$$

Cioè $\nabla f_1(p), \ldots, \nabla f_m(p)$ sono una *base* per *l'ortogonale* della giacitura del piano tangente.

Ovvero:

la giacitura del $\underline{\text{piano } \textit{tangente ad un luogo di zeri } \{x : f(x) - f(p) = \vec{0}_{\mathbf{R}^m}\} \text{ in } p}$
è il $\underline{\textit{luogo di zeri del differenziale}\text{: Ker}Jf(p).}$

Dimostrazione corollario: - poichè Z localmente coincide con un grafico di una funzione di $(M-1)-$variabili a valori reali, ha piano tangente $M-1$ dimensionale in p.
- - Per il teorema 3 la giacitura di tale spazio tangente sono le velocità $v = (v_1, \ldots, v_M)$ dei cammini in Z passanti per p.
- - D'altronde tali velocità devono soddisfare la condizione $\langle v \cdot \nabla f(p)\rangle = 0$.
- - Poichè $\nabla f(p) \neq \vec{0}_{\mathbf{R}^M}$ tale relazione definisce un sottospazio $M-1$ dimensionale che quindi coincide con la giacitura dello spazio tangente, che a priori è un suo sottospazio di egual dimensione.
- - Traslando in p si ha che l'equazione dello spazio tangente è $\langle (x-p) \cdot \nabla f(p)\rangle = 0$.
- Il caso di f a valori in $\mathbf{R}^m$, $m < M$, con $\nabla f(p)$ di rango massimo m, è del tutto analogo. Infatti il sistema definisce un sottospazio $M - m$ dimensionale che contiene il tangente anch'esso $M - m$ dimensionale.

Osservazione 17: - l'insieme di livello $Z = \{x : f(x) = f(p)\}$ è l'intersezione degli insiemi di livello $Z_i = \{x : f_i(x) = f_i(p)\}$, $1 \leq i \leq m$;
- il suo tangente $M - m$ dimensionale in p è l'intersezione dei tangenti $M - 1$ dimensionali in p agli m luoghi di zeri Z_i, $1 \leq i \leq m$.

Osservazione 18: in realtà il *teorema del Dini delle funzioni implicite*, garantisce che, nelle ipotesi fatte su f (che sia C^1 in un intorno di p, con $\nabla f(p) \neq \vec{0}$), il livello Z è effettivamente in un intorno del punto p grafico di una funzione.
Quindi è ben definito il piano tangente ($(M-1)$-dimensionale). Coincidendo la sua giacitura con le velocità dei cammini in Z e passanti per p, $\nabla f(p)$ è ortogonale ad essa.
Osservazione 19: più in generale il teorema delle funzioni implicite garantisce che il luogo di zeri Z di una funzione vettoriale $f : D \subseteq \mathbf{R}^M \to \mathbf{R}^m$, $m < M$, C^1 , con l'ipotesi $\nabla f(p)$ di *rango massimo* m, ha intersezione con un intorno di p che è grafico di una funzione differenziabile di $(M-m)$-variabili a valori in qualche sottospazio di $\mathbf{R}^M$ di dimensione m.
Quindi $\nabla f_1(p), \ldots, \nabla f_m(p)$ dà una *base* per *l'ortogonale* alla giacitura $M - m$ dimensionale del tangente a Z in p.
Quindi nelle ipotesi fatte la giacitura del tangente ad un luogo di zeri in p è il luogo di zeri del differenziale $\mathrm{Ker} Jf(p)$

Un'altra importante proprietà, quando non nullo, del gradiente in un punto p, come vettore del dominio, è quella di individuare in p la direzione di massima crescita della funzione:
Massima pendenza necessaria. Se $f : D \subseteq \mathbf{R}^M \to \mathbf{R}$ è differenziabile in p e $\nabla f(p) \neq \vec{0}_{\mathbf{R}^M}$ allora: $\dfrac{\nabla f(p)}{|\nabla f(p)|_M}$ è la direzione di massima crescita di f in p. Ovvero per ogni v, $|v|_M = 1$:

$$\frac{\partial f}{\partial v}(p) \leq |\nabla f(p)|_M = \left\langle \frac{\nabla f(p)}{|\nabla f(p)|_M} \cdot \nabla f(p) \right\rangle = \frac{\partial f}{\partial \frac{\nabla f(p)}{|\nabla f(p)|_M}}(p).$$

Dimostrazione: per differenziabilità (cfr. teorema 2-5): $\dfrac{\partial f}{\partial v}(p) = \langle v \cdot \nabla f(p) \rangle$, quindi se $|v|_M = 1$ per la diseguaglianza di Cauchy-Schwarz (cfr. FT 2) $\dfrac{\partial f}{\partial v}(p) \leq |\nabla f(p)|_M$.

Vale in un certo senso il viceversa, di dimostrazione più impegnativa , che dà una condizione sufficiente per la differenziabilità in un punto p usando le derivate direzionali solo nel punto p:

Massima pendenza sufficiente. Sia f è Lipschitziana di costante L: $\dfrac{|f(x) - f(y)|}{|x - y|_M} \leq L$.
Se esiste v, $|v|_M = 1$ per cui esiste $\dfrac{\partial f}{\partial v}(p) = L$, allora: f è differenziabile in p e $\nabla f(p) = Lv$.

2.7 Tangente ad immagini

Osservazione 20: - oltre che come preimmagini (luoghi di zeri o insiemi di livello) di funzioni conviene spesso descrivere un sottoinsieme di $\mathbf{R}^m$ come *immagine* di una funzione vettoriale. Per esempio il segmento in $\mathbf{R}^3$ di estremi $(1,2,3)$ ed $(4,5,6)$ può essere descritto come immagine della funzione lineare affine $(1,2,3)+t((4,5,6)-1,2,3))=(1,2,3)+t(3,3,3)$, ristretta ai $t \in [0;1]$. O più in generale una "curva materiale" può essere vista come immagine (sostegno) di un cammino.
- Per tali insiemi visti come immagine, come per i luoghi di zeri, è utile individuare in termini delle funzioni in gioco e delle loro derivate parziali i piani tangenti e ortogonali all'insieme.

Jacobiano nel codomino: per $m > M$, se $f=(f_1,\dots,f_m): D \subseteq \mathbf{R}^M \to \mathbf{R}^m$ è differenziabile in p, con $Jf(p)$ di rango massimo M,
per ogni $\Gamma:(a;b)\to D$ con $\Gamma(c)=p$, $\Gamma'(c)=v \neq \vec{0}_{\mathbf{R}^M}$ si ha che

$$\frac{\partial f}{\partial v}(p) \neq \vec{0}_{\mathbf{R}^m} \text{ è vettore tangente al cammino } t \mapsto f(\Gamma(t)) \text{ in } f(p).$$

Dimostrazione: è lo stesso argomento usato nel teorema 3, per veder i vettori tangenti a un grafico, basato sulla regola della catena per la composizione con cammini (proposizione 2). Si ha infatti

$$(f\circ\Gamma)'(c)=\Gamma_1(c)\frac{\partial f}{\partial x_1}(p)+\cdots+\Gamma_M(c)\frac{\partial f}{\partial x_M}(p)=\frac{\partial f}{\partial v}(p)=Jf(p)\Gamma'(c)\neq \vec{0}_{\mathbf{R}^m} \text{ per iniettività.}$$

Piano tangente ad un immagine: per $m > M$, sia $f=(f_1,\dots,f_m): D \subseteq \mathbf{R}^M \to \mathbf{R}^m$ differenziabile in p, con $Jf(p)$ di rango massimo M.
Se l'immagine di f intersecata un intorno di $f(p)$ è il grafico di una funzione differenziabile di M variabili a valori in qualche sottospazio di $\mathbf{R}^m$ di dimensione $m-M$, *essendo ben definito il piano tangente* M dimensionale ad $\mathrm{Im}f$ in $f(p)$, si ha che una *base della sua giacitura* è data dalle M derivate parziali in p di f:

$$\frac{\partial f}{\partial x_1}(p),\dots,\frac{\partial f}{\partial x_M}(p)$$

ovvero dalle colonne di $Jf(p)$ ovvero dalle righe di $\nabla f(p)$.
Ovvero

la giacitura del piano *tangente all' immagine* $\mathrm{Im}f$ in $f(p)$
è *l'immagine del differenziale in p*: $\mathrm{Im}Jf(p)$.

Quindi la forma parametrica del piano tangente all'immagine è:

$$f(p)\ +\ s_1\frac{\partial f}{\partial x_1}(p)+\cdots+s_M\frac{\partial f}{\partial x_M}(p)=Jf(p)s,\quad s=(s_1,\dots,s_M)\in\mathbf{R}^M$$

Dimostrazione: - Essendo p interno a D vi è $r>0$ per cui $\Gamma^{(j)}(t)=p+te_j$, $|t|<r$, $1\le j\le M$, sono a valori in D: quindi $f(\Gamma^{(j)}(t))$ sono ben definiti e $f(\Gamma^{(j)}(0))=p$, $\left(f\circ\Gamma^{(j)}\right)'(0)=\dfrac{\partial f}{\partial x_j}(p)$.
- Per la proposizione 3 tali vettori velocità, e lo spazio da essi generato, sono contenuti nella giacitura dello spazio tangente in $f(p)$ ad $\mathrm{Im}\,f$. Avendo, per ipotesi, essa dimensione M, viene a coicindere con lo spazio generato dalle derivate parziali in p: infatti
- per ipotesi gli M vettori $\frac{\partial f}{\partial x_1}(p),\dots,\frac{\partial f}{\partial x_M}(p)\in\mathbf{R}^m$ sono lineramente indipendenti e generano un sottospazio di $\mathbf{R}^m$ di dimensione M.

Osservazione 21: per le immagini invece del teorema del Dini vi è il *teorema del rango*: se $f: D\subseteq\mathbf{R}^M\to\mathbf{R}^m$, $M<m$, è differenziabile in p, con $Jf(p)$ di rango massimo M, allora f *ristretta ad un intorno U di* p ha immagine che è il grafico di una funzione differenziabile di M variabili a valori in qualche sottospazio di $\mathbf{R}^m$ di dimensione $m-M$, quindi *è ben definito il piano tangente* M dimensionale ad $\mathrm{Im}_U f$ in $f(p)$, e una *base della sua giaciutura* è quindi data dalle M derivate parziali in p di f.

2.8 Il differenziale del determinante

- **Matrice dei cofattori, matrice aggiunta** Se matrice A è una matrice quadrata $m \times m$, si definiscono: la matrice $cofA$ dei *cofattori* $(cofA)^k_h = (-1)^{h+k} detA^{\cancel{k}}_{\cancel{h}}$
e la sua trasposta, la matrice *aggiunta* $adgA$, $(adjA)^k_h = (-1)^{h+k} detA^{\cancel{h}}_{\cancel{k}}$.
Se A è invertibile $(detA)A^{-1} = adjA$.

- **Prodotti scalari tra matrici** identificando lo spazio $\mathcal{M}_{h\times m}$ delle matrici $h \times m$ con $\mathbf{R}^{hm}$, *e.g.* allineando verticalmente le colonne per ottenere vettori colonna $A = (A^j_i) \mapsto {}^t(A^1_1, \ldots, A^1_h, A^2_1, \ldots, A^m_1, \ldots, A^m_h)$, e orizzontalemente le righe per ottenere vettori riga $A \mapsto (A^1_1, \ldots, A^m_1, A^1_2, \ldots, A^1_h, \ldots, A^m_h)$, il prodotto scalare in $\mathbf{R}^{hm}$ diventa un prodotto scalare $\cdot_{\mathcal{M}}$ tra matrici $h \times m$: $\langle A \cdot B\rangle_{\mathbf{R}^{hm}} = \sum_i \sum_r A^i_r B^i_r = tr\,{}^tAB =: A \cdot_{\mathcal{M}} B$
corrispondente al prodotto righe per colonne

$$\begin{aligned} tr\,{}^tAB &= (({}^tA)^1_1, \ldots, ({}^tA)^h_1, ({}^tA)^1_2, \ldots, ({}^tA)^1_m, \ldots, ({}^tA)^h_m)\ {}^t(B^1_1, \ldots, B^1_h, B^2_1, \ldots, B^m_1, \ldots, B^m_h) = \\ &= (A^1_1, \ldots, A^1_h, A^2_1, \ldots, A^m_1, \ldots, A^m_h)\ {}^t(B^1_1, \ldots, B^1_h, B^2_1, \ldots, B^m_1, \ldots, B^m_h) \end{aligned}.$$

- **Derivate del determinante di un cammino di matrici:** sia $A : I \to \mathcal{M}_{m\times m}$ cammino di matrici, con $A^j_i(t)$ derivabili. Essendo $\det A(t)$ un polinomio nelle funzione coefficienti, è derivabile e: $$(detA)' = cofA \cdot_{\mathcal{M}} A' = tr(adjA\ A').$$

Differenziale del determinante di una matrice La funzione polinomiale

$$\det : \mathcal{M}_{m\times m} \to \mathbf{R},\ A = \{A^j_i\},\ \det(A) = \sum_{j_p \neq j_q \text{ se } p\neq q,\ 1}^{m} \text{segno}(j_1, \ldots, j_m) A^{j_1}_1 \cdot \cdots \cdot A^{j_m}_m$$

è differenziabile e, considerando le matrici vettori colonna $H \sim {}^t(H^1_1, \ldots, H^1_m, \ldots, H^m_1, \ldots, H^m_m)$:

$$\begin{aligned} (J\det(A))\, H &= tr(adjAH) = cofA \cdot_{\mathcal{M}} H = \nabla\det(A) \cdot_{\mathcal{M}} H \\ (J\det(A)) &= adjA = ((adjA)^1_1, \ldots, (adjA)^m_1, \ldots, (adjA)^1_m, \ldots, (adjA)^1_1) \text{ come riga,} \\ (\nabla\det(A)) &= cofA = {}^t((cofA)^1_1, \ldots, (cofA)^m_1, \ldots, (cofA)^1_m, \ldots, (cofA)^1_1) \text{ come colonna} \\ \frac{\partial\det}{\partial A^j_i}(A) &= (cofA)^j_i = (adjA)^i_j. \end{aligned}$$

Dimostrazione: - per la differenziabilità dei polinomi in più variabili conviene usare il teorema del differenziale totale. Per una dimostrazione induttiva usando solo la defінzione di differenziabilità, basta mostrare che il prodotto di due funzioni coordinate è differenziabile:
$(x_i + h_i)(x_j + h_j) = x_ix_j + x_ih_j + h_ix_j + h_ih_j$, $x_ih_j + h_ix_j$ è lineare in h, e $h_ih_j = o(|h|_m)$.
- Si usa il corollario alla proposizione 2, componendo det, con un cammino di matrici, e la formula della derivata del determinante di una cammino di matrici.
Alternativa: la matrice jacobiana $1 \times m^2$ (riga) di det, e la sua trasposta, matrice gradiente $m^2 \times 1$ (colonna), sono determinate dalle derivate parziali rispetto alla base canonica di $\mathbf{R}^{m^2}$, corrisponenti alle m^2 matrici $B(i,j)$, $m \times m$, con solo la componente $B(i,j)^j_i$ eguale ad 1 e le altre nulle, ovvero con solo la colonna j^a non nulla uguale a $e^i_{\mathbf{R}^m}$ (base canonica di $\mathbf{R}^m$):

$$\begin{aligned} \frac{\partial\det}{\partial A^j_i}(A) &= \frac{d}{dt}\left(\det(A + tB(i,j))\right)_{t=0} = tr(adjAB(i,j)) = tr(\vec{0}|\ldots|\vec{0}|(adjA)^i{}_{\text{colonna } j^a}|\vec{0}|\ldots|\vec{0}) = \\ &= (adjA)^i_j \end{aligned}$$

<u>Esercizio:</u> sia $(F_1, \ldots, F_m) = F = F(t,x) : \mathbf{R}\times\mathbf{R}^m \to \mathbf{R}^m$, differenziabile, $F_i \in C^2(\mathbf{R}\times\mathbf{R}^m)$, $1 \le i \le m$, con $F(0,x) = x$. Posto $v(x) =: \dfrac{\partial F}{\partial t}(0,x) = (v_1(x), \ldots, v_m(x))$, $J_xF(t,x)$ la matrice jacobiana delle derivate rispetto alle x, si calcoli $\dfrac{\partial\det J_xF}{\partial t}(0,x)$ in termini delle derivate parziali prime di v.

2.9 Differenziale e gradiente tangenziali 1:

Notazioni per le matrici: 1- Se A è una matrice $h \times k$ con A_i^j si indica la sua componente di i-esima riga e j-esima colonna: considerando gli elementi di $\mathbf{R}^h$ e di $\mathbf{R}^k$ come colonne, ed indicando le rispettive basi canoniche con $\mathbf{e}_i^{\mathbf{R}^h}$ e $\mathbf{e}_j^{\mathbf{R}^k}$, si ha $A_i^j = {}^t\mathbf{e}_i^{\mathbf{R}^h} A \mathbf{e}_j^{\mathbf{R}^k}$.

2- Con A_i si indica la i-esima riga, con A^j la j-esima colonna: $A_i = {}^t\mathbf{e}_i^{\mathbf{R}^h} A$, $A^j = A\mathbf{e}_j^{\mathbf{R}^k}$.

3- Siano dati: A matrice $h \times k$, due *multi-indici crescenti* di numeri naturali non nulli: $I = (i_1, \dots, i_r)$, $r \le h$, $1 \le i_1 < \dots < i_r \le h$ e, rispettivamente, $J = (j_1, \dots, j_s)$, $s \le k$, $1 \le j_1 < \dots < j_s \le k$.

- - Con A_I o $A_{i_1,\dots,i_r}$ rispettivamente A^J o $A^{j_1,\dots,j_s}$, si indicano le matrici $r \times k$ ed $h \times s$ ottenute selezionando le righe $i_1 \dots i_r$ rispettivamente le colonne $j_1 \dots j_s$;

- - con A_I^J o $A_{i_1,\dots,i_r}^{j_1,\dots,j_s}$ la matrice $r \times s$ ottenuta selezionando sia le r righe che le s colonne;

- - con $A_{\not I}$ o $A_{\not i_1,\dots,\not i_r}$, $A^{\not J}$ o $A^{\not j_1,\dots,\not j_s}$, $A_{\not I}^{\not J}$ o $A_{\not i_1,\dots,\not i_r}^{\not j_1,\dots,\not j_s}$ le matrici $(h-r) \times k$, $h \times (k-s)$, $(h-r) \times (k-s)$ ottenute nei vari casi scartando le righe o le colonne degli indici barrati.

- - Quindi per una matrice quadrata A invertibile si ha $\left[A^{-1}\right]_i^j = (-1)^{i+j} \dfrac{\det A_{\not j}^{\not i}}{\det A}$.

4- Per la sostituzione con una colonna v, della colonna j^a di una matrice A, $h \times k$ si usa la notazione $A\,[v/A^j]$, o $(\dots|A^{j-1}|V|A^{j+1}\dots)$. Analogamente per le righe.

- Quindi la regola di Cramer per le soluzioni u del sistema lineare $Au = v$ $h \times h$, ovvero $u_i = \displaystyle\sum_{j=1}^{h} (-1)^{i+j} v_j \det A_{\not j}^{\not i} (\det A)^{-1}$ diventa, essendo la sommatoria lo sviluppo del determinante per la i^a colonna della matrice $A[v/A^i]$:

$u_i = \dfrac{\det A[v/A^i]}{\det \mathrm{A}}$. Per un sistema matriciale $AU = V$, $U = A^{-1}V$: $U_i^j = \dfrac{\det A[V^j/A^i]}{\det A}$.

Notazione: - usandosi spesso le notazioni $z = f(x)$, ovvero $\begin{cases} z_1 = f_1(x_1, \dots, x_M) \\ \quad\vdots \\ z_m = f_m(x_1, \dots, x_M) \end{cases}$, per indicare una funzione $f : A \subseteq \mathbf{R}^M \to \mathbf{R}^m$, e $\dfrac{dz}{dx}$, per la derivata prima di f rispetto a x, nel caso in cui f sia di una variabile reale, è suggestivo usare per la matrice Jacobiana, similmente per la matrice gradiente, e le loro sottomatrici, le notazioni: $Jf(p) = \dfrac{\partial z}{\partial x} = \dfrac{\partial(z_1, \dots, z_m)}{\partial(x_1, \dots, x_M)}$,

$$(Jf(p))_{i_1 \dots i_h}^{j_1 \dots j_k} = \frac{\partial z_{i_1 \dots i_h}}{\partial x_{j_1 \dots j_k}} \ , \quad (Jf(p))_{\not i_1 \dots \not i_h}^{\not j_1 \dots \not j_k} = \frac{\partial z_{\not i_1 \dots \not i_h}}{\partial x_{\not j_1 \dots \not j_k}} = \frac{\partial(z_{\not i_1}, \dots, z_{\not i_h})}{\partial(x_{\not j_1}, \dots, x_{\not j_1})},$$

in breve per $I = \{i_1 < \dots < i_h]\}$, $J = \{j_1 < \dots < j_k\}$: $(Jf(p))_I^J = \dfrac{\partial z_I}{\partial x_J}$, $\quad (Jf(p))_{\not I}^{\not J} = \dfrac{\partial z_{\not I}}{\partial x_{\not J}}$.

- Selezionando le righe $I = i_1 < \dots < i_h$ non si considerano tutte le funzioni componenti di f, ma solo le proiezioni di f sugli h assi coordinati scelti I: $(Jf(p))_I = Jf_I(p)$.

- La selezione delle colonne con $J = j_1 < \dots < j_k$ piuttosto significa che si sta considerando

- - la matrice associata alla restrizione di $D_p f$ al sottospazio generato da $e_{j_1}, \dots, e_{j_k}$,

- - o meglio, *coerentemente alla definizione di derivata parziale*, lo Jacobiano della funzione di k variabili ottenuta *componendo* f con la parametrizzazione del sottospazio affine $p + x_{j_1} e_{j_1} + \dots + x_{j_k} e_{j_k}$: $(x_{j_1}, \dots, x_{j_k}) \mapsto f(p + x_{j_1} e_{j_1} + \dots + x_{j_k} e_{j_k})$.

Differenziale tangenziale 1: - Sia W un sottospazio vettoriale di dimensione k di $\mathbf{R}^M$: $f : A = \overset{o}{A} \subseteq \mathbf{R}^M \to \mathbf{R}^m$ si dice *differenziabile in* $p \in A$ *lungo* W se la sua restrizione $f_{|W}$ al sottospazio affine $p + W$, munito della norma indotta da quella di $\mathbf{R}^M$, è differenziabile in p.
- Il suo differenziale si dice differenziale tangenziale di f in p lungo W.
Si indica con $D^W f(p) : W \to \mathbf{R}^m$, o con $\dfrac{\partial f}{\partial W}(p)$. Se $W = \{tw\}_{t\in\mathbf{R}}$, $|w|_M = 1$: $\dfrac{\partial f}{\partial W} \sim \dfrac{\partial f}{\partial w}$.
- Se poi f è differenziabile in p, il suo differenziale lungo W è la restrizione a W di $D_p f$.

Osservazione 22: se una funzione g coincide con f su $U \cap W$, ove U è un intorno di p, allora anch'essa è differenziabile in p lungo W e $D^W_p g = D^W_p f$.

Osservazione 23: - se non si introduce una base su W *non ha senso parlare di matrice Jacobiana di* f *o di gradiente* lungo W. Può essere utile, nella pratica, assumendo che f abbia un'estensione ad un intorno in $\mathbf{R}^M$ di p differenziabile in p, riferirsi alle coordinate dello spazio $\mathbf{R}^M$ ambiente, ed usando P^W *la proiezione ortogonale su* W, $({}^tP = P\ PP = P, P_{|W} = Id_W)$ definire due matrici, rispettivamente $m \times M$ e $M \times m$, come segue:

Matrice Jacobiana ambiente e gradiente ambiente: sono rispettivamente:

$$J^W f(p) =: Jf(p)P^W \sim D_p f P^W \quad (D_p f P^W_{|W} = D_p f_{|W} = D^W_p f)\ , \quad \nabla^W f(p) =: P^W \nabla f(p).$$

Osservazione 24: - data una base $\beta = \{w_1, \dots, w_k\} \subset \mathbf{R}^M$ di W, lo si identifica con $\mathbf{R}^k$.
- Si ha che f è differenziabile in p lungo W se la funzione $f^\beta \sim f_{|W}$ da $\mathbf{R}^k$ in $\mathbf{R}^m$ data da $p + s_1 w_1 + \cdots + s_k w_k \mapsto f(p + s_1 w_1 + \cdots + s_k w_k)$ è differenziabile in $\vec{0}_{\mathbf{R}^k}$.
- La matrice Jacobiana $J_{0_{\mathbf{R}^k}} f^\beta$ di f^β è la matrice Jacobiana $J^\beta_p f$ associata nella base β a $D^W f(p)$. È una matrice $m \times k$.
- La relazione tra $J^\beta_p f$ e $D^W_p f$ è la seguente: dato $w = s_1 w_1 + \cdots + s_k w_k$, $s = {}^t(s_1, \dots, s_k)$
$J^\beta_p f s = f^\beta(s) - f^\beta(0_{\mathbf{R}^k}) + o(|s|_k) = f(p+w) - f(p) - o(|w|_M) = D^W_p f w = D^W_p f(w_1|\dots|w_k)s$,

quindi
$$J^\beta f(p) = D^W_p f(w_1|\dots|w_k) = \big(D^W_p f w_1|\dots|D^W_p f w_k\big), \text{ matrice } m \times k.$$

- Se poi β è *ortonormale* per il prodotto scalare su W dato da quello di $\mathbf{R}^M$ ha senso parlare di

gradiente tangenziale $\nabla^\beta f(p) = \nabla f^\beta(0_{\mathbf{R}^k}) = \begin{pmatrix} {}^tD^W_p f w_1 \\ \vdots \\ {}^tD^W_p f w_k \end{pmatrix} = \begin{pmatrix} {}^tw_1 \\ \vdots \\ {}^tw_k \end{pmatrix} \nabla f(p)$, matrice $k \times m$.