UNIVERSITÀ DI SIENA 1240

Prof. Lonzi Marco

Dispense per il Corso di

METODI QUANTITATIVI PER LE APPLICAZIONI ECONOMICHE E FINANZIARIE

MATEMATICA PER LE APPLICAZIONI ECONOMICHE E FINANZIARIE

Volume 2

Calcolo differenziale per funzioni vettoriali di variabile vettoriale

AA. 2022/23

CALCOLO DIFFERENZIALE PER FUNZIONI VETTORIALI DI VARIABILE VETTORIALE

Sia $\mathbb{R}^n$ lo spazio vettoriale di dimensione n, avente per elementi tutte le possibili n-uple di numeri reali, o vettori, risultante dal prodotto cartesiano dell'insieme $\mathbb{R}$ per sè stesso n volte.
Una qualsiasi funzione f avrà dominio e codominio contenuti in opportuni spazi vettoriali, $\mathbb{R}^n$ e $\mathbb{R}^m$, con $n \geq 1$ e $m \geq 1$; sia la variabile indipendente che quella dipendente potranno allora assumere valori reali oppure vettoriali. Diremo allora:
$f : \mathbb{R} \to \mathbb{R}$, funzione reale di variabile reale;
$f : \mathbb{R} \to \mathbb{R}^n$, funzione vettoriale di variabile reale;
$f : \mathbb{R}^n \to \mathbb{R}$, funzione reale di variabile vettoriale;
$f : \mathbb{R}^n \to \mathbb{R}^m$, funzione vettoriale di variabile vettoriale.

TOPOLOGIA IN $\mathbb{R}^n$
Un elemento $\mathbb{X} = (x_1, x_2, ..., x_n) \in \mathbb{R}^n$ viene detto anche punto o **vettore**.
Dati due punti $\mathbb{X}, \mathbb{Y} \in \mathbb{R}^n$, $\mathbb{X} = (x_1, x_2, ..., x_n)$, $\mathbb{Y} = (y_1, y_2, ..., y_n)$, abbiamo la:
Definizione 1 : Dicesi **distanza** (Euclidea) tra $\mathbb{X}$ e $\mathbb{Y}$ il modulo della loro differenza:

$$\text{d}(\mathbb{X},\mathbb{Y}) = \|\mathbb{X} - \mathbb{Y}\| = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}\,.$$

Molte sono le distanze che si possono definire in $\mathbb{R}^n$; qui useremo solo quella Euclidea.
Definizione 2 : Dicesi **intorno** del punto $\mathbb{X}_0 \in \mathbb{R}^n$ di ampiezza ε l'insieme:
$\mathfrak{J}(\mathbb{X}_0, \varepsilon) = \{\mathbb{X} \in \mathbb{R}^n : \text{d}(\mathbb{X}, \mathbb{X}_0) < \varepsilon\} = \{\mathbb{X} \in \mathbb{R}^n : \|\mathbb{X} - \mathbb{X}_0\| < \varepsilon\}$.
Un intorno $\mathfrak{J}(\mathbb{X}_0, \varepsilon)$ in $\mathbb{R}^2$ è costituito dai punti interni ad un cerchio di centro $\mathbb{X}_0$ e raggio ε; un intorno $\mathfrak{J}(\mathbb{X}_0, \varepsilon)$ in $\mathbb{R}^3$ è costituito dai punti interni ad una sfera di centro $\mathbb{X}_0$ e raggio ε.
Le definizioni topologiche dei vari tipi di punto in $\mathbb{R}^n$ sono analoghe a quelle date in $\mathbb{R}$; preso un punto $\mathbb{X}_0$ ed un insieme $\mathbb{A} \subset \mathbb{R}^n$ si hanno le seguenti:
Definizione 3 : Il punto $\mathbb{X}_0$ è detto di **accumulazione** per l'insieme $\mathbb{A}$ se un qualunque intorno di $\mathbb{X}_0$ ha intersezione con $\mathbb{A}$ non vuota e diversa dal solo punto $\mathbb{X}_0$, ovvero se:
$\forall\, \varepsilon > 0 : \Big\{\mathfrak{J}(\mathbb{X}_0, \varepsilon)/\{\mathbb{X}_0\}\Big\} \cap \mathbb{A} \neq \emptyset$;
Definizione 4 : Il punto $\mathbb{X}_0 \in \mathbb{A}$ è detto **isolato** per l'insieme $\mathbb{A}$ se esiste almeno un intorno di $\mathbb{X}_0$ che non ha punti in comune con $\mathbb{A}$, eccettuato il punto $\mathbb{X}_0$, ovvero se:
$\exists\, \varepsilon > 0 :\ \mathfrak{J}(\mathbb{X}_0, \varepsilon) \cap \mathbb{A} = \{\mathbb{X}_0\}$.
Definizione 5 : Il punto $\mathbb{X}_0 \in \mathbb{A}$ è detto **interno** ad $\mathbb{A}$ se esiste almeno un intorno di $\mathbb{X}_0$ tutto contenuto in $\mathbb{A}$, ovvero se:
$\exists\, \varepsilon > 0 :\ \mathfrak{J}(\mathbb{X}_0, \varepsilon) \subset \mathbb{A}$.
Definizione 6 : Il punto $\mathbb{X}_0$ è detto **esterno** ad $\mathbb{A}$ se è interno a $\mathcal{C}(\mathbb{A})$, il complementare di $\mathbb{A}$.
Definizione 7 : Il punto $\mathbb{X}_0$ è detto di **frontiera** per $\mathbb{A}$ se qualunque intorno di $\mathbb{X}_0$ ha intersezione non vuota sia con $\mathbb{A}$ che con $\mathcal{C}(\mathbb{A})$, ovvero se:
$\forall\, \varepsilon > 0 :\ \mathfrak{J}(\mathbb{X}_0, \varepsilon) \cap \mathbb{A} \neq \emptyset$ **e** $\mathfrak{J}(\mathbb{X}_0, \varepsilon) \cap \mathcal{C}(\mathbb{A}) \neq \emptyset$.
Per essere punto isolato o punto interno per $\mathbb{A}$, $\mathbb{X}_0$ deve appartenere ad $\mathbb{A}$; ciò non è invece richiesto per essere punto d'accumulazione o punto di frontiera.
Se un punto è isolato per $\mathbb{A}$ allora esso è di frontiera per $\mathbb{A}$; se un punto è interno ad $\mathbb{A}$, allora è di accumulazione per $\mathbb{A}$.
Dalle definizioni topologiche di punto seguono quelle per gli insiemi:
Definizione 8 : Un insieme $\mathbb{A} \subset \mathbb{R}^n$ si dice **aperto** se tutti i suoi punti sono interni.
Quindi un insieme è aperto se nessuno dei suoi punti di frontiera gli appartiene.
Definizione 9 : Un insieme $\mathbb{A} \subset \mathbb{R}^n$ si dice **chiuso** se il suo complementare è aperto.

Un insieme $\mathbb{A} \subset \mathbb{R}^n$ è quindi chiuso se gli appartengono tutti i suoi punti di frontiera (o se gli appartengono tutti i suoi punti d'accumulazione).
Un **intervallo** in $\mathbb{R}^n$ è dato dal prodotto cartesiano di n intervalli di $\mathbb{R}$.
Se tutti gli intervalli sono chiusi : $\left[x_i^a, x_i^b\right]$, l'intervallo $\prod_{i=1}^{n}\left[x_i^a, x_i^b\right]$ sarà chiuso, se tutti gli intervalli sono aperti : $\left]x_i^a, x_i^b\right[$, l'intervallo $\prod_{i=1}^{n}\left]x_i^a, x_i^b\right[$ sarà aperto.
Definizione 10 : Un insieme $\mathbb{A} \subset \mathbb{R}^n$ si dice **limitato** se esiste un intorno $\mathfrak{J}(\mathbb{X}_0, \varepsilon)$, di centro un punto opportuno $\mathbb{X}_0$ e di ampiezza opportuna ε, tale che $\mathbb{A} \subset \mathfrak{J}(\mathbb{X}_0, \varepsilon)$.
Definizione 11 : Un insieme $\mathbb{A} \subset \mathbb{R}^n$ che sia chiuso e limitato si dice **compatto**.

FUNZIONI VETTORIALI DI VARIABILE REALE $f : \mathbb{R} \to \mathbb{R}^n$
Si consideri un vettore $\mathbb{X} \in \mathbb{R}^n$, ciascuna delle cui componenti sia una funzione $\mathbb{R} \to \mathbb{R}$ della variabile reale t.
Scriviamo $\mathbb{X}(t) = (x_1(t), x_2(t), ..., x_n(t))$ oppure $f(t) = (f_1(t), f_2(t), ..., f_n(t))$ per indicare una tale funzione. Ciascuna delle funzioni $x_i(t) = f_i(t)$ è una funzione $f_i : \mathbb{R} \to \mathbb{R}$, $1 \leq i \leq n$. La funzione $f : t \to \mathbb{X}(t)$ è detta **funzione vettoriale di variabile reale**, in quanto associa alla variabile reale t come immagine il vettore $\mathbb{X}(t) \in \mathbb{R}^n$. Una funzione di questo tipo è detta anche **curva** in $\mathbb{R}^n$. Per questo tipo di funzioni il **grafico** coincide con il codominio. La linea che si determina è detta anche **sostegno** della curva. Il campo d'esistenza è dato dall'intersezione dei campi d'esistenza delle n funzioni $x_i(t) = f_i(t)$.

Esempio 1 : La funzione $f_1 : \mathbb{R} \to \mathbb{R}^2, t \to (\cos t, \operatorname{sen} t)$, ha per grafico la circonferenza di equazione cartesiana $x^2 + y^2 = 1$, di centro l'origine e raggio pari a 1.
La funzione $f_2 : \left[0, 2\pi\right[\to \mathbb{R}^2, t \to (\cos t, \operatorname{sen} t)$, ha per grafico la stessa circonferenza, percorsa una sola volta in senso antiorario partendo, per $t = 0$, dal punto $(1, 0)$, mentre la f_1 percorre la circonferenza infinite volte, in senso antiorario, quando t assume valori da $-\infty$ a $+\infty$.
La funzione $f_3 : \left[0, 2\pi\right[\to \mathbb{R}^2, t \to (\operatorname{sen} t, \cos t)$, ha per grafico la solita circonferenza, percorsa una volta in senso orario partendo, per $t = 0$, dal punto $(0, 1)$.

Come si vede da questo esempio, non basta descrivere geometricamente i punti di un grafico, ma occorre anche vedere come, quante volte e in quale verso tale grafico venga percorso.

Esempio 2 : La funzione $f_1 : \mathbb{R} \to \mathbb{R}^2, t \to \left(t, t^2\right)$, ha per grafico la parabola $y = x^2$; la funzione $f_2 : \mathbb{R} \to \mathbb{R}^2, t \to \left(t^2, t^4\right)$, ha per grafico la sola parte destra della parabola $y = x^2$, percorsa due volte, da destra a sinistra per $t < 0$ e da sinistra verso destra per $t \geq 0$. La funzione $f_3 : \mathbb{R} \to \mathbb{R}^2, t \to \left(\operatorname{sen} t, \operatorname{sen}^2 t\right)$, ha per grafico la parte della parabola $y = x^2$ compresa tra il punto $(-1, 1)$ e il punto $(1, 1)$, percorsa oscillando infinite volte tra i punti estremi.

Esempio 3 : La funzione $f : \mathbb{R} \to \mathbb{R}^3, t \to (\cos t, \operatorname{sen} t, t)$, ha per grafico un'elica che si avvolge, salendo se $t \geq 0$, scendendo se $t < 0$, lungo l'asse della terza variabile.

Supponiamo di avere una funzione in forma cartesiana $f : \mathbb{R} \to \mathbb{R}, x \to y = f(x)$; se esiste una funzione $F : \mathbb{R} \to \mathbb{R}^2, t \to (x(t), y(t))$, tale che $y(t) = f(x(t)), \forall\, t \in D_f$, si dice che $F(t)$ è una **rappresentazione in forma parametrica** di $y = f(x)$.

Se abbiamo una funzione $F : \mathbb{R} \to \mathbb{R}^2, t \to (x(t), y(t))$, posto $x = g(t)$ e $y = h(t)$, se $g(t)$ risulta invertibile, dalla $x = g(t)$ otteniamo $t = g^{-1}(x)$, e quindi $y = h\big(g^{-1}(x)\big) = f(x)$, ovvero si ottiene la rappresentazione cartesiana della funzione.

Esempio 4 : La funzione $f : \mathbb{R} \to \mathbb{R}^2, t \to (x_0 + \rho\cos t, y_0 + \rho \operatorname{sen} t)$ è una rappresentazione parametrica della circonferenza di centro (x_0, y_0) e raggio $\rho > 0$.

Esempio 5 : La funzione $f : \mathbb{R} \to \mathbb{R}^2, t \to (a\cos t, b \operatorname{sen} t) = (x(t), y(t)), a, b \in \mathbb{R}_+$ è una rappresentazione parametrica dell'ellisse di centro $(0, 0)$, assi paralleli agli assi coordinati, e semiassi a e b. Infatti : $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = \dfrac{(a\cos t)^2}{a^2} + \dfrac{(b \operatorname{sen} t)^2}{b^2} = 1$.
La funzione $f : \mathbb{R} \to \mathbb{R}^2, t \to (x_0 + a\cos t, y_0 + b \operatorname{sen} t)$ è invece una rappresentazione parametrica dell'ellisse di centro (x_0, y_0), assi paralleli agli assi x e y, e semiassi a e b.

LIMITI E CONTINUITA' PER FUNZIONI $f : \mathbb{R} \to \mathbb{R}^n$
Data $f : \mathbb{R} \to \mathbb{R}^n, t \to f(t) = (f_1(t), f_2(t), ..., f_n(t))$, sia t_0 punto di accumulazione per D_f.
Definizione 12 : Il $\lim\limits_{t\to t_0} f(t) = \lim\limits_{t\to t_0} (f_1(t), f_2(t), ..., f_n(t))$ è definito come:
$\lim\limits_{t\to t_0} f(t) = \Big(\lim\limits_{t\to t_0} f_1(t), \lim\limits_{t\to t_0} f_2(t), ..., \lim\limits_{t\to t_0} f_n(t)\Big)$,
ovvero il limite di un vettore è definito come il vettore avente per componenti i limiti, per $t \to t_0$, di ciascuna delle componenti di $f(t)$.
La funzione $f(t)$ ammette quindi limite per $t \to t_0$ se ciascuna delle sue componenti ammette limite per $t \to t_0$. Calcolare il limite di una funzione $f : \mathbb{R} \to \mathbb{R}^n$ consiste quindi nel calcolare n limiti di funzioni $\mathbb{R} \to \mathbb{R}$.
Definizione 13 : Diremo che $\lim\limits_{t\to t_0} f(t) = \mathfrak{L} \in \mathbb{R}^n$, con $t_0 \in \mathbb{R}$, se:
$\forall\, \varepsilon > 0\ \exists\, \delta(\varepsilon) : 0 < |t - t_0| < \delta(\varepsilon) \Rightarrow \|f(t) - \mathfrak{L}\| < \varepsilon$.
Definizione 14 : Diremo che $\lim\limits_{t\to -\infty} f(t) = \mathfrak{L} \in \mathbb{R}^n$ se:
$\forall\, \varepsilon > 0\ \exists\, \delta(\varepsilon) : t < \delta(\varepsilon) \Rightarrow \|f(t) - \mathfrak{L}\| < \varepsilon$.
Definizione 15 : Diremo che $\lim\limits_{t\to +\infty} f(t) = \mathfrak{L} \in \mathbb{R}^n$ se:
$\forall\, \varepsilon > 0\ \exists\, \delta(\varepsilon) : t > \delta(\varepsilon) \Rightarrow \|f(t) - \mathfrak{L}\| < \varepsilon$.
Non trattiamo qui il concetto di limite infinito.
Sia ora t_0 punto d'accumulazione per D_f, con $t_0 \in D_f$. Diamo allora la
Definizione 16 : La funzione $f : \mathbb{R} \to \mathbb{R}^n, t \to f(t) = (f_1(t), f_2(t), ..., f_n(t))$ si dice **continua** nel punto t_0 se $\lim\limits_{t\to t_0} f(t) = f(t_0)$.
Questa definizione equivale a richiedere che sia:
$\lim\limits_{t\to t_0} f_1(t) = f_1(t_0), \lim\limits_{t\to t_0} f_2(t) = f_2(t_0), ..., \lim\limits_{t\to t_0} f_n(t) = f_n(t_0)$, ovvero che ciascuna delle componenti $f_i(t)$ sia continua in t_0.
Avremo quindi, in forma metrica, che $f(t)$ è continua in t_0 se:
$\forall\, \varepsilon > 0\ \exists\, \delta(\varepsilon) : |t - t_0| < \delta(\varepsilon) \Rightarrow \|f(t) - f(t_0)\| < \varepsilon$.
Se t_0 è punto isolato per D_f, allora per definizione la funzione si dice continua in t_0.

DERIVATA PER FUNZIONI $f : \mathbb{R} \to \mathbb{R}^n$
Data $f : \mathbb{R} \to \mathbb{R}^n, t \to f(t) = (f_1(t), f_2(t), ..., f_n(t)) = \mathbb{X}(t) = (x_1(t), x_2(t), ..., x_n(t))$, e preso un punto $t_0 \in D_f$, definiamo la **derivata** della funzione $f(t)$ nel punto t_0. Vale la
Definizione 17 : Si dice che $f(t)$ è derivabile in t_0 se esiste finito il :
$\lim\limits_{t\to t_0} \dfrac{f(t) - f(t_0)}{t - t_0} = f'(t_0) = \mathbb{X}'(t_0)$.

Dicendo che tale limite esiste finito intendiamo affermare che $f'(t_0) = \mathbb{X}'(t_0) \in \mathbb{R}^n$.
Vista la definizione 12 di limite, dato che $\frac{1}{t - t_0} \in \mathbb{R}$ mentre $(f(t) - f(t_0)) \in \mathbb{R}^n$, risulta:

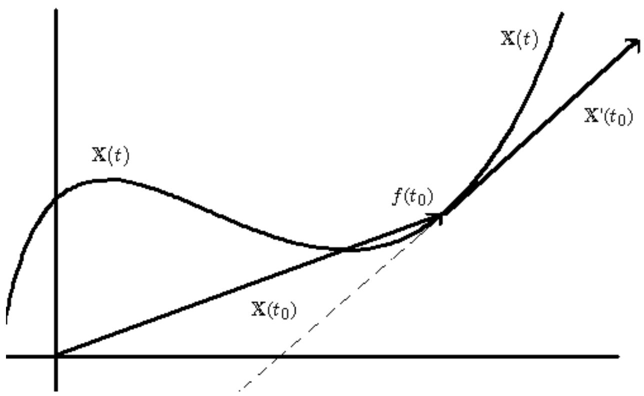
$$\lim_{t \to t_0} \frac{f(t) - f(t_0)}{t - t_0} = \lim_{t \to t_0} \frac{1}{t - t_0} \cdot (f(t) - f(t_0)) =$$
$$= \lim_{t \to t_0} \frac{1}{t - t_0} \cdot (f_1(t) - f_1(t_0), f_2(t) - f_2(t_0), ..., f_n(t) - f_n(t_0)) =$$
$$= \left(\lim_{t \to t_0} \frac{f_1(t) - f_1(t_0)}{t - t_0}, \lim_{t \to t_0} \frac{f_2(t) - f_2(t_0)}{t - t_0}, ..., \lim_{t \to t_0} \frac{f_n(t) - f_n(t_0)}{t - t_0} \right) =$$
$$= (f_1'(t_0), f_2'(t_0), ..., f_n'(t_0)) .$$

Quindi una funzione $f : \mathbb{R} \to \mathbb{R}^n$ risulta derivabile nel punto t_0 se ciascuna delle sue componenti $f_i(t)$ risulta derivabile nel punto t_0.
Dal punto di vista del calcolo pratico, il risultato trovato ci dice che la derivata di un vettore è ancora un vettore, avente per componenti le derivate delle componenti.

Esempio 6 : Data $f : \mathbb{R} \to \mathbb{R}^2, t \to (t, t^2)$, si ha : $f'(t) = (1, 2t)$.
Data $f : \mathbb{R} \to \mathbb{R}^2, t \to (t^2, t^4)$, si ha : $f'(t) = (2t, 4t^3)$.
Data $f : \mathbb{R} \to \mathbb{R}^3, t \to f(t) = (\cos t, \operatorname{sen} t, t)$, si ha : $f'(t) = (-\operatorname{sen} t, \cos t, 1)$.

Vale il
Teorema 1 : Se $f(t)$ è derivabile in t_0, allora $f(t)$ è continua in t_0.
Dimostrazione : Vediamo che $\lim_{t \to t_0} f(t) = f(t_0)$ ovvero che $\lim_{t \to t_0} (f(t) - f(t_0)) = \mathbb{O}$. Ma $\lim_{t \to t_0} (f(t) - f(t_0)) = \lim_{t \to t_0} (t - t_0) \cdot \frac{f(t) - f(t_0)}{t - t_0} = 0 \cdot f'(t_0) = \mathbb{O}$, ovvero la tesi.●

La definizione di funzione derivabile nel punto t_0 può essere scritta anche come: $\lim_{t \to t_0} \frac{\mathbb{X}(t) - \mathbb{X}(t_0)}{t - t_0} = \mathbb{X}'(t_0)$; il vettore $\mathbb{X}'(t_0) \in \mathbb{R}^n$ prende il nome di **vettore tangente** alla curva $\mathbb{X}(t)$ nel punto t_0. Tutto questo ha geometricamente senso purchè risulti $\mathbb{X}'(t_0) \neq \mathbb{O}$, vettore nullo, ovvero se non si annullano contemporaneamente in t_0 tutte le componenti di $\mathbb{X}'(t_0)$.



Come illustrato in figura nel caso di una $f : \mathbb{R} \to \mathbb{R}^2$, se vogliamo scrivere l'equazione della **retta tangente** alla curva $\mathbb{X}(t)$ nel punto $f(t_0) = \mathbb{X}(t_0)$, occorre anzitutto portarsi nel punto di tangenza con il vettore $\mathbb{X}(t_0)$ per poi, moltiplicando per lo scalare t, generare tutti i punti che stanno sulla retta avente la direzione di $\mathbb{X}'(t_0)$.
Otteniamo una funzione $r : \mathbb{R} \to \mathbb{R}^n, t \to r(t) = \mathbb{X}(t_0) + t \cdot \mathbb{X}'(t_0)$, ovvero una funzione $t \to (f_1(t_0), f_2(t_0), ..., f_n(t_0)) + t \cdot (f_1'(t_0), f_2'(t_0), ..., f_n'(t_0))$, che rappresenta appunto l'equazione della retta tangente in t_0.

Esempio 7 : Sia $f:\mathbb{R}\to\mathbb{R}^3, f(t)=\left(t^2,t^3,e^t\right)$ e sia $t_0=1$. Si ha $f(t_0)=(1,1,e)$, $f'(t)=\left(2t,3t^2,e^t\right)$ e quindi $f'(t_0)=(2,3,e)$. La funzione che esprime la retta tangente in t_0 è allora la seguente: $t\to(1,1,e)+t\cdot(2,3,e)=(1+2t,1+3t,e+et)$.

Dato che il calcolo della derivata di una funzione vettoriale è ricondotto al calcolo di n derivate di funzioni $\mathbb{R}\to\mathbb{R}$, indicata con $\mathcal{D}f(t)$ la derivata di $f(t)$, valgono i seguenti:
Teorema 2 : Se $f(t)$ e $g(t)$ sono derivabili in t, allora $\mathcal{D}(f(t)\pm g(t))=\mathcal{D}f(t)\pm\mathcal{D}g(t)$.
Ovvero, la derivata di una somma (di una differenza) di funzioni è uguale alla somma (alla differenza) delle derivate.
Per quanto riguarda il prodotto, dato che ora le immagini sono vettori, dobbiamo considerare due casi, quello del prodotto per uno scalare e quello del prodotto scalare; valgono i seguenti:
Teorema 3 : Siano $f:\mathbb{R}\to\mathbb{R}^n, t\to f(t)=(f_1(t),f_2(t),...,f_n(t))$ e $g:\mathbb{R}\to\mathbb{R}, t\to g(t)$, funzioni derivabili in t. Allora $\mathcal{D}(g(t)\cdot f(t))=\mathcal{D}g(t)\cdot f(t)+g(t)\cdot\mathcal{D}f(t)$.

Esempio 8 : Siano $f:\mathbb{R}\to\mathbb{R}^3, f(t)=\left(\cos t,\text{sen}\, t,e^t\right)$ e $g:\mathbb{R}\to\mathbb{R}, g(t)=t^2$. Allora:
$g(t)\cdot f(t)=t^2\left(\cos t,\text{sen}\, t,e^t\right)=\left(t^2\cos t,t^2\,\text{sen}\, t,t^2\,e^t\right)$. Risulta quindi:
$\mathcal{D}(g(t)\cdot f(t))=\left(2t\cos t-t^2\,\text{sen}\, t,2t\,\text{sen}\, t+t^2\cos t,2t\,e^t+t^2\,e^t\right)=$
$=2t\left(\cos t,\text{sen}\, t,e^t\right)+t^2\left(-\text{sen}\, t,\cos t,e^t\right)=\mathcal{D}g(t)\cdot f(t)+g(t)\cdot\mathcal{D}f(t)$.

Teorema 4 : Siano f e g : $\mathbb{R}\to\mathbb{R}^n$, funzioni derivabili in t. Allora:
$\mathcal{D}(f(t)\cdot g(t))=\mathcal{D}f(t)\cdot g(t)+f(t)\cdot\mathcal{D}g(t)$.

Esempio 9 : Siano $f:\mathbb{R}\to\mathbb{R}^3, f(t)=\left(t,\text{sen}\, t,e^t\right)$ e $g:\mathbb{R}\to\mathbb{R}^3, g(t)=\left(t^2,e^t,t^3\right)$.
Allora: $f(t)\cdot g(t)=\left(t,\text{sen}\, t,e^t\right)\cdot\left(t^2,e^t,t^3\right)=t^3+e^t\,\text{sen}\, t+t^3\,e^t$. Risulta quindi:
$\mathcal{D}(f(t)\cdot g(t))=3t^2+e^t\,\text{sen}\, t+e^t\cos t+3t^2\,e^t+t^3\,e^t=$
$=\left(1,\cos t,e^t\right)\cdot\left(t^2,e^t,t^3\right)+\left(t,\text{sen}\, t,e^t\right)\cdot\left(2t,e^t,3t^2\right)=\mathcal{D}f(t)\cdot g(t)+f(t)\cdot\mathcal{D}g(t)$.

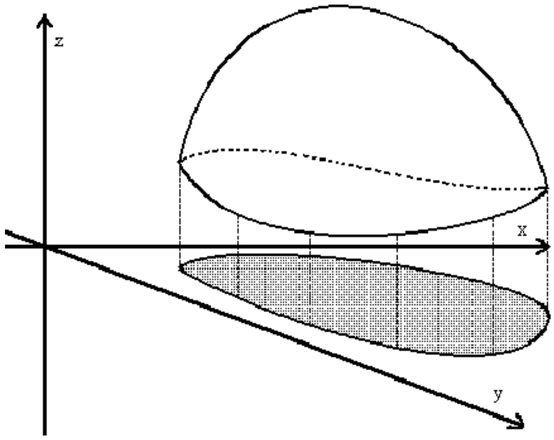
In ambedue i casi quindi, indipendentemente dal prodotto effettuato, vale la regola secondo la quale la derivata di un prodotto si ottiene mediante la somma di due termini, ciascuno dei quali è il prodotto della derivata di un fattore per l'altro non derivato.
Per quanto riguarda la composizione di funzioni, possiamo per ora trattare solo questo caso:
Teorema 5 : Sia $g:\mathbb{R}\to\mathbb{R}, t\to g(t)$ una funzione derivabile nel punto t e sia $f:\mathbb{R}\to\mathbb{R}^n$, $t\to f(t)=(f_1(t),f_2(t),...,f_n(t))$ una funzione derivabile nel punto $g(t)$. Allora la funzione composta, $\mathbb{R}\xrightarrow{g}\mathbb{R}\xrightarrow{f}\mathbb{R}^n, t\to f(g(t))=(f_1(g(t)),f_2(g(t)),...,f_n(g(t)))$ è derivabile in t e si ha: $\mathcal{D}(f(g(t)))=f'(g(t))\cdot g'(t)$.
Dimostrazione : Dato che $\mathcal{D}f(g(t))=(\mathcal{D}f_1(g(t)),\ \mathcal{D}f_2(g(t)),...,\mathcal{D}f_n(g(t)))$, avremo:
$\mathcal{D}f(g(t))=(f_1'(g(t))\cdot g'(t),f_2'(g(t))\cdot g'(t),...,f_n'(g(t))\cdot g'(t))=f'(g(t))\cdot g'(t)$.●

Esempio 10 : Siano $g:\mathbb{R}\to\mathbb{R}, g(t)=t^2$, e $f:\mathbb{R}\to\mathbb{R}^3, f(t)=\left(\cos t,\text{sen}\, t,e^t\right)$. Allora:
$f(g(t))=\left(\cos t^2,\text{sen}\, t^2,e^{t^2}\right)$. Risulta quindi:
$\mathcal{D}(f(g(t)))=\left(-\text{sen}\, t^2\cdot 2t,\cos t^2\cdot 2t,e^{t^2}\cdot 2t\right)=$
$=\left(-\text{sen}\, t^2,\cos t^2,e^{t^2}\right)\cdot(2t)=f'(g(t))\cdot g'(t)$.

Data la natura vettoriale del codominio, non ha senso parlare del reciproco e del quoziente per funzioni $\mathbb{R}\to\mathbb{R}^n$.

FUNZIONI REALI DI VARIABILE VETTORIALE $f:\mathbb{R}^n\to\mathbb{R}$

Dato che ora il dominio è dato da $\mathbb{R}^n$ o da un suo opportuno sottoinsieme, la variabile indipendente assume la natura di vettore, e per questo motivo le funzioni $f:\mathbb{R}^n \to \mathbb{R}$ vengono dette anche **funzioni di più variabili**.
Verranno rappresentate nella forma $y = f(\mathbb{X}) = f(x_1, x_2, ..., x_n),\, x_i \in \mathbb{R}$. Se la dimensione n è piccola ($n = 2$ o $n = 3$) si usano anche notazioni del tipo $z = f(x, y)$, per $f:\mathbb{R}^2 \to \mathbb{R}$, mentre si usa scrivere $w = f(x, y, z)$ per $f:\mathbb{R}^3 \to \mathbb{R}$.



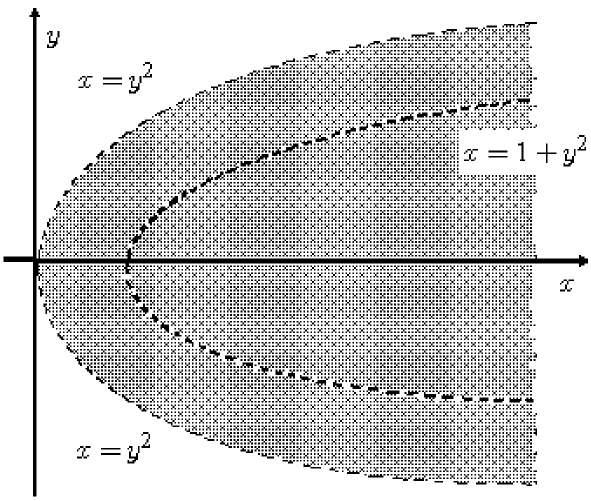
Per **grafico** di una funzione $f:\mathbb{R}^n \to \mathbb{R},\, y = f(\mathbb{X}) = f(x_1, x_2, ..., x_n)$, si intende il sottoinsieme di $\mathbb{R}^n \times \mathbb{R} = \mathbb{R}^{n+1}$ costituito dai punti $(x_1, x_2, ..., x_n, y)$ con $y = f(x_1, x_2, ..., x_n)$. Nel caso $n = 2$, essendo il dominio una parte di piano, il grafico di $z = f(x, y)$ costituisce una superficie bidimensionale giacente in $\mathbb{R}^3$, che può essere rappresentata come nella figura precedente, schiacciando, in prospettiva, il piano delle variabili indipendenti x e y.
Se $n > 2$ si dice che il grafico costituisce una **ipersuperficie** n-dimensionale in $\mathbb{R}^{n+1}$.
Il campo d'esistenza di una funzione di due variabili è un sottoinsieme di $\mathbb{R}^2$, e può quindi essere rappresentato graficamente.

Esempio 11 : Sia $f:\mathbb{R}^2 \to \mathbb{R},\, f(x, y) = \dfrac{1}{\log(x - y^2)}$.

Determiniamo e rappresentiamo il suo campo d'esistenza. Dovrà essere $\begin{cases} x - y^2 > 0 \\ \log(x - y^2) \neq 0 \end{cases}$

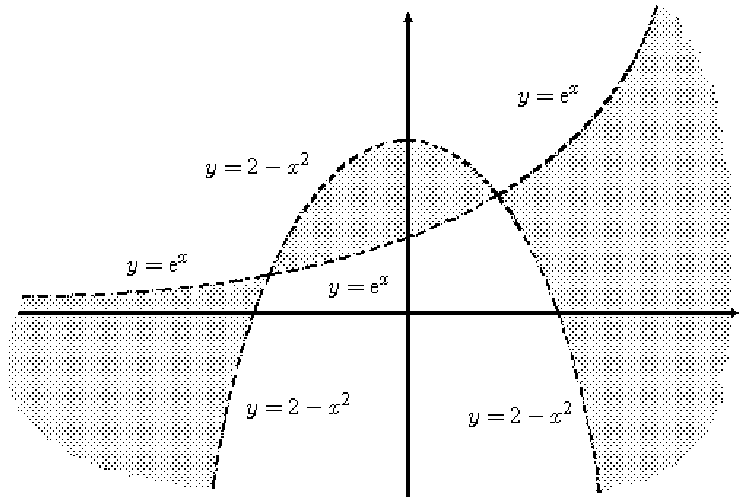
ovvero $\begin{cases} x > y^2 \\ x - y^2 \neq 1 \Rightarrow x \neq 1 + y^2 \end{cases}$.

Il campo d'esistenza della funzione è rappresentato dalla regione scura della figura di sinistra, ovvero dai punti a destra della parabola $x = y^2$, tolti i punti della parabola $x = 1 + y^2$.



Esempio 12 : Sia $f : \mathbb{R}^2 \to \mathbb{R}, f(x,y) = \log\left(\dfrac{y-e^x}{2-x^2-y}\right)$. Determiniamo e rappresentiamo il suo campo d'esistenza. Dovrà essere $\dfrac{y-e^x}{2-x^2-y} > 0$, che risulta soddisfatta se $\begin{cases} y-e^x > 0 \\ 2-x^2-y > 0 \end{cases}$ oppure se $\begin{cases} y-e^x < 0 \\ 2-x^2-y < 0 \end{cases}$, che sono rispettivamente equivalenti alla $\begin{cases} y > e^x \\ y < 2-x^2 \end{cases}$ ed alla $\begin{cases} y < e^x \\ y > 2-x^2 \end{cases}$.

Il campo d'esistenza della funzione data è rappresentato dall'unione delle regioni scure; quella centrale rappresenta la soluzione del primo sistema, le due rimanenti, quella a sinistra e quella a destra, rappresentano la soluzione del secondo. I bordi delle zone sono tratteggiati in quanto non appartengono al campo d'esistenza, dato che le disequazioni sono strette.



Esempio 13 : Sia $f : \mathbb{R}^2 \to \mathbb{R}, f(x,y) = \dfrac{1}{\text{sen}\,(x-y)}$. Determiniamo il suo campo d'esistenza. Dovrà essere $\text{sen}\,(x-y) \neq 0$, ovvero: $x-y \neq k\pi \Rightarrow y \neq x-k\pi, \forall\, k \in \mathbb{Z}$. Si tratta quindi di togliere dal piano $\mathbb{R}^2$ tutte le rette, parallele alla bisettrice $y = x$, di equazione $y = x - k\pi$.

Definizione 18 : Una funzione $f : \mathbb{R}^n \to \mathbb{R}, y = f(\mathbb{X})$, si dice **limitata** in $\mathbb{A} \subseteq \mathbb{R}^n$ se esistono due valori, $y_1, y_2 \in \mathbb{R}$, tali che : $y_1 \leq f(\mathbb{X}) \leq y_2, \forall\, \mathbb{X} \in \mathbb{A}$.
Definizione 19 : Un punto $\mathbb{X}_0 \in \mathbb{R}^n$ si dice di **massimo (di minimo) relativo** se esiste un intorno $\mathfrak{J}(\mathbb{X}_0, \varepsilon)$ nel quale si ha che $f(\mathbb{X}) \leq f(\mathbb{X}_0)$ $(f(\mathbb{X}) \geq f(\mathbb{X}_0))$ $\forall\, \mathbb{X} \in \mathfrak{J}(\mathbb{X}_0, \varepsilon)$. Se tali disuguaglianze valgono per ogni $\mathbb{X}$ appartenente al dominio della funzione, allora il punto $\mathbb{X}_0$ si dirà di **massimo (di minimo) assoluto**. Si parla di massimo (di minimo) **stretto** (o forte), sia relativo che assoluto, se le disuguaglianze valgono in senso stretto ($f(\mathbb{X}) < f(\mathbb{X}_0)$ oppure $f(\mathbb{X}) > f(\mathbb{X}_0)$).

LIMITI PER FUNZIONI $f : \mathbb{R}^n \to \mathbb{R}$

Sia $\mathbb{X}_0 \in \mathbb{R}^n$ un punto di accumulazione per il dominio della funzione $y = f(x_1, x_2, ..., x_n)$. Posto $\mathbb{X} = (x_1, x_2, ..., x_n)$ e $\mathbb{X}_0 = \left(x_1^0, x_2^0, ..., x_n^0\right)$, sia $l \in \mathbb{R}$.
Definizione 20 : Si dice che:
$\lim\limits_{\mathbb{X} \to \mathbb{X}_0} f(x_1, x_2, ..., x_n) = l$ se : $\forall\, \varepsilon > 0\ \exists\, \delta(\varepsilon) : 0 < \|\mathbb{X} - \mathbb{X}_0\| < \delta \Rightarrow |f(\mathbb{X}) - l| < \varepsilon$;
Definizione 21 : Si dice che:
$\lim\limits_{\mathbb{X} \to \mathbb{X}_0} f(x_1, x_2, ..., x_n) = +\infty$ se : $\forall\, \varepsilon\ \exists\, \delta(\varepsilon) : 0 < \|\mathbb{X} - \mathbb{X}_0\| < \delta \Rightarrow f(\mathbb{X}) > \varepsilon$;

Definizione 22 : Si dice che:
$\lim_{\mathbb{X} \to \mathbb{X}_0} f(x_1, x_2, ..., x_n) = -\infty$ se : $\forall\, \varepsilon\, \exists\, \delta(\varepsilon) : 0 < \|\mathbb{X} - \mathbb{X}_0\| < \delta \Rightarrow f(\mathbb{X}) < \varepsilon$.
Per i limiti di funzioni $f : \mathbb{R}^n \to \mathbb{R}$ valgono i Teoremi cosiddetti di unicità del limite, della permanenza del segno e del confronto, già visti per le funzioni $\mathbb{R} \to \mathbb{R}$.
Dire che $\mathbb{X} \to \mathbb{X}_0$, ovvero che $\mathbb{X}$ assume valori sempre più vicini a $\mathbb{X}_0$, viene espresso al solito mediante l'appartenenza ad un intorno di centro $\mathbb{X}_0$ e raggio δ; essendo $\mathbb{X}_0 \in \mathbb{R}^n$, questo richiede, nella definizione di limite, l'utilizzo della norma Euclidea $\|\mathbb{X} - \mathbb{X}_0\|$, mentre non si hanno variazioni per quanto riguarda il codominio, essendo questo contenuto in $\mathbb{R}$.
Nei limiti di funzioni di una sola variabile l'avvicinarsi di x a x_0 lungo la retta reale può avvenire al massimo in due modi : da sinistra e da destra, comportando così il calcolo del limite sinistro e del limite destro in x_0. Questo non vale per le funzioni $f : \mathbb{R}^n \to \mathbb{R}$. Limitandoci, ad esempio, alle funzioni $\mathbb{R}^2 \to \mathbb{R}$, quando diciamo che $(x, y) \to (x_0, y_0)$ intendiamo che questo avvenga mediante tutti i possibili percorsi (o curve) continui (ovvero senza interruzioni o salti) che da (x, y) portano in (x_0, y_0). L'eventuale risultato del limite deve essere sempre lo stesso qualunque sia il percorso utilizzato per poter affermare che il limite esiste. Se percorsi diversi portano a risultati diversi, la conclusione sarà che il limite non esiste.

Esempio 14 : Calcoliamo $\lim_{(x,y)\to(0,0)} \frac{x\,y^2}{x^2+y^2}$. La funzione è definita in $\mathbb{R}^2 \setminus \{(0,0)\}$. Dato che a numeratore abbiamo un polinomio infinitesimo di 3° grado, mentre quello a denominatore è di 2° grado, ipotizziamo che il limite valga 0 e proviamo a verificarlo mediante la definizione. Dovrà risultare $\left|\frac{x\,y^2}{x^2+y^2} - 0\right| < \varepsilon$ in un intorno di $(0,0)$.
Ma $\left|\frac{x\,y^2}{x^2+y^2}\right| = |x| \cdot \frac{y^2}{x^2+y^2} \le |x|$, in quanto $\frac{y^2}{x^2+y^2} \le 1, \forall\, (x,y) \in \mathbb{R}^2 \setminus \{(0,0)\}$.
Basta allora imporre $|x| < \varepsilon$ affinchè risulti $\left|\frac{x\,y^2}{x^2+y^2}\right| < \varepsilon$.
Ma $|x| < \varepsilon \Leftrightarrow -\varepsilon < x < \varepsilon$, ovvero si determina una striscia verticale all'interno della quale è sempre possibile ricavare un intorno di $(0,0)$: basta prendere $\delta < \varepsilon$.
E' quindi verificato che $\lim_{(x,y)\to(0,0)} \frac{x\,y^2}{x^2+y^2} = 0$.

Esempio 15 : Data $f(x,y) = \begin{cases} x \operatorname{sen} \frac{1}{y} & : y \ne 0 \\ 0 & : y = 0 \end{cases}$, studiamo $\lim_{(x,y)\to(0,0)} f(x,y)$. Dato che x tende a 0, mentre $\operatorname{sen} \frac{1}{y}$ è una quantità limitata, anche ora ipotizziamo che il limite valga 0, coerentemente, oltretutto, con il comportamento della funzione lungo l'asse x.
Verifichiamo allora che risulta : $\left|x \operatorname{sen} \frac{1}{y} - 0\right| < \varepsilon$ in un intorno di $(0,0)$.
Ma $\left|x \operatorname{sen} \frac{1}{y}\right| = |x| \cdot \left|\operatorname{sen} \frac{1}{y}\right| \le |x|$, in quanto $\left|\operatorname{sen} \frac{1}{y}\right| \le 1$. Imponendo che sia $|x| < \varepsilon$ troviamo la soluzione come nell'esempio precedente. Per i punti che stanno sull'asse x avremo invece : $|f(x,y) - 0| = |0 - 0| < \varepsilon$, che è sempre verificata.
Quindi $\lim_{(x,y)\to(0,0)} f(x,y) = 0$.

Esempio 16 : Studiamo il $\lim\limits_{(x,y)\to(0,0)} \dfrac{x^2}{x^2+y^2}$. Dato che $(x,y)\to(0,0)$, usiamo come percorsi di avvicinamento le rette passanti per l'origine, di equazione $y = m\,x$. Studiare il limite lungo questi percorsi significa, operata la sostituzione $y = m\,x$, calcolare :

$\lim\limits_{x\to 0} \dfrac{x^2}{x^2+m^2\,x^2} = \lim\limits_{x\to 0} \dfrac{1}{1+m^2} = \dfrac{1}{1+m^2}$. Il risultato dipende da m, variando al variare della retta utilizzata. Quindi abbiamo che $\lim\limits_{(x,y)\to(0,0)} \dfrac{x^2}{x^2+y^2}$ non esiste.

Esempio 17 : Studiamo il $\lim\limits_{(x,y)\to(0,0)} \dfrac{x^2}{x^2+y^4}$. Operando come nell'esempio precedente, posto $y = m\,x$, calcoliamo $\lim\limits_{x\to 0} \dfrac{x^2}{x^2+m^4\,x^4} = \lim\limits_{x\to 0} \dfrac{1}{1+m^4\,x^2} = 1$. Quindi percorrendo una qualsiasi retta passante per l'origine il limite vale 1. Da questo non è assolutamente lecito dedurre che $\lim\limits_{(x,y)\to(0,0)} \dfrac{x^2}{x^2+y^4} = 1$.

Utilizziamo infatti percorsi diversi per avvicinarci a $(0,0)$, quali le parabole aventi per asse l'asse delle x, e quindi di equazione $x = k\,y^2$.

Studiando il limite su queste restrizioni avremo : $\lim\limits_{y\to 0} \dfrac{k^2\,y^4}{k^2\,y^4+y^4} = \lim\limits_{y\to 0} \dfrac{k^2}{k^2+1} = \dfrac{k^2}{k^2+1}$.

Variando parabola varia il risultato, e quindi $\lim\limits_{(x,y)\to(0,0)} \dfrac{x^2}{x^2+y^4}$ non esiste.

LIMITI ITERATI

E' bene infine notare come, per il calcolo del $\lim\limits_{(x,y)\to(x_0,y_0)} f(x,y)$, non sia lecito ricorrere al calcolo dei due limiti : $\lim\limits_{x\to x_0}\lim\limits_{y\to y_0} f(x,y)$ e $\lim\limits_{y\to y_0}\lim\limits_{x\to x_0} f(x,y)$. Questi due limiti, detti **limiti iterati**, consistono nel calcolo successivo di due limiti di funzioni di una sola variabile, tenendo l'altra come costante. Se anche ambedue esistessero e fossero uguali, questo non consentirebbe comunque di concludere nulla sul $\lim\limits_{(x,y)\to(x_0,y_0)} f(x,y)$. I due limiti iterati corrispondono infatti a due modi particolari di far tendere (x,y) a (x_0,y_0), consistendo in percorsi paralleli agli assi x e y. Se $\lim\limits_{(x,y)\to(x_0,y_0)} f(x,y)$ esiste e se i due limiti iterati esistono, allora hanno tutti lo stesso valore. Se anche fosse $\lim\limits_{x\to x_0}\lim\limits_{y\to y_0} f(x,y) = \lim\limits_{y\to y_0}\lim\limits_{x\to x_0} f(x,y) = l$, non è comunque lecito attribuire a $\lim\limits_{(x,y)\to(x_0,y_0)} f(x,y)$ il valore trovato l.

Esempio 18 : Si è già visto (Esempio 16) che $\lim\limits_{(x,y)\to(0,0)} \dfrac{x^2}{x^2+y^2}$ non esiste. Si ha anche:

$\lim\limits_{x\to 0}\lim\limits_{y\to 0} \dfrac{x^2}{x^2+y^2} = \lim\limits_{x\to 0} \dfrac{x^2}{x^2} = \lim\limits_{x\to 0} 1 = 1$;

$\lim\limits_{y\to 0}\lim\limits_{x\to 0} \dfrac{x^2}{x^2+y^2} = \lim\limits_{y\to 0} \dfrac{0}{y^2} = \lim\limits_{y\to 0} 0 = 0$,

coerentemente con la non esistenza del limite.

Esempio 19 : Si è visto (Esempio 15) che per $f(x,y) = \begin{cases} x\,\mathrm{sen}\,\dfrac{1}{y} & : y \neq 0 \\ 0 & : y = 0 \end{cases}$ risulta:

$\lim\limits_{(x,y)\to(0,0)} f(x,y) = 0$. Eppure si ha che:

$\lim\limits_{y\to 0}\lim\limits_{x\to 0} x \operatorname{sen}\dfrac{1}{y} = \lim\limits_{y\to 0} 0 = 0$, mentre invece $\lim\limits_{x\to 0}\lim\limits_{y\to 0} x \operatorname{sen}\dfrac{1}{y}$ non esiste. L'esistenza del $\lim\limits_{(x,y)\to(0,0)} f(x,y)$ non implica infatti l'esistenza del limite lungo tutti i possibili percorsi da (x,y) a $(0,0)$.

Esempio 20 : Applichiamo i limiti iterati a $\lim\limits_{(x,y)\to(0,0)} \dfrac{x\,y}{x^2+y^2}$. Avremo:

$\lim\limits_{x\to 0}\lim\limits_{y\to 0} \dfrac{x\,y}{x^2+y^2} = \lim\limits_{x\to 0}\dfrac{0}{x^2} = \lim\limits_{x\to 0} 0 = 0$, e

$\lim\limits_{y\to 0}\lim\limits_{x\to 0} \dfrac{x\,y}{x^2+y^2} = \lim\limits_{y\to 0}\dfrac{0}{y^2} = \lim\limits_{y\to 0} 0 = 0$.

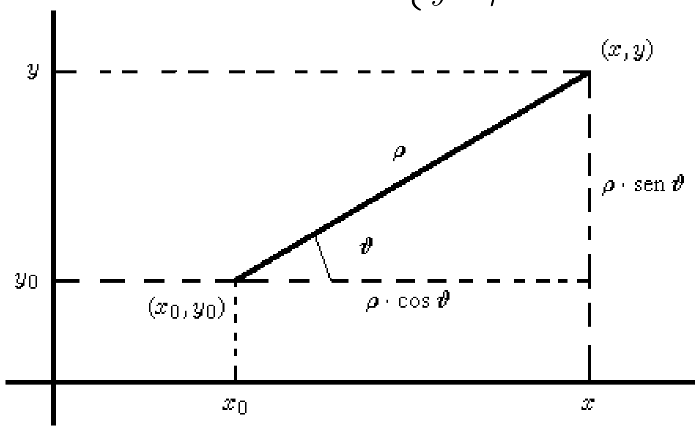
Se usiamo le rette passanti per l'origine $y = m\,x$, si ha invece $\lim\limits_{x\to 0}\dfrac{m\,x^2}{x^2+m^2\,x^2} = \dfrac{m}{1+m^2}$, e quindi il limite dato non esiste, anche se i due limiti iterati esistono ed hanno lo stesso valore.

LIMITI IN COORDINATE POLARI PER FUNZIONI $f : \mathbb{R}^2 \to \mathbb{R}$

Può essere utile, limitatamente però al solo caso di funzioni di due variabili, passare da coordinate cartesiane a coordinate polari per il calcolo di un limite. Ogni punto $(x,y)\in\mathbb{R}^2$ può essere espresso, facendo riferimento ad un dato punto (x_0,y_0), come: $\begin{cases} x = x_0 + \rho\cos\vartheta \\ y = y_0 + \rho\operatorname{sen}\vartheta \end{cases}$.

Diremo queste le **coordinate polari** del punto (x,y) rispetto al punto (x_0,y_0).

Se $(x_0,y_0) = (0,0)$, si ha, come caso particolare, $\begin{cases} x = \rho\cos\vartheta \\ y = \rho\operatorname{sen}\vartheta \end{cases}$.



Operando allora la sostituzione, avremo:

$\lim\limits_{(x,y)\to(x_0,y_0)} f(x,y) = \lim\limits_{\rho\to 0} f(\rho\cos\vartheta, \rho\operatorname{sen}\vartheta) = \lim\limits_{\rho\to 0} F(\rho,\vartheta)$.

Il limite per $(x,y)\to(x_0,y_0)$ si trasforma in un limite nella sola variabile ρ, con $\rho\to 0$ per esprimere l'avvicinarsi di (x,y) a (x_0,y_0), dato che $\rho = \|(x,y)-(x_0,y_0)\|$.

Ma il valore del limite non deve dipendere dal particolare percorso utilizzato, e questo si traduce nella richiesta che $\lim\limits_{\rho\to 0} F(\rho,\vartheta)$, se esiste, non deve dipendere dalla particolare direzione, cioè da ϑ; diciamo che la convergenza (o la divergenza) del $\lim\limits_{\rho\to 0} F(\rho,\vartheta)$ deve avvenire in modo uniforme rispetto a ϑ, ovvero che la determinazione del δ avviene in funzione di ε, indipendentemente da ϑ.

Avremo quindi, usando le coordinate polari, le seguenti definizioni di limite:

Definizione 23 : Si dice che:

$\lim\limits_{\rho\to 0} F(\rho,\vartheta) = l \in \mathbb{R}$ se : $\forall\, \varepsilon > 0\ \exists\, \delta(\varepsilon) : \forall\, \vartheta,\, 0 < \rho < \delta \Rightarrow |F(\rho,\vartheta) - l| < \varepsilon$;

Definizione 24 : Si dice che:

$\lim\limits_{\rho\to 0} F(\rho,\vartheta) = +\infty$ se : $\forall\, \varepsilon\ \exists\, \delta(\varepsilon) : \forall\, \vartheta,\, 0 < \rho < \delta \Rightarrow F(\rho,\vartheta) > \varepsilon$;

Definizione 25 : Si dice che:

$\lim\limits_{\rho\to 0} F(\rho,\vartheta) = -\infty$ se : $\forall\, \varepsilon\ \exists\, \delta(\varepsilon) : \forall\, \vartheta,\, 0 < \rho < \delta \Rightarrow F(\rho,\vartheta) < \varepsilon$.

L'uniformità è espressa nella: $\forall\, \varepsilon > 0\ \exists\, \delta(\varepsilon) : \forall\, \vartheta$, che esprime appunto l'indipendenza da ϑ nella scelta del $\delta(\varepsilon)$. Se quindi $\lim\limits_{\rho\to 0} F(\rho,\vartheta) = l$ (o $+\infty$ o $-\infty$) uniformemente rispetto a ϑ, allora si ha che: $\lim\limits_{(x,y)\to(x_0,y_0)} f(x,y) = \lim\limits_{\rho\to 0} F(\rho,\vartheta)$.

Esempio 21 : Studiamo $\lim\limits_{(x,y)\to(0,0)} \dfrac{x^4}{x^2+y^2}$ mediante passaggio a coordinate polari.

Posto $\begin{cases} x = \rho\cos\vartheta \\ y = \rho\,\mathrm{sen}\,\vartheta \end{cases}$, avremo che $\lim\limits_{\rho\to 0} \dfrac{(\rho\cos\vartheta)^4}{(\rho\cos\vartheta)^2 + (\rho\,\mathrm{sen}\,\vartheta)^2} = \lim\limits_{\rho\to 0} \rho^2\cos^4\vartheta = 0$.

Vediamo ora se la convergenza a 0 del limite risulta uniforme rispetto a ϑ.

Per la definizione di limite dovrà essere: $\left|\rho^2\cos^4\vartheta\right| < \varepsilon$ in $\mathfrak{I}(0,\ \delta(\varepsilon))$.

Ma $\left|\rho^2\cos^4\vartheta\right| \le \rho^2$, in quanto $\cos^4\vartheta \le 1$, per cui, posto $\rho^2 < \varepsilon$, cioè $-\sqrt{\varepsilon} < \rho < \sqrt{\varepsilon}$, (ovvero $0 < \rho < \sqrt{\varepsilon}$, dato che ρ è sempre positivo), posto $\delta(\varepsilon) = \sqrt{\varepsilon}$, risulta:

$0 < \rho < \delta(\varepsilon) = \sqrt{\varepsilon} \Rightarrow \left|\rho^2\cos^4\vartheta\right| < \varepsilon$, ovvero $\lim\limits_{\rho\to 0} F(\rho,\vartheta) = 0$ e la convergenza del limite è uniforme. Quindi $\lim\limits_{(x,y)\to(0,0)} \dfrac{x^4}{x^2+y^2} = 0$.

Esempio 22 : Studiamo $\lim\limits_{(x,y)\to(0,0)} \dfrac{x^2\,y}{x^4+y^2}$ mediante le coordinate polari.

Usando le rette $y = m\,x$ si ha: $\lim\limits_{(x,y)\to(0,0)} \dfrac{x^2\,y}{x^4+y^2} = \lim\limits_{x\to 0} \dfrac{m\,x^3}{x^4+m^2\,x^2} = \lim\limits_{x\to 0} \dfrac{m\,x}{x^2+m^2} = 0$, ma sappiamo che questo non basta a garantire l'esistenza del limite.

Posto $\begin{cases} x = \rho\cos\vartheta \\ y = \rho\,\mathrm{sen}\,\vartheta \end{cases}$, calcoliamo allora:

$\lim\limits_{\rho\to 0} \dfrac{\rho^2\cos^2\vartheta\ \cdot\rho\,\mathrm{sen}\,\vartheta}{\rho^4\cos^4\vartheta + \rho^2\,\mathrm{sen}^2\,\vartheta} = \lim\limits_{\rho\to 0} \rho\cdot\dfrac{\cos^2\vartheta\cdot\mathrm{sen}\,\vartheta}{\rho^2\cos^4\vartheta + \mathrm{sen}^2\,\vartheta} = 0$, in quanto il primo fattore, ρ, tende a 0 mentre il secondo tende a $\dfrac{\cos^2\vartheta}{\mathrm{sen}\,\vartheta}$. Quindi $\lim\limits_{\rho\to 0} F(\rho,\vartheta) = 0$, ma non è vero che $\lim\limits_{(x,y)\to(0,0)} \dfrac{x^2\,y}{x^4+y^2} = 0$, in quanto la convergenza non è uniforme. Infatti, per la definizione di limite deve risultare: $\left|\rho\cdot\dfrac{\cos^2\vartheta\cdot\mathrm{sen}\,\vartheta}{\rho^2\cos^4\vartheta + \mathrm{sen}^2\,\vartheta} - 0\right| = \rho\cdot\dfrac{\cos^2\vartheta\cdot|\mathrm{sen}\,\vartheta|}{\rho^2\cos^4\vartheta + \mathrm{sen}^2\,\vartheta} < \varepsilon$. La non uniformità della convergenza si spiega osservando che la quantità $\dfrac{\cos^2\vartheta\cdot|\mathrm{sen}\,\vartheta|}{\rho^2\cos^4\vartheta + \mathrm{sen}^2\,\vartheta}$ può assumere valori arbitrariamente grandi, per valori di ϑ prossimi a 0 o a π, quando $\rho\to 0$; quindi non si può trovare una maggiorazione per $\dfrac{\cos^2\vartheta\cdot|\mathrm{sen}\,\vartheta|}{\rho^2\cos^4\vartheta + \mathrm{sen}^2\,\vartheta}$ che risulti indipendente da ϑ.

Quindi $\lim\limits_{(x,y)\to(0,0)} \dfrac{x^2\,y}{x^4+y^2}$ non esiste, e vedremo nel seguito un'altra procedura atta a giustificare questa conclusione.

CURVE DI LIVELLO

Data $y = f(x_1, x_2, ..., x_n)$, se poniamo $y = k \in \mathbb{R}$, otteniamo $f(x_1, x_2, ..., x_n) = k$, ovvero tutti i punti $(x_1, x_2, ..., x_n)$ del dominio in cui la funzione assume un dato valore costante k. Un tale insieme viene detto **curva di livello**. Anche le curve di livello possono essere utilizzate per dimostrare la non esistenza del limite.

Esempio 23 : Studiamo $\lim_{(x,y)\to(0,0)} f(x,y)$, con $f(x,y) = \begin{cases} \dfrac{x^2+y^2}{y} & y \neq 0 \\ 0 & y = 0 \end{cases}$.

Usando le rette $y = m\,x$, avremo:

$\lim_{x\to 0} \dfrac{x^2+m^2\,x^2}{m\,x} = 0$ per $m \neq 0$ mentre $\lim_{x\to 0} 0 = 0$ per $m = 0$, cioè se $y = 0$. Questo limite però non esiste. Vediamolo utilizzando le curve di livello: posto $y \neq 0$, se imponiamo:

$f(x,y) = \dfrac{x^2+y^2}{y} = k \in \mathbb{R}$, ricaviamo: $x^2 + y^2 = k\,y$, ovvero $x^2 + y^2 - k\,y = 0$.

Questa è l'equazione di una circonferenza nel piano $\mathbb{R}^2$ avente centro nel punto $\left(0, \dfrac{k}{2}\right)$ e passante, questo è importante, per $(0,0)$, cioè il punto verso il quale calcoliamo il limite. In ogni intorno di $(0,0)$ ci sono infinite circonferenze, tutte passanti per $(0,0)$, su ciascuna delle quali la funzione $f(x,y) = \dfrac{x^2+y^2}{y}$ assume valore costante. Ma allora il valore del limite dovrebbe essere uguale a k, e potendo la costante k assumere infiniti valori, per il teorema di unicità del limite, si ha che $\lim_{(x,y)\to(0,0)} f(x,y)$ non esiste.

Esempio 24 : Riprendiamo il $\lim_{(x,y)\to(0,0)} \dfrac{x^2\,y}{x^4+y^2}$, ed usiamo le curve di livello per vedere che tale limite non esiste. Posto $\dfrac{x^2\,y}{x^4+y^2} = k$ otteniamo: $k\,y^2 - x^2\,y + k\,x^4 = 0$, che ha per soluzioni: $y = \dfrac{x^2 \pm \sqrt{x^4\,(1-4k^2)}}{2k} = \dfrac{\left(1 \pm \sqrt{1-4k^2}\right)}{2k} x^2 = m\,x^2$.

Se $1 - 4k^2 \geq 0$, cioè se $-\dfrac{1}{2} \leq k \leq \dfrac{1}{2}$, tale equazione rappresenta infinite parabole passanti per il punto $(0,0)$, sulle quali la funzione $f(x,y) = \dfrac{x^2\,y}{x^4+y^2}$ è costante e vale k. Data l'arbitrarietà di k, per il Teorema di unicità del limite si ha che $\lim_{(x,y)\to(0,0)} \dfrac{x^2\,y}{x^4+y^2}$ non esiste.

LIMITI ALL'INFINITO

Data $f : \mathbb{R}^n \to \mathbb{R}, y = f(\mathbb{X})$, essendo la variabile indipendente $\mathbb{X}$ un vettore, dobbiamo definire cosa significhi che un vettore $\mathbb{X}$ tende all'infinito. Intendiamo allora che $\|\mathbb{X}\| \to +\infty$, ovvero che il modulo di $\mathbb{X}$ diviene illimitatamente grande. Avremo allora altre tre definizioni di limite:

Definizione 26 : Si dice che:

$\lim_{\mathbb{X}\to\infty} f(\mathbb{X}) = l \in \mathbb{R}$ se : $\forall\, \varepsilon > 0\ \exists\, \delta(\varepsilon) : \|\mathbb{X}\| > \delta \Rightarrow |f(\mathbb{X}) - l| < \varepsilon$;

Definizione 27 : Si dice che:

$\lim\limits_{\mathbb{X}\to\infty} f(\mathbb{X}) = +\infty$ se : $\forall\,\varepsilon\;\exists\,\delta(\varepsilon) : \|\mathbb{X}\| > \delta \Rightarrow f(\mathbb{X}) > \varepsilon\,;$

Definizione 28 : Si dice che:
$\lim\limits_{\mathbb{X}\to\infty} f(\mathbb{X}) = -\infty$ se : $\forall\,\varepsilon\;\exists\,\delta(\varepsilon) : \|\mathbb{X}\| > \delta \Rightarrow f(\mathbb{X}) < \varepsilon\,.$

Anche nei limiti per $\mathbb{X}\to\infty$ possono usarsi le coordinate polari, purchè sia $f:\mathbb{R}^2\to\mathbb{R}$.

Esempio 25 : Calcoliamo $\lim\limits_{\mathbb{X}\to\infty} e^{-x^2-y^2}$. Verifichiamo che tale limite vale 0. Infatti:

$\left|e^{-x^2-y^2}-0\right| < \varepsilon \Rightarrow e^{-x^2-y^2} < \varepsilon \Rightarrow -\left(x^2+y^2\right) < \log\varepsilon \Rightarrow x^2+y^2 > \log\dfrac{1}{\varepsilon}\,.$

Ma $x^2+y^2 = \|\mathbb{X}\|^2$, e quindi, posto $\delta(\varepsilon) = \sqrt{\log\dfrac{1}{\varepsilon}}$, se $\|\mathbb{X}\| > \delta \Rightarrow \left|e^{-x^2-y^2}\right| < \varepsilon$, ovvero è verificato che il limite vale 0.

Se usiamo invece le coordinate polari $\begin{cases} x=\rho\cos\vartheta \\ y=\rho\,\mathrm{sen}\,\vartheta\end{cases}$, avremo $\lim\limits_{\rho\to+\infty} e^{-\rho^2} = 0$, sicuramente in modo uniforme in quanto nell'espressione $F(\rho,\vartheta)$ non compare ϑ.

FUNZIONI CONTINUE

Analogamente al caso di $f:\mathbb{R}\to\mathbb{R}$, anche per le funzioni $f:\mathbb{R}^n\to\mathbb{R}$ diamo la seguente:
Definizione 29 : Sia $\mathbb{X}_0\in\mathbb{R}^n$ punto di accumulazione appartenente al dominio della funzione $f(\mathbb{X})$; si dice che la funzione $f(\mathbb{X})$ è **continua** in $\mathbb{X}_0$ se $\lim\limits_{\mathbb{X}\to\mathbb{X}_0} f(\mathbb{X}) = f(\mathbb{X}_0)$.

Esempio 26 : Verifichiamo che $\lim\limits_{(x,y)\to(0,0)} \dfrac{x^3}{x^2+y^2} = 0$. Infatti, essendo:

$\left|\dfrac{x^3}{x^2+y^2}\right| = |x|\cdot\left|\dfrac{x^2}{x^2+y^2}\right| \le |x|$, se $|x| < \delta = \varepsilon$, segue $\left|\dfrac{x^3}{x^2+y^2}-0\right| < \varepsilon$, e quindi il limite vale 0. Possiamo allora definire la funzione $f(x,y) = \begin{cases} \dfrac{x^3}{x^2+y^2} & (x,y)\neq(0,0) \\ 0 & (x,y)=(0,0)\end{cases}$.

Dato che $\lim\limits_{(x,y)\to(0,0)} f(x,y) = f(0,0) = 0$, la funzione $f(x,y)$ è continua nel punto $(0,0)$.

Per le funzioni continue $f:\mathbb{R}^n\to\mathbb{R}$ valgono Teoremi analoghi a quelli stabiliti per le funzioni continue di una sola variabile, ovvero:
- Sommando, sottraendo e moltiplicando funzioni continue si ottengono funzioni continue;
- Il reciproco ed il quoziente di funzioni continue (con denominatore non infinitesimo) sono funzioni continue;
- Componendo funzioni continue si ottengono funzioni continue.

Vale anche il:
Teorema 6 : (di Weierstrass) Se $f:\mathbb{R}^n\to\mathbb{R}$ è continua in un compatto, allora ammette massimo e minimo assoluti.

DERIVATE PARZIALI

Per una funzione $f:\mathbb{R}\to\mathbb{R}$ ha senso cercare di determinare la tendenza istantanea al cambiamento (ovvero la derivata) in un punto x_0 facendo assumere alla variabile indipendente x valori sulla sinistra e sulla destra di x_0, visto che il dominio di f è contenuto in $\mathbb{R}$, spazio ad una sola dimensione. Sappiamo che la derivata, indicata con $f'(x_0)$, è definita come il limite, se esiste ed è finito, del rapporto incrementale calcolato in x_0 :

$\lim\limits_{x\to x_0}\dfrac{f(x)-f(x_0)}{x-x_0}=\lim\limits_{h\to 0}\dfrac{f(x_0+h)-f(x_0)}{h}=f'(x_0)$, e ci fornisce il coefficiente angolare della retta tangente nel punto $(x_0,\, f(x_0))$ al grafico della funzione.

Per le funzioni $f:\mathbb{R}^n\to\mathbb{R}$, essendo il dominio contenuto in uno spazio n-dimensionale, come già visto per l'operazione di limite, preso $\mathbb{X}_0$ punto interno al dominio, occorrerà scegliere una tra le infinite direzioni (rette) che passano per il punto $\mathbb{X}_0$ per poi determinare la tendenza istantanea al cambiamento della funzione f nel punto $\mathbb{X}_0$ relativamente alla direzione scelta. Se il punto $\mathbb{X}_0$ è interno a D_f sarà possibile sviluppare tale analisi in ogni direzione, mentre invece se $\mathbb{X}_0$ è sulla frontiera del D_f questo sarà possibile solo in alcune direzioni.

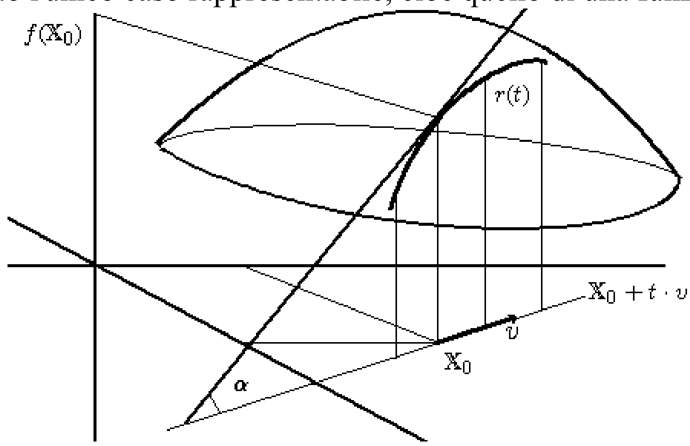
Quanto detto ci introduce alla definizione di **derivata direzionale** (o derivata in una direzione). Sia allora $\mathbb{X}_0$ un punto interno a D_f. Abbiamo la:

Definizione 30 : Scelto un versore $v\in\mathbb{R}^n$ ($\|v\|=1$), si dice che la funzione f è derivabile in $\mathbb{X}_0$ nella direzione v se esiste finito il limite: $\lim\limits_{t\to 0}\dfrac{f(\mathbb{X}_0+t\,v)-f(\mathbb{X}_0)}{t}=\mathcal{D}_v f(\mathbb{X}_0)$.

Notiamo che in questa definizione $t\in\mathbb{R}$.

Dato che il grafico di f è una (iper)superficie, essendo $\mathbb{X}_0+t\,v$ un segmento che, al variare di t, partendo da $\mathbb{X}_0$ porta nella direzione di v, la proiezione, mediante f, di tale segmento genera una curva $r(t)$, giacente sulla (iper)superficie rappresentante il grafico; a tale curva è possibile tracciare la retta tangente nel punto $(\mathbb{X}_0, f(\mathbb{X}_0))$, che formerà un angolo α con la retta passante per $\mathbb{X}_0$ nella direzione di v. Risulta $\mathcal{D}_v f(\mathbb{X}_0)=\text{tg}\,\alpha$, ovvero ne rappresenta la pendenza.

Nel caso particolare $f:\mathbb{R}^2\to\mathbb{R}$, $z=f(x,y)$, ogni versore $v\in\mathbb{R}^2$ può essere espresso nella forma $v=(\cos\alpha,\text{sen}\,\alpha)$, con $0\le\alpha<2\pi$. La derivata direzionale $\mathcal{D}_v f(\mathbb{X}_0)$ può essere allora espressa come: $\mathcal{D}_v f(\mathbb{X}_0)=\lim\limits_{t\to 0}\dfrac{f(x_0+t\cos\alpha, y_0+t\,\text{sen}\,\alpha)-f(x_0,y_0)}{t}$. Nella figura viene illustrato l'unico caso rappresentabile, cioè quello di una funzione $f:\mathbb{R}^2\to\mathbb{R}$.



Sia $E=(e_1,e_2,...,e_n)$ la base canonica di $\mathbb{R}^n$, dove e_i è il versore avente tutte le componenti uguali a 0 eccettuata l'i-esima uguale a 1. Tra le infinite direzioni v queste ultime rivestono un ruolo particolare; esse sono le direzioni parallele ad uno degli assi coordinati e sono espresse da uno dei vettori di E. Esaminare la tendenza istantanea al cambiamento in una delle direzioni e_i conduce al concetto di **derivata parziale** fatta rispetto alla variabile x_i.

Abbiamo quindi la

Definizione 31 : Scelto $e_i\in E$, si dice che la funzione f ammette in $\mathbb{X}_0$ la derivata parziale rispetto alla variabile x_i se esiste ed è finito: $\lim\limits_{t\to 0}\dfrac{f(\mathbb{X}_0+t\,e_i)-f(\mathbb{X}_0)}{t}=\dfrac{\partial f(\mathbb{X}_0)}{\partial x_i}$.

Tale limite può essere scritto per esteso, posto $\mathbb{X}_0=(x_1,x_2,...,x_n)$, anche come:

$\lim\limits_{h\to 0}\dfrac{f(x_1,x_2,...,x_i+h,...,x_n)-f(x_1,x_2,...,x_i,...,x_n)}{h}=\dfrac{\partial f(\mathbb{X}_0)}{\partial x_i}$.

Si usano pure altre notazioni quali: $\dfrac{\partial f(\mathbb{X}_0)}{\partial x_i} = f'_{x_i}(\mathbb{X}_0) = f'_i(\mathbb{X}_0) = \mathcal{D}_{x_i} f(\mathbb{X}_0) = \mathcal{D}_i f(\mathbb{X}_0)$.

Muoversi in una direzione parallela ad uno degli assi significa incrementare una sola variabile (quella relativa all'asse considerato) e mantenere tutte le altre costanti. Se consideriamo una funzione di due variabili $f(x, y)$, avremo due possibili derivate parziali nel punto (x_0, y_0):

-la derivata parziale fatta rispetto a x, definita come:

$$\lim_{h \to 0} \frac{f(x_0 + h, y_0) - f(x_0, y_0)}{h} = \frac{\partial f(x_0, y_0)}{\partial\, x} = f_x{}'(x_0, y_0)$$

-la derivata parziale fatta rispetto a y, definita come:

$$\lim_{h \to 0} \frac{f(x_0, y_0 + h) - f(x_0, y_0)}{h} = \frac{\partial f(x_0, y_0)}{\partial\, y} = f_y{}'(x_0, y_0)$$

purchè tali limiti esistano finiti.

Per una funzione di tre variabili $f(x, y, z)$ avremo tre possibili derivate parziali, definite come:

$$\lim_{h \to 0} \frac{f(x_0 + h, y_0, z_0) - f(x_0, y_0, z_0)}{h} = \frac{\partial f(x_0, y_0, z_0)}{\partial\, x} = f_x{}'(x_0, y_0, z_0)$$

$$\lim_{h \to 0} \frac{f(x_0, y_0 + h, z_0) - f(x_0, y_0, z_0)}{h} = \frac{\partial f(x_0, y_0, z_0)}{\partial\, y} = f_y{}'(x_0, y_0, z_0)$$

$$\lim_{h \to 0} \frac{f(x_0, y_0, z_0 + h) - f(x_0, y_0, z_0)}{h} = \frac{\partial f(x_0, y_0, z_0)}{\partial\, z} = f_z{}'(x_0, y_0, z_0).$$

Per una funzione $f : \mathbb{R}^n \to \mathbb{R}$ in un punto $\mathbb{X}_0$ interno al dominio si possono quindi avere, se esistono, esattamente n derivate parziali. Se in $\mathbb{X}_0$ la funzione risulta derivabile rispetto a tutte le sue variabili, allora esiste un vettore, detto **Gradiente** della funzione, indicato con $\nabla f(\mathbb{X}_0)$, che ha per componenti le derivate parziali della funzione calcolate nel punto:

$\nabla f(\mathbb{X}_0) = \left(\dfrac{\partial\, f(\mathbb{X}_0)}{\partial\, x_1}, \dfrac{\partial\, f(\mathbb{X}_0)}{\partial\, x_2}, ..., \dfrac{\partial\, f(\mathbb{X}_0)}{\partial\, x_n} \right)$. Il simbolo ∇f si legge "del f ".

L'essere le derivate parziali definite mediante il limite di un rapporto incrementale nel quale una sola variabile si incrementa mentre tutte le altre rimangono costanti ha un'importante conseguenza per il calcolo pratico delle derivate parziali: basta applicare le usuali regole di derivazione stabilite per le funzioni di una sola variabile, quella rispetto a cui si deriva, trattando tutte le altre variabili come costanti; potremo così calcolare sia le derivate parziali in un punto che le funzioni derivate parziali.

Esempio 27 : Sia data $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = x^y$. Avremo allora:

$\dfrac{\partial f}{\partial x} = y\, x^{y-1}$ (derivata di una potenza, in quanto x è variabile e y è costante)

$\dfrac{\partial f}{\partial y} = x^y \log x$ (derivata di una esponenziale, in quanto y è variabile e x è costante).

Esempio 28 : Sia data $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = x \operatorname{arctg} \dfrac{x}{y}$. Avremo allora:

$$\frac{\partial f}{\partial x} = 1 \cdot \operatorname{arctg} \frac{x}{y} + x \cdot \frac{1}{1 + \left(\dfrac{x}{y}\right)^2} \cdot \frac{1}{y} \;;$$

$$\frac{\partial f}{\partial y} = x \cdot \frac{1}{1 + \left(\dfrac{x}{y}\right)^2} \cdot x \cdot \left(- \frac{1}{y^2} \right).$$

Esempio 29 : Sia data $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = y^{x-y}$. Avremo allora.

$\dfrac{\partial f}{\partial x} = y^{x-y} \cdot \log y$ (derivata di una esponenziale, in quanto x è variabile e y è costante)

$\dfrac{\partial f}{\partial y} = \mathcal{D}_y\, e^{(x-y)\log y} = y^{x-y}\left(-1 \cdot \log y + \dfrac{x-y}{y}\right)$ (derivata di una $f(y)^{g(y)}$).

Esempio 30 : Sia data $f : \mathbb{R}^3 \to \mathbb{R}$, $f(x,y,z) = z - \text{sen}\left(\dfrac{x-z}{x^2}\right)$. Avremo allora:

$\dfrac{\partial f}{\partial x} = 0 - \cos\left(\dfrac{x-z}{x^2}\right) \cdot \dfrac{x^2 - 2x\,(x-z)}{x^4}$;

$\dfrac{\partial f}{\partial y} = 0$ in quanto la funzione data è costante rispetto a y ;

$\dfrac{\partial f}{\partial z} = 1 - \cos\left(\dfrac{x-z}{x^2}\right) \cdot \dfrac{1}{x^2} \cdot (-1)$.

Esempio 31 : Data $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x,y) = e^{x+y}$, calcoliamone la derivata direzionale nel punto (x,y) nella direzione del vettore $w = (1,1)$.

Essendo $\|w\| = \sqrt{2}$, risulta $v = \left(\dfrac{1}{\sqrt{2}}, \dfrac{1}{\sqrt{2}}\right)$ e quindi avremo:

$$\lim_{t\to 0}\frac{f((x,y)+t\,v) - f(x,y)}{t} = \lim_{t\to 0}\frac{f\left(x + \dfrac{1}{\sqrt{2}}\,t, y + \dfrac{1}{\sqrt{2}}\,t\right) - f(x,y)}{t} =$$

$$= \lim_{t\to 0}\frac{e^{x+\frac{1}{\sqrt{2}}t+y+\frac{1}{\sqrt{2}}t} - e^{x+y}}{t} = \lim_{t\to 0}\frac{e^{x+y}\left(e^{\sqrt{2}\,t} - 1\right)}{t} =$$

$$= e^{x+y} \cdot \sqrt{2} \cdot \lim_{t\to 0}\frac{e^{\sqrt{2}\,t} - 1}{\sqrt{2}\,t} = e^{x+y} \cdot \sqrt{2} \cdot \lim_{w\to 0}\frac{e^w - 1}{w} = \sqrt{2}\,e^{x+y} .$$

DERIVABILITA' E CONTINUITA'

Per le funzioni di una sola variabile sappiamo che la continuità è condizione necessaria per la derivabilità, e quindi anche che la derivabilità è condizione sufficiente per la continuità. Vediamo che questo legame non vale invece per le funzioni $\mathbb{R}^n \to \mathbb{R}$.

Esempio 32 : Data la funzione $f(x,y) = \begin{cases} 1 : & x\,y = 0 \\ 0 : & x\,y \neq 0 \end{cases}$, calcoliamone le derivate parziali nel punto $(0,0)$. Risulta:

$\dfrac{\partial f(0,0)}{\partial x} = \lim\limits_{h\to 0}\dfrac{f(0+h,0) - f(0,0)}{h} = \lim\limits_{h\to 0}\dfrac{1-1}{h} = 0$,

$\dfrac{\partial f(0,0)}{\partial y} = \lim\limits_{h\to 0}\dfrac{f(0,0+h) - f(0,0)}{h} = \lim\limits_{h\to 0}\dfrac{1-1}{h} = 0$.

Quindi in $(0,0)$ la funzione è derivabile (parzialmente) pur essendo palesemente discontinua, visto che in ogni intorno di $(0,0)$ ci sono punti in cui $f(x,y) = 1$ e punti in cui $f(x,y) = 0$.

Esempio 33 : Sia data la funzione $f(x,y) = \begin{cases} \dfrac{x^2\,y}{x^4 + y^2} & (x,y) \neq (0,0) \\ 0 & (x,y) = (0,0) \end{cases}$. Di questa funzione è già stato calcolato il $\lim\limits_{(x,y)\to(0,0)} f(x,y)$, Esempio 22, e si è visto che tale limite non esiste, per cui la funzione non è continua in $(0,0)$. Vediamo comunque se in $(0,0)$ la funzione data risulta derivabile in una qualche direzione $v = (\cos\alpha, \text{sen}\,\alpha)$. Dovremo allora calcolare:

$$\lim_{t\to 0}\frac{f((0,0)+t\,v)-f(0,0)}{t}=\lim_{t\to 0}\frac{f(t\cos\alpha,t\,\mathrm{sen}\,\alpha)-0}{t}=$$
$$=\lim_{t\to 0}\frac{1}{t}\cdot\frac{t^2\cos^2\alpha\cdot t\,\mathrm{sen}\,\alpha}{t^4\cos^4\alpha+t^2\,\mathrm{sen}^2\,\alpha}=\lim_{t\to 0}\frac{t^3\cos^2\alpha\cdot\mathrm{sen}\,\alpha}{t^3\,(t^2\cos^4\alpha+\mathrm{sen}^2\,\alpha)}=\frac{\cos^2\alpha\cdot\mathrm{sen}\,\alpha}{\mathrm{sen}^2\,\alpha}=\frac{\cos^2\alpha}{\mathrm{sen}\,\alpha},$$
purchè sia $\mathrm{sen}\,\alpha\neq 0$, cioè $\alpha\neq 0$ e $\alpha\neq\pi$.
Se $\alpha=0$ o $\alpha=\pi$, la direzione è quella dell'asse x, quindi la derivata direzionale è la derivata parziale rispetto a x, per la quale risulta: $\lim\limits_{t\to 0}\dfrac{f(0+t,0)-f(0,0)}{t}=\lim\limits_{t\to 0}\dfrac{1}{t}\cdot\dfrac{t^2\cdot 0}{t^4+0}=0$.

Quindi la funzione dell'esempio 33 è derivabile in ogni direzione pur non essendo continua nel punto $(0,0)$. Non è più quindi necessario essere funzione continua per essere funzione derivabile (parzialmente o in una direzione). Il legame trovato per le funzioni di una sola variabile verrà ristabilito con la classe delle funzioni differenziabili.

FUNZIONI DIFFERENZIABILI

Richiamiamo il concetto di funzione differenziabile per funzioni $f:\mathbb{R}\to\mathbb{R}$, per estenderlo alle funzioni $f:\mathbb{R}^n\to\mathbb{R}$.
Data $f:\mathbb{R}\to\mathbb{R}$, e dato x_0 punto interno al dominio di f, la funzione $f(x)$ si dice differenziabile in x_0 se esiste una costante $\alpha\in\mathbb{R}$ per la quale vale la relazione:
$f(x)=f(x_0)+\alpha\,(x-x_0)+o(x-x_0)$.
Essere differenziabile significa essere approssimabile in modo lineare, ovvero mediante una retta. Valgono alcuni importanti Teoremi:
-se una funzione è differenziabile in x_0 allora è continua in x_0;
-se una funzione è differenziabile in x_0 allora è derivabile in x_0, e risulta $\alpha=f'(x_0)$;
-una funzione è differenziabile in x_0 se e solo se è derivabile in x_0.
Si può quindi riformulare la differenziabilità in x_0 mediante l'uguaglianza:
$f(x)=f(x_0)+f'(x_0)\,(x-x_0)+o(x-x_0)$.
Essere differenziabile significa quindi essere approssimabile mediante la retta tangente al grafico di $f(x)$ nel punto x_0, con un errore, $o(x-x_0)$, che è trascurabile rispetto all'incremento $x-x_0$, ovvero tale che $\lim\limits_{x\to x_0}\dfrac{o(x-x_0)}{x-x_0}=0$.
Passiamo ora alla differenziabilità per funzioni $f:\mathbb{R}^n\to\mathbb{R}$. Diamo la seguente:
Definizione 32 : Data $f:\mathbb{R}^n\to\mathbb{R}$, e dato $\mathbb{X}_0$ punto interno al dominio di f, la funzione si dice **differenziabile** in $\mathbb{X}_0$ se esiste un vettore a termini costanti $\mathbb{K}\in\mathbb{R}^n$ per il quale vale l'uguaglianza: $f(\mathbb{X})=f(\mathbb{X}_0)+\mathbb{K}\cdot(\mathbb{X}-\mathbb{X}_0)+o(\|\mathbb{X}-\mathbb{X}_0\|)$,
dove il prodotto $\mathbb{K}\cdot(\mathbb{X}-\mathbb{X}_0)$ è il prodotto scalare di due vettori di $\mathbb{R}^n$.
L'errore commesso con tale approssimazione deve risultare trascurabile rispetto a $\|\mathbb{X}-\mathbb{X}_0\|$, modulo dell'incremento $\mathbb{X}-\mathbb{X}_0$.
Vista la definizione di "o piccolo", la definizione di funzione differenziabile può essere formulata anche nella forma del $\lim\limits_{\mathbb{X}\to\mathbb{X}_0}\dfrac{f(\mathbb{X})-f(\mathbb{X}_0)-\mathbb{K}\cdot(\mathbb{X}-\mathbb{X}_0)}{\|\mathbb{X}-\mathbb{X}_0\|}=0$.
Notiamo come $\mathbb{K}\cdot(\mathbb{X}-\mathbb{X}_0)$ rappresenti un'applicazione lineare $\mathbb{R}^n\to\mathbb{R}$.
Vediamo subito la relazione che intercorre tra la differenziabilità e la continuità, la derivabilità parziale e quella direzionale. Vale anzitutto il seguente:
Teorema 7 : Se f è differenziabile in $\mathbb{X}_0$ allora f è continua in $\mathbb{X}_0$.
Dimostrazione : Essendo per ipotesi f differenziabile in $\mathbb{X}_0$ sarà:
$f(\mathbb{X})-f(\mathbb{X}_0)=\mathbb{K}\cdot(\mathbb{X}-\mathbb{X}_0)+o(\|\mathbb{X}-\mathbb{X}_0\|)$.
Dobbiamo far vedere che $\lim\limits_{\mathbb{X}\to\mathbb{X}_0}f(\mathbb{X})=f(\mathbb{X}_0)$, ovvero che $\lim\limits_{\mathbb{X}\to\mathbb{X}_0}f(\mathbb{X})-f(\mathbb{X}_0)=0$.
Ma $\lim\limits_{\mathbb{X}\to\mathbb{X}_0}f(\mathbb{X})-f(\mathbb{X}_0)=\lim\limits_{\mathbb{X}\to\mathbb{X}_0}(\mathbb{K}\cdot(\mathbb{X}-\mathbb{X}_0)+o(\|\mathbb{X}-\mathbb{X}_0\|))=0$, in quanto:

$\lim_{\mathbb{X}\to\mathbb{X}_0} \mathbb{K}\cdot(\mathbb{X}-\mathbb{X}_0) = 0$, dato che $\mathbb{K}$ è costante mentre $\mathbb{X}\to\mathbb{X}_0$, mentre $o(\|\mathbb{X}-\mathbb{X}_0\|)\to 0$ in conseguenza della definizione di "o piccolo", e quindi il Teorema è dimostrato.●

Essere funzione continua è quindi ancora condizione necessaria per essere funzione differenziabile. Vale poi il seguente:

Teorema 8 : Se f è differenziabile in $\mathbb{X}_0$, punto interno al dominio, allora f ammette in $\mathbb{X}_0$ tutte le derivate parziali, quindi esiste $\nabla f(\mathbb{X}_0)$, e risulta $\mathbb{K} = \nabla f(\mathbb{X}_0)$. Inoltre f è derivabile in $\mathbb{X}_0$ in tutte le direzioni v e risulta: $\mathcal{D}_v f(\mathbb{X}_0) = \nabla f(\mathbb{X}_0)\cdot v$.

Dimostrazione : Vediamo anzitutto che f ammette in $\mathbb{X}_0$ tutte le derivate parziali.

Poniamo $\mathbb{K} = (k_1, k_2, ..., k_n)$ ed avremo:

$$\frac{\partial f(\mathbb{X}_0)}{\partial x_i} = \lim_{t\to 0}\frac{f(\mathbb{X}_0 + t\,e_i) - f(\mathbb{X}_0)}{t} = \lim_{t\to 0}\frac{\mathbb{K}\cdot t\,e_i + o(\|t\,e_i\|)}{t} =$$

$$= \lim_{t\to 0}\frac{t\cdot[(k_1,k_2,...,k_i,...,k_n)\cdot(0,0,...,1_i,...,0)] + o(|t|\cdot\|e_i\|)}{t} = \lim_{t\to 0}\frac{t\,k_i + o(|t|)}{t} =$$

$$= \lim_{t\to 0}\left(k_i + \frac{o(|t|)}{t}\right) = k_i\text{, in quanto } \lim_{t\to 0}\frac{o(|t|)}{t} = 0 \text{ per definizione.}$$

Dato che $k_i\in\mathbb{R}$, ne segue che il limite esiste finito e vale k_i, ovvero risulta $\frac{\partial f(\mathbb{X}_0)}{\partial x_i} = k_i$ e, data l'arbitrarietà di i, si ha pure che $\mathbb{K} = \nabla f(\mathbb{X}_0)$.

Quindi la differenziabilità di f in $\mathbb{X}_0$ può essere espressa come:

$f(\mathbb{X}) = f(\mathbb{X}_0) + \nabla f(\mathbb{X}_0)\cdot(\mathbb{X}-\mathbb{X}_0) + o(\|\mathbb{X}-\mathbb{X}_0\|)$.

L'equazione $y = f(\mathbb{X}_0) + \nabla f(\mathbb{X}_0)\cdot(\mathbb{X}-\mathbb{X}_0)$ è quella del piano (se $f:\mathbb{R}^2\to\mathbb{R}$) o dell'iperpiano (se $n>2$) tangente alla superficie in $(\mathbb{X}_0, f(\mathbb{X}_0))$.

Essere differenziabile significa quindi essere approssimabile mediante il piano (o l'iperpiano) tangente con un errore che risulta trascurabile rispetto a $\|\mathbb{X}-\mathbb{X}_0\|$.

Passiamo infine a calcolare la derivata di f in $\mathbb{X}_0$ in una qualsiasi direzione v.

Dobbiamo calcolare:

$$\mathcal{D}_v f(\mathbb{X}_0) = \lim_{t\to 0}\frac{f(\mathbb{X}_0 + t\,v) - f(\mathbb{X}_0)}{t} = \lim_{t\to 0}\frac{\mathbb{K}\cdot(\mathbb{X}_0 + t\,v - \mathbb{X}_0) + o(\|\mathbb{X}_0 + t\,v - \mathbb{X}_0\|)}{t} =$$

$$= \lim_{t\to 0}\frac{t\,\nabla f(\mathbb{X}_0)\cdot v + o(|t|)}{t} = \lim_{t\to 0}\left(\nabla f(\mathbb{X}_0)\cdot v + \frac{o(|t|)}{t}\right) = \nabla f(\mathbb{X}_0)\cdot v = \mathcal{D}_v f(\mathbb{X}_0).$$

Quindi la funzione risulta derivabile in ogni direzione v, e vediamo anche che per calcolare derivate direzionali non occorre operare il limite previsto dalla definizione, ma basta invece eseguire il prodotto scalare tra il gradiente di f nel punto $\mathbb{X}_0$ ed il versore v; quindi basta conoscere le n derivate parziali di f in $\mathbb{X}_0$. Questo naturalmente se la funzione è differenziabile in $\mathbb{X}_0$.●

SIGNIFICATO DEL GRADIENTE

Dalla formula di Schwarz sappiamo che $\mathbb{X}\cdot\mathbb{Y} = \|\mathbb{X}\|\cdot\|\mathbb{Y}\|\cdot\cos\alpha$, dove $\alpha: 0\le\alpha<\pi$, rappresenta l'angolo compreso tra i vettori $\mathbb{X}$ e $\mathbb{Y}$. Se f è differenziabile in $\mathbb{X}_0$, essendo v un versore si ha: $\mathcal{D}_v f(\mathbb{X}_0) = \nabla f(\mathbb{X}_0)\cdot v = \|\nabla f(\mathbb{X}_0)\|\cdot\|v\|\cdot\cos\alpha = \|\nabla f(\mathbb{X}_0)\|\cdot\cos\alpha$.

Se $\alpha = 0$, allora $\cos\alpha = 1$ e risulta $\mathcal{D}_v f(\mathbb{X}_0) = \|\nabla f(\mathbb{X}_0)\|$, e questo è il massimo valore che $\mathcal{D}_v f(\mathbb{X}_0)$ può assumere; ma $\alpha = 0$ significa che $\nabla f(\mathbb{X}_0)$ e v stanno sulla stessa retta e sono orientati dalla stessa parte; dato che v è la direzione in cui si deriva, e dato che $\nabla f(\mathbb{X}_0)$ esprime la stessa direzione, possiamo concludere che $\nabla f(\mathbb{X}_0)$ esprime la direzione di massimo accrescimento per f in $\mathbb{X}_0$.

Analogamente, se $\alpha = \pi$, $\nabla f(\mathbb{X}_0)$ e v stanno sulla stessa retta ma sono orientati in direzioni opposte, ed essendo $\cos\alpha = -1$, abbiamo $\mathcal{D}_v f(\mathbb{X}_0) = -\|\nabla f(\mathbb{X}_0)\|$, e questo è il minimo valore che $\mathcal{D}_v f(\mathbb{X}_0)$ può assumere.

CONDIZIONI PER LA DIFFERENZIABILITA'

L'esistenza del gradiente in un punto è condizione necessaria, ma non sufficiente, ad assicurare la differenziabilità della funzione nel punto stesso.

La funzione $f(x,y)=\begin{cases}1: & x\,y=0\\ 0: & x\,y\neq 0\end{cases}$ dell'Esempio 32 ci fornisce un esempio di funzione che in $(0,0)$ ammette le derivate parziali, e quindi il gradiente, senza essere continua, e quindi, per il Teorema 7, senza essere differenziabile.

Esempio 34 : Vediamo se $f(x,y)=\sqrt{|x\,y|}$ risulta differenziabile in $(0,0)$. Per prima cosa calcoliamo $\nabla f(0,0)$ ed abbiamo:

$$\frac{\partial f(0,0)}{\partial x}=\lim_{h\to 0}\frac{f(0+h,0)-f(0,0)}{h}=\lim_{h\to 0}\frac{\sqrt{|(0+h)\cdot 0|}-0}{h}=0\,,$$
$$\frac{\partial f(0,0)}{\partial y}=\lim_{h\to 0}\frac{f(0,0+h)-f(0,0)}{h}=\lim_{h\to 0}\frac{\sqrt{|0\cdot(0+h)|}-0}{h}=0\,.$$

Quindi $\nabla f(0,0)=(0,0)$. Se la funzione fosse differenziabile in $(0,0)$ dovrà essere:

$$f(x,y)-f(0,0)=\frac{\partial f(0,0)}{\partial x}\cdot(x-0)+\frac{\partial f(0,0)}{\partial y}\cdot(y-0)+o\left(\sqrt{(x-0)^2+(y-0)^2}\right)$$

ovvero dovrà essere $\displaystyle\lim_{(x,y)\to(0,0)}\frac{\sqrt{|x\,y|}}{\sqrt{x^2+y^2}}=0$; passando a coordinate polari avremo:

$\displaystyle\lim_{\rho\to 0}\frac{\sqrt{|\rho^2\cos\vartheta\cdot\operatorname{sen}\vartheta|}}{\sqrt{\rho^2}}=\lim_{\rho\to 0}\sqrt{|\cos\vartheta\cdot\operatorname{sen}\vartheta|}=\sqrt{|\cos\vartheta\cdot\operatorname{sen}\vartheta|}$, che vale 0 solo nel caso che sia $\vartheta=k\,\dfrac{\pi}{2}$. Anche se esiste $\nabla f(0,0)$, la funzione non è quindi differenziabile in $(0,0)$.

Vediamo ora come l'esistenza con l'aggiunta della continuità delle derivate parziali in $\mathbb{X}_0$, risulti sufficiente a garantire la differenziabilità della funzione in $\mathbb{X}_0$. Vale infatti il seguente:

Teorema 9 (del differenziale totale) : Se la funzione ammette in $\mathbb{X}_0$ tutte le derivate parziali, ovvero se esiste $\nabla f(\mathbb{X}_0)$, e se le funzioni derivate parziali sono continue in $\mathbb{X}_0$, allora f è differenziabile in $\mathbb{X}_0$.

Per Dimostrazione e controesempio vedere Appendice n.1

Esempio 35 : Data $f(x,y)=e^{x-y}$, calcoliamo $\mathcal{D}_v f(0,0)$, con $v=\left(\dfrac{1}{\sqrt{2}}\,,\,-\dfrac{1}{\sqrt{2}}\right)$.

La funzione è data dalla composizione di una esponenziale con il polinomio $x-y$, quindi è continua e derivabile $\forall\,(x,y)\in\mathbb{R}^2$. Essendo $\dfrac{\partial f(x,y)}{\partial x}=e^{x-y}$ e $\dfrac{\partial f(x,y)}{\partial y}=-\,e^{x-y}$, anche le due funzioni derivate parziali sono continue $\forall\,(x,y)\in\mathbb{R}^2$, e quindi la funzione data è differenziabile in tutto $\mathbb{R}^2$. Essendo $\dfrac{\partial f(0,0)}{\partial x}=1$ e $\dfrac{\partial f(0,0)}{\partial y}=-\,1$, otteniamo:

$$\mathcal{D}_v f(0,0)=\nabla f(0,0)\cdot\left(\frac{1}{\sqrt{2}}\,,\,-\frac{1}{\sqrt{2}}\right)=(1,-1)\cdot\left(\frac{1}{\sqrt{2}}\,,\,-\frac{1}{\sqrt{2}}\right)=\sqrt{2}\,.$$

DERIVATE PARZIALI E DIREZIONALI SUCCESSIVE

Come visto negli esempi, se $f:\mathbb{R}^n\to\mathbb{R}$ è derivabile, essa ammette n funzioni derivate parziali, ciascuna delle quali è ancora una funzione $\dfrac{\partial f(\mathbb{X})}{\partial x_i}:\mathbb{R}^n\to\mathbb{R}$, che, se risulta derivabile, può essere derivata rispetto a ciascuna delle sue n variabili: $\dfrac{\partial}{\partial x_j}\left(\dfrac{\partial f(\mathbb{X})}{\partial x_i}\right)$. Si ottengono co-

sì $n^2 = n \cdot n$ derivate parziali seconde che, nell'ipotesi di derivabilità, saranno a loro volta derivabili rispetto alle n variabili, dando luogo a $n^3 = n^2 \cdot n$ derivate parziali terze e così via.
Possiamo quindi comporre il seguente schema, valido per $f : \mathbb{R}^2 \to \mathbb{R}, (x,y) \to f(x,y)$:

$$f(x,y) \Rightarrow \begin{cases} \dfrac{\partial f(x,y)}{\partial x} \Rightarrow \begin{cases} \dfrac{\partial^2 f(x,y)}{\partial x^2} = f''_{xx} \Rightarrow \begin{cases} \dfrac{\partial^3 f(x,y)}{\partial x\, \partial x\, \partial x} = \dfrac{\partial^3 f(x,y)}{\partial x^3} = f'''_{xxx} \\ \dfrac{\partial^3 f(x,y)}{\partial x\, \partial x\, \partial y} = \dfrac{\partial^3 f(x,y)}{\partial x^2\; \partial y} = f'''_{xxy} \end{cases} \\ \dfrac{\partial^2 f(x,y)}{\partial x\, \partial y} = f''_{xy} \Rightarrow \begin{cases} \dfrac{\partial^3 f(x,y)}{\partial x\, \partial y\, \partial x} = f'''_{xyx} \\ \dfrac{\partial^3 f(x,y)}{\partial x\, \partial y\, \partial y} = \dfrac{\partial^3 f(x,y)}{\partial x\, \partial y^2} = f'''_{xyy} \end{cases} \end{cases} \\ \dfrac{\partial f(x,y)}{\partial y} \Rightarrow \begin{cases} \dfrac{\partial^2 f(x,y)}{\partial y\, \partial x} = f''_{yx} \Rightarrow \begin{cases} \dfrac{\partial^3 f(x,y)}{\partial y\, \partial x\, \partial x} = \dfrac{\partial^3 f(x,y)}{\partial y\, \partial x^2} = f'''_{yxx} \\ \dfrac{\partial^3 f(x,y)}{\partial y\, \partial x\, \partial y} = f'''_{yxy} \end{cases} \\ \dfrac{\partial^2 f(x,y)}{\partial y^2} = f''_{yy} \Rightarrow \begin{cases} \dfrac{\partial^3 f(x,y)}{\partial y\, \partial y\, \partial x} = \dfrac{\partial^3 f(x,y)}{\partial y^2\, \partial x} = f'''_{yyx} \\ \dfrac{\partial^3 f(x,y)}{\partial y\, \partial y\, \partial y} = \dfrac{\partial^3 f(x,y)}{\partial y^3} = f'''_{yyy} \end{cases} \end{cases} \end{cases}$$

Si noti la diversa posizione dello pseudo-esponente 2: in $\dfrac{\partial^2 f(x,y)}{\partial x^2}$ esso significa derivare 2 volte $(\partial^2 f)$ la funzione f rispetto alla variabile x ambedue le volte (∂x^2), spiegando così il diverso posizionamento.

Le derivate fatte rispetto alla stessa variabile, $\dfrac{\partial^2 f(x,y)}{\partial x^2}$ e $\dfrac{\partial^2 f(x,y)}{\partial y^2}$, vengono dette **pure**, mentre $\dfrac{\partial^2 f(x,y)}{\partial x\, \partial y}$ e $\dfrac{\partial^2 f(x,y)}{\partial y\, \partial x}$, quelle in cui cambia la variabile di derivazione, vengono dette **miste**. Tornando al caso generale, $f : \mathbb{R}^n \to \mathbb{R}$, analogamente alle derivate parziali del secondo e degli ordini successivi, si possono definire le derivate direzionali del secondo e degli ordini successivi. Se vogliamo introdurre formalmente una derivata parziale seconda la dobbiamo definire come:
$\dfrac{\partial^2 f(\mathbb{X}_0)}{\partial x_i\, \partial x_j} = \lim\limits_{t \to 0} \dfrac{1}{t} \cdot \left(\dfrac{\partial f(\mathbb{X}_0 + t\, e_j)}{\partial x_i} - \dfrac{\partial f(\mathbb{X}_0)}{\partial x_i} \right)$, ovvero come il limite, se esiste ed è finito, del rapporto incrementale di una derivata parziale prima. Analogamente, una derivata direzionale del secondo ordine viene definita come il limite, se esiste ed è finito, del rapporto incrementale di una derivata direzionale prima:
$\mathcal{D}^2_{v\,w} f(\mathbb{X}_0) = \lim\limits_{t \to 0} \dfrac{\mathcal{D}_v f(\mathbb{X}_0 + t\, w) - \mathcal{D}_v f(\mathbb{X}_0)}{t} = \mathcal{D}_w(\mathcal{D}_v f(\mathbb{X}_0))$.
Similmente si definiscono le derivate direzionali degli ordini successivi.
In generale non è vero che $\dfrac{\partial^2 f(\mathbb{X}_0)}{\partial x_i\, \partial x_j} = \dfrac{\partial^2 f(\mathbb{X}_0)}{\partial x_j\, \partial x_i}$, nè che $\mathcal{D}^2_{v\,w} f(\mathbb{X}_0) = \mathcal{D}^2_{w\,v} f(\mathbb{X}_0)$, come mostrato nel seguente:

Esempio 36 : Sia $f(x,y) = \begin{cases} x\, y \cdot \dfrac{x^2 - y^2}{x^2 + y^2} & (x,y) \neq (0,0) \\ 0 & (x,y) = (0,0) \end{cases}$. Vediamo che la funzione è continua in $(0,0)$; infatti, usando le coordinate polari si ha:

$\lim_{\rho\to 0}\left(\rho^2\cos\vartheta\cdot\operatorname{sen}\vartheta\right)\cdot\dfrac{\rho^2\cos^2\vartheta-\rho^2\operatorname{sen}^2\vartheta}{\rho^2\cos^2\vartheta+\rho^2\operatorname{sen}^2\vartheta}=\lim_{\rho\to 0}\left(\rho^2\cos\vartheta\cdot\operatorname{sen}\vartheta\cdot\cos 2\vartheta\right)=0$, in modo uniforme in quanto $\left|\rho^2\cos\vartheta\cdot\operatorname{sen}\vartheta\cdot\cos 2\vartheta\right|\le\rho^2\left|\cos\vartheta\cdot\operatorname{sen}\vartheta\cdot\cos 2\vartheta\right|\le\rho^2$, per cui se prendiamo $\rho\le\delta(\varepsilon)=\sqrt{\varepsilon}$ abbiamo la verifica del limite.

La funzione ammette poi le derivate parziali nel punto $(0,0)$; infatti:

$$\frac{\partial f(0,0)}{\partial x}=\lim_{h\to 0}\frac{(0+h)\cdot 0\cdot\dfrac{(0+h)^2-0^2}{(0+h)^2+0^2}-0}{h}=\lim_{h\to 0}\frac{0}{h}=0\,,$$

$$\frac{\partial f(0,0)}{\partial y}=\lim_{h\to 0}\frac{0\cdot(0+h)\cdot\dfrac{0^2-(0+h)^2}{0^2+(0+h)^2}-0}{h}=\lim_{h\to 0}\frac{0}{h}=0\,.$$

Passando alle funzioni derivate parziali prime avremo:

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

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

Da queste ricaviamo infine:

$$\frac{\partial^2 f(0,0)}{\partial x\,\partial y}=\lim_{h\to 0}\frac{\dfrac{\partial f(0{,}0+h)}{\partial x}-\dfrac{\partial f(0,0)}{\partial x}}{h}=\lim_{h\to 0}\frac{h\cdot\dfrac{0-h^4+0}{h^4}-0}{h}=\,-\,1\,,\text{ e}$$

$$\frac{\partial^2 f(0,0)}{\partial y\,\partial x}=\lim_{h\to 0}\frac{\dfrac{\partial f(0+h{,}0)}{\partial y}-\dfrac{\partial f(0,0)}{\partial y}}{h}=\lim_{h\to 0}\frac{h\cdot\dfrac{h^4-0-0}{h^4}-0}{h}=\,+\,1\,.$$

Quindi per questa funzione si ha che $\dfrac{\partial^2 f(0,0)}{\partial x\,\partial y}\neq\dfrac{\partial^2 f(0,0)}{\partial y\,\partial x}$.

Vale però il seguente:

Teorema 10 (di Schwarz) : Se $f:\mathbb{R}^n\to\mathbb{R}$ ammette in un intorno del punto $\mathbb{X}_0$ le derivate seconde miste $\dfrac{\partial^2 f(\mathbb{X})}{\partial x_i\,\partial x_j}$ e $\dfrac{\partial^2 f(\mathbb{X})}{\partial x_j\,\partial x_i}$, e se queste sono continue nel punto $\mathbb{X}_0$, allora sono uguali: $\dfrac{\partial^2 f(\mathbb{X}_0)}{\partial x_i\,\partial x_j}=\dfrac{\partial^2 f(\mathbb{X}_0)}{\partial x_j\,\partial x_i}$.

Si omette la dimostrazione di questo Teorema.

Esiste anche una forma più generale di questo Teorema, le cui ipotesi prevedono esistenza e continuità di una sola delle due derivate seconde miste; si dimostra allora che l'altra derivata seconda mista esiste, è continua ed è uguale alla prima.

Esempio 37 : Sia $f(x,y)=x\,e^y-y\,e^x$. Avremo allora:

$$f(x,y)=x\,e^y-y\,e^x\Rightarrow\begin{cases}f'_x=e^y-y\,e^x\Rightarrow\begin{cases}f''_{xx}=\,-\,y\,e^x\\ f''_{xy}=e^y-e^x\end{cases}\\ f'_y=x\,e^y-e^x\Rightarrow\begin{cases}f''_{yx}=e^y-e^x\\ f''_{yy}=x\,e^y\end{cases}\end{cases}\text{, e quindi } f''_{xy}=f''_{yx}\,.$$

Il Teorema di Schwarz esprime comunque una condizione sufficiente, e non necessaria, per l'uguaglianza delle derivate seconde miste.

Per una funzione di n variabili le derivate seconde sono in numero di n^2. Dato che n sono pure, saranno $n^2 - n$ quelle miste, solo la metà delle quali dovrà essere calcolata, se vale il Teorema di Schwarz, per cui le derivate seconde si riducono a $n^2 - \dfrac{n^2 - n}{2} = \dfrac{n^2 + n}{2}$.

Osservazione 1 : Il Teorema di Schwarz non vale solo per le derivate parziali seconde, ma vale per le derivate miste di qualsiasi ordine, dato che una derivata di ordine n è pur sempre la derivata seconda di una derivata di ordine $n - 2$: $\partial^n f = \partial^2 \left(\partial^{n-2} f\right)$. Adeguando le ipotesi del Teorema di Schwarz alla continuità delle derivate miste di ordine opportuno, scriveremo allora, nel caso di derivazioni fatte rispetto a due sole variabili, $\dfrac{\partial^m f(\mathbb{X})}{\partial x_i^p \, \partial x_j^q}$, con $p + q = m$, per indicare la derivata parziale mista di ordine m, ottenuta derivando p volte rispetto a x_i e q volte rispetto a x_j, con $p + q = m$, senza bisogno di specificare l'ordine con cui si è derivato rispetto a x_i e a x_j, essendo questo ininfluente alla luce del Teorema di Schwarz.

Esempio 38 : Data $f : \mathbb{R}^n \to \mathbb{R}$, se le sue funzioni derivate terze sono continue, risulta:

$$\frac{\partial^3 f(\mathbb{X})}{\partial x_i^2 \, \partial x_j} = \frac{\partial^3 f(\mathbb{X})}{\partial x_i \, \partial x_i \, \partial x_j} = \frac{\partial^3 f(\mathbb{X})}{\partial x_i \, \partial x_j \, \partial x_i} = \frac{\partial^3 f(\mathbb{X})}{\partial x_j \, \partial x_i \, \partial x_i} = \frac{\partial^3 f(\mathbb{X})}{\partial x_j \, \partial x_i^2}.$$

Osservazione 2 : Il Teorema di Schwarz non vale solo per le derivate parziali, ma per tutte le derivate direzionali almeno del secondo ordine. Se $\mathcal{D}^2_{v\,w} f(\mathbb{X})$ esiste ed è continua in $\mathbb{X}_0$, allora $\mathcal{D}^2_{w\,v} f(\mathbb{X}_0)$ esiste e risulta $\mathcal{D}^2_{v\,w} f(\mathbb{X}_0) = \mathcal{D}^2_{w\,v} f(\mathbb{X}_0)$. Analoghe conclusioni valgono per le derivate direzionali degli ordini superiori.

DIFFERENZIABILITA' DEL SECONDO E DEGLI ORDINI SUCCESSIVI

Sia data $f : \mathbb{R}^n \to \mathbb{R}$ e supponiamo che tale funzione sia differenziabile $\forall\, \mathbb{X} \in \mathbb{A}$, $\mathbb{A} \subseteq D_f$. Esistono quindi in $\mathbb{X}_0 \in \mathbb{A}$ le n funzioni derivate parziali $\dfrac{\partial f(\mathbb{X})}{\partial x_i} : \mathbb{R}^n \to \mathbb{R}$. Diamo la seguente:

Definizione 33 : La funzione $f(\mathbb{X})$ si dice **differenziabile due volte** in $\mathbb{X}_0$ se ciascuna delle sue funzioni derivate prime $\dfrac{\partial f(\mathbb{X})}{\partial x_i}$ risulta differenziabile in $\mathbb{X}_0$.

Essere differenziabile due volte significa quindi avere derivate prime differenziabili. E' facile quindi estendere la definizione al caso di una funzione differenziabile k volte:

Definizione 34 : La funzione $f(\mathbb{X})$ si dice **differenziabile k volte** in $\mathbb{X}_0$ se ciascuna delle sue funzioni derivate di ordine $k - 1$ risulta differenziabile in $\mathbb{X}_0$.

Per le funzioni differenziabili due volte vale il seguente Teorema, che ci limitiamo ad enunciare:

Teorema 11 : Se una funzione è differenziabile due volte in $\mathbb{X}_0$, punto interno a D_f, allora esistono in $\mathbb{X}_0$ tutte le derivate direzionali del secondo ordine. Inoltre le derivate miste (parziali e direzionali) in $\mathbb{X}_0$ risultano uguali: $\dfrac{\partial^2 f(\mathbb{X})}{\partial x_i \, \partial x_j} = \dfrac{\partial^2 f(\mathbb{X})}{\partial x_j \, \partial x_i}$ e $\mathcal{D}^2_{v\,w} f(\mathbb{X}_0) = \mathcal{D}^2_{w\,v} f(\mathbb{X}_0)$.

Quest'ultimo Teorema fornisce un'ulteriore condizione sufficiente, e non necessaria, per l'uguaglianza delle derivate miste in aggiunta a quella stabilita con il Teorema di Schwarz.

Analogamente a quanto visto per la differenziabilità del primo ordine, non è necessario, ma solo sufficiente, che le derivate seconde siano continue affinchè la funzione risulti differenziabile due volte. Questo spiega le due diverse ipotesi che rendono $\dfrac{\partial^2 f(\mathbb{X})}{\partial x_i\,\partial x_j} = \dfrac{\partial^2 f(\mathbb{X})}{\partial x_j\,\partial x_i}$ oppure $\mathcal{D}^2_{v\,w} f(\mathbb{X}_0) = \mathcal{D}^2_{w\,v} f(\mathbb{X}_0)$.

DIFFERENZIALI TOTALI DEL PRIMO E DEGLI ORDINI SUCCESSIVI

Abbiamo visto che l'essere funzione differenziabile in $\mathbb{X}_0$ può esprimersi nella forma:
$f(\mathbb{X}) = f(\mathbb{X}_0) + \nabla f(\mathbb{X}_0) \cdot (\mathbb{X} - \mathbb{X}_0) + o(\|\mathbb{X} - \mathbb{X}_0\|)$.
In questa il termine $\nabla f(\mathbb{X}_0) \cdot (\mathbb{X} - \mathbb{X}_0)$ prende il nome di differenziale totale del primo ordine, e viene indicato anche con $\mathrm{d}f(\mathbb{X}_0) = \nabla f(\mathbb{X}_0) \cdot (\mathbb{X} - \mathbb{X}_0)$.
Posto $\nabla f(\mathbb{X}_0) = \left(\dfrac{\partial f(\mathbb{X}_0)}{\partial x_1}, \dfrac{\partial f(\mathbb{X}_0)}{\partial x_2}, ..., \dfrac{\partial f(\mathbb{X}_0)}{\partial x_n} \right) = (f_1'(\mathbb{X}_0), f_2'(\mathbb{X}_0), \dots, f_n'(\mathbb{X}_0))$ e posto anche $(\mathbb{X} - \mathbb{X}_0) = \left(x_1 - x_1^0, x_2 - x_2^0, ..., x_n - x_n^0\right) = (\mathrm{d}x_1, \mathrm{d}x_2, \dots, \mathrm{d}x_n)$, con queste notazioni possiamo scrivere:

$$\mathrm{d}f(\mathbb{X}_0) = (f_1'(\mathbb{X}_0), f_2'(\mathbb{X}_0), \dots, f_n'(\mathbb{X}_0)) \cdot (\mathrm{d}x_1, \mathrm{d}x_2, \dots, \mathrm{d}x_n) = \sum_{i=1}^{n} f_i'(\mathbb{X}_0) \cdot \mathrm{d}x_i .$$

Se non interessa specificare il punto $\mathbb{X}_0$, potremo scrivere anche:

$$\mathrm{d}f = (f_1', f_2', .., f_n') \cdot (\mathrm{d}x_1, \mathrm{d}x_2, \dots, \mathrm{d}x_n) = \sum_{i=1}^{n} f_i' \mathrm{d}x_i = f_1'\,\mathrm{d}x_1 + f_2'\,\mathrm{d}x_2 + \ \dots\ + f_n'\,\mathrm{d}x_n .$$

Quest'ultima espressione definisce il **differenziale totale** del I ordine per una funzione di n variabili. Nel caso di $f : \mathbb{R}^2 \to \mathbb{R}$, otteniamo $\mathrm{d}f = f_1'\,\mathrm{d}x_1 + f_2'\,\mathrm{d}x_2$, se le variabili sono indicate con x_1 e x_2, oppure $\mathrm{d}f = f_x'\,\mathrm{d}x + f_y'\,\mathrm{d}y$, se invece le indichiamo con x e y.
Per una funzione di tre variabili $f(x, y, z)$ avremo $\mathrm{d}f = f_x'\,\mathrm{d}x + f_y'\,\mathrm{d}y + f_z'\,\mathrm{d}z$.
Passiamo ora a definire il differenziale totale del II ordine per $f : \mathbb{R}^2 \to \mathbb{R}$. Avremo:
$\mathrm{d}^2 f = \mathrm{d}(\mathrm{d}f) = \dfrac{\partial(\mathrm{d}f)}{\partial x}\,\mathrm{d}x + \dfrac{\partial(\mathrm{d}f)}{\partial y}\,\mathrm{d}y$. Tenendo presente che f_x' e f_y' sono funzioni di x e di y, mentre $\mathrm{d}x$ e $\mathrm{d}y$ sono invece costanti rispetto a x ed a y, otteniamo:

$$\mathrm{d}^2 f = \frac{\partial\left(f_x'\,\mathrm{d}x + f_y'\,\mathrm{d}y\right)}{\partial\,x}\,\mathrm{d}x + \frac{\partial\left(f_x'\,\mathrm{d}x + f_y'\,\mathrm{d}y\right)}{\partial\,y}\,\mathrm{d}y =$$

$$= \left(f_{xx}''\,\mathrm{d}x + f_{yx}''\,\mathrm{d}y\right)\mathrm{d}x + \left(f_{xy}''\,\mathrm{d}x + f_{yy}''\,\mathrm{d}y\right)\mathrm{d}y = f_{xx}''\,(\mathrm{d}x)^2 + 2\,f_{xy}''\,\mathrm{d}x\,\mathrm{d}y + f_{yy}''\,(\mathrm{d}y)^2 ,$$

se supponiamo che la funzione sia differenziabile due volte, per cui $f_{xy}'' = f_{yx}''$.
Nel caso di $f : \mathbb{R}^3 \to \mathbb{R}$ analogamente otteniamo:

$$\mathrm{d}^2 f = f_{xx}''\,(\mathrm{d}x)^2 + f_{yy}''\,(\mathrm{d}y)^2 + f_{zz}''\,(\mathrm{d}z)^2 + 2f_{xy}''\,\mathrm{d}x\,\mathrm{d}y + 2f_{xz}''\,\mathrm{d}x\,\mathrm{d}z + 2f_{yz}''\,\mathrm{d}y\,\mathrm{d}z .$$

Per una funzione $f : \mathbb{R}^n \to \mathbb{R}, y = f(x_1, x_2, ..., x_n)$, scriveremo invece, in forma compatta:

$$\mathrm{d}^2 f = \mathrm{d}^2 y = \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 f(\mathbb{X})}{\partial x_i\,\partial x_j}\,\mathrm{d}x_i\,\mathrm{d}x_j = \sum_{i,j=1}^{n} \frac{\partial^2 f(\mathbb{X})}{\partial x_i\,\partial x_j}\,\mathrm{d}x_i\,\mathrm{d}x_j .$$

Se la funzione è differenziabile due volte ci saranno da sommare poi tra loro i termini uguali.
Si notino le analogie (non certo l'identità) tra il differenziale secondo di una funzione di due variabili ed il quadrato di un binomio, tra il differenziale secondo di una funzione di tre variabili ed il quadrato di un trinomio, e così via per il differenziale secondo di una funzione di n variabili, analogo con il quadrato di un n-omio. Queste analogie si ritrovano anche nei differenziali totali di ordine superiore al secondo, dovendosi però allora parlare di potenza terza, quarta ecc. di un binomio, trinomio, ecc..
Se calcoliamo infatti il differenziale totale terzo di una funzione $f : \mathbb{R}^2 \to \mathbb{R}$ avremo:

$$\mathrm{d}^3 f = f_{xxx}'''\,(\mathrm{d}x)^3 + 3\,f_{xxy}'''\,(\mathrm{d}x)^2\,\mathrm{d}y + 3\,f_{xyy}'''\,\mathrm{d}x\,(\mathrm{d}y)^2 + f_{yyy}'''\,(\mathrm{d}y)^3 .$$

E per il differenziale totale quarto di $f : \mathbb{R}^2 \to \mathbb{R}$ sarà:

$$\mathrm{d}^4 f = f_{xxxx}^{(4)}\,(\mathrm{d}x)^4 + 4f_{xxxy}^{(4)}\,(\mathrm{d}x)^3\mathrm{d}y + 6f_{xxyy}^{(4)}\,(\mathrm{d}x)^2(\mathrm{d}y)^2 + 4f_{xyyy}^{(4)}\,\mathrm{d}x(\mathrm{d}y)^3 + f_{yyyy}^{(4)}\,(\mathrm{d}y)^4 .$$

Si usa formalizzare quanto detto scrivendo, per il differenziale di ordine m di una funzione di n variabili : $\mathrm{d}^m f = \left(\sum_{i=1}^{n} \frac{\partial}{\partial x_i} \cdot \mathrm{d}x_i\right)^m (f(\mathbb{X}))$, che per quanto detto in precedenza ha solo un significato simbolico.

FORMA VETTORIALE-MATRICIALE DEL DIFFERENZIALE SECONDO

Il differenziale totale secondo di una funzione di un numero qualsiasi di variabili può essere espresso anche in forma vettoriale-matriciale, mediante la cosiddetta matrice Hessiana $\mathbb{H}$.

La **matrice Hessiana** è la matrice costituita dalle derivate parziali seconde, ordinate per riga rispetto alla prima variabile di derivazione e per colonna rispetto alla seconda, e nel caso di una funzione di n variabili assume la forma: $\mathbb{H} = \left\| \begin{matrix} f''_{11} & f''_{12} & ... & ... & f''_{1n} \\ f''_{21} & f''_{22} & ... & ... & f''_{2n} \\ ... & ... & ... & ... & ... \\ ... & ... & ... & ... & ... \\ f''_{n1} & f''_{n2} & ... & ... & f''_{nn} \end{matrix} \right\|$.

Se valgono le ipotesi del Teorema di Schwarz oppure se la funzione è differenziabile due volte, la matrice Hessiana è una matrice simmetrica. Eseguendo il prodotto vettore per matrice per vettore, posto $\mathrm{d}\mathbb{X} = \|\mathrm{d}x_1, \mathrm{d}x_2, ... , \mathrm{d}x_n\|$, si verifica la seguente uguaglianza:

$$\mathrm{d}^2 f = \mathrm{d}\mathbb{X} \cdot \mathbb{H} \cdot \mathrm{d}\mathbb{X}^{\mathrm{T}} = \|\mathrm{d}x_1, \mathrm{d}x_2, ... , \mathrm{d}x_n\| \cdot \left\| \begin{matrix} f''_{11} & f''_{12} & ... & ... & f''_{1n} \\ f''_{21} & f''_{22} & ... & ... & f''_{2n} \\ ... & ... & ... & ... & ... \\ ... & ... & ... & ... & ... \\ f''_{n1} & f''_{n2} & ... & ... & f''_{nn} \end{matrix} \right\| \cdot \left\| \begin{matrix} \mathrm{d}x_1 \\ \mathrm{d}x_2 \\ ... \\ ... \\ \mathrm{d}x_n \end{matrix} \right\|.$$

Verificandola solo nel caso più semplice, quello di una funzione di due sole variabili, si ha:

$$\mathrm{d}^2 f = \|\mathrm{d}x, \mathrm{d}y\| \cdot \left\| \begin{matrix} f''_{xx} & f''_{xy} \\ f''_{yx} & f''_{yy} \end{matrix} \right\| \cdot \left\| \begin{matrix} \mathrm{d}x \\ \mathrm{d}y \end{matrix} \right\| = \left\| \begin{matrix} f''_{xx}\,\mathrm{d}x + f''_{yx}\,\mathrm{d}y \\ f''_{xy}\,\mathrm{d}x + f''_{yy}\,\mathrm{d}y \end{matrix} \right\| \cdot \left\| \begin{matrix} \mathrm{d}x \\ \mathrm{d}y \end{matrix} \right\| =$$
$$= f''_{xx}\,(\mathrm{d}x)^2 + 2\,f''_{xy}\,\mathrm{d}x\,\mathrm{d}y + f''_{yy}\,(\mathrm{d}y)^2.$$

Non esiste invece la possibilità per i differenziali di ordine superiore al secondo di poter essere espressi in forma vettoriale-matriciale.

Se una funzione risulta differenziabile due volte in $\mathbb{X}_0$, sappiamo che esistono in $\mathbb{X}_0$ tutte le derivate direzionali del secondo ordine. Con la matrice Hessiana, detti $v = (v_1, v_2, ..., v_n)$ e $w = (w_1, w_2, ..., w_n)$ due generici versori di $\mathbb{R}^n$, per quanto riguarda il calcolo pratico di una derivata direzionale del secondo ordine vale un risultato simile a quello trovato per le derivate direzionali del primo ordine: $\mathcal{D}_v f(\mathbb{X}_0) = \nabla f(\mathbb{X}_0) \cdot v$. Vale infatti il seguente:

Teorema 12 : Sia $f(\mathbb{X})$ differenziabile due volte in $\mathbb{X}_0$. Allora:
$\mathcal{D}^2_{v\,w} f(\mathbb{X}_0) = v \cdot \mathbb{H}(\mathbb{X}_0) \cdot w^{\mathrm{T}} = w \cdot \mathbb{H}(\mathbb{X}_0) \cdot v^{\mathrm{T}} = \mathcal{D}^2_{w\,v} f(\mathbb{X}_0)$.

Esempio 39 : Sia $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x,y) = e^{x-y}$. Calcoliamo $\mathcal{D}^2_{v\,w} f(x,y)$, mediante due procedure diverse, lasciando generici v, w e (x,y).

Posto $v = (\cos\alpha, \operatorname{sen}\alpha)$ e $w = (\cos\beta, \operatorname{sen}\beta)$, vediamo come la funzione $f(x,y) = e^{x-y}$, essendo composizione di una esponenziale e di un polinomio, che sono funzioni continue e derivabili con derivate continue di qualsiasi ordine, sia differenziabile due volte in tutto $\mathbb{R}^2$. Quindi $\mathcal{D}^2_{v\,w} f(x,y) = v \cdot \mathbb{H}(x,y) \cdot w^{\mathrm{T}}$.

Avremo allora:

$\frac{\partial f}{\partial x} = e^{x-y}$; $\frac{\partial f}{\partial y} = -\,e^{x-y}$; $\frac{\partial^2 f}{\partial x^2} = e^{x-y}$; $\frac{\partial^2 f}{\partial x\,\partial y} = \frac{\partial^2 f}{\partial y\,\partial x} = -\,e^{x-y}$; $\frac{\partial^2 f}{\partial y^2} = e^{x-y}$, da cui:

$$\mathcal{D}^2_{v\,w} f(x,y) = \|\cos\alpha \;\; \operatorname{sen}\alpha\| \cdot \left\| \begin{matrix} e^{x-y} & -\,e^{x-y} \\ -\,e^{x-y} & e^{x-y} \end{matrix} \right\| \cdot \left\| \begin{matrix} \cos\beta \\ \operatorname{sen}\beta \end{matrix} \right\| =$$

$$= \|\cos\alpha \ \ \mathrm{sen}\,\alpha\| \cdot \left\| \begin{matrix} e^{x-y}\cos\beta - e^{x-y}\,\mathrm{sen}\,\beta \\ -\,e^{x-y}\cos\beta + e^{x-y}\,\mathrm{sen}\,\beta \end{matrix} \right\| =$$
$$= e^{x-y}\cos\beta\cdot\cos\alpha - e^{x-y}\,\mathrm{sen}\,\beta\cdot\cos\alpha - e^{x-y}\cos\beta\cdot\mathrm{sen}\,\alpha + e^{x-y}\,\mathrm{sen}\,\beta\cdot\mathrm{sen}\,\alpha =$$
$$= e^{x-y}\cos(\alpha-\beta) - e^{x-y}\,\mathrm{sen}\,(\alpha+\beta) = e^{x-y}\left(\cos(\alpha-\beta) - \mathrm{sen}\,(\alpha+\beta)\right).$$

In alternativa, possiamo calcolare la derivata direzionale seconda come derivata direzionale della derivata direzionale prima, e quindi avremo:

$$\mathcal{D}_v f(x,y) = \nabla f(x,y)\cdot v = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right)\cdot v = \left(e^{x-y},\ -\,e^{x-y}\right)\cdot(\cos\alpha, \mathrm{sen}\,\alpha) =$$
$$= e^{x-y}\cos\alpha - e^{x-y}\,\mathrm{sen}\,\alpha\,.$$

Calcoliamo ora $\nabla(\mathcal{D}_v f(x,y))$ e poi calcoleremo $\nabla(\mathcal{D}_v f(x,y))\cdot w$. Si ha:

$$\frac{\partial(\mathcal{D}_v f(x,y))}{\partial x} = e^{x-y}\cos\alpha - e^{x-y}\,\mathrm{sen}\,\alpha \quad \text{e} \quad \frac{\partial(\mathcal{D}_v f(x,y))}{\partial y} = -\,e^{x-y}\cos\alpha + e^{x-y}\,\mathrm{sen}\,\alpha\,.$$

Quindi:

$$\mathcal{D}^2_{vw} f(x,y) = \nabla(\mathcal{D}_v f(x,y))\cdot w = \left(\frac{\partial(\mathcal{D}_v f(x,y))}{\partial x}, \frac{\partial(\mathcal{D}_v f(x,y))}{\partial y}\right)\cdot(\cos\beta, \mathrm{sen}\,\beta) =$$
$$= \left(e^{x-y}\cos\alpha - e^{x-y}\,\mathrm{sen}\,\alpha;\ -\,e^{x-y}\cos\alpha + e^{x-y}\,\mathrm{sen}\,\alpha\right)\cdot(\cos\beta, \mathrm{sen}\,\beta) =$$
$$= e^{x-y}\cos\alpha\cdot\cos\beta - e^{x-y}\,\mathrm{sen}\,\alpha\cdot\cos\beta - e^{x-y}\cos\alpha\cdot\mathrm{sen}\,\beta + e^{x-y}\,\mathrm{sen}\,\alpha\cdot\mathrm{sen}\,\beta\,,$$

ovvero lo stesso risultato ottenuto con l'altra procedura.

POLINOMIO DI TAYLOR E MAC LAURIN

Anche per le funzioni $f:\mathbb{R}^n\to\mathbb{R}$ è possibile fornire una approssimazione migliore di quella ottenuta con la formula della differenziabilità, ovvero costruire un polinomio (in n variabili) di grado m opportunamente scelto, per il quale valga l'uguaglianza:

$$f(\mathbb{X}) - \mathbb{P}_m(\mathbb{X},\mathbb{X}_0) = o(\|\mathbb{X}-\mathbb{X}_0\|^m)\,.$$

Per una funzione $f:\mathbb{R}\to\mathbb{R}$ il Polinomio di Taylor nel punto x_0 ha la seguente espressione:

$$\mathbb{P}_n(x,x_0) = f(x_0) + f'(x_0)\,(x-x_0) + \frac{f''(x_0)}{2\,!}\,(x-x_0)^2 + \frac{f'''(x_0)}{3\,!}\,(x-x_0)^3 + \ldots$$

Affinchè esista il Polinomio di grado m occorre che la funzione sia derivabile m volte in un intorno di x_0. Per una funzione $f:\mathbb{R}^n\to\mathbb{R}$ vale invece il seguente:

Teorema 13 : Sia la funzione $f(\mathbb{X})$, $\mathbb{R}^n\to\mathbb{R}$, differenziabile fino all'ordine m in un intorno del punto $\mathbb{X}_0$; allora esiste ed è unico il Polinomio di grado m tale che:
$f(\mathbb{X}) - \mathbb{P}_m(\mathbb{X},\mathbb{X}_0) = o(\|\mathbb{X}-\mathbb{X}_0\|^m)$. Tale polinomio ha la seguente espressione:

$$\mathbb{P}_m(\mathbb{X},\mathbb{X}_0) = f(\mathbb{X}_0) + \mathrm{d}f(\mathbb{X}_0) + \frac{\mathrm{d}^2 f(\mathbb{X}_0)}{2\,!} + \frac{\mathrm{d}^3 f(\mathbb{X}_0)}{3\,!} + \ldots + \frac{\mathrm{d}^m f(\mathbb{X}_0)}{m\,!}\,.$$

Come si vede, il Polinomio è costituito mediante i differenziali totali, da quello del primo fino a quello di ordine m. Se $\mathbb{X}_0 = \mathbb{O}$, vettore nullo, il Polinomio è detto di MacLaurin. Il Polinomio di secondo grado può essere espresso, per quanto già visto, in forma vettoriale-matriciale, ed avremo la seguente espressione:

$$\mathbb{P}_2(\mathbb{X},\mathbb{X}_0) = f(\mathbb{X}_0) + \nabla f(\mathbb{X}_0)\cdot\mathrm{d}\mathbb{X} + \frac{1}{2}\,\mathrm{d}\mathbb{X}\cdot\mathbb{H}(\mathbb{X}_0)\cdot\mathrm{d}\mathbb{X}^{\mathrm{T}}\,,$$

nella quale $\mathrm{d}\mathbb{X} = \mathbb{X}-\mathbb{X}_0$, e $\mathrm{d}\mathbb{X}^{\mathrm{T}}$ ne rappresenta il trasposto, e quindi per la funzione $f(\mathbb{X})$ risulta valida la seguente approssimazione:

$$f(\mathbb{X}) = f(\mathbb{X}_0) + \nabla f(\mathbb{X}_0)\cdot(\mathbb{X}-\mathbb{X}_0) + \frac{1}{2}\,(\mathbb{X}-\mathbb{X}_0)\cdot\mathbb{H}(\mathbb{X}_0)\cdot(\mathbb{X}-\mathbb{X}_0)^{\mathrm{T}} + o\big(\|\mathbb{X}-\mathbb{X}_0\|^2\big)\,.$$

Esempio 40 : Sia $f(x,y,z) = x^2\,\mathrm{sen}\,(y-z)$; determiniamo l'espressione del Polinomio di Taylor di secondo grado nel punto $(1,1,1)$.

Risulta anzitutto $f(\mathbb{X}_0) = f(1,1,1) = 0$. Sarà poi:

$f'_x = 2x\,\mathrm{sen}\,(y-z) \Rightarrow f'_x(1,1,1) = 0\,;$
$f'_y = x^2\cos(y-z) \Rightarrow f'_y(1,1,1) = 1\,;$

$f_z' = -x^2 \cos(y-z) \Rightarrow f_z'(1,1,1) = -1\,;$
$f_{xx}'' = 2 \operatorname{sen}(y-z) \Rightarrow f_{xx}''(1,1,1) = 0\,;$
$f_{yy}'' = -x^2 \operatorname{sen}(y-z) \Rightarrow f_{yy}''(1,1,1) = 0\,;$
$f_{zz}'' = -x^2 \operatorname{sen}(y-z) \Rightarrow f_{zz}''(1,1,1) = 0\,;$
$f_{xy}'' = 2x \cos(y-z) \Rightarrow f_{xy}''(1,1,1) = 2\,;$
$f_{xz}'' = -2x \cos(y-z) \Rightarrow f_{xz}''(1,1,1) = -2\,;$
$f_{yz}'' = x^2 \operatorname{sen}(y-z) \Rightarrow f_{yz}''(1,1,1) = 0\,.$
Avremo quindi, in forma analitica:

$$\begin{aligned}&\mathbb{P}_2(\mathbb{X},(1,1,1)) = 0 + 0\cdot(x-1) + 1\cdot(y-1) - 1\cdot(z-1) + \\ &+ \frac{1}{2}\left(0\cdot(x-1)^2 + 0\cdot(y-1)^2 + 0\cdot(z-1)^2\right) + \\ &+ \frac{1}{2}\left(2\cdot 2\,(x-1)\,(y-1) + 2\,(-2)\,(x-1)\,(z-1) + 2\cdot 0\cdot(y-1)\,(z-1)\right) = \\ &= z - y + 2x\,y - 2x\,z\,.\end{aligned}$$

In forma vettoriale invece, ponendo $\mathrm{d}x = x-1,\ \mathrm{d}y = y-1,\ \mathrm{d}z = z-1$, abbiamo:

$$\mathbb{P}_2(\mathbb{X},(1,1,1)) = 0 + (0,1,-1)\cdot(\mathrm{d}x,\mathrm{d}y,\mathrm{d}z) + \frac{1}{2}\,\|\mathrm{d}x\ \ \mathrm{d}y\ \ \mathrm{d}z\| \cdot \begin{Vmatrix} 0 & 2 & -2 \\ 2 & 0 & 0 \\ -2 & 0 & 0 \end{Vmatrix} \cdot \begin{Vmatrix} \mathrm{d}x \\ \mathrm{d}y \\ \mathrm{d}z \end{Vmatrix}$$

FUNZIONI CONVESSE E CONCAVE

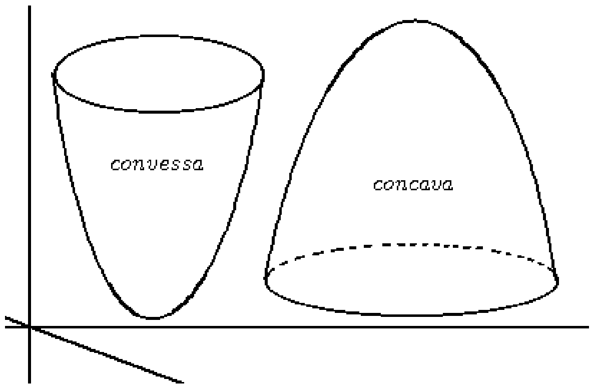
E' bene anzitutto notare come per le funzioni $f:\mathbb{R}^n \to \mathbb{R}$ non sia definito il concetto di funzione crescente o decrescente. Tale concetto potrebbe essere recuperato se si parlasse di crescenza o di decrescenza in una certa direzione, ovvero riportandolo ad una analisi di tipo monodimensionale, che però non è, in generale, utile per trarre conclusioni di tipo globale. Non esistono, di conseguenza, criteri analoghi a quello dello studio del segno di $f'(x)$, valido per funzioni $f:\mathbb{R}\to\mathbb{R}$. Rimane invece valido e utile per le applicazioni, principalmente per lo studio dei massimi e dei minimi, il concetto di funzione convessa. Diamo anzitutto la

Definizione 35 : Un insieme $\mathbb{A} \subseteq \mathbb{R}^n$ si dice **convesso** se presi due qualunque punti $\mathbb{X}_1 \in \mathbb{A}$, $\mathbb{X}_2 \in \mathbb{A}$, il segmento che li congiunge è tutto contenuto in $\mathbb{A}$. In forma analitica:
$\forall\, \mathbb{X}_1, \mathbb{X}_2 \in \mathbb{A} : \forall\, \alpha \in [0,1] \ \Rightarrow\ \alpha\cdot\mathbb{X}_1 + (1-\alpha)\cdot\mathbb{X}_2 \in \mathbb{A}\,,$
dove $\alpha\cdot\mathbb{X}_1 + (1-\alpha)\cdot\mathbb{X}_2,\ \alpha\in\mathbb{R}$ rappresenta la retta passante per i due punti $\mathbb{X}_1$ e $\mathbb{X}_2$.
Se un insieme $\mathbb{A}\subseteq\mathbb{R}^n$ non risulta convesso, allora viene detto concavo. Diamo poi la

Definizione 36 : Data $f:\mathbb{R}^n\to\mathbb{R}$, si dice **epigrafico** della funzione, relativo ad $\mathbb{A}\subseteq\mathbb{R}^n$, l'insieme: $\mathcal{E}(f) = \left\{(\mathbb{X},y)\in\mathbb{R}^{n+1}, \mathbb{X}\in\mathbb{A} :\ y \geq f(\mathbb{X})\right\}$.
L'epigrafico è quindi la parte di spazio situata al di sopra del grafico, grafico compreso. Da questa segue la:

Definizione 37 : Data $f:\mathbb{R}^n\to\mathbb{R}$, sia $\mathbb{A}\subseteq\mathbb{R}^n$ un insieme convesso; la funzione f si dice **convessa** in $\mathbb{A}$ se il suo epigrafico (relativamente ad $\mathbb{A}$) è un insieme convesso.
Notiamo che la definizione di funzione convessa non viene data in un insieme qualsiasi, ma solo in un dominio anch'esso convesso. Come per le funzioni $\mathbb{R}\to\mathbb{R}$, contrariamente a quanto definito per gli insiemi, una funzione che non risulti convessa non viene detta concava.
Vale invece la:

Definizione 38 : Data $f:\mathbb{R}^n\to\mathbb{R}$, sia $\mathbb{A}\subseteq\mathbb{R}^n$ un insieme convesso; la funzione f si dice **concava** in $\mathbb{A}$ se risulta convessa in $\mathbb{A}$ la funzione $-f$, opposta di f.



Le funzioni concave sono quindi le simmetriche, rispetto al (iper)piano delle variabili indipendenti, delle funzioni convesse, ed esistono funzioni che, in un dato insieme, non sono nè concave nè convesse, mentre invece per gli insiemi o si è concavi o si è convessi.
Limitandoci per ovvii motivi al caso di funzioni $f : \mathbb{R}^2 \to \mathbb{R}$, vediamo in figura un esempio di funzione convessa e uno di concava.
Enunciamo infine due Teoremi che legano la convessità di una funzione alla differenziabilità del primo e del secondo ordine, nonchè a considerazioni di tipo geometrico.
Teorema 14 : Sia $f : \mathbb{R}^n \to \mathbb{R}$ differenziabile in $\mathbb{A} \subseteq \mathbb{R}^n$ insieme convesso. Allora f è convessa in $\mathbb{A}$ se e solo se $\forall\, \mathbb{X}, \mathbb{X}_0 \in \mathbb{A} : f(\mathbb{X}) \geq f(\mathbb{X}_0) + \mathrm{d}f(\mathbb{X}_0)$, oppure, equivalentemente:
$f(\mathbb{X}) \geq f(\mathbb{X}_0) + \nabla f(\mathbb{X}_0) \cdot (\mathbb{X} - \mathbb{X}_0)$.
Questa formulazione ci permette di esprimere la convessità di una funzione differenziabile come equivalente ad affermare che il grafico della funzione non sta al di sotto del (iper)piano tangente in un qualunque punto $\mathbb{X}_0 \in \mathbb{A}$.
Utilizzando il Polinomio di Taylor si può poi provare che vale anche il:
Teorema 15 : Sia $f : \mathbb{R}^n \to \mathbb{R}$ differenziabile due volte in $\mathbb{A} \subseteq \mathbb{R}^n$ insieme convesso. Allora f è convessa in $\mathbb{A}$ se e solo se $\mathrm{d}^2 f(\mathbb{X}) \geq 0, \forall\, \mathbb{X} \in \mathbb{A}$.
Infatti, essendo f differenziabile due volte, risulta:
$f(\mathbb{X}) = f(\mathbb{X}_0) + \mathrm{d}f(\mathbb{X}_0) + \dfrac{1}{2}\mathrm{d}^2 f(\mathbb{X}_0) + o\big(\|\mathbb{X} - \mathbb{X}_0\|^2\big)$, da cui:
$f(\mathbb{X}) - [f(\mathbb{X}_0) + \mathrm{d}f(\mathbb{X}_0)] = \dfrac{1}{2}\mathrm{d}^2 f(\mathbb{X}_0) + o\big(\|\mathbb{X} - \mathbb{X}_0\|^2\big)$, e se $\mathrm{d}^2 f(\mathbb{X}_0) \geq 0$ segue che:
$f(\mathbb{X}) \geq f(\mathbb{X}_0) + \mathrm{d}f(\mathbb{X}_0)$.
Utilizzando invece la forma vettoriale, il Teorema precedente conduce alla:
$f(\mathbb{X}) - [f(\mathbb{X}_0) + \nabla f(\mathbb{X}_0) \cdot (\mathbb{X} - \mathbb{X}_0)] = \dfrac{1}{2}(\mathbb{X} - \mathbb{X}_0) \cdot \mathbb{H}(\mathbb{X}_0) \cdot (\mathbb{X} - \mathbb{X}_0)^{\mathrm{T}} + o\big(\|\mathbb{X} - \mathbb{X}_0\|^2\big)$.
La quantità $\mathrm{d}^2 f(\mathbb{X}_0) = \mathrm{d}\mathbb{X} \cdot \mathbb{H}(\mathbb{X}_0) \cdot \mathrm{d}\mathbb{X}^{\mathrm{T}} = (\mathbb{X} - \mathbb{X}_0) \cdot \mathbb{H}(\mathbb{X}_0) \cdot (\mathbb{X} - \mathbb{X}_0)^{\mathrm{T}}$ rappresenta una forma quadratica nelle n variabili $\mathrm{d}x_1, \mathrm{d}x_2, \ldots, \mathrm{d}x_n$; se $\mathrm{d}^2 f(\mathbb{X}_0) > 0$ (≥ 0) essa è detta forma quadratica definita (semidefinita) positiva. Lo studio della convessità e della concavità di una funzione viene ricondotto allo studio del segno della forma quadratica $\mathrm{d}^2 f(\mathbb{X}_0)$, e questo argomento verrà trattato nel seguito, all'interno della parte riguardante la ricerca dei punti di massimo e di minimo relativi.
Se fosse $f(\mathbb{X}) > f(\mathbb{X}_0) + \mathrm{d}f(\mathbb{X}_0)$ oppure se fosse $\mathrm{d}^2 f(\mathbb{X}_0) > 0$, parleremo di funzione strettamente convessa.

FUNZIONI $f : \mathbb{R}^n \to \mathbb{R}^m$
Estendiamo i concetti fin qui svolti alle funzioni vettoriali di variabile vettoriale, ovvero alle funzioni $\mathbb{Y} = f(\mathbb{X})$, con $\mathbb{X} = (x_1, x_2, ..., x_n) \in \mathbb{R}^n$, $\mathbb{Y} = (y_1, y_2, ..., y_m) \in \mathbb{R}^m$, cioè alle funzioni che associano ad un vettore come immagine un altro vettore, anche appartenenti a spazi diversi. Un primo esempio di funzione $f : \mathbb{R}^n \to \mathbb{R}^m$ è costituito dalle Applicazioni li-

neari, ovvero dalle funzioni esprimibili come $\mathbb{Y} = f(\mathbb{X}) = \mathbb{A} \cdot \mathbb{X}$, dove $\mathbb{A}$ è una matrice $m \cdot n$ ad elementi costanti.
In caso di bisogno, la notazione $\mathbb{Y} = f(\mathbb{X})$ potrà essere scritta in forma estesa come:
$(y_1, y_2, ..., y_m) = (f_1(x_1, x_2, ..., x_n), f_2(x_1, x_2, ..., x_n), ..., f_m(x_1, x_2, ..., x_n))$.
Quindi una funzione $f : \mathbb{R}^n \to \mathbb{R}^m$ si può vedere come una m-upla di funzioni $f_i : \mathbb{R}^n \to \mathbb{R}$.
Basterà ripetere quanto fatto per le funzioni $f : \mathbb{R} \to \mathbb{R}^m$, adeguando la teoria alla luce di quanto visto per le funzioni $f : \mathbb{R}^n \to \mathbb{R}$, per ottenere l'estensione delle principali definizioni e proprietà relative alle funzioni $f : \mathbb{R}^n \to \mathbb{R}^m$.

LIMITI, CONTINUITA', DERIVABILITA', DIFFERENZIABILITA'

Definizione 39 : La funzione $f : \mathbb{R}^n \to \mathbb{R}^m, \mathbb{Y} = f(\mathbb{X})$ ammette limite per $\mathbb{X} \to \mathbb{X}_0$, con $\mathbb{X}_0$ punto di accumulazione per D_f, se ciascuna delle componenti $f_i : \mathbb{R}^n \to \mathbb{R}, 1 \leq i \leq m$, ammette limite per $\mathbb{X} \to \mathbb{X}_0$.

Definizione 40 : Preso $\mathbb{X}_0$ punto di accumulazione appartenente a D_f, si dirà che la funzione $f : \mathbb{R}^n \to \mathbb{R}^m$ è continua in $\mathbb{X}_0$ se $\lim\limits_{\mathbb{X} \to \mathbb{X}_0} f(\mathbb{X}) = f(\mathbb{X}_0)$, ovvero se ciascuna delle m componenti f_i è continua in $\mathbb{X}_0$, ovvero se $\lim\limits_{\mathbb{X} \to \mathbb{X}_0} f_i(\mathbb{X}) = f_i(\mathbb{X}_0), 1 \leq i \leq m$.

Per quanto concerne la derivabilità, dato che ogni componente è una funzione $f_i : \mathbb{R}^n \to \mathbb{R}$, possiamo definire le derivate parziali e le derivate direzionali.
Preso $\mathbb{X}_0$ punto interno a D_f, e preso un versore $v \in \mathbb{R}^n$, abbiamo la

Definizione 41 : La funzione $f : \mathbb{R}^n \to \mathbb{R}^m, \mathbb{Y} = f(\mathbb{X})$, ammette in $\mathbb{X}_0$ derivata nella direzione v se esiste finito il $\lim\limits_{t \to 0} \dfrac{f(\mathbb{X}_0 + t\,v) - f(\mathbb{X}_0)}{t} = \mathcal{D}_v f(\mathbb{X}_0)$. Questo equivale a richiedere che esistano finiti gli m limiti: $\lim\limits_{t \to 0} \dfrac{f_i(\mathbb{X}_0 + t\,v) - f_i(\mathbb{X}_0)}{t} = \mathcal{D}_v f_i(\mathbb{X}_0)$, ovvero che ciascuna delle m funzioni $f_i : \mathbb{R}^n \to \mathbb{R}$ risulti in $\mathbb{X}_0$ derivabile nella direzione v.

Anche le derivate parziali di $f : \mathbb{R}^n \to \mathbb{R}^m$ vengono definite come le derivate parziali delle componenti $f_i : \mathbb{R}^n \to \mathbb{R}$. Una funzione $f : \mathbb{R}^n \to \mathbb{R}^m$ risulta quindi derivabile rispetto alla variabile x_j se ciascuna componente f_i risulta derivabile rispetto a x_j. Si vengono così a determinare $m \cdot n$ derivate parziali, derivando la funzione f_i, $1 \leq i \leq m$, rispetto alla variabile x_j, $1 \leq j \leq n$. Si rappresentano tutte le derivate parziali prime mediante la cosiddetta **matrice Jacobiana** $m \cdot n$:

$$J_f(\mathbb{X}) = \frac{\partial(f_1, f_2, ..., f_m)}{\partial(x_1, x_2, ..., x_n)} = \begin{Vmatrix} \dfrac{\partial f_1}{\partial x_1} & \dfrac{\partial f_1}{\partial x_2} & .... & \dfrac{\partial f_1}{\partial x_n} \\ \dfrac{\partial f_2}{\partial x_1} & \dfrac{\partial f_2}{\partial x_2} & .... & \dfrac{\partial f_2}{\partial x_n} \\ .... & .... & .... & .... \\ \dfrac{\partial f_m}{\partial x_1} & \dfrac{\partial f_m}{\partial x_2} & .... & \dfrac{\partial f_m}{\partial x_n} \end{Vmatrix}$$

, nella quale l'elemento di posto (i, j) è dato dalla derivata parziale di f_i fatta rispetto a x_j. Notiamo anche come ogni riga della matrice Jacobiana sia un gradiente: la riga i-esima è infatti il gradiente $\nabla f_i(\mathbb{X})$.
Per quanto riguarda la differenziabilità di una funzione $f : \mathbb{R}^n \to \mathbb{R}^m, f = (f_1, f_2, ..., f_m)$, possiamo seguire la procedura delle definizioni precedenti, ovvero:

Definizione 42 : La funzione $f : \mathbb{R}^n \to \mathbb{R}^m, \mathbb{Y} = f(\mathbb{X})$ si dice differenziabile nel punto $\mathbb{X}_0$ se in tale punto risulta differenziabile ciascuna delle sue componenti $y_i = f_i(\mathbb{X})$, ovvero se:
$f_i(\mathbb{X}) = f_i(\mathbb{X}_0) + \mathbb{K}_i \cdot (\mathbb{X} - \mathbb{X}_0) + o(\|\mathbb{X} - \mathbb{X}_0\|)$, con $\mathbb{K}_i \in \mathbb{R}^n, \forall\, i : 1 \leq i \leq m$.
Equivalentemente, si può chiedere che esista una matrice a termini costanti $\mathbb{M}_{m,n}$, per cui valga la: $f(\mathbb{X}) = f(\mathbb{X}_0) + \mathbb{M} \cdot (\mathbb{X} - \mathbb{X}_0) + o(\|\mathbb{X} - \mathbb{X}_0\|)$.
La matrice $\mathbb{M}$ ha per righe i vettori $\mathbb{K}_i$.

Ma, per quanto visto per le funzioni $f:\mathbb{R}^n \to \mathbb{R}$, se f_i risulta differenziabile nel punto $\mathbb{X}_0$ si ha che $\mathbb{K}_i = \nabla f_i(\mathbb{X}_0)$, e la differenziabilità di f_i può essere espressa come:
$f_i(\mathbb{X}) = f_i(\mathbb{X}_0) + \nabla f_i(\mathbb{X}_0)\cdot(\mathbb{X}-\mathbb{X}_0) + o(\|\mathbb{X}-\mathbb{X}_0\|), \forall\, i : 1 \le i \le m$.
Queste m uguaglianze possono infine essere compattate, osservato che $\mathbb{M} = J_f(\mathbb{X}_0)$, nella:
$f(\mathbb{X}) = f(\mathbb{X}_0) + J_f(\mathbb{X}_0)\cdot(\mathbb{X}-\mathbb{X}_0) + o(\|\mathbb{X}-\mathbb{X}_0\|)$,
dove $f(\mathbb{X}), f(\mathbb{X}_0) \in \mathbb{R}^m, o(\|\mathbb{X}-\mathbb{X}_0\|) \in \mathbb{R}^m$ mentre $(\mathbb{X}-\mathbb{X}_0) \in \mathbb{R}^n$ e $J_f(\mathbb{X}_0)$ è una matrice $(m\cdot n)$.
Per calcolare $\mathcal{D}_v f(\mathbb{X}_0)$, se la funzione è differenziabile in $\mathbb{X}_0$, basterà calcolare $J_f(\mathbb{X}_0)\cdot v$, per ottenere il vettore $(\mathcal{D}_v f_1(\mathbb{X}_0), \mathcal{D}_v f_2(\mathbb{X}_0), \ldots, \mathcal{D}_v f_m(\mathbb{X}_0)) \in \mathbb{R}^m$.

La matrice Jacobiana esprime e sintetizza il concetto di derivata nel caso più generale, cioè quello di $f:\mathbb{R}^n \to \mathbb{R}^m$, comprendendo come casi particolari tutti i tipi di derivate incontrati finora.
Se $f:\mathbb{R}\to\mathbb{R}$, la derivata $f'(x)$ è una matrice Jacobiana $(1\cdot 1)$; se $f:\mathbb{R}^n\to\mathbb{R}$ si ha il gradiente, che è una matrice Jacobiana $(1\cdot n)$, ovvero una Jacobiana costituita da una sola riga. Per le curve, ovvero per funzioni $f:\mathbb{R}\to\mathbb{R}^n$, si ha il vettore tangente che, se visto come vettore $(n\cdot 1)$, risulta una Jacobiana costituita da una sola colonna.

Esempio 41 : Se $f:\mathbb{R}^n\to\mathbb{R}^m$ è una applicazione lineare, $\mathbb{Y} = f(\mathbb{X}) = \mathbb{A}\cdot\mathbb{X}$, è facile vedere che risulta $J_f(\mathbb{X}_0) = \mathbb{A}$, $\forall\, \mathbb{X}_0 \in \mathbb{R}^n$.

Esempio 42 : Se $f:\mathbb{R}^2\to\mathbb{R}^2$ esprime un cambiamento di coordinate, da cartesiane (x,y) a polari (ρ,ϑ): $f(\rho,\vartheta) = (x_0 + \rho\cos\vartheta, y_0 + \rho\,\mathrm{sen}\,\vartheta)$, otteniamo:
$$J_f(\rho,\vartheta) = \frac{\partial(x,y)}{\partial(\rho,\vartheta)} = \begin{Vmatrix} \dfrac{\partial x}{\partial\rho} & \dfrac{\partial x}{\partial\vartheta} \\ \dfrac{\partial y}{\partial\rho} & \dfrac{\partial y}{\partial\vartheta} \end{Vmatrix} = \begin{Vmatrix} \cos\vartheta & -\rho\,\mathrm{sen}\,\vartheta \\ \mathrm{sen}\,\vartheta & \rho\cos\vartheta \end{Vmatrix}.$$

REGOLA DI DERIVAZIONE DELLE FUNZIONI COMPOSTE
Scopo di questo paragrafo è quello di formulare la regola di derivata di una funzione composta nel caso più generale: $\mathbb{R}^m \xrightarrow{g} \mathbb{R}^n \xrightarrow{f} \mathbb{R}^p, \mathbb{T} \xrightarrow{g} \mathbb{X} \xrightarrow{f} \mathbb{Y}$, dove:
$\mathbb{T} = (t_1, t_2, \ldots, t_m)$; $\mathbb{X} = (x_1, x_2, \ldots, x_n)$; $\mathbb{Y} = (y_1, y_2, \ldots, y_p)$;
$\mathbb{X} = g(\mathbb{T}) = (g_1(\mathbb{T}), g_2(\mathbb{T}), .., g_n(\mathbb{T})) = (g_1(t_1, .., t_m), g_2(t_1, .., t_m), .., g_n(t_1, .., t_m))$;
$\mathbb{Y} = f(\mathbb{X}) = (f_1(\mathbb{X}), f_2(\mathbb{X}), .., f_p(\mathbb{X})) = (f_1(x_1, \ldots, x_n), f_2(x_1, \ldots, x_n), .., f_p(x_1, \ldots, x_n))$
per avere quindi $\mathbb{Y} = f(\mathbb{X}) = f(g(\mathbb{T}))$.

Nel caso $m = n = p = 1$ risulta $\mathcal{D}(f(g(x_0))) = f'(g(x_0))\cdot g'(x_0)$, cioè la derivata della funzione composta è data dal prodotto di due numeri: $f'(g(x))$ per $g'(x)$.

Nel caso $\mathbb{R} \xrightarrow{g} \mathbb{R} \xrightarrow{f} \mathbb{R}^p, t \xrightarrow{g} x \xrightarrow{f} \mathbb{Y}, \mathbb{Y} = f(g(t))$, vedi Teorema 5, risulta:
$\mathcal{D}(f(g(t_0))) = \dfrac{\mathrm{d}\mathbb{Y}(x(t_0))}{\mathrm{d}t} = \mathbb{Y}'(x(t_0))\cdot x'(t_0)$,
ovvero la derivata della funzione composta è data dal prodotto del vettore tangente $\mathbb{Y}'(x(t_0))$ per lo scalare $x'(t_0)$.
Determiniamo ora la regola nel caso $\mathbb{R} \xrightarrow{g} \mathbb{R}^n \xrightarrow{f} \mathbb{R}, t \xrightarrow{g} \mathbb{X} \xrightarrow{f} y$, $n > 1$.
Sia allora $y = f(\mathbb{X}) = f(x_1, x_2, \ldots, x_n) = f(g(t)) = f(x_1(t), x_2(t), \ldots, x_n(t))$. Vale il:

Teorema 16 : Sia $\mathbb{X}(t)=g(t)=(x_1(t),x_2(t),...,x_n(t)), \mathbb{R}\to\mathbb{R}^n$ derivabile in $t=t_0$ e sia $y=f(\mathbb{X}), \mathbb{R}^n\to\mathbb{R}$ differenziabile in $\mathbb{X}(t_0)=g(t_0)$. Allora la funzione $y=f(g(t))$ è derivabile in t_0 e risulta:

$$\mathcal{D}(f(g(t_0)))=\frac{\mathrm{d}y(t_0)}{\mathrm{d}t}=\nabla f(\mathbb{X}(t_0))\cdot\mathbb{X}'(t_0)\,,$$

ovvero la derivata della funzione composta è data dal prodotto scalare del vettore gradiente $\nabla f(\mathbb{X}(t_0))$ per il vettore tangente $\mathbb{X}'(t_0)=(x_1'(t_0),x_2'(t_0),...,x_n'(t_0))$.

Dimostrazione : Applicando la definizione di derivata, avremo:

$$\frac{\mathrm{d}y(t_0)}{\mathrm{d}t}=\lim_{t\to t_0}\frac{f(g(t))-f(g(t_0))}{t-t_0}=\lim_{t\to t_0}\frac{f(\mathbb{X}(t))-f(\mathbb{X}(t_0))}{t-t_0}\,.$$

Essendo f differenziabile in $\mathbb{X}_0$ si ha:

$$f(\mathbb{X}(t))-f(\mathbb{X}(t_0))=\nabla f(\mathbb{X}(t_0))\cdot(\mathbb{X}(t)-\mathbb{X}(t_0))+o(\|\mathbb{X}(t)-\mathbb{X}(t_0)\|)\,,$$

dalla quale, sostituendo, otteniamo:

$$\lim_{t\to t_0}\frac{f(\mathbb{X}(t))-f(\mathbb{X}(t_0))}{t-t_0}=\lim_{t\to t_0}\frac{\nabla f(\mathbb{X}(t_0))\cdot(\mathbb{X}(t)-\mathbb{X}(t_0))+o(\|\mathbb{X}(t)-\mathbb{X}(t_0)\|)}{t-t_0}=$$

$$=\lim_{t\to t_0}\nabla f(\mathbb{X}(t_0))\cdot\frac{\mathbb{X}(t)-\mathbb{X}(t_0)}{t-t_0}+\lim_{t\to t_0}\frac{o(\|\mathbb{X}(t)-\mathbb{X}(t_0)\|)}{\|\mathbb{X}(t)-\mathbb{X}(t_0)\|}\cdot\left(\pm\left\|\frac{\mathbb{X}(t)-\mathbb{X}(t_0)}{t-t_0}\right\|\right).$$

Ma $\nabla f(\mathbb{X}(t_0))$ è un vettore costante, mentre

$$\lim_{t\to t_0}\frac{\mathbb{X}(t)-\mathbb{X}(t_0)}{t-t_0}=\left(\lim_{t\to t_0}\frac{x_1(t)-x_1(t_0)}{t-t_0},\dots,\lim_{t\to t_0}\frac{x_n(t)-x_n(t_0)}{t-t_0}\right)=$$

$$=(x_1'(t_0),x_2'(t_0),...,x_n'(t_0))=\mathbb{X}'(t_0)$$

in quanto per ipotesi la funzione $\mathbb{X}(t)$ è derivabile in t_0. Infine

$\lim_{t\to t_0}\frac{o(\|\mathbb{X}(t)-\mathbb{X}(t_0)\|)}{\|\mathbb{X}(t)-\mathbb{X}(t_0)\|}=0$ per la definizione di "o piccolo" mentre

$\lim_{t\to t_0}\left\|\frac{\mathbb{X}(t)-\mathbb{X}(t_0)}{t-t_0}\right\|=\|\mathbb{X}'(t_0)\|$, valore finito in quanto $\mathbb{X}(t)$ è derivabile in t_0. Quindi

$$\mathcal{D}(f(g(t_0)))=\frac{\mathrm{d}y(t_0)}{\mathrm{d}t}=\nabla f(\mathbb{X}(t_0))\cdot\mathbb{X}'(t_0)\,.\bullet$$

Il risultato trovato può essere espresso anche nella forma:

$\frac{\mathrm{d}y}{\mathrm{d}t}=\left(\frac{\partial f}{\partial x_1},\frac{\partial f}{\partial x_2},...,\frac{\partial f}{\partial x_n}\right)\cdot\left(\frac{\mathrm{d}x_1}{\mathrm{d}t},\frac{\mathrm{d}x_2}{\mathrm{d}t},...,\frac{\mathrm{d}x_n}{\mathrm{d}t}\right)$ e quindi come:

$$\frac{\mathrm{d}y}{\mathrm{d}t}=\frac{\partial f}{\partial x_1}\cdot\frac{\mathrm{d}x_1}{\mathrm{d}t}+\frac{\partial f}{\partial x_2}\cdot\frac{\mathrm{d}x_2}{\mathrm{d}t}+...+\frac{\partial f}{\partial x_n}\cdot\frac{\mathrm{d}x_n}{\mathrm{d}t}=\sum_{i=1}^{n}\frac{\partial f}{\partial x_i}\cdot\frac{\mathrm{d}x_i}{\mathrm{d}t}\,.$$

Utilizzando questo risultato, passiamo al caso $\mathbb{R}\overset{g}{\to}\mathbb{R}^n\overset{f}{\to}\mathbb{R}^p,\ t\overset{g}{\to}\mathbb{X}\overset{f}{\to}\mathbb{Y}$.
Sarà ora $\mathbb{Y}=(y_1,y_2,...,y_p)$, con $y_i=f_i(x_1,x_2,...,x_n)$, $1\le i\le p$, con $x_j=g_j(t)$.
Supponendo che ciascuna $f_i(x_1,x_2,...,x_n)=f_i(\mathbb{X})$ sia differenziabile in $\mathbb{X}(t_0)$ e supponendo $\mathbb{X}(t)=(x_1(t),x_2(t),...,x_n(t))$ derivabile in t_0, per quanto visto nel Teorema precedente sarà:

$\frac{\mathrm{d}y_i(t_0)}{\mathrm{d}t}=\nabla f_i(\mathbb{X}(t_0))\cdot\mathbb{X}'(t_0)$ ovvero:

$$\frac{\mathrm{d}y_i}{\mathrm{d}t}=\frac{\partial f_i}{\partial x_1}\cdot\frac{\mathrm{d}x_1}{\mathrm{d}t}+\frac{\partial f_i}{\partial x_2}\cdot\frac{\mathrm{d}x_2}{\mathrm{d}t}+...+\frac{\partial f_i}{\partial x_n}\cdot\frac{\mathrm{d}x_n}{\mathrm{d}t}\,,\ 1\le i\le p\,.$$

Queste p uguaglianze possono essere scritte in forma matriciale come:

$$\left\|\begin{array}{c} \dfrac{\mathrm{d}y_1}{\mathrm{d}t} \\ \dfrac{\mathrm{d}y_2}{\mathrm{d}t} \\ \dots. \\ \dfrac{\mathrm{d}y_p}{\mathrm{d}t} \end{array}\right\| = \left\|\begin{array}{cccc} \dfrac{\partial f_1}{\partial x_1} & \dfrac{\partial f_1}{\partial x_2} & \dots. & \dfrac{\partial f_1}{\partial x_n} \\ \dfrac{\partial f_2}{\partial x_1} & \dfrac{\partial f_2}{\partial x_2} & \dots. & \dfrac{\partial f_2}{\partial x_n} \\ \dots. & \dots. & \dots. & \dots. \\ \dfrac{\partial f_p}{\partial x_1} & \dfrac{\partial f_p}{\partial x_2} & \dots. & \dfrac{\partial f_p}{\partial x_n} \end{array}\right\| \cdot \left\|\begin{array}{c} \dfrac{\mathrm{d}x_1}{\mathrm{d}t} \\ \dfrac{\mathrm{d}x_2}{\mathrm{d}t} \\ \dots. \\ \dfrac{\mathrm{d}x_n}{\mathrm{d}t} \end{array}\right\|$$, ovvero come:

$\mathcal{D}(f(g(t_0))) = \dfrac{\mathrm{d}\mathbb{Y}(\mathbb{X}(t_0))}{\mathrm{d}t} = \dfrac{\partial(f_1, f_2, ..., f_p)(\mathbb{X}(t_0))}{\partial(x_1, x_2, ..., x_n)} \cdot \mathbb{X}'(t_0)$ ed anche :
$\mathbb{Y}'(\mathbb{X}(t_0)) = J_f(\mathbb{X}(t_0)) \cdot \mathbb{X}'(t_0)$.
Il vettore tangente $\mathbb{Y}'(\mathbb{X}(t_0))$ è quindi dato dal prodotto della matrice Jacobiana $J_f(\mathbb{X}(t_0))$ per il vettore tangente $\mathbb{X}'(t_0)$.

Estendiamo quanto visto finora al caso generale $\mathbb{R}^m \xrightarrow{g} \mathbb{R}^n \xrightarrow{f} \mathbb{R}^p, \mathbb{T} \xrightarrow{g} \mathbb{X} \xrightarrow{f} \mathbb{Y}$. Vale il **Teorema** 17 : Siano $\mathbb{T} = (t_1, t_2, \ ..., t_m)$, $\mathbb{X} = (x_1, x_2, ..., x_n)$, $\mathbb{Y} = (y_1, y_2, ..., y_p)$ con:
$x_i = g_i(t_1, t_2, \ ..., t_m)$; $y_j = f_j(x_1, x_2, ..., x_n)$.
Posto $\mathbb{Y}_0 = f(\mathbb{X}(\mathbb{T}_0)) = f(g(\mathbb{T}_0))$, sia $\mathbb{X} = g(\mathbb{T})$ differenziabile in $\mathbb{T}_0$ e sia $\mathbb{Y} = f(\mathbb{X})$ differenziabile in $\mathbb{X}(\mathbb{T}_0)$. Risulta allora:
$\mathcal{D}(f(g(\mathbb{T}_0))) = \dfrac{\partial(y_1, y_2, ..., y_p)(\mathbb{X}(\mathbb{T}_0))}{\partial(t_1, t_2, \ ..., t_m)} = \dfrac{\partial(f_1, f_2, ..., f_p)(\mathbb{X}(t_0))}{\partial(x_1, x_2, ..., x_n)} \cdot \dfrac{\partial(x_1, x_2, ..., x_n)(\mathbb{T}_0)}{\partial(t_1, t_2, \ ..., t_m)}$.
Risulta quindi:
$J_{f(g)}(\mathbb{T}_0) = J_f(\mathbb{X}(\mathbb{T}_0)) \cdot J_g(\mathbb{T}_0)$, che, in forma generica, può essere espressa come:
$\dfrac{\partial(y_1, y_2, ..., y_p)}{\partial(t_1, t_2, \ ..., t_m)} = \dfrac{\partial(y_1, y_2, ..., y_p)}{\partial(x_1, x_2, ..., x_n)} \cdot \dfrac{\partial(x_1, x_2, ..., x_n)}{\partial(t_1, t_2, \ ..., t_m)}$.
Per esteso, in forma matriciale, avremo:

$$\left\|\begin{array}{cccc} \dfrac{\partial y_1}{\partial t_1} & \dfrac{\partial y_1}{\partial t_2} & \dots. & \dfrac{\partial y_1}{\partial t_m} \\ \dfrac{\partial y_2}{\partial t_1} & \dfrac{\partial y_2}{\partial t_2} & \dots. & \dfrac{\partial y_2}{\partial t_m} \\ \dots. & \dots. & \dots. & \dots. \\ \dfrac{\partial y_p}{\partial t_1} & \dfrac{\partial y_p}{\partial t_2} & \dots. & \dfrac{\partial y_p}{\partial t_m} \end{array}\right\| = \left\|\begin{array}{cccc} \dfrac{\partial y_1}{\partial x_1} & \dfrac{\partial y_1}{\partial x_2} & \dots. & \dfrac{\partial y_1}{\partial x_n} \\ \dfrac{\partial y_2}{\partial x_1} & \dfrac{\partial y_2}{\partial x_2} & \dots. & \dfrac{\partial y_2}{\partial x_n} \\ \dots. & \dots. & \dots. & \dots. \\ \dfrac{\partial y_p}{\partial x_1} & \dfrac{\partial y_p}{\partial x_2} & \dots. & \dfrac{\partial y_p}{\partial x_n} \end{array}\right\| \cdot \left\|\begin{array}{cccc} \dfrac{\partial x_1}{\partial t_1} & \dfrac{\partial x_1}{\partial t_2} & \dots. & \dfrac{\partial x_1}{\partial t_m} \\ \dfrac{\partial x_2}{\partial t_1} & \dfrac{\partial x_2}{\partial t_2} & \dots. & \dfrac{\partial x_2}{\partial t_m} \\ \dots. & \dots. & \dots. & \dots. \\ \dfrac{\partial x_n}{\partial t_1} & \dfrac{\partial x_n}{\partial t_2} & \dots. & \dfrac{\partial x_n}{\partial t_m} \end{array}\right\|$$

E' facile ora verificare come tutti i casi di derivata di funzione composta incontrati in precedenza siano compresi nella $J_{f(g)}(\mathbb{T}_0) = J_f(g(\mathbb{T}_0)) \cdot J_g(\mathbb{T}_0)$, con $J_{f(g)}(\mathbb{T}_0)$ matrice $p \cdot m$, $J_f(g(\mathbb{T}_0))$ matrice $p \cdot n$, e $J_g(\mathbb{T}_0)$ matrice $n \cdot m$. Per una funzione $\mathbb{R} \to \mathbb{R}^n$ la matrice Jacobiana diviene un vettore colonna $n \cdot 1$, ovvero il vettore tangente; nel caso di una funzione $\mathbb{R}^n \to \mathbb{R}$ la matrice Jacobiana diviene un vettore riga $1 \cdot n$, ovvero il vettore gradiente.

Se vogliamo evidenziare l'espressione della singola derivata di $f(g(\mathbb{T}))$ avremo:
$\dfrac{\partial y_i}{\partial t_j} = \dfrac{\partial f_i}{\partial x_1} \cdot \dfrac{\partial x_1}{\partial t_j} + \dfrac{\partial f_i}{\partial x_2} \cdot \dfrac{\partial x_2}{\partial t_j} + ... + \dfrac{\partial f_i}{\partial x_n} \cdot \dfrac{\partial x_n}{\partial t_j}, 1 \le i \le p, 1 \le j \le m$, o anche:
$\dfrac{\partial y_i}{\partial t_j} = \dfrac{\partial y_i}{\partial x_1} \cdot \dfrac{\partial x_1}{\partial t_j} + \dfrac{\partial y_i}{\partial x_2} \cdot \dfrac{\partial x_2}{\partial t_j} + ... + \dfrac{\partial y_i}{\partial x_n} \cdot \dfrac{\partial x_n}{\partial t_j}, 1 \le i \le p, 1 \le j \le m$.
Nel caso della composizione di tre funzioni: $\mathbb{R}^k \xrightarrow{h} \mathbb{R}^m \xrightarrow{g} \mathbb{R}^n \xrightarrow{f} \mathbb{R}^p, \mathbb{W} \xrightarrow{h} \mathbb{T} \xrightarrow{g} \mathbb{X} \xrightarrow{f} \mathbb{Y}$, per avere le derivate di $\mathbb{Y} = f(g(h(\mathbb{W})))$ dovremo calcolare:
$J_{f(g(h))}(\mathbb{W}_0) = J_f(\mathbb{X}_0) \cdot J_g(\mathbb{T}_0) \cdot J_h(\mathbb{W}_0) = J_f(g(h(\mathbb{W}_0))) \cdot J_g(h(\mathbb{W}_0)) \cdot J_h(\mathbb{W}_0)$, ovvero:

$$\frac{\partial(y_1,y_2,...,y_p)}{\partial(w_1,w_2,...,w_k)}=\frac{\partial(y_1,y_2,...,y_p)}{\partial(x_1,x_2,...,x_n)}\cdot\frac{\partial(x_1,x_2,...,x_n)}{\partial(t_1,t_2,...,t_m)}\cdot\frac{\partial(t_1,t_2,...,t_m)}{\partial(w_1,w_2,...,w_k)}.$$

Non essendo il prodotto tra matrici commutativo, è bene sottolineare l'importanza del giusto ordine di tale prodotto.

Esempio 43 : Sia $z=f(x,y)$; operiamo un passaggio da coordinate cartesiane (x,y) a coordinate polari (ρ,ϑ) mediante la $g(\rho,\vartheta)=(x_0+\rho\cos\vartheta, y_0+\rho\,\mathrm{sen}\,\vartheta)$. Otteniamo una funzione composta $z=f(g(\rho,\vartheta))$. Supponendo f differenziabile avremo:

$J_{f(g)}(\rho,\vartheta)=J_f(x(\rho,\vartheta),y(\rho,\vartheta))\cdot J_g(\rho,\vartheta)$, ovvero: $\dfrac{\partial z}{\partial(\rho,\vartheta)}=\dfrac{\partial z}{\partial(x,y)}\cdot\dfrac{\partial(x,y)}{\partial(\rho,\vartheta)}$ da cui:

$$\left\|\begin{matrix}\dfrac{\partial z}{\partial\rho} & \dfrac{\partial z}{\partial\vartheta}\end{matrix}\right\|=\left\|\begin{matrix}\dfrac{\partial z}{\partial x} & \dfrac{\partial z}{\partial y}\end{matrix}\right\|\cdot\left\|\begin{matrix}\dfrac{\partial x}{\partial\rho} & \dfrac{\partial x}{\partial\vartheta}\\ \dfrac{\partial y}{\partial\rho} & \dfrac{\partial y}{\partial\vartheta}\end{matrix}\right\|=\|\,z'_x \quad z'_y\,\|\cdot\left\|\begin{matrix}\cos\vartheta & -\rho\,\mathrm{sen}\,\vartheta\\ \mathrm{sen}\,\vartheta & \rho\cos\vartheta\end{matrix}\right\|=$$

$$=\|\,z'_x\cos\vartheta+z'_y\,\mathrm{sen}\,\vartheta \quad -z'_x\,\rho\,\mathrm{sen}\,\vartheta+z'_y\,\rho\cos\vartheta\,\|.$$

Esempio 44 : Siano $g:\mathbb{R}^3\to\mathbb{R}, x=g(t_1,t_2,t_3)$, $f:\mathbb{R}\to\mathbb{R}^2,(y_1,y_2)=(f_1(x),f_2(x))$ funzioni differenziabili. Allora avremo: $J_{f(g)}(\mathbb{T})=J_f(g(\mathbb{T}))\cdot J_g(\mathbb{T})$, esprimibile come:

$\dfrac{\partial(y_1,y_2)}{\partial(t_1,t_2,t_3)}=\dfrac{\partial(y_1,y_2)}{\partial x}\cdot\dfrac{\partial x}{\partial(t_1,t_2,t_3)}$ e quindi:

$$\left\|\begin{matrix}\dfrac{\partial y_1}{\partial t_1} & \dfrac{\partial y_1}{\partial t_2} & \dfrac{\partial y_1}{\partial t_3}\\ \dfrac{\partial y_2}{\partial t_1} & \dfrac{\partial y_2}{\partial t_2} & \dfrac{\partial y_2}{\partial t_3}\end{matrix}\right\|=\left\|\begin{matrix}\dfrac{\partial y_1}{\partial x}\\ \dfrac{\partial y_2}{\partial x}\end{matrix}\right\|\cdot\left\|\begin{matrix}\dfrac{\partial x}{\partial t_1} & \dfrac{\partial x}{\partial t_2} & \dfrac{\partial x}{\partial t_3}\end{matrix}\right\| \text{ da cui otteniamo:}$$

$$\left\|\begin{matrix}\dfrac{\partial y_1}{\partial t_1} & \dfrac{\partial y_1}{\partial t_2} & \dfrac{\partial y_1}{\partial t_3}\\ \dfrac{\partial y_2}{\partial t_1} & \dfrac{\partial y_2}{\partial t_2} & \dfrac{\partial y_2}{\partial t_3}\end{matrix}\right\|=\left\|\begin{matrix}\dfrac{\partial y_1}{\partial x}\cdot\dfrac{\partial x}{\partial t_1} & \dfrac{\partial y_1}{\partial x}\cdot\dfrac{\partial x}{\partial t_2} & \dfrac{\partial y_1}{\partial x}\cdot\dfrac{\partial x}{\partial t_3}\\ \dfrac{\partial y_2}{\partial x}\cdot\dfrac{\partial x}{\partial t_1} & \dfrac{\partial y_2}{\partial x}\cdot\dfrac{\partial x}{\partial t_2} & \dfrac{\partial y_2}{\partial x}\cdot\dfrac{\partial x}{\partial t_3}\end{matrix}\right\|.$$

Esempio 45 : Siano $g:\mathbb{R}^2\to\mathbb{R}^3,(t_1,t_2)\to(x_1,x_2,x_3)=\left(2t_1-t_2^2,t_1\,t_2,t_1^2\right)$ e sia inoltre: $f:\mathbb{R}^3\to\mathbb{R}^2,(x_1,x_2,x_3)\to(y_1,y_2)=(x_1+x_2,x_1\,x_3)$. Sarà allora:

$\dfrac{\partial(y_1,y_2)}{\partial(t_1,t_2)}=\dfrac{\partial(y_1,y_2)}{\partial(x_1,x_2,x_3)}\cdot\dfrac{\partial(x_1,x_2,x_3)}{\partial(t_1,t_2)}$ e quindi:

$$\left\|\begin{matrix}\dfrac{\partial y_1}{\partial t_1} & \dfrac{\partial y_1}{\partial t_2}\\ \dfrac{\partial y_2}{\partial t_1} & \dfrac{\partial y_2}{\partial t_2}\end{matrix}\right\|=\left\|\begin{matrix}\dfrac{\partial y_1}{\partial x_1} & \dfrac{\partial y_1}{\partial x_2} & \dfrac{\partial y_1}{\partial x_3}\\ \dfrac{\partial y_2}{\partial x_1} & \dfrac{\partial y_2}{\partial x_2} & \dfrac{\partial y_2}{\partial x_3}\end{matrix}\right\|\cdot\left\|\begin{matrix}\dfrac{\partial x_1}{\partial t_1} & \dfrac{\partial x_1}{\partial t_2}\\ \dfrac{\partial x_2}{\partial t_1} & \dfrac{\partial x_2}{\partial t_2}\\ \dfrac{\partial x_3}{\partial t_1} & \dfrac{\partial x_3}{\partial t_2}\end{matrix}\right\| \text{, da cui, sostituendo:}$$

$$\left\|\begin{matrix}\dfrac{\partial y_1}{\partial t_1} & \dfrac{\partial y_1}{\partial t_2}\\ \dfrac{\partial y_2}{\partial t_1} & \dfrac{\partial y_2}{\partial t_2}\end{matrix}\right\|=\left\|\begin{matrix}1 & 1 & 0\\ x_3 & 0 & x_1\end{matrix}\right\|\cdot\left\|\begin{matrix}2 & -2t_2\\ t_2 & t_1\\ 2t_1 & 0\end{matrix}\right\|=\left\|\begin{matrix}2+t_2 & t_1-2t_2\\ 2x_3+2x_1t_1 & -2x_3t_2\end{matrix}\right\|=$$

sostituendo a x_1 e x_3 le loro espressioni:

$$=\left\|\begin{matrix}2+t_2 & t_1-2t_2\\ 6t_1^2-2t_1t_2^2 & -2t_1^2t_2\end{matrix}\right\|.$$

DERIVATE SECONDE DELLE FUNZIONI COMPOSTE

Iniziamo trattando, come caso introduttivo, il seguente:
Siano $g:\mathbb{R}\to\mathbb{R}^2, t\to(x_1,x_2)$ e $f:\mathbb{R}^2\to\mathbb{R},(x_1,x_2)\to y$, ovvero:
$y=f(x_1,x_2)=f(x_1(t),x_2(t))$. Supponiamo che g sia derivabile due volte e che f sia differenziabile due volte. Posto $\mathbb{X}=(x_1,x_2)$, da cui $\mathbb{X}'=\dfrac{\mathrm{d}\mathbb{X}}{\mathrm{d}t}=(x_1',x_2')$, per la regola di derivazione di funzione composta avremo:
$\dfrac{\mathrm{d}y}{\mathrm{d}t}=\nabla f(\mathbb{X}(t))\cdot\mathbb{X}'(t)=\dfrac{\partial f}{\partial x_1}\cdot\dfrac{\mathrm{d}x_1}{\mathrm{d}t}+\dfrac{\partial f}{\partial x_2}\cdot\dfrac{\mathrm{d}x_2}{\mathrm{d}t}$. Ora dobbiamo calcolare:
$\dfrac{\mathrm{d}^2y}{\mathrm{d}t^2}=\dfrac{\mathrm{d}}{\mathrm{d}t}\left(\dfrac{\mathrm{d}y}{\mathrm{d}t}\right)=\dfrac{\mathrm{d}}{\mathrm{d}t}\left(\dfrac{\partial f}{\partial x_1}\cdot\dfrac{\mathrm{d}x_1}{\mathrm{d}t}+\dfrac{\partial f}{\partial x_2}\cdot\dfrac{\mathrm{d}x_2}{\mathrm{d}t}\right)$.
Ma $\dfrac{\partial f}{\partial x_1}$ e $\dfrac{\partial f}{\partial x_2}$ sono funzioni di x_1 e x_2 e quindi, applicando ancora la regola di derivazione di funzione composta, nonchè la derivata di somma e di prodotto, avremo:
$\dfrac{\mathrm{d}^2y}{\mathrm{d}t^2}=\dfrac{\mathrm{d}}{\mathrm{d}t}\left(\dfrac{\partial f}{\partial x_1}\right)\cdot\dfrac{\mathrm{d}x_1}{\mathrm{d}t}+\dfrac{\partial f}{\partial x_1}\cdot\dfrac{\mathrm{d}}{\mathrm{d}t}\left(\dfrac{\mathrm{d}x_1}{\mathrm{d}t}\right)+\dfrac{\mathrm{d}}{\mathrm{d}t}\left(\dfrac{\partial f}{\partial x_2}\right)\cdot\dfrac{\mathrm{d}x_2}{\mathrm{d}t}+\dfrac{\partial f}{\partial x_2}\cdot\dfrac{\mathrm{d}}{\mathrm{d}t}\left(\dfrac{\mathrm{d}x_2}{\mathrm{d}t}\right)$.
Per la derivata di funzione composta sarà:
$\dfrac{\mathrm{d}}{\mathrm{d}t}\left(\dfrac{\partial f}{\partial x_1}\right)=\dfrac{\partial}{\partial x_1}\left(\dfrac{\partial f}{\partial x_1}\right)\cdot\dfrac{\mathrm{d}x_1}{\mathrm{d}t}+\dfrac{\partial}{\partial x_2}\left(\dfrac{\partial f}{\partial x_1}\right)\cdot\dfrac{\mathrm{d}x_2}{\mathrm{d}t}=\dfrac{\partial^2 f}{\partial x_1^2}\cdot\dfrac{\mathrm{d}x_1}{\mathrm{d}t}+\dfrac{\partial^2 f}{\partial x_1\partial x_2}\cdot\dfrac{\mathrm{d}x_2}{\mathrm{d}t}$, e
$\dfrac{\mathrm{d}}{\mathrm{d}t}\left(\dfrac{\partial f}{\partial x_2}\right)=\dfrac{\partial}{\partial x_1}\left(\dfrac{\partial f}{\partial x_2}\right)\cdot\dfrac{\mathrm{d}x_1}{\mathrm{d}t}+\dfrac{\partial}{\partial x_2}\left(\dfrac{\partial f}{\partial x_2}\right)\cdot\dfrac{\mathrm{d}x_2}{\mathrm{d}t}=\dfrac{\partial^2 f}{\partial x_2\partial x_1}\cdot\dfrac{\mathrm{d}x_1}{\mathrm{d}t}+\dfrac{\partial^2 f}{\partial x_2^2}\cdot\dfrac{\mathrm{d}x_2}{\mathrm{d}t}$,
dalle quali otteniamo:

$$\dfrac{\mathrm{d}^2y}{\mathrm{d}t^2}=\left(\dfrac{\partial^2 f}{\partial x_1^2}\cdot\dfrac{\mathrm{d}x_1}{\mathrm{d}t}+\dfrac{\partial^2 f}{\partial x_1\partial x_2}\cdot\dfrac{\mathrm{d}x_2}{\mathrm{d}t}\right)\cdot\dfrac{\mathrm{d}x_1}{\mathrm{d}t}+\dfrac{\partial f}{\partial x_1}\cdot\dfrac{\mathrm{d}^2x_1}{\mathrm{d}t^2}+$$
$$+\left(\dfrac{\partial^2 f}{\partial x_2\partial x_1}\cdot\dfrac{\mathrm{d}x_1}{\mathrm{d}t}+\dfrac{\partial^2 f}{\partial x_2^2}\cdot\dfrac{\mathrm{d}x_2}{\mathrm{d}t}\right)\cdot\dfrac{\mathrm{d}x_2}{\mathrm{d}t}+\dfrac{\partial f}{\partial x_2}\cdot\dfrac{\mathrm{d}^2x_2}{\mathrm{d}t^2}=$$
$$=\dfrac{\partial^2 f}{\partial x_1^2}\left(\dfrac{\mathrm{d}x_1}{\mathrm{d}t}\right)^2+2\,\dfrac{\partial^2 f}{\partial x_1\partial x_2}\,\dfrac{\mathrm{d}x_1}{\mathrm{d}t}\,\dfrac{\mathrm{d}x_2}{\mathrm{d}t}+\dfrac{\partial^2 f}{\partial x_2^2}\left(\dfrac{\mathrm{d}x_2}{\mathrm{d}t}\right)^2+\dfrac{\partial f}{\partial x_1}\,\dfrac{\mathrm{d}^2x_1}{\mathrm{d}t^2}+\dfrac{\partial f}{\partial x_2}\,\dfrac{\mathrm{d}^2x_2}{\mathrm{d}t^2},$$

avendo supposto soddisfatte le ipotesi per le quali $\dfrac{\partial^2 f}{\partial x_1\partial x_2}=\dfrac{\partial^2 f}{\partial x_2\partial x_1}$.
In notazione abbreviata l'uguaglianza precedente può essere scritta come:
$y''=f_{11}''\,(x_1')^2+2\,f_{12}''\,x_1'\,x_2'+f_{22}''\,(x_2')^2+f_1'\,x_1''+f_2'\,x_2''$ ed anche:
$y''=y_{11}''\,(x_1')^2+2\,y_{12}''\,x_1'\,x_2'+y_{22}''\,(x_2')^2+y_1'\,x_1''+y_2'\,x_2''$.

Da $\mathbb{X}=(x_1,x_2)$ e da $\mathbb{X}'=\dfrac{\mathrm{d}\mathbb{X}}{\mathrm{d}t}=(x_1',x_2')$ segue $\mathbb{X}''=\dfrac{\mathrm{d}^2\mathbb{X}}{\mathrm{d}t^2}=(x_1'',x_2'')$, ed essendo poi $\nabla f(x_1,x_2)=(f_1',f_2')=(y_1',y_2')$ e $\mathbb{H}(f(x_1,x_2))=\begin{Vmatrix} y_{11}'' & y_{12}'' \\ y_{12}'' & y_{22}'' \end{Vmatrix}=\begin{Vmatrix} f_{11}'' & f_{12}'' \\ f_{12}'' & f_{22}'' \end{Vmatrix}$, la precedente uguaglianza può essere espressa anche come:
$y''=\mathbb{X}'(t)\cdot\mathbb{H}(f(\mathbb{X}(t)))\cdot(\mathbb{X}'(t))^{\mathrm{T}}+\nabla f(\mathbb{X}(t))\cdot\mathbb{X}''(t)$,
dove $\mathbb{X}'(t),\mathbb{X}''(t),\nabla f(\mathbb{X}(t))\in\mathbb{R}^2$ e $\mathbb{H}(f(\mathbb{X}(t)))$ è una matrice 2×2.

La precente uguaglianza può essere generalizzata. Vediamo i vari casi.

Posto $\mathbb{X}=(x_1,x_2,...,x_n)$, siano:
$g:\mathbb{R}\to\mathbb{R}^n, t\to(x_1,x_2,...,x_n)$ e $f:\mathbb{R}^n\to\mathbb{R},(x_1,x_2,...,x_n)\to y$, ovvero:
$\mathbb{R}\to\mathbb{R}^n\to\mathbb{R}, t\to\mathbb{X}\to y, y=f(\mathbb{X}(t))$.
Con calcoli analoghi ai precedenti si ottiene ancora:

$$\frac{\mathrm{d}^2 y}{\mathrm{d}t^2} = \mathbb{X}'(t) \cdot \mathbb{H}(f(\mathbb{X}(t))) \cdot (\mathbb{X}'(t))^{\mathrm{T}} + \nabla f(\mathbb{X}(t)) \cdot \mathbb{X}''(t),$$
dove $\mathbb{X}'(t), \mathbb{X}''(t), \nabla f(\mathbb{X}(t)) \in \mathbb{R}^n$ e $\mathbb{H}(f(\mathbb{X}(t)))$ è una matrice $n \times n$.

Posto $\mathbb{X} = (x_1, x_2, ..., x_n)$, $\mathbb{Y} = (y_1, y_2, ..., y_p)$ e $y_i = f_i(\mathbb{X})$, siano:
$g : \mathbb{R} \to \mathbb{R}^n, t \to (x_1, x_2, ..., x_n)$ e $f : \mathbb{R}^n \to \mathbb{R}^p, (x_1, x_2, ..., x_n) \to (y_1, y_2, ..., y_p)$, ovvero:
$\mathbb{R} \to \mathbb{R}^n \to \mathbb{R}^p, t \to \mathbb{X} \to \mathbb{Y}, \mathbb{Y} = f(\mathbb{X}(t)) = (f_1(\mathbb{X}(t)), ..., f_p(\mathbb{X}(t)))$.
Con calcoli analoghi ai precedenti si ottengono p uguaglianze:
$$\frac{\mathrm{d}^2 y_i}{\mathrm{d}t^2} = \mathbb{X}'(t) \cdot \mathbb{H}(f_i(\mathbb{X}(t))) \cdot (\mathbb{X}'(t))^{\mathrm{T}} + \nabla f_i(\mathbb{X}(t)) \cdot \mathbb{X}''(t), \ 1 \le i \le p$$
dove $\mathbb{X}'(t), \mathbb{X}''(t), \nabla f(\mathbb{X}(t)) \in \mathbb{R}^n$ e $\mathbb{H}(f_i(\mathbb{X}(t)))$ è una matrice $n \times n$.

Posto infine $\mathbb{T} = (t_1, t_2, ..., t_m)$, $\mathbb{X} = (x_1, x_2, ..., x_n)$ e $\mathbb{Y} = (y_1, y_2, ..., y_p)$, vale il:
Teorema 18 : Siano date le due funzioni $g : \mathbb{R}^m \to \mathbb{R}^n, (t_1, t_2, ..., t_m) \to (x_1, x_2, ..., x_n)$,
$f : \mathbb{R}^n \to \mathbb{R}^p, (x_1, x_2, ..., x_n) \to (y_1, y_2, ..., y_p)$, ovvero:
$\mathbb{R}^m \to \mathbb{R}^n \to \mathbb{R}^p, \mathbb{T} \to \mathbb{X} \to \mathbb{Y}, \mathbb{Y} = f(\mathbb{X}(\mathbb{T})) = (f_1(\mathbb{X}(\mathbb{T})), ..., f_p(\mathbb{X}(\mathbb{T}))) = f(g(\mathbb{T}))$
ambedue differenziabili due volte. Vale allora la seguente formula generale:
$$\frac{\partial^2 y_i}{\partial t_j \partial t_k} = \frac{\partial \mathbb{X}(\mathbb{T})}{\partial t_j} \cdot \mathbb{H}(f_i(\mathbb{X}(\mathbb{T}))) \cdot \left(\frac{\partial \mathbb{X}(\mathbb{T})}{\partial t_k}\right)^{\mathrm{T}} + \nabla f_i(\mathbb{X}(\mathbb{T})) \cdot \frac{\partial^2 \mathbb{X}(\mathbb{T})}{\partial t_j \partial t_k},$$
$1 \le i \le p$, $1 \le j, k \le m$, costituita da $p \cdot \dfrac{m^2 + m}{2}$ uguaglianze, dove:
$\dfrac{\partial \mathbb{X}(\mathbb{T})}{\partial t_j}, \dfrac{\partial^2 \mathbb{X}(\mathbb{T})}{\partial t_j \partial t_k}, \nabla f_i(\mathbb{X}(\mathbb{T})) \in \mathbb{R}^n$ e $\mathbb{H}(f_i(\mathbb{X}(\mathbb{T})))$ è una matrice $n \times n$.

Esempio 46 : Siano $g : \mathbb{R}^3 \to \mathbb{R}, x = g(t_1, t_2, t_3)$, $f : \mathbb{R} \to \mathbb{R}^2, (y_1, y_2) = (f_1(x), f_2(x))$ funzioni differenziabili due volte.
Essendo: $\dfrac{\partial x(\mathbb{T})}{\partial t_i} = \dfrac{\partial x}{\partial t_i}$, $\mathbb{H}(f_1(\mathbb{X}(\mathbb{T}))) = \dfrac{\mathrm{d}^2 y_1}{\mathrm{d}x^2}$, $\nabla f_1(\mathbb{X}(\mathbb{T})) = \dfrac{\mathrm{d}y_1}{\mathrm{d}x}$, $\dfrac{\partial^2 x(\mathbb{T})}{\partial t_1 \partial t_3} = \dfrac{\partial^2 x}{\partial t_1 \partial t_3}$,
avremo:
$$\frac{\partial^2 y_1}{\partial t_1 \partial t_3} = \frac{\partial \mathbb{X}(\mathbb{T})}{\partial t_1} \cdot \mathbb{H}(f_1(\mathbb{X}(\mathbb{T}))) \cdot \left(\frac{\partial \mathbb{X}(\mathbb{T})}{\partial t_3}\right)^{\mathrm{T}} + \nabla f_1(\mathbb{X}(\mathbb{T})) \cdot \frac{\partial^2 \mathbb{X}(\mathbb{T})}{\partial t_1 \partial t_3} \text{ da cui:}$$
$$\frac{\partial^2 y_1}{\partial t_1 \partial t_3} = \frac{\partial x}{\partial t_1} \cdot \frac{\mathrm{d}^2 y_1}{\mathrm{d}x^2} \cdot \frac{\partial x}{\partial t_3} + \frac{\mathrm{d}y_1}{\mathrm{d}x} \cdot \frac{\partial^2 x}{\partial t_1 \partial t_3}.$$

Esempio 47 : Siano $g : \mathbb{R}^2 \to \mathbb{R}^2, (x_1, x_2) = g(t_1, t_2)$, $f : \mathbb{R}^2 \to \mathbb{R}, y = f(x_1, x_2)$ funzioni differenziabili due volte. Essendo:
$$\frac{\partial \mathbb{X}(\mathbb{T})}{\partial t_i} = \left(\frac{\partial x_1}{\partial t_i}, \frac{\partial x_2}{\partial t_i}\right), \ \mathbb{H}(f(\mathbb{X}(\mathbb{T}))) = \begin{Vmatrix} \dfrac{\partial^2 y}{\partial x_1^2} & \dfrac{\partial^2 y}{\partial x_1 \partial x_2} \\ \dfrac{\partial^2 y}{\partial x_1 \partial x_2} & \dfrac{\partial^2 y}{\partial x_2^2} \end{Vmatrix} = \begin{Vmatrix} y''_{11} & y''_{12} \\ y''_{12} & y''_{22} \end{Vmatrix},$$
$$\nabla f(\mathbb{X}(\mathbb{T})) = \left(\frac{\partial y}{\partial x_1}, \frac{\partial y}{\partial x_2}\right) = (y'_1, y'_2), \ \frac{\partial^2 \mathbb{X}(\mathbb{T})}{\partial t_1^2} = \left(\frac{\partial^2 x_1}{\partial t_1^2}, \frac{\partial^2 x_2}{\partial t_1^2}\right), \text{ avremo:}$$
$$\frac{\partial^2 y}{\partial t_1^2} = \begin{Vmatrix} \dfrac{\partial x_1}{\partial t_1}, \dfrac{\partial x_2}{\partial t_1} \end{Vmatrix} \cdot \begin{Vmatrix} y''_{11} & y''_{12} \\ y''_{12} & y''_{22} \end{Vmatrix} \cdot \begin{Vmatrix} \dfrac{\partial x_1}{\partial t_1} \\ \dfrac{\partial x_2}{\partial t_1} \end{Vmatrix} + \left(\frac{\partial y}{\partial x_1}, \frac{\partial y}{\partial x_2}\right) \cdot \left(\frac{\partial^2 x_1}{\partial t_1^2}, \frac{\partial^2 x_2}{\partial t_1^2}\right) =$$
$$\frac{\partial^2 y}{\partial t_1^2} = \left(\frac{\partial x_1}{\partial t_1}\right)^2 y''_{11} + 2\,\frac{\partial x_1}{\partial t_1}\,\frac{\partial x_2}{\partial t_1}\,y''_{12} + \left(\frac{\partial x_2}{\partial t_1}\right)^2 y''_{22} + y'_1\,\frac{\partial^2 x_1}{\partial t_1^2} + y'_2\,\frac{\partial^2 x_2}{\partial t_1^2};$$

$$\frac{\partial^2 y}{\partial t_1 \partial t_2} = \left\| \frac{\partial x_1}{\partial t_1}, \frac{\partial x_2}{\partial t_1} \right\| \cdot \left\| \begin{matrix} y''_{11} & y''_{12} \\ y''_{12} & y''_{22} \end{matrix} \right\| \cdot \left\| \begin{matrix} \dfrac{\partial x_1}{\partial t_2} \\ \dfrac{\partial x_2}{\partial t_2} \end{matrix} \right\| + \left(\frac{\partial y}{\partial x_1}, \frac{\partial y}{\partial x_2} \right) \cdot \left(\frac{\partial^2 x_1}{\partial t_1 \partial t_2}, \frac{\partial^2 x_2}{\partial t_1 \partial t_2} \right) =$$

$$\frac{\partial^2 y}{\partial t_1 \partial t_2} = \frac{\partial x_1}{\partial t_1}\frac{\partial x_1}{\partial t_2} y''_{11} + \left(\frac{\partial x_1}{\partial t_2}\frac{\partial x_2}{\partial t_1} + \frac{\partial x_1}{\partial t_1}\frac{\partial x_2}{\partial t_2} \right) y''_{12} + \frac{\partial x_2}{\partial t_1}\frac{\partial x_2}{\partial t_2} y''_{22} +$$

$$+ y'_1 \frac{\partial^2 x_1}{\partial t_1 \partial t_2} + y'_2 \frac{\partial^2 x_2}{\partial t_1 \partial t_2} ;$$

$$\frac{\partial^2 y}{\partial t_2^2} = \left\| \frac{\partial x_1}{\partial t_2}, \frac{\partial x_2}{\partial t_2} \right\| \cdot \left\| \begin{matrix} y''_{11} & y''_{12} \\ y''_{12} & y''_{22} \end{matrix} \right\| \cdot \left\| \begin{matrix} \dfrac{\partial x_1}{\partial t_2} \\ \dfrac{\partial x_2}{\partial t_2} \end{matrix} \right\| + \left(\frac{\partial y}{\partial x_1}, \frac{\partial y}{\partial x_2} \right) \cdot \left(\frac{\partial^2 x_1}{\partial t_2^2}, \frac{\partial^2 x_2}{\partial t_2^2} \right) =$$

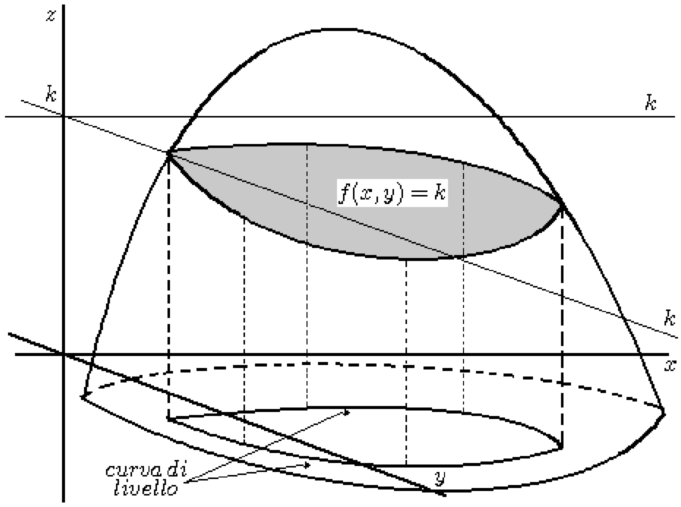
$$\frac{\partial^2 y}{\partial t_2^2} = \left(\frac{\partial x_1}{\partial t_2} \right)^2 y''_{11} + 2\frac{\partial x_1}{\partial t_2}\frac{\partial x_2}{\partial t_2} y''_{12} + \left(\frac{\partial x_2}{\partial t_2} \right)^2 y''_{22} + y'_1 \frac{\partial^2 x_1}{\partial t_2^2} + y'_2 \frac{\partial^2 x_2}{\partial t_2^2} .$$

FUNZIONI IN FORMA IMPLICITA

Sia $f : \mathbb{R}^n \to \mathbb{R}^m$, $n \geq 1, m \geq 1$. La funzione è detta "in forma esplicita" quando si conosce la legge f che permette di associare ad ogni elemento del dominio la sua unica immagine.
Tutte le funzioni utilizzate fino ad ora sono funzioni in forma esplicita.
Consideriamo ora l'equazione $x^2 + y^2 = 1$. I punti $(x, y) \in \mathbb{R}^2$ che la soddisfano costituiscono la circonferenza di centro l'origine e raggio pari a 1, ovvero la circonferenza trigonometrica. Tale circonferenza non è però il grafico di una funzione $y = f(x)$, in quanto ad ogni valore x corrispondono due valori y (tolti i casi $x = \pm 1$) così come ad ogni y (tolti i casi $y = \pm 1$) corrispondono due valori della x. Se risolviamo algebricamente l'equazione possiamo ottenere $y = \pm\sqrt{1 - x^2}$, cioè l'espressione esplicita di due possibili funzioni $y = y(x)$, aventi per grafico rispettivamente la semicirconferenza superiore e quella inferiore, oppure possiamo ottenere $x = \pm\sqrt{1 - y^2}$, cioè due possibili espressioni esplicite di funzioni $x = x(y)$, dove si sono scambiati i ruoli delle variabili indipendente e dipendente; queste due ultime funzioni hanno per grafico, rispettivamente, la semicirconferenza di destra e quella di sinistra. Si vede facilmente che $x^2 + (y(x))^2 = (x(y))^2 + y^2 = 1$. Potremmo comunque scomporre la circonferenza in un numero opportuno di tratti che non si sovrappongano, per formare altre possibili funzioni che però, a differenza delle quattro precedenti, non risulterebbero continue.
Consideriamo ora invece l'equazione $f(x, y) = x\,e^y + x^2\,y^2 = 1$.
Vediamo che il punto $P_0 = (1, 0)$ la soddisfa; tale equazione non è risolvibile rispetto a y, mentre può esserlo rispetto a x, in quanto costituisce un polinomio di II grado:
$y^2\,x^2 + e^y\,x - 1 = 0$ che ha soluzioni: $x = \dfrac{-\,e^y \pm \sqrt{e^{2y} + 4y^2}}{2y^2}$, che sono le due possibili forme di funzione esplicita $x = x(y)$ ricavabili dall'equazione data.
Facilmente si verifica che $f(x(y), y) = 1$.
Consideriamo infine l'equazione $f(x, y) = x\,e^y + x^3\,y^2 = 1$.
Anche ora il punto $P_0 = (1, 0)$ la soddisfa, ma tale equazione non è risolvibile nè rispetto a y nè rispetto a x. Ci chiediamo se mediante quest'ultima equazione sia possibile garantire l'esistenza di funzioni $y = y(x)$ oppure $x = x(y)$, che diremo, nell'impossibilità di ricavarne l'espressione esplicita, funzioni in forma implicita. Esamineremo di seguito una serie di esempi introduttivi, aumentando di volta in volta prima il numero delle variabili e poi quello delle e-

quazioni, per giungere infine a stabilire un Teorema valido per il caso generale. Iniziamo dal caso più semplice, cioè quello di una equazione in due variabili.

FUNZIONI DEFINITE IMPLICITAMENTE DA UNA EQUAZIONE
I Caso: Equazione $f(x,y)=k$: funzione implicita $\mathbb{R}\to\mathbb{R}$
Supponiamo di avere una generica equazione in due variabili $f(x,y)=k, k\in\mathbb{R}$.
Ci poniamo anzitutto alla ricerca di un punto $\mathrm{P}_0=(x_0,y_0)$ che la soddisfi: $f(x_0,y_0)=k$.
Trovato P_0 vogliamo vedere se, in un intorno di x_0, l'insieme dei punti (x,y) tali che $f(x,y)=k$ costituisce il grafico di una funzione $y=y(x)$, della quale non è detto che si riesca a trovare l'espressione esplicita, ma della quale si cerca di sapere comunque se è continua, derivabile e quale sia la sua derivata (si può anche vedere se in un intorno di y_0 è definibile una funzione $x=x(y)$ con analoghe proprietà).
In caso positivo, ovvero se $y_0=y(x_0)$ e $f(x,y(x))=k$, $\forall\, x\in\mathfrak{I}(x_0)$ diremo che la funzione $y=y(x)$ è definita implicitamente dall'equazione $f(x,y)=k$.



Questo problema può essere visto anche in modo geometrico attraverso le cosiddette curve di livello. Limitandoci al caso di $f:\mathbb{R}^2\to\mathbb{R}$ per poter rappresentare graficamente la situazione, posto $w=f(x,y)$, l'equazione $f(x,y)=k$ equivale a determinare le coppie (x,y) per le quali $w=k$, e ciò equivale a tagliare la superficie $w=f(x,y)$ con un piano, parallelo al piano $(x;y)$, ad altezza pari a k. L'intersezione tra la superficie ed il piano origina una curva, che è la proiezione mediante f di una curva giacente nel piano $(x;y)$: la curva di livello.
Possiamo allora vedere questa curva di livello come una curva nel vero senso del termine, ovvero come una funzione $g:\mathbb{R}\to\mathbb{R}^2, t\to(x(t),y(t))$.
Otteniamo quindi la seguente composizione di funzioni:
$\mathbb{R}\xrightarrow{g}\mathbb{R}^2\xrightarrow{f}\mathbb{R}, t\xrightarrow{g}(x(t),y(t))\xrightarrow{f}w$, dove però $w=f(x,y)=k$ costante.
Avremo allora, derivando:
$$\frac{\mathrm{d}w}{\mathrm{d}t}=\nabla f(x,y)\cdot\frac{\mathrm{d}(x,y)}{\mathrm{d}t}=\frac{\partial f}{\partial x}\cdot\frac{\mathrm{d}x}{\mathrm{d}t}+\frac{\partial f}{\partial y}\cdot\frac{\mathrm{d}y}{\mathrm{d}t}=f_x'\cdot x'(t)+f_y'\cdot y'(t),$$
dove $(x'(t),y'(t))$ rappresenta il vettore tangente alla curva data.
Ma $\dfrac{\mathrm{d}w}{\mathrm{d}t}=0$, in quanto w è costante, e quindi $\nabla f(x,y)\cdot(x'(t),y'(t))=0$, ovvero il gradiente di f ed il vettore tangente risultano perpendicolari, in un qualunque punto della curva di livello. Utilizzeremo nel seguito questa proprietà.
Torniamo al problema dell'esistenza della funzione implicita $y=y(x)$ mediante l'equazione $f(x,y)=k$.

Esistenza e proprietà di una tale funzione implicita sono stabilite nel seguente:
Teorema 19 (U. Dini) : Sia $f : \mathbb{R}^2 \to \mathbb{R}$ una funzione continua con derivata f'_y continua in $\mathbb{A} \subseteq \mathbb{R}^2$; sia $f(x_0, y_0) = k$ e sia $f'_y(x_0, y_0) \neq 0$. Allora esistono un intorno $\mathfrak{I}(x_0)$ e un'unica funzione continua $y = y(x)$, tale che $y_0 = y(x_0)$ e $f(x, y(x)) = k$, $\forall\, x \in \mathfrak{I}(x_0)$.
Inoltre, se anche f'_x è continua in $\mathbb{A}$, allora $y(x)$ è derivabile in $\mathfrak{I}(x_0)$ e risulta:
$y'(x) = \dfrac{\mathrm{d}y}{\mathrm{d}x} = -\dfrac{f'_x}{f'_y}$, $\forall\, x \in \mathfrak{I}(x_0)$. La funzione $y'(x)$ infine è continua $\forall\, x \in \mathfrak{I}(x_0)$.
Non diamo la dimostrazione di questo Teorema, ma verificheremo, mediante la derivata di funzione composta, il risultato trovato relativamente alla derivata $y'(x)$.
Notiamo come le ipotesi del Teorema forniscano una condizione sufficiente, e non necessaria, per l'esistenza della funzione implicita, come vedremo negli esempi che seguono; inoltre, se le ipotesi fossero tutte soddisfatte, esse implicano la differenziabilità di $f(x, y)$ in $\mathbb{A}$.
Con opportuni scambi nelle ipotesi si possono poi dedurre esistenza, continuità e derivabilità di una funzione implicita $x = x(y)$ in $\mathfrak{I}(y_0)$.

Esempio 48 : Consideriamo l'equazione $f(x, y) = x^2 - y = 0$. Nel punto $(x_0, y_0) = (0, 0)$ risulta $f(0, 0) = 0$; $f(x, y)$ è continua con derivate continue in tutto $\mathbb{R}^2$ e risulta inoltre $f'_y = -1 \neq 0$ mentre $f'_x = 2x$, che si annulla per $x = 0$.
La funzione definita da $f(x, y) = x^2 - y = 0$ è esplicitabile come $y = y(x) = x^2$, è continua e derivabile in tutto $\mathbb{R}$. In un intorno di $y = 0$ non è invece possibile definire una funzione $x = x(y)$, in quanto ogni $y > 0$ ha due corrispondenti: $x = \pm\sqrt{y}$. La ricerca della funzione inversa porta infatti a $y = \sqrt{x}$ oppure a $y = -\sqrt{x}$, a seconda che si restringa il dominio di $g(x) = x^2$ al solo $\mathbb{R}_+$ o al solo $\mathbb{R}_-$.

Esempio 49 : Sia $f(x, y) = x^3 - y = 0$. In $(x_0, y_0) = (0, 0)$ si ha $f(0, 0) = 0$; inoltre $f(x, y)$ è continua con derivate continue in $\mathbb{R}^2$ e risulta $f'_y = -1 \neq 0$ mentre $f'_x = 3x^2$, che si annulla per $x = 0$.
La funzione definita da $f(x, y) = x^3 - y = 0$ è esplicitabile come $y = y(x) = x^3$, è continua e derivabile in $\mathbb{R}$, dove risulta pure invertibile, con inversa $x = \sqrt[3]{y}$.
Anche se $f'_x(0, 0) = 0$ quindi, in un intorno di $y = 0$ esiste la funzione (esplicita), a conferma del fatto che la condizione $f'_y \neq 0$ (o $f'_x \neq 0$) è solo sufficiente e non necessaria.
Geometricamente, la funzione $y = x^3$ ha in $x = 0$ la retta tangente orizzontale, ma questo avviene in un punto di flesso, e non in un punto di minimo, come nell'esempio precedente, e questo permette l'esistenza della funzione $x = x(y)$.

Funzione implicita $\mathbb{R} \to \mathbb{R}$: **Derivata prima**
Se $f(x, y) = k$ e se definiamo implicitamente y come $y = y(x)$, otteniamo questa composizione di funzioni:
$\mathbb{R} \to \mathbb{R}^2 \xrightarrow{f} \mathbb{R}$, $x \to (x, y(x)) \xrightarrow{f} w = f(x, y) = k$. Possiamo quindi calcolare:
$$\frac{\mathrm{d}w}{\mathrm{d}x} = \frac{\partial(f)}{\partial(x, y)} \cdot \frac{\partial(x, y)}{\partial(x)} = \nabla f(x, y) \cdot \left(\frac{\mathrm{d}x}{\mathrm{d}x}, \frac{\mathrm{d}y}{\mathrm{d}x}\right) = f'_x \cdot 1 + f'_y \cdot y'(x) = 0,$$
dato che f è costante, e quindi si ricava:
$$y'(x) = \frac{\mathrm{d}y}{\mathrm{d}x} = -\frac{f'_x}{f'_y}.$$
Se si fosse definita implicitamente una funzione $x = x(y)$ si sarebbe invece ottenuto:
$\mathbb{R} \to \mathbb{R}^2 \xrightarrow{f} \mathbb{R}$, $y \to (x(y), y) \xrightarrow{f} w = f(x, y) = k$, e da questa:

$$\frac{\mathrm{d}w}{\mathrm{d}y} = \frac{\partial(f)}{\partial(x,y)} \cdot \frac{\partial(x,y)}{\partial(y)} = \nabla f(x,y) \cdot \left(\frac{\mathrm{d}x}{\mathrm{d}y}, \frac{\mathrm{d}y}{\mathrm{d}y}\right) = f'_x \cdot x'(y) + f'_y \cdot 1 = 0\,,$$

e quindi: $x'(y) = \dfrac{\mathrm{d}x}{\mathrm{d}y} = -\dfrac{f'_y}{f'_x}$.

Dato che $\dfrac{f'_y}{f'_x}$ è il reciproco di $\dfrac{f'_x}{f'_y}$, ritroviamo la regola di derivazione della funzione inversa:

$y'(x) = \dfrac{1}{x'(y)}$, nell'ipotesi che sia $f'_x \neq 0$ e $f'_y \neq 0$.

Funzione implicita $\mathbb{R} \to \mathbb{R}$: Equazione retta tangente

Riprendendo la $f(x,y) = k$, nell'ipotesi che con essa venga definita la $y = y(x)$, possiamo scrivere l'**equazione della retta tangente** alla curva $y = y(x)$ nel punto x_0, che sarà la seguente: $y - y(x_0) = y'(x_0)\,(x - x_0)$, da cui: $y - y_0 = -\dfrac{f'_x(x_0,y_0)}{f'_y(x_0,y_0)} \cdot (x - x_0)$,

che si scrive anche come: $f'_x(x_0,y_0)\,(x - x_0) + f'_y(x_0,y_0)\,(y - y_0) = 0$ ovvero: $\nabla f(x_0,y_0) \cdot (x - x_0, y - y_0) = 0$, cioè si ritrova la perpendicolarità tra il gradiente di f ed il vettore tangente alla curva di livello.

Esempio 50 : Consideriamo l'equazione $f(x,y) = x\,e^y + x^2\,y^2 = 1$.

Risulta $f(1,0) = 1$; inoltre $f'_x = e^y + 2x\,y^2$ mentre $f'_y = x\,e^y + 2x^2\,y$. Quindi, essendo $f'_x(1,0) = 1$ e $f'_y(1,0) = 1 \neq 0$, esiste, in un intorno del punto $x = 1$, la funzione implicita $y = y(x)$ e risulta: $y'(1) = -\dfrac{f'_x(1,0)}{f'_y(1,0)} = -1$.

Si ha anche, nello stesso intorno di $x = 1$: $y'(x) = -\dfrac{e^y + 2x\,y^2}{x\,e^y + 2x^2\,y}$.

Per l'equazione della retta tangente alla $y = y(x)$ in $x = 1$, essendo $y(1) = y_0 = 0$, avremo: $y = -1 \cdot (x - 1) = 1 - x$.

Funzione implicita $\mathbb{R} \to \mathbb{R}$: Derivata seconda; Polinomio di Taylor

Supponendo, data $w = f(x,y) = k$ con $f(x_0,y_0) = k$, che f sia derivabile due volte con derivate continue, vediamo come calcolare la derivata seconda della funzione implicita $y = y(x)$, nell'ipotesi che sia $f'_y(x_0,y_0) \neq 0$. Possiamo seguire vari modi.

Da $\dfrac{\mathrm{d}w}{\mathrm{d}x} = f'_x + f'_y \cdot y' = 0$, derivando ancora rispetto a x, essendo f'_x e f'_y funzioni sia di x che di y, per la regola di derivazione delle funzioni composte otteniamo, posto $\dfrac{\mathrm{d}y}{\mathrm{d}x} = y'$:

$$\frac{\mathrm{d}^2w}{\mathrm{d}x^2} = \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\mathrm{d}w}{\mathrm{d}x}\right) = \frac{\mathrm{d}}{\mathrm{d}x}\left(f'_x + f'_y\,y'\right) = \frac{\mathrm{d}}{\mathrm{d}x}(f'_x) + \frac{\mathrm{d}}{\mathrm{d}x}\left(f'_y\right) \cdot y' + f'_y \cdot \frac{\mathrm{d}}{\mathrm{d}x}(y') =$$
$$= \frac{\partial}{\partial x}(f'_x) \cdot \frac{\mathrm{d}x}{\mathrm{d}x} + \frac{\partial}{\partial y}(f'_x) \cdot \frac{\mathrm{d}y}{\mathrm{d}x} + \left(\frac{\partial}{\partial x}\left(f'_y\right) \cdot \frac{\mathrm{d}x}{\mathrm{d}x} + \frac{\partial}{\partial y}\left(f'_y\right) \cdot \frac{\mathrm{d}y}{\mathrm{d}x}\right) \cdot y' + f'_y \cdot \frac{\mathrm{d}}{\mathrm{d}x}(y') =$$
$$= f''_{xx} + f''_{xy} \cdot y' + \left(f''_{yx} + f''_{yy} \cdot y'\right) \cdot y' + f'_y \cdot y'' =$$
$$= f''_{xx} + 2f''_{xy} \cdot y' + f''_{yy} \cdot (y')^2 + f'_y \cdot y'' = 0\,,$$

in quanto, essendo w costante, la sua derivata seconda è nulla, e per la differenziabilità del secondo ordine risulta pure $f''_{xy} = f''_{yx}$. Risolvendo rispetto a y'' otteniamo:

$y'' = -\dfrac{f''_{xx} + 2f''_{xy}\,y' + f''_{yy}\,(y')^2}{f'_y}$ dalla quale, sostituendo $y' = -\dfrac{f'_x}{f'_y}$, otteniamo infine:

$$y'' = -\frac{f''_{xx}\left(f'_y\right)^2 - 2f''_{xy}\, f'_x\, f'_y + f''_{yy}\left(f'_x\right)^2}{\left(f'_y\right)^3}\,.$$

Si perviene allo stesso risultato derivando rispetto a x il quoziente $-\dfrac{f'_x}{f'_y}$, tenendo al solito presente che si tratta di derivare le funzioni composte f'_x e f'_y, per cui avremo:

$$y'' = \frac{\mathrm{d}}{\mathrm{d}x}(y') = -\frac{\left(f''_{xx}\cdot\dfrac{\mathrm{d}x}{\mathrm{d}x} + f''_{xy}\cdot\dfrac{\mathrm{d}y}{\mathrm{d}x}\right) f'_y - \left(f''_{yx}\cdot\dfrac{\mathrm{d}x}{\mathrm{d}x} + f''_{yy}\cdot\dfrac{\mathrm{d}y}{\mathrm{d}x}\right) f'_x}{\left(f'_y\right)^2} =$$

$$= -\frac{f''_{xx}\, f'_y + f''_{xy}\, y'\, f'_y - f''_{yx}\, f'_x - f''_{yy}\, y'\, f'_x}{\left(f'_y\right)^2} = -\frac{f''_{xx}\left(f'_y\right)^2 - 2f''_{xy}\, f'_x\, f'_y + f''_{yy}\left(f'_x\right)^2}{\left(f'_y\right)^3}\,,$$

avendo fatto la sostituzione $y' = -\dfrac{f'_x}{f'_y}$.

Si può infine anche applicare il Teorema 18, nel caso $x \to (x,y) \to w = f(x,y) = k$ dove ora $\mathbb{X} = \mathbb{X}(x) = (x,y)$, da cui $\mathbb{X}'(x) = (1, y')$ e $\mathbb{X}''(x) = (0, y'')$ ottenendo:
$w'' = \mathbb{X}'(x)\cdot \mathbb{H}(f(\mathbb{X}(x)))\cdot(\mathbb{X}'(x))^{\mathrm{T}} + \nabla f(\mathbb{X}(x))\cdot \mathbb{X}''(x) = 0$ da cui:
$\|\, 1 \quad y'\,\| \cdot \left\|\begin{matrix} f''_{xx} & f''_{xy} \\ f''_{xy} & f''_{yy}\end{matrix}\right\| \cdot \left\|\begin{matrix} 1 \\ y'\end{matrix}\right\| + \left(f'_x, f'_y\right)\cdot(0, y'') = 0$ ed eseguendo i calcoli, si ricava infine: $y'' = -\dfrac{f''_{xx} + 2f''_{xy}\, y' + f''_{yy}\,(y')^2}{f'_y}$, per poi fare la sostituzione $y' = -\dfrac{f'_x}{f'_y}$.

Conoscendo la derivata seconda $y''(x_0)$ possiamo determinare allora, per la funzione implicita $y = y(x)$, anche l'espressione del polinomio di Taylor di II grado in $x = x_0$, che sarà:
$\mathrm{P}_2(x, x_0) = y_0 - \dfrac{f'_x(x_0,y_0)}{f'_y(x_0,y_0)}\,(x - x_0) + \dfrac{1}{2}\, y''(x_0)\,(x - x_0)^2$,
dove $y''(x_0)$ va calcolata con la formula trovata.

Esempio 51 : Calcoliamo $y''(1)$ dalla $f(x,y) = x\, e^y + x^2\, y^2 = 1$.
Essendo $f'_x = e^y + 2x\, y^2$ e $f'_y = x\, e^y + 2x^2\, y$, avremo allora:
$f''_{xx} = 2y^2$, $f''_{xy} = f''_{yx} = e^y + 4x\, y$, $f''_{yy} = x\, e^y + 2x^2$, dalle quali otteniamo:
$f''_{xx}(1,0) = 0$, $f''_{xy}(1,0) = 1$, $f''_{yy}(1,0) = 3$, da cui, essendo $f'_x(1,0) = 1$ e $f'_y(1,0) = 1$, si ha: $y''(1) = -\dfrac{0\cdot 1 - 2\cdot 1\cdot 1\cdot 1 + 3\cdot 1}{1^3} = -1$.
Possiamo scrivere poi l'espressione del polinomio di Taylor di II grado in $x = 1$, che sarà:
$\mathrm{P}_2(x, 1) = y(1) + y'(1)\,(x-1) + \dfrac{1}{2}\, y''(1)\,(x-1)^2 = 0 - (x-1) - \dfrac{1}{2}\,(x-1)^2$.

II Caso: Equazione $f(x_1, x_2, y) = k$: funzione implicita $\mathbb{R}^2 \to \mathbb{R}$

Supponiamo ora di avere una equazione in tre variabili $f(x_1, x_2, y) = k$, $k \in \mathbb{R}$. Esista un punto (x_1^0, x_2^0, y_0) che la soddisfa: $f\left(x_1^0, x_2^0, y_0\right) = k$, e supponiamo che f sia derivabile con derivate continue. La presenza di una sola equazione ci permette di ricavare implicitamente (o esplicitamente) una variabile, diciamo y, in funzione delle altre due: $y = y(x_1, x_2)$, ottenendo così una funzione implicita $\mathbb{R}^2 \to \mathbb{R}$.
Il Teorema del Dini applicato a questo caso ci dice che otterremo una funzione implicita continua e derivabile, rispetto a x_1 e rispetto a x_2, nell'ipotesi che risulti $f'_y\left(x_1^0, x_2^0, y_0\right) \neq 0$.

Funzione implicita $\mathbb{R}^2 \to \mathbb{R}$: **Derivata prima**
Posto $w = f(x_1, x_2, y) = k$ e $(\mathbb{X}|\, y) = (x_1, x_2, y)$, abbiamo la composizione di funzioni:
$\mathbb{R}^2 \to \mathbb{R}^3 \xrightarrow{f} \mathbb{R}, \mathbb{X} \to (\mathbb{X}|\, y) \xrightarrow{f} w = f(\mathbb{X}|\, y) = k$.
Posto $y'_{x_i} = y'_i,\ f'_{x_i} = f'_i$, possiamo allora calcolare:
$\dfrac{\partial(w)}{\partial(x_1, x_2)} = \dfrac{\partial(w)}{\partial(\mathbb{X}|\, y)} \cdot \dfrac{\partial(\mathbb{X}|\, y)}{\partial(x_1, x_2)} = 0$ essendo f costante, e dato che $\dfrac{\partial x_i}{\partial x_i} = 1$, mentre $\dfrac{\partial x_i}{\partial x_j} = 0$, in quanto x_1 e x_2 sono variabili tra loro indipendenti, otteniamo:

$$\left\|\begin{matrix} \dfrac{\partial w}{\partial x_1} & \dfrac{\partial w}{\partial x_2} \end{matrix}\right\| = \nabla f(\mathbb{X}|\, y) \cdot \left\|\begin{matrix} 1 & 0 \\ 0 & 1 \\ y'_1 & y'_2 \end{matrix}\right\| = \mathbb{O} \text{ ovvero:}$$

$$\begin{cases} \dfrac{\partial w}{\partial x_1} = \left(f'_1, f'_2, f'_y\right) \cdot (1, 0, y'_1) = f'_1 + f'_y \cdot y'_1 = 0 \\ \dfrac{\partial w}{\partial x_2} = \left(f'_1, f'_2, f'_y\right) \cdot (0, 1, y'_2) = f'_2 + f'_y \cdot y'_2 = 0 \end{cases},$$

dalle quali otteniamo: $(y'_1, y'_2) = \left(-\dfrac{f'_1}{f'_y}, -\dfrac{f'_2}{f'_y} \right)$ e quindi:

$$\nabla y\big(x_1^0, x_2^0\big) = \left(-\frac{f'_1(x_1^0, x_2^0, y_0)}{f'_y(x_1^0, x_2^0, y_0)}, -\frac{f'_2(x_1^0, x_2^0, y_0)}{f'_y(x_1^0, x_2^0, y_0)} \right).$$

Funzione implicita $\mathbb{R}^2 \to \mathbb{R}$: **Equazione piano tangente**
Possiamo scrivere l'**equazione del piano tangente** alla superficie $y = y(x_1, x_2)$ nel punto (x_1^0, x_2^0), con $y\big(x_1^0, x_2^0\big) = y_0$:

$$y - y_0 = -\frac{f'_1(x_1^0, x_2^0, y_0)}{f'_y(x_1^0, x_2^0, y_0)} \cdot \big(x_1 - x_1^0\big) - \frac{f'_2(x_1^0, x_2^0, y_0)}{f'_y(x_1^0, x_2^0, y_0)} \cdot \big(x_2 - x_2^0\big),$$

che può, essendo $f'_y\big(x_1^0, x_2^0, y_0\big) \neq 0$, essere scritta anche come:

$$f'_1\big(x_1^0, x_2^0, y_0\big)\big(x_1 - x_1^0\big) + f'_2\big(x_1^0, x_2^0, y_0\big)\big(x_2 - x_2^0\big) + f'_y\big(x_1^0, x_2^0, y_0\big)(y - y_0) = 0,$$

ovvero $\nabla f\big(x_1^0, x_2^0, y_0\big) \cdot \big(x_1 - x_1^0, x_2 - x_2^0, y - y_0\big) = 0$.
Aumentando la dimensione del problema, abbiamo ora che il vettore gradiente risulta perpendicolare al piano tangente alla superficie (e non più alla curva) di livello.

Funzione implicita $\mathbb{R}^2 \to \mathbb{R}$: **Derivate seconde**
Data la composizione di funzioni: $\mathbb{R}^2 \to \mathbb{R}^3 \xrightarrow{f} \mathbb{R}$:
$\mathbb{X} \to (\mathbb{X}|\, y) \xrightarrow{f} w = f(\mathbb{X}|\, y) = k$, applicando il Teorema 18, posto: $y''_{x_i x_j} = y''_{ij}$, $f''_{x_i x_j} = f''_{ij}$ ed essendo $(\mathbb{X}|\, y) = (x_1, x_2, y)$, da cui:
$\dfrac{\partial(\mathbb{X}|\, y)}{\partial x_1} = (1, 0, y'_1)$ e $\dfrac{\partial(\mathbb{X}|\, y)}{\partial x_2} = (0, 1, y'_2)$, avremo anche:
$\dfrac{\partial^2(\mathbb{X}|\, y)}{\partial x_1^2} = (0, 0, y''_{11})$, $\dfrac{\partial^2(\mathbb{X}|\, y)}{\partial x_1\, \partial x_2} = \dfrac{\partial^2(\mathbb{X}|\, y)}{\partial x_2\, \partial x_1} = (0, 0, y''_{12})$ e $\dfrac{\partial^2(\mathbb{X}|\, y)}{\partial x_2^2} = (0, 0, y''_{22})$,
e quindi:

$$\frac{\partial^2 w}{\partial x_1^2} = \|1 \quad 0 \quad y'_1\| \cdot \left\|\begin{matrix} f''_{11} & f''_{12} & f''_{1y} \\ f''_{12} & f''_{22} & f''_{2y} \\ f''_{1y} & f''_{2y} & f''_{yy} \end{matrix}\right\| \cdot \left\|\begin{matrix} 1 \\ 0 \\ y'_1 \end{matrix}\right\| + \left(f'_1, f'_2, f'_y\right) \cdot (0, 0, y''_{11}) = 0, \text{ da cui:}$$

$$y''_{11} = -\frac{f''_{11} + 2f''_{1y}\,y'_1 + f''_{yy}\,(y'_1)^2}{f'_y}\,;$$

$$\frac{\partial^2 w}{\partial x_1\,\partial x_2} = \|\,1 \quad 0 \quad y'_1\,\| \cdot \left\|\begin{matrix} f''_{11} & f''_{12} & f''_{1y} \\ f''_{12} & f''_{22} & f''_{2y} \\ f''_{1y} & f''_{2y} & f''_{yy} \end{matrix}\right\| \cdot \left\|\begin{matrix} 0 \\ 1 \\ y'_2 \end{matrix}\right\| + \left(f'_1, f'_2, f'_y\right) \cdot (0, 0, y''_{12}) = 0\,, \text{ da cui:}$$

$$y''_{12} = -\frac{f''_{12} + f''_{1y}\,y'_2 + f''_{2y}\,y'_1 + f''_{yy}\,y'_1\,y'_2}{f'_y}\,;$$

$$\frac{\partial^2 w}{\partial x_2^2} = \|\,0 \quad 1 \quad y'_2\,\| \cdot \left\|\begin{matrix} f''_{11} & f''_{12} & f''_{1y} \\ f''_{12} & f''_{22} & f''_{2y} \\ f''_{1y} & f''_{2y} & f''_{yy} \end{matrix}\right\| \cdot \left\|\begin{matrix} 0 \\ 1 \\ y'_2 \end{matrix}\right\| + \left(f'_1, f'_2, f'_y\right) \cdot (0, 0, y''_{22}) = 0\,, \text{ da cui:}$$

$$y''_{22} = -\frac{f''_{22} + 2f''_{2y}\,y'_2 + f''_{yy}\,(y'_2)^2}{f'_y}\,.$$

Funzione implicita $\mathbb{R}^2 \to \mathbb{R}$: Differenziali I° e II°; Polinomio di Taylor

Essendo la funzione di variabile vettoriale, per un'eventuale analisi dei suoi punti di massimo e di minimo sono più utili, come vedremo nel seguito, i differenziali totali I° e II°. Vediamo allora come calcolare, partendo dall'equazione $f(x_1, x_2, y) = k$, i differenziali primo e secondo di una funzione $y = y(x_1, x_2)$ definita implicitamente.

Essendo $y'_1 = -\dfrac{f'_1}{f'_y}$ e $y'_2 = -\dfrac{f'_2}{f'_y}$, sostituendo si ha:

$\mathrm{d}y = y'_1\,\mathrm{d}x_1 + y'_2\,\mathrm{d}x_2 = -\dfrac{f'_1}{f'_y}\,\mathrm{d}x_1 - \dfrac{f'_2}{f'_y}\,\mathrm{d}x_2$, che si poteva anche calcolare differenziando:

$\mathrm{d}w = f'_1\,\mathrm{d}x_1 + f'_2\,\mathrm{d}x_2 + f'_y\,\mathrm{d}y = 0$, e da questa ricavare $\mathrm{d}y = -\dfrac{f'_1}{f'_y}\,\mathrm{d}x_1 - \dfrac{f'_2}{f'_y}\,\mathrm{d}x_2$, nell'ipotesi $f'_y \neq 0$.

Per calcolare $\mathrm{d}^2 y$ differenziamo nuovamente $\mathrm{d}w = f'_1\,\mathrm{d}x_1 + f'_2\,\mathrm{d}x_2 + f'_y\,\mathrm{d}y = 0$, tenendo presente che f'_1, f'_2 e f'_y sono funzioni di x_1, x_2 e y, ma che ora anche $\mathrm{d}y$ dipende da x_1 e da x_2, essendo y funzione implicita di x_1 e x_2; avremo quindi:

$$\mathrm{d}(\mathrm{d}w) = 0 = \mathrm{d}\big(f'_1\,\mathrm{d}x_1 + f'_2\,\mathrm{d}x_2 + f'_y\,\mathrm{d}y\big) = \frac{\partial}{\partial x_1}\big(f'_1\,\mathrm{d}x_1 + f'_2\,\mathrm{d}x_2 + f'_y\,\mathrm{d}y\big)\,\mathrm{d}x_1 +$$

$$+\frac{\partial}{\partial x_2}\big(f'_1\,\mathrm{d}x_1 + f'_2\,\mathrm{d}x_2 + f'_y\,\mathrm{d}y\big)\,\mathrm{d}x_2 + \frac{\partial}{\partial y}\big(f'_1\,\mathrm{d}x_1 + f'_2\,\mathrm{d}x_2 + f'_y\,\mathrm{d}y\big)\,\mathrm{d}y =$$

$$= \left(f''_{11}\,\mathrm{d}x_1 + f''_{21}\,\mathrm{d}x_2 + f''_{y1}\,\mathrm{d}y + f'_y\,\frac{\partial}{\partial x_1}(\mathrm{d}y)\right)\mathrm{d}x_1 +$$

$$+ \left(f''_{12}\,\mathrm{d}x_1 + f''_{22}\,\mathrm{d}x_2 + f''_{y2}\,\mathrm{d}y + f'_y\,\frac{\partial}{\partial x_2}(\mathrm{d}y)\right)\mathrm{d}x_2 +$$

$$+ \left(f''_{1y}\,\mathrm{d}x_1 + f''_{2y}\,\mathrm{d}x_2 + f''_{yy}\,\mathrm{d}y + f'_y\,\frac{\partial}{\partial y}(\mathrm{d}y)\right)\mathrm{d}y = 0\,. \text{ Ma}$$

$$\big(f''_{11}\,\mathrm{d}x_1 + f''_{21}\,\mathrm{d}x_2 + f''_{y1}\,\mathrm{d}y\big)\,\mathrm{d}x_1 + \big(f''_{12}\,\mathrm{d}x_1 + f''_{22}\,\mathrm{d}x_2 + f''_{y2}\,\mathrm{d}y\big)\,\mathrm{d}x_2 +$$
$$+ \big(f''_{1y}\,\mathrm{d}x_1 + f''_{2y}\,\mathrm{d}x_2 + f''_{yy}\,\mathrm{d}y\big)\,\mathrm{d}y = \mathrm{d}^2 f(\mathbb{X}\,|\,y) \text{ mentre}$$

$$f'_y \cdot \frac{\partial}{\partial x_1}(\mathrm{d}y)\,\mathrm{d}x_1 + f'_y \cdot \frac{\partial}{\partial x_2}(\mathrm{d}y)\,\mathrm{d}x_2 + f'_y \cdot \frac{\partial}{\partial y}(\mathrm{d}y)\,\mathrm{d}y = f'_y\,\mathrm{d}(\mathrm{d}y) = f'_y\,\mathrm{d}^2 y\,, \text{ e quindi si ha:}$$

$\mathrm{d}^2 f(\mathbb{X}\,|\,y) + f'_y\,\mathrm{d}^2 y = 0$, da cui ricaviamo infine:

$$\mathrm{d}^2 y = -\frac{\mathrm{d}^2 f(x_1, x_2, y)}{f'_y} = -\frac{\mathrm{d}^2 f(\mathbb{X}\,|\,y)}{f'_y}\,.$$

Possiamo ora utilizzare le espressioni trovate per $\mathrm{d}y$ e d^2y per scrivere l'espressione del polinomio di Taylor di II grado della funzione implicita $y = y(x_1, x_2)$ che sarà dato da:

$$\mathrm{P}_2(x_1, x_2) = y_0 + \mathrm{d}y + \frac{1}{2}\,\mathrm{d}^2y = y_0 - \frac{f_1'}{f_y'}\,\mathrm{d}x_1 - \frac{f_2'}{f_y'}\,\mathrm{d}x_2 - \frac{1}{2}\,\frac{\mathrm{d}^2f(\mathbb{X}\,|\,y)}{f_y'}\,.$$

Esempio 52 : Sia $f(x, y, z) = x\,e^{x\,(y-z)} + y - z = 0$. Risulta $f(0,0,0) = 0$. Essendo:

$f_x' = e^{x\,(y-z)} + x\,e^{x\,(y-z)}\,(y - z)$, da cui $f_x'(0,0,0) = 1 \neq 0$;
$f_y' = x^2\,e^{x\,(y-z)} + 1$, da cui $f_y'(0,0,0) = 1 \neq 0$,
$f_z' = -\,x^2\,e^{x\,(y-z)} - 1$, da cui $f_z'(0,0,0) = -\,1 \neq 0$,

possiamo definire una a scelta tra le tre possibili funzioni implicite: $x = x(y,z)$, $y = y(x,z)$ oppure $z = z(x,y)$. Scegliamo la terza possibilità ed avremo:

$$\frac{\partial z(0,0)}{\partial x} = -\,\frac{f_x'(0,0,0)}{f_z'(0,0,0)} = -\,\frac{1}{-\,1} = 1\,,\quad \frac{\partial z(0,0)}{\partial y} = -\,\frac{f_y'(0,0,0)}{f_z'(0,0,0)} = -\,\frac{1}{-\,1} = 1$$

per cui :
$\mathrm{d}z(0,0) = 1\cdot\mathrm{d}x + 1\cdot\mathrm{d}y = \mathrm{d}x + \mathrm{d}y$. Essendo poi:

$f_{xx}'' = 2\,(y - z)\,e^{x\,(y-z)} + x\,e^{x\,(y-z)}\,(y - z)^2$ segue $f_{xx}''(0,0,0) = 0$,
$f_{yy}'' = x^3\,e^{x\,(y-z)}$ segue $f_{yy}''(0,0,0) = 0$,
$f_{zz}'' = x^3\,e^{x\,(y-z)}$ segue $f_{zz}''(0,0,0) = 0$,
$f_{xy}'' = f_{yx}'' = 2x\,e^{x\,(y-z)} + x^2\,e^{x\,(y-z)}\,(y - z)$ segue $f_{xy}''(0,0,0) = 0$,
$f_{xz}'' = f_{zx}'' = -\,2x\,e^{x\,(y-z)} - x^2\,e^{x\,(y-z)}\,(y - z)$ segue $f_{xz}''(0,0,0) = 0$,
$f_{yz}'' = f_{zy}'' = -\,x^3\,e^{x\,(y-z)}$ segue $f_{yz}''(0,0,0) = 0$,

da cui otteniamo: $\mathrm{d}^2z = -\,\dfrac{\mathrm{d}^2f}{f_z'} = -\,\dfrac{0}{-\,1} = 0$, e quindi:

$\mathrm{P}_2(x,y,0,0) = z_0 + \mathrm{d}z + \dfrac{1}{2}\,\mathrm{d}^2z = 0 + \mathrm{d}x + \mathrm{d}y + 0 = \mathrm{d}x + \mathrm{d}y$; essendo:
$\mathrm{d}x = x - 0 = x$ e $\mathrm{d}y = y - 0 = y$, avremo infine $\mathrm{P}_2(x,y,0,0) = x + y$.

Usiamo ora invece l'equazione $f(x,y,z) = x\,e^{x\,(y-z)} + y - z = 1$. Risulta $f(1,0,0) = 0$; inoltre, dato che non cambiano le derivate prime e seconde della funzione, avremo:
$f_x'(1,0,0) = 1 \neq 0$; $f_y'(1,0,0) = 2 \neq 0$ e $f_z'(1,0,0) = -\,2 \neq 0$, quindi rimangono valide le stesse tre possibilità di funzione implicita; scegliamo ancora $z = z(x,y)$, ed avremo:

$\mathrm{d}z = -\,\dfrac{f_x'(1,0,0)}{f_z'(1,0,0)}\,\mathrm{d}x - \dfrac{f_y'(1,0,0)}{f_z'(1,0,0)}\,\mathrm{d}y = \dfrac{1}{2}\,\mathrm{d}x + \mathrm{d}y$. Sarà poi:
$f_{xx}''(1,0,0) = 0$, $f_{yy}''(1,0,0) = 1$, $f_{zz}''(1,0,0) = 1$, $f_{xy}''(1,0,0) = 2$, $f_{xz}''(1,0,0) = -\,2$, $f_{yz}''(1,0,0) = -\,1$, e quindi avremo:
$\mathrm{d}^2f(1,0,0) = (\mathrm{d}y)^2 + (\mathrm{d}z)^2 + 4\mathrm{d}x\,\mathrm{d}y - 4\mathrm{d}x\,\mathrm{d}z - 2\mathrm{d}y\,\mathrm{d}z$, dalla quale ricaviamo:

$$\mathrm{d}^2z(1,0,0) = -\,\frac{\mathrm{d}^2f(1,0,0)}{f_z'(1,0,0)} = -\,\frac{(\mathrm{d}y)^2 + (\mathrm{d}z)^2 + 4\mathrm{d}x\,\mathrm{d}y - 4\mathrm{d}x\,\mathrm{d}z - 2\mathrm{d}y\,\mathrm{d}z}{-\,2};$$

essendo $\mathrm{d}z = \dfrac{1}{2}\,\mathrm{d}x + \mathrm{d}y$, sostituendo, ricaviamo:

$$\mathrm{d}^2z = \frac{1}{2}\left[(\mathrm{d}y)^2 + \left(\frac{1}{2}\,\mathrm{d}x + \mathrm{d}y\right)^2 + 4\mathrm{d}x\,\mathrm{d}y - 4\mathrm{d}x\left(\frac{1}{2}\,\mathrm{d}x + \mathrm{d}y\right) - 2\mathrm{d}y\left(\frac{1}{2}\,\mathrm{d}x + \mathrm{d}y\right)\right]$$

e quindi $\mathrm{d}^2z(1,0,0) = -\,\dfrac{7}{8}\,(\mathrm{d}x)^2$, per cui il polinomio di Taylor di II grado avrà la forma:

$\mathrm{P}_2(x,y,0,0) = z_0 + \mathrm{d}z + \dfrac{1}{2}\,\mathrm{d}^2z = 0 + \dfrac{1}{2}\,\mathrm{d}x + \mathrm{d}y - \dfrac{7}{16}\,(\mathrm{d}x)^2$; essendo:
$\mathrm{d}x = x - 1$ e $\mathrm{d}y = y - 0 = y$, avremo infine:

$$\mathrm{P}_2(x,y,0,0)=\frac{1}{2}\,(x-1)+y-\frac{7}{16}\,(x-1)^2\,.$$

III Caso: Equazione $f(x_1,x_2,...,x_n,y)=k$: **funzione implicita** $\mathbb{R}^n\to\mathbb{R}$
Generalizzando i due casi precedenti, supponiamo di avere una sola equazione in $n+1$ variabili: $f(x_1,x_2,...,x_n,y)=k$, $k\in\mathbb{R}$, con $f:\mathbb{R}^{n+1}\to\mathbb{R}$ funzione derivabile con derivate continue. Posto $\mathbb{X}=(x_1,x_2,...,x_n)$ e $(\mathbb{X}|\,y)=(x_1,x_2,...,x_n,y)$, sia $f(\mathbb{X}_0|\,y_0)=k$ e sia $f'_y(\mathbb{X}_0|\,y_0)\neq 0$. Esiste allora la funzione implicita $y=y(x_1,x_2,...,x_n)$ in un intorno $\mathfrak{I}(\mathbb{X}_0)$, con $y_0=y(\mathbb{X}_0)$, ed avremo la seguente composizione di funzioni:
$\mathbb{R}^n\to\mathbb{R}^{n+1}\overset{f}{\to}\mathbb{R}$, $\mathbb{X}\to(\mathbb{X}\,|\,y)\overset{f}{\to}w=f(\mathbb{X}\,|\,y)=k$.

Funzione implicita $\mathbb{R}^n\to\mathbb{R}$: **Derivate prime**
Da $\mathbb{X}\to(\mathbb{X}\,|\,y)\overset{f}{\to}w=f(\mathbb{X}\,|\,y)=k$, derivando rispetto alla generica variabile x_i si ha:

$$\frac{\partial w}{\partial x_i}=\frac{\partial(w)}{\partial(\mathbb{X}\,|\,y)}\cdot\frac{\partial(\mathbb{X}\,|\,y)}{\partial x_i}=\nabla f(\mathbb{X}\,|\,y)\cdot\frac{\partial(\mathbb{X}\,|\,y)}{\partial x_i}=\frac{\partial f}{\partial x_i}+\frac{\partial f}{\partial y}\cdot\frac{\partial y}{\partial x_i}=f'_i+f'_y\cdot\frac{\partial y}{\partial x_i}=0\,,$$

essendo $\dfrac{\partial x_i}{\partial x_i}=1$ e $\dfrac{\partial x_j}{\partial x_i}=0$ se $i\neq j$, in quanto le variabili x_i sono tra loro indipendenti, per cui ricaviamo: $\dfrac{\partial y}{\partial x_i}=-\dfrac{f'_i}{f'_y}$, $1\leq i\leq n$.

Funzione implicita $\mathbb{R}^n\to\mathbb{R}$: **Equazione iperpiano tangente**
Risulta $\nabla y=\left(-\dfrac{f'_1}{f'_y},\,-\dfrac{f'_2}{f'_y},...,\,-\dfrac{f'_n}{f'_y}\right)$ e l'**equazione dell'iperpiano tangente** all'ipersuperficie $y=y(x_1,x_2,...,x_n)$ nel punto $(x_1^0,x_2^0,...,x_n^0)$ sarà data da:
$y-y_0=\displaystyle\sum_{i:1}^{n}\left(-\frac{f'_i}{f'_y}\right)\cdot\left(x_i-x_i^0\right)$, cioè $\displaystyle\sum_{i:1}^{n}f'_i\cdot\left(x_i-x_i^0\right)+f'_y\cdot(y-y_0)=0$, ovvero:
$\nabla f\left(x_1^0,x_2^0,...,x_n^0,y_0\right)\cdot\left(x_1-x_1^0,x_2-x_2^0,...,x_n-x_n^0,y-y_0\right)=0$, che esprime la solita relazione di perpendicolarità tra il gradiente della funzione f e, stavolta, l'iperpiano tangente.

Funzione implicita $\mathbb{R}^n\to\mathbb{R}$: **Derivate seconde**
Dalla composizione di funzioni:
$\mathbb{R}^n\to\mathbb{R}^{n+1}\overset{f}{\to}\mathbb{R}\,:\,\mathbb{X}\to(\mathbb{X}|\,y)\overset{f}{\to}w=f(\mathbb{X}|\,y)=k$, derivando si ha:
$\dfrac{\partial(\mathbb{X}|\,y)}{\partial x_i}=(0,..,1_i,..,0,y'_i)$ dalle quali otteniamo anche:
$\dfrac{\partial^2(\mathbb{X}|\,y)}{\partial x_i\,\partial x_j}=\left(0,..,0,y''_{ij}\right)$, e quindi, applicando il Teorema 18, avremo:

$$\frac{\partial^2 w}{\partial x_i\,\partial x_j}=\frac{\partial(\mathbb{X}|\,y)}{\partial x_i}\cdot\mathbb{H}(f(\mathbb{X}|\,y))\cdot\left(\frac{\partial(\mathbb{X}|\,y)}{\partial x_j}\right)^{\mathrm{T}}+\left(f'_1,..,f'_n,f'_y\right)\cdot\left(0,..,0,y''_{ij}\right)=0$$

che permette di ricavare le derivate parziali seconde:

$$y''_{ij}=\frac{\partial^2 y}{\partial x_i\,\partial x_j}=-\frac{1}{f'_y}\cdot\frac{\partial(\mathbb{X}|\,y)}{\partial x_i}\cdot\mathbb{H}(f(\mathbb{X}|\,y))\cdot\left(\frac{\partial(\mathbb{X}|\,y)}{\partial x_j}\right)^{\mathrm{T}},\ 1\leq i,j\leq n\ .$$

Funzione implicita $\mathbb{R}^n\to\mathbb{R}$: **Differenziali I° e II°**
Operando in maniera analoga a quanto visto nel caso di funzione implicita $\mathbb{R}^2\to\mathbb{R}$, dalla composizione $\mathbb{X}\to(\mathbb{X}|\,y)\overset{f}{\to}w=f(\mathbb{X}|\,y)=k$ otteniamo:

$\mathrm{d}w = f_1' \,\mathrm{d}x_1 + f_2' \,\mathrm{d}x_2 + .. + f_{x_n}' \,\mathrm{d}x_n + f_y' \,\mathrm{d}y = 0$, da cui ricaviamo $\mathrm{d}y$, e differenziando una seconda volta, otteniamo $\mathrm{d}^2 f(\mathbb{X}|\, y) + f_y' \,\mathrm{d}^2 y = 0$ da cui ricaviamo $\mathrm{d}^2 y = -\dfrac{\mathrm{d}^2 f(\mathbb{X}|\, y)}{f_y'}$.

Con questi differenziali possiamo costruire il Polinomio di Taylor.

Possiamo sintetizzare i casi visti sino ad ora, tutti accumunati dal fatto di avere una sola equazione in due o più variabili, dicendo che è possibile definire una funzione implicita, avente una sola variabile dipendente, mentre tutte le altre rimangono indipendenti, se il gradiente della funzione (dell'equazione) nel punto considerato è diverso dal vettore nullo, ovvero se ha almeno una componente diversa da zero. La variabile di tale derivazione può essere allora assunta come variabile dipendente. Dato che il gradiente è comunque una matrice Jacobiana, anche se formata da una sola riga, possiamo dire che tale Jacobiana deve avere caratteristica pari ad 1, ovvero massima, e questa sarà la condizione che vedremo valere nel caso generale.

FUNZIONI DEFINITE IMPLICITAMENTE DA UN SISTEMA DI EQUAZIONI

I Caso: Sistema $\begin{cases} f(x, y_1, y_2) = k_1 \\ g(x, y_1, y_2) = k_2 \end{cases}$: **funzione implicita** $\mathbb{R} \to \mathbb{R}^2$

Aumentiamo ora non il numero delle variabili ma quello delle equazioni, avendo, come caso minimo, un sistema di due equazioni in tre variabili: $\begin{cases} f(x, y_1, y_2) = k_1 \\ g(x, y_1, y_2) = k_2 \end{cases}$, $k_1, k_2 \in \mathbb{R}$.

La presenza di due equazioni potrebbe consentire di esplicitare due variabili, diciamo y_1 e y_2, in funzione della rimanente x. Se y_1 e y_2 non fossero esplicitabili, avremo comunque una funzione implicita $\mathbb{R} \to \mathbb{R}^2, x \to (y_1(x), y_2(x))$, se le opportune ipotesi che vedremo saranno soddisfatte. Il numero delle equazioni ci dice quante possono essere le variabili dipendenti; le variabili rimanenti avranno allora il ruolo di indipendenti. Vediamo allora cosa si ottiene per quanto riguarda le derivate di una funzione $\mathbb{R} \to \mathbb{R}^2, x \to (y_1(x), y_2(x))$ definita mediante un sistema di equazioni come quello dato.

Funzione implicita $\mathbb{R} \to \mathbb{R}^2$: **Derivate prime**

Posto $\begin{cases} w_1 = f(x, y_1, y_2) = k_1 \\ w_2 = g(x, y_1, y_2) = k_2 \end{cases}$ e posto $(x|\,\mathbb{Y}) = (x, y_1, y_2)$, avremo queste due composizioni di funzioni:

$\mathbb{R} \to \mathbb{R}^3 \xrightarrow{f} \mathbb{R},\ x \to (x|\,\mathbb{Y}) \xrightarrow{f} w_1 = f(x|\,\mathbb{Y}) = k_1$, e

$\mathbb{R} \to \mathbb{R}^3 \xrightarrow{g} \mathbb{R},\ x \to (x|\,\mathbb{Y}) \xrightarrow{g} w_2 = g(x|\,\mathbb{Y}) = k_2$,

dalle quali, derivando rispetto a x, essendo w_1 e w_2 costanti, otteniamo :

$\dfrac{\partial(w_1, w_2)}{\partial(x)} = \dfrac{\partial(f, g)}{\partial(x)} = \dfrac{\partial(f, g)}{\partial(x|\,\mathbb{Y})} \cdot \dfrac{\partial(x|\,\mathbb{Y})}{\partial(x)} = \mathbb{O}$, ovvero il sistema:

$\begin{cases} \dfrac{\mathrm{d}w_1}{\mathrm{d}x} = \nabla f(x|\,\mathbb{Y}) \cdot \dfrac{\partial(x|\,\mathbb{Y})}{\partial x} = 0 \\ \dfrac{\mathrm{d}w_2}{\mathrm{d}x} = \nabla g(x|\,\mathbb{Y}) \cdot \dfrac{\partial(x|\,\mathbb{Y})}{\partial x} = 0 \end{cases}$ equivalente a:

$\begin{cases} \dfrac{\partial f}{\partial x} \cdot \dfrac{\mathrm{d}x}{\mathrm{d}x} + \dfrac{\partial f}{\partial y_1} \cdot \dfrac{\mathrm{d}y_1}{\mathrm{d}x} + \dfrac{\partial f}{\partial y_2} \cdot \dfrac{\mathrm{d}y_2}{\mathrm{d}x} = 0 \\ \dfrac{\partial g}{\partial x} \cdot \dfrac{\mathrm{d}x}{\mathrm{d}x} + \dfrac{\partial g}{\partial y_1} \cdot \dfrac{\mathrm{d}y_1}{\mathrm{d}x} + \dfrac{\partial g}{\partial y_2} \cdot \dfrac{\mathrm{d}y_2}{\mathrm{d}x} = 0 \end{cases} \Rightarrow \begin{cases} f_{y_1}' \cdot y_1' + f_{y_2}' \cdot y_2' = -f_x' \\ g_{y_1}' \cdot y_1' + g_{y_2}' \cdot y_2' = -g_x' \end{cases}$, che risulta un sistema lineare di due equazioni nelle due incognite y_1' e y_2', e può essere scritto in forma matriciale come: $\begin{Vmatrix} f'_{y_1} & f'_{y_2} \\ g'_{y_1} & g'_{y_2} \end{Vmatrix} \cdot \begin{Vmatrix} y'_1 \\ y'_2 \end{Vmatrix} = - \begin{Vmatrix} f'_x \\ g'_x \end{Vmatrix}$, ed anche, utilizzando le matrici Jacobiane, come: $\dfrac{\partial(f,g)}{\partial(y_1,y_2)} \cdot \dfrac{\partial(y_1,y_2)}{\partial(x)} = - \dfrac{\partial(f,g)}{\partial(x)}$.

Per il teorema di Cramer, se $\begin{Vmatrix} f'_{y_1} & f'_{y_2} \\ g'_{y_1} & g'_{y_2} \end{Vmatrix} = \left| \dfrac{\partial(f,g)}{\partial(y_1,y_2)} \right| \neq 0$ il sistema lineare ammette un'unica soluzione $\left(\dfrac{\partial y_1}{\partial x}, \dfrac{\partial y_2}{\partial x} \right)$.

Il fatto che la matrice Jacobiana delle funzioni f e g rispetto alle variabili dipendenti y_1 e y_2 sia non singolare, ovvero che abbia caratteristica massima e, in questo caso, pari a 2, è la condizione che ci permette di enunciare il teorema del Dini relativo a questo caso; vale infatti il seguente:

Teorema 20 : Dato il sistema $\begin{cases} f(x,y_1,y_2) = k_1 \\ g(x,y_1,y_2) = k_2 \end{cases}$, siano f e g, $\mathbb{R}^3 \to \mathbb{R}$, funzioni derivabili con derivate continue, sia poi (x_0, y_1^0, y_2^0) un punto che soddisfa il sistema, e risulti infine $\left| \dfrac{\partial(f,g)(x_0,y_1^0,y_2^0)}{\partial(y_1,y_2)} \right| \neq 0$. Allora esiste un intorno $\mathfrak{I}(x_0)$ nel quale è definita una funzione implicita $\mathbb{R} \to \mathbb{R}^2, x \to (y_1(x), y_2(x))$, che risulta continua e derivabile $\forall\, x \in \mathfrak{I}(x_0)$, le cui derivate sono date dalla:

$$\begin{Vmatrix} y'_1 \\ y'_2 \end{Vmatrix} = \frac{\partial(y_1,y_2)}{\partial(x)} = - \begin{Vmatrix} \dfrac{\partial(f,g)}{\partial(y_1,y_2)} \end{Vmatrix}^{-1} \cdot \frac{\partial(f,g)}{\partial(x)} = - \begin{Vmatrix} f'_{y_1} & f'_{y_2} \\ g'_{y_1} & g'_{y_2} \end{Vmatrix}^{-1} \cdot \begin{Vmatrix} f'_x \\ g'_x \end{Vmatrix}.$$

Oltre alla soluzione globale espressa mediante l'inversa della matrice Jacobiana, esiste anche un'altra procedura, conseguenza del Teorema di Cramer sui sistemi lineari, che ci consente di calcolare singolarmente ciascuna incognita. Ognuna di esse è infatti esprimibile come un quoziente, a denominatore del quale c'è il determinante della matrice dei coefficienti delle incognite (lo Jacobiano) mentre a numeratore c'è il determinante della matrice ottenuta sostituendo nella Jacobiana la colonna dei termini noti al posto della colonna dei coefficienti dell'incognita cercata. Avremo allora, nel caso che stiamo trattando:

$$\frac{\mathrm{d}y_1}{\mathrm{d}x} = - \frac{\begin{vmatrix} f'_x & f'_{y_2} \\ g'_x & g'_{y_2} \end{vmatrix}}{\begin{vmatrix} f'_{y_1} & f'_{y_2} \\ g'_{y_1} & g'_{y_2} \end{vmatrix}} \;;\; \frac{\mathrm{d}y_2}{\mathrm{d}x} = - \frac{\begin{vmatrix} f'_{y_1} & f'_x \\ g'_{y_1} & g'_x \end{vmatrix}}{\begin{vmatrix} f'_{y_1} & f'_{y_2} \\ g'_{y_1} & g'_{y_2} \end{vmatrix}}.$$

Esempio 53 : Dato il sistema $\begin{cases} f(x,y_1,y_2) = x - y_1 + \operatorname{sen}(y_1 - y_2) + \cos(x - y_2) = 1 \\ g(x,y_1,y_2) = x + y_1 - \operatorname{sen}(x - y_2) - \cos(y_2 - y_1) = 1 \end{cases}$, vediamo che è soddisfatto dal punto $\mathrm{P}_0 = (x,y_1,y_2) = (1,1,1)$, e che le funzioni f e g sono differenziabili con derivate continue in $\mathbb{R}^2$. Calcoliamo la matrice Jacobiana:

$$\frac{\partial(f,g)}{\partial(x,y_1,y_2)} = \begin{Vmatrix} 1 - \operatorname{sen}(x - y_2) & \cos(y_1 - y_2) - 1 & \operatorname{sen}(x - y_2) - \cos(y_1 - y_2) \\ 1 - \cos(x - y_2) & 1 - \operatorname{sen}(y_2 - y_1) & \cos(x - y_2) + \operatorname{sen}(y_2 - y_1) \end{Vmatrix}.$$

Calcolando lo Jacobiano in P_0 otteniamo $\dfrac{\partial(f,g)(1,1,1)}{\partial(x,y_1,y_2)} = \begin{Vmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \end{Vmatrix}$, che ha caratteristica massima e pari a 2 in quanto $\begin{vmatrix} 0 & -1 \\ 1 & 1 \end{vmatrix} \neq 0$, e quindi in un intorno di $x = 1$ è definita una funzione implicita $\mathbb{R} \to \mathbb{R}^2, x \to (y_1(x), y_2(x))$, continua e derivabile, le cui derivate sono date dalla:

$$\frac{\mathrm{d}y_1}{\mathrm{d}x} = -\frac{\begin{vmatrix} 1-\operatorname{sen}(x-y_2) & \operatorname{sen}(x-y_2)-\cos(y_1-y_2) \\ 1-\cos(x-y_2) & \cos(x-y_2)+\operatorname{sen}(y_2-y_1) \end{vmatrix}}{\begin{vmatrix} \cos(y_1-y_2)-1 & \operatorname{sen}(x-y_2)-\cos(y_1-y_2) \\ 1-\operatorname{sen}(y_2-y_1) & \cos(x-y_2)+\operatorname{sen}(y_2-y_1) \end{vmatrix}};$$

$$\frac{\mathrm{d}y_2}{\mathrm{d}x} = -\frac{\begin{vmatrix} \cos(y_1-y_2)-1 & 1-\operatorname{sen}(x-y_2) \\ 1-\operatorname{sen}(y_2-y_1) & 1-\cos(x-y_2) \end{vmatrix}}{\begin{vmatrix} \cos(y_1-y_2)-1 & \operatorname{sen}(x-y_2)-\cos(y_1-y_2) \\ 1-\operatorname{sen}(y_2-y_1) & \cos(x-y_2)+\operatorname{sen}(y_2-y_1) \end{vmatrix}}.$$

Nel punto $x = 1$ avremo: $\dfrac{\mathrm{d}y_1(1)}{\mathrm{d}x} = -\dfrac{\begin{vmatrix} 1 & -1 \\ 0 & 1 \end{vmatrix}}{\begin{vmatrix} 0 & -1 \\ 1 & 1 \end{vmatrix}} = -1$ e $\dfrac{\mathrm{d}y_2(1)}{\mathrm{d}x} = -\dfrac{\begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix}}{\begin{vmatrix} 0 & -1 \\ 1 & 1 \end{vmatrix}} = 1$.

Esempio 54 : Dato il sistema: $\begin{cases} z = f(x,y) \\ g(x,y) = 0 \end{cases}$, vediamolo come $\begin{cases} F(x,y,z) = f(x,y) - z \\ g(x,y) = 0 \end{cases}$.
Sia P_0 un punto che soddisfa tale sistema, con f e g ovunque derivabili con derivate continue. Sarà $\dfrac{\partial(F,g)}{\partial(x,y,z)} = \left\|\begin{matrix} f'_x & f'_y & -1 \\ g'_x & g'_y & 0 \end{matrix}\right\|$. Se $\left|\dfrac{\partial(F,g)(\mathrm{P}_0)}{\partial(y,z)}\right| = \begin{vmatrix} f'_y & -1 \\ g'_y & 0 \end{vmatrix} = g'_y \neq 0$, le ipotesi del Teorema del Dini sono soddisfatte e quindi tale sistema definisce una funzione implicita $\mathbb{R} \to \mathbb{R}^2, x \to (y(x), z(x))$. Per calcolare le derivate $y'(x)$ e $z'(x)$ si può procedere in due modi:

1) da $g(x,y) = 0$, derivando rispetto a x, si ha: $g'_x \cdot 1 + g'_y \dfrac{\mathrm{d}y}{\mathrm{d}x} = 0$, da cui $\dfrac{\mathrm{d}y}{\mathrm{d}x} = -\dfrac{g'_x}{g'_y}$.
Derivando rispetto a x la prima equazione $f(x,y) - z = 0$ si ha:
$f'_x \cdot 1 + f'_y \dfrac{\mathrm{d}y}{\mathrm{d}x} + (-1)\dfrac{\mathrm{d}z}{\mathrm{d}x} = 0$, da cui: $f'_x + f'_y\left(-\dfrac{g'_x}{g'_y}\right) - \dfrac{\mathrm{d}z}{\mathrm{d}x} = 0$ e quindi:
$\dfrac{\mathrm{d}z}{\mathrm{d}x} = f'_x - f'_y \dfrac{g'_x}{g'_y} = \dfrac{f'_x g'_y - f'_y g'_x}{g'_y}$.

2) usando le matrici Jacobiane, abbiamo: $\left\|\begin{matrix} f'_y & -1 \\ g'_y & 0 \end{matrix}\right\| \cdot \left\|\begin{matrix} \dfrac{\mathrm{d}y}{\mathrm{d}x} \\ \dfrac{\mathrm{d}z}{\mathrm{d}x} \end{matrix}\right\| = -\left\|\begin{matrix} f'_x \\ g'_x \end{matrix}\right\|$, da cui si ha:

$$\frac{\mathrm{d}y}{\mathrm{d}x} = -\frac{\begin{vmatrix} f'_x & -1 \\ g'_x & 0 \end{vmatrix}}{\begin{vmatrix} f'_y & -1 \\ g'_y & 0 \end{vmatrix}} = -\frac{g'_x}{g'_y}, \quad \frac{\mathrm{d}z}{\mathrm{d}x} = -\frac{\begin{vmatrix} f'_y & f'_x \\ g'_y & g'_x \end{vmatrix}}{\begin{vmatrix} f'_y & -1 \\ g'_y & 0 \end{vmatrix}} = \frac{f'_x g'_y - f'_y g'_x}{g'_y}.$$

Esempio 55 : Dato il sistema: $\begin{cases} z = f(x,y) \\ g(x,y,z) = 0 \end{cases}$, ovvero $\begin{cases} F(x,y,z) = f(x,y) - z \\ g(x,y,z) = 0 \end{cases}$, siano soddisfatte le opportune ipotesi come nell'esempio precedente. Anche ora una variabile, z, è già in forma esplicita, mentre la seconda equazione ci permette di ricavare, anche solo implicitamente, un'altra variabile, diciamo y. Essendo: $\dfrac{\partial(F,g)}{\partial(x,y,z)} = \left\|\begin{matrix} f'_x & f'_y & -1 \\ g'_x & g'_y & g'_z \end{matrix}\right\|$, se nel punto P_0 risulta $\begin{vmatrix} f'_y & -1 \\ g'_y & g'_z \end{vmatrix} \neq 0$, abbiamo una funzione $\mathbb{R} \to \mathbb{R}^2, x \to (y(x), z(x))$, le cui derivate sono date dalla:

$\begin{Vmatrix} f'_y & -1 \\ g'_y & g'_z \end{Vmatrix} \cdot \begin{Vmatrix} \dfrac{\mathrm{d}y}{\mathrm{d}x} \\ \dfrac{\mathrm{d}z}{\mathrm{d}x} \end{Vmatrix} = -\begin{Vmatrix} f'_x \\ g'_x \end{Vmatrix}$, dalla quale otteniamo:

$$\frac{\mathrm{d}y}{\mathrm{d}x} = -\frac{\begin{vmatrix} f'_x & -1 \\ g'_x & g'_z \end{vmatrix}}{\begin{vmatrix} f'_y & -1 \\ g'_y & g'_z \end{vmatrix}} = -\frac{f'_x\, g'_z + g'_x}{f'_y\, g'_z + g'_y}\,,\quad \frac{\mathrm{d}z}{\mathrm{d}x} = -\frac{\begin{vmatrix} f'_y & f'_x \\ g'_y & g'_x \end{vmatrix}}{\begin{vmatrix} f'_y & -1 \\ g'_y & g'_z \end{vmatrix}} = -\frac{f'_y\, g'_x - f'_x\, g'_y}{f'_y\, g'_z + g'_y}\,.$$

I risultati, come si vede, sono diversi da quelli dell'esempio precedente, e dipendono dal fatto che la seconda equazione non è costante rispetto a z ma dipende invece da tutte le variabili.

Derivate seconde e Differenziali I° e II di funzioni implicite $\mathbb{R} \longrightarrow \mathbb{R}^2$ in Appendice n.2

II Caso: Sistema $\begin{cases} f(x_1,x_2,y_1,y_2) = k_1 \\ g(x_1,x_2,y_1,y_2) = k_2 \end{cases}$: funzione implicita $\mathbb{R}^2 \to \mathbb{R}^2$

Passiamo ora a considerare un sistema di due equazioni ma in quattro variabili: $\begin{cases} f(x_1,x_2,y_1,y_2) = k_1 \\ g(x_1,x_2,y_1,y_2) = k_2 \end{cases}$, $k_1, k_2 \in \mathbb{R}$. Le due equazioni del sistema ci permettono, sotto le opportune ipotesi, di definire una funzione $\mathbb{R}^2 \to \mathbb{R}^2, (x_1,x_2) \to (y_1(x_1,x_2), y_2(x_1,x_2))$.

Funzione implicita $\mathbb{R}^2 \to \mathbb{R}^2$: Derivate prime

Posto $\begin{cases} w_1 = f(x_1,x_2,y_1,y_2) = k_1 \\ w_2 = g(x_1,x_2,y_1,y_2) = k_2 \end{cases}$, e posto $(\mathbb{X}|\,\mathbb{Y}) = (x_1,x_2,y_1,y_2)$, avremo queste due composizioni di funzioni:

$\mathbb{R}^2 \to \mathbb{R}^4 \xrightarrow{f} \mathbb{R},\ \mathbb{X} \to (\mathbb{X}|\,\mathbb{Y}) \xrightarrow{f} w_1 = f(\mathbb{X}|\,\mathbb{Y}) = k_1$, e

$\mathbb{R}^2 \to \mathbb{R}^4 \xrightarrow{g} \mathbb{R},\ \mathbb{X} \to (\mathbb{X}|\,\mathbb{Y}) \xrightarrow{g} w_2 = g(\mathbb{X}|\,\mathbb{Y}) = k_2$ dalle quali, derivando rispetto a x_1 e x_2, essendo w_1 e w_2 costanti, otteniamo il sistema:

$$\begin{cases} \dfrac{\partial w_1}{\partial x_1} = \nabla f(\mathbb{X}|\,\mathbb{Y}) \cdot \dfrac{\partial(\mathbb{X}|\,\mathbb{Y})}{\partial x_1} = 0 \\ \dfrac{\partial w_2}{\partial x_1} = \nabla g(\mathbb{X}|\,\mathbb{Y}) \cdot \dfrac{\partial(\mathbb{X}|\,\mathbb{Y})}{\partial x_1} = 0 \\ \dfrac{\partial w_1}{\partial x_2} = \nabla f(\mathbb{X}|\,\mathbb{Y}) \cdot \dfrac{\partial(\mathbb{X}|\,\mathbb{Y})}{\partial x_2} = 0 \\ \dfrac{\partial w_2}{\partial x_2} = \nabla g(\mathbb{X}|\,\mathbb{Y}) \cdot \dfrac{\partial(\mathbb{X}|\,\mathbb{Y})}{\partial x_2} = 0 \end{cases}.$$

Essendo $\dfrac{\partial(\mathbb{X}|\,\mathbb{Y})}{\partial x_1} = \left(1, 0, \dfrac{\partial y_1}{\partial x_1}, \dfrac{\partial y_2}{\partial x_1}\right)$ e $\dfrac{\partial(\mathbb{X}|\,\mathbb{Y})}{\partial x_2} = \left(0, 1, \dfrac{\partial y_1}{\partial x_2}, \dfrac{\partial y_2}{\partial x_2}\right)$ otteniamo il sistema nelle incognite $\dfrac{\partial y_1}{\partial x_1}$, $\dfrac{\partial y_2}{\partial x_1}$, $\dfrac{\partial y_1}{\partial x_2}$ e $\dfrac{\partial y_2}{\partial x_2}$:

$$\begin{cases} \dfrac{\partial f}{\partial y_1} \cdot \dfrac{\partial y_1}{\partial x_1} + \dfrac{\partial f}{\partial y_2} \cdot \dfrac{\partial y_2}{\partial x_1} = -\dfrac{\partial f}{\partial x_1} \\ \dfrac{\partial g}{\partial y_1} \cdot \dfrac{\partial y_1}{\partial x_1} + \dfrac{\partial g}{\partial y_2} \cdot \dfrac{\partial y_2}{\partial x_1} = -\dfrac{\partial g}{\partial x_1} \\ \dfrac{\partial f}{\partial y_1} \cdot \dfrac{\partial y_1}{\partial x_2} + \dfrac{\partial f}{\partial y_2} \cdot \dfrac{\partial y_2}{\partial x_2} = -\dfrac{\partial f}{\partial x_2} \\ \dfrac{\partial g}{\partial y_1} \cdot \dfrac{\partial y_1}{\partial x_2} + \dfrac{\partial g}{\partial y_2} \cdot \dfrac{\partial y_2}{\partial x_2} = -\dfrac{\partial g}{\partial x_2} \end{cases}$$

, che in forma matriciale diviene:

$$\left\|\begin{array}{cccc} \frac{\partial f}{\partial y_1} & \frac{\partial f}{\partial y_2} & 0 & 0 \\ \frac{\partial g}{\partial y_1} & \frac{\partial g}{\partial y_2} & 0 & 0 \\ 0 & 0 & \frac{\partial f}{\partial y_1} & \frac{\partial f}{\partial y_2} \\ 0 & 0 & \frac{\partial g}{\partial y_1} & \frac{\partial g}{\partial y_2} \end{array}\right\| \cdot \left\|\begin{array}{c} \frac{\partial y_1}{\partial x_1} \\ \frac{\partial y_2}{\partial x_1} \\ \frac{\partial y_1}{\partial x_2} \\ \frac{\partial y_2}{\partial x_2} \end{array}\right\| = -\left\|\begin{array}{c} \frac{\partial f}{\partial x_1} \\ \frac{\partial g}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \frac{\partial g}{\partial x_2} \end{array}\right\|$$, che può essere scritto come:

$$\left\|\begin{array}{cc} \frac{\partial f}{\partial y_1} & \frac{\partial f}{\partial y_2} \\ \frac{\partial g}{\partial y_1} & \frac{\partial g}{\partial y_2} \end{array}\right\| \cdot \left\|\begin{array}{cc} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_1}{\partial x_2} \\ \frac{\partial y_2}{\partial x_1} & \frac{\partial y_2}{\partial x_2} \end{array}\right\| = -\left\|\begin{array}{cc} \frac{\partial f}{\partial x_1} & \frac{\partial f}{\partial x_2} \\ \frac{\partial g}{\partial x_1} & \frac{\partial g}{\partial x_2} \end{array}\right\|$$ e, mediante le matrici Jacobiane,

come: $\frac{\partial(f,g)}{\partial(y_1,y_2)} \cdot \frac{\partial(y_1,y_2)}{\partial(x_1,x_2)} = -\frac{\partial(f,g)}{\partial(x_1,x_2)}$.

Se il punto $(x_1^0, x_2^0, y_1^0, y_2^0)$ soddisfa il sistema, se f e g sono derivabili con derivate continue, e se $\left|\frac{\partial(f,g)(x_1^0, x_2^0, y_1^0, y_2^0)}{\partial(y_1,y_2)}\right| \neq 0$ ovvero se $\frac{\partial(f,g)}{\partial(y_1,y_2)}$ ha caratteristica massima, il Teorema del Dini, in questo caso, assicura l'esistenza, in un intorno del punto (x_1^0, x_2^0) della funzione implicita $\mathbb{R}^2 \to \mathbb{R}^2, (x_1,x_2) \to (y_1(x_1,x_2), y_2(x_1,x_2))$, le cui derivate sono ottenibili come: $\frac{\partial(y_1,y_2)}{\partial(x_1,x_2)} = -\left\|\frac{\partial(f,g)}{\partial(y_1,y_2)}\right\|^{-1} \cdot \frac{\partial(f,g)}{\partial(x_1,x_2)}$, che è la formula generale mediante la quale otteniamo le derivate di y_1 e y_2 rispetto a x_1 e x_2. Rimane valida anche l'altra metodologia di calcolo, cioè quella derivante dal Teorema di Cramer sui sistemi lineari, dalla quale otteniamo:

$$\frac{\partial y_1}{\partial x_1} = -\frac{\left|\begin{array}{cc} \frac{\partial f}{\partial x_1} & \frac{\partial f}{\partial y_2} \\ \frac{\partial g}{\partial x_1} & \frac{\partial g}{\partial y_2} \end{array}\right|}{\left|\begin{array}{cc} \frac{\partial f}{\partial y_1} & \frac{\partial f}{\partial y_2} \\ \frac{\partial g}{\partial y_1} & \frac{\partial g}{\partial y_2} \end{array}\right|} = -\frac{\left|\frac{\partial(f,g)}{\partial(x_1,y_2)}\right|}{\left|\frac{\partial(f,g)}{\partial(y_1,y_2)}\right|} \; ; \; \frac{\partial y_1}{\partial x_2} = -\frac{\left|\begin{array}{cc} \frac{\partial f}{\partial x_2} & \frac{\partial f}{\partial y_2} \\ \frac{\partial g}{\partial x_2} & \frac{\partial g}{\partial y_2} \end{array}\right|}{\left|\begin{array}{cc} \frac{\partial f}{\partial y_1} & \frac{\partial f}{\partial y_2} \\ \frac{\partial g}{\partial y_1} & \frac{\partial g}{\partial y_2} \end{array}\right|} = -\frac{\left|\frac{\partial(f,g)}{\partial(x_2,y_2)}\right|}{\left|\frac{\partial(f,g)}{\partial(y_1,y_2)}\right|}$$

$$\frac{\partial y_2}{\partial x_1} = -\frac{\left|\begin{array}{cc} \frac{\partial f}{\partial y_1} & \frac{\partial f}{\partial x_1} \\ \frac{\partial g}{\partial y_1} & \frac{\partial g}{\partial x_1} \end{array}\right|}{\left|\begin{array}{cc} \frac{\partial f}{\partial y_1} & \frac{\partial f}{\partial y_2} \\ \frac{\partial g}{\partial y_1} & \frac{\partial g}{\partial y_2} \end{array}\right|} = -\frac{\left|\frac{\partial(f,g)}{\partial(y_1,x_1)}\right|}{\left|\frac{\partial(f,g)}{\partial(y_1,y_2)}\right|} \; ; \; \frac{\partial y_2}{\partial x_2} = -\frac{\left|\begin{array}{cc} \frac{\partial f}{\partial y_1} & \frac{\partial f}{\partial x_2} \\ \frac{\partial g}{\partial y_1} & \frac{\partial g}{\partial x_2} \end{array}\right|}{\left|\begin{array}{cc} \frac{\partial f}{\partial y_1} & \frac{\partial f}{\partial y_2} \\ \frac{\partial g}{\partial y_1} & \frac{\partial g}{\partial y_2} \end{array}\right|} = -\frac{\left|\frac{\partial(f,g)}{\partial(y_1,x_2)}\right|}{\left|\frac{\partial(f,g)}{\partial(y_1,y_2)}\right|}$$

Esempio 56 : Il punto $P_0 = (1, 1, -1, -1)$ soddisfa il sistema:

$$\begin{cases} f(x_1,x_2,y_1,y_2) = (x_1 + x_2)\, e^{y_2 - y_1} + (y_2 + y_1)\, e^{x_1 - x_2} = 0 \\ g(x_1,x_2,y_1,y_2) = (2x_2 - y_1)\, e^{x_1 + y_2} - (x_1 - y_2)\, e^{x_2 + y_1} = 1 \end{cases}.$$

Le funzioni f e g sono ovunque derivabili con derivate continue. Abbiamo poi:

$\frac{\partial f}{\partial x_1} = e^{y_2 - y_1} + (y_2 + y_1)\, e^{x_1 - x_2}$; $\quad \frac{\partial g}{\partial x_1} = (2x_2 - y_1)\, e^{x_1 + y_2} - e^{x_2 + y_1}$;

$$\frac{\partial f}{\partial x_2}=e^{y_2-y_1}-(y_2+y_1)\,e^{x_1-x_2}\,;\qquad \frac{\partial g}{\partial x_2}=2\,e^{x_1+y_2}-(x_1-y_2)\,e^{x_2+y_1}\,;$$
$$\frac{\partial f}{\partial y_1}=-(x_1+x_2)\,e^{y_2-y_1}+e^{x_1-x_2}\,;\qquad \frac{\partial g}{\partial y_1}=-e^{x_1+y_2}-(x_1-y_2)\,e^{x_2+y_1}\,;$$
$$\frac{\partial f}{\partial y_2}=(x_1+x_2)\,e^{y_2-y_1}+e^{x_1-x_2}\,;\qquad \frac{\partial g}{\partial y_2}=(2x_2-y_1)\,e^{x_1+y_2}+e^{x_2+y_1}\,.$$

Si ha $\dfrac{\partial(f,g)(1,1,-1,-1)}{\partial(x_1,x_2,y_1,y_2)}=\left\|\begin{matrix}-1 & 3 & -1 & 3\\ 2 & 0 & -3 & 4\end{matrix}\right\|$, ed essendo $\left|\begin{matrix}-1 & 3\\ -3 & 4\end{matrix}\right|=5\neq 0$, con il sistema dato si può definire una funzione implicita:
$\mathbb{R}^2\to\mathbb{R}^2,(x_1,x_2)\to(y_1(x_1,x_2),y_2(x_1,x_2))$, in un intorno di $(x_1,x_2)=(1,1)$.
Per le derivate di questa funzione, calcolate in $(1,1)$, dalla:
$\dfrac{\partial(y_1,y_2)}{\partial(x_1,x_2)}=-\left\|\dfrac{\partial(f,g)}{\partial(y_1,y_2)}\right\|^{-1}\cdot\dfrac{\partial(f,g)}{\partial(x_1,x_2)}$ otteniamo:

$$\left\|\begin{matrix}\dfrac{\partial y_1(1,1)}{\partial x_1} & \dfrac{\partial y_1(1,1)}{\partial x_2}\\ \dfrac{\partial y_2(1,1)}{\partial x_1} & \dfrac{\partial y_2(1,1)}{\partial x_2}\end{matrix}\right\|=-\left\|\begin{matrix}-1 & 3\\ -3 & 4\end{matrix}\right\|^{-1}\cdot\left\|\begin{matrix}-1 & 3\\ 2 & 0\end{matrix}\right\|=$$
$$=-\left\|\begin{matrix}\dfrac{4}{5} & -\dfrac{3}{5}\\ \dfrac{3}{5} & -\dfrac{1}{5}\end{matrix}\right\|\cdot\left\|\begin{matrix}-1 & 3\\ 2 & 0\end{matrix}\right\|,\text{ per cui: }\left\|\begin{matrix}\dfrac{\partial y_1(1,1)}{\partial x_1} & \dfrac{\partial y_1(1,1)}{\partial x_2}\\ \dfrac{\partial y_2(1,1)}{\partial x_1} & \dfrac{\partial y_2(1,1)}{\partial x_2}\end{matrix}\right\|=\left\|\begin{matrix}2 & -\dfrac{12}{5}\\ 1 & -\dfrac{9}{5}\end{matrix}\right\|.$$

Derivate seconde e Differenziali I° e II di funzioni implicite $\mathbb{R}^2\longrightarrow\mathbb{R}^2$ in Appendice n.3

TEOREMA DEL DINI NEL CASO GENERALE

Sistema di m equazioni in $n+m$ incognite: funzione implicita $\mathbb{R}^n\to\mathbb{R}^m$

Concludiamo trattando il caso generale, cioè quello di un sistema di m equazioni in $m+n$ variabili. Enunciando il Teorema del Dini in questo caso generale ritroveremo, come casi particolari, tutti i casi trattati in precedenza. Sia dato allora il sistema:

$$\begin{cases} f_1(x_1,x_2,...,x_n,y_1,y_2,...,y_m)=k_1\\ f_2(x_1,x_2,...,x_n,y_1,y_2,...,y_m)=k_2\\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\ f_m(x_1,x_2,...,x_n,y_1,y_2,...,y_m)=k_m\end{cases}.$$

Un tale sistema consente, soddisfatte le opportune ipotesi, di definire una funzione implicita: $\mathbb{R}^n\to\mathbb{R}^m,(x_1,x_2,...,x_n)\to(y_1,y_2,...,y_m)$, con $y_i=y_i(x_1,x_2,...,x_n)$. Vediamo che il numero delle equazioni corrisponde al numero delle variabili dipendenti, e quindi come, per differenza, si ottenga quello delle variabili indipendenti: m equazioni comportano m variabili dipendenti e di conseguenza $m+n-m=n$ saranno le variabili indipendenti. Vale allora il **Teorema** 21 : Sia $\mathrm{P}_0=\left(x_1^0,x_2^0,...,x_n^0,y_1^0,y_2^0,...,y_m^0\right)$ un punto che soddisfa tale sistema, siano le funzioni f_i derivabili con derivate continue in un intorno di P_0, e sia inoltre $\left|\dfrac{\partial(f_1,f_2,..,f_m)(\mathrm{P}_0)}{\partial(y_1,y_2,...,y_m)}\right|\neq 0$. Allora il sistema definisce, in un intorno di $(x_1^0,x_2^0,...,x_n^0)$, una funzione implicita , continua con derivate continue:
$\mathbb{R}^n\to\mathbb{R}^m,(x_1,x_2,...,x_n)\to(y_1,y_2,...,y_m)$, con $y_i=y_i(x_1,x_2,...,x_n)$, la cui matrice Jacobiana è data da $\dfrac{\partial(y_1,y_2,...,y_m)}{\partial(x_1,x_2,...,x_n)}=-\left\|\dfrac{\partial(f_1,f_2,..,f_m)}{\partial(y_1,y_2,...,y_m)}\right\|^{-1}\cdot\dfrac{\partial(f_1,f_2,..,f_m)}{\partial(x_1,x_2,...,x_n)}$.

Ogni derivata $\dfrac{\partial y_i}{\partial x_k}$ può essere espressa, per il Teorema di Cramer, anche come un quoziente:

$$\frac{\partial y_i}{\partial x_k} = -\frac{\begin{vmatrix} \frac{\partial f_1}{\partial y_1} & \cdots & \frac{\partial f_1}{\partial y_{i-1}} & \frac{\partial f_1}{\partial x_k} & \frac{\partial f_1}{\partial y_{i+1}} & \cdots & \frac{\partial f_1}{\partial y_m} \\ \frac{\partial f_2}{\partial y_1} & \cdots & \frac{\partial f_2}{\partial y_{i-1}} & \frac{\partial f_2}{\partial x_k} & \frac{\partial f_2}{\partial y_{i+1}} & \cdots & \frac{\partial f_2}{\partial y_m} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \frac{\partial f_m}{\partial y_1} & \cdots & \frac{\partial f_m}{\partial y_{i-1}} & \frac{\partial f_m}{\partial x_k} & \frac{\partial f_m}{\partial y_{i+1}} & \cdots & \frac{\partial f_m}{\partial y_m} \end{vmatrix}}{\begin{vmatrix} \frac{\partial f_1}{\partial y_1} & \cdots & \frac{\partial f_1}{\partial y_{i-1}} & \frac{\partial f_1}{\partial y_i} & \frac{\partial f_1}{\partial y_{i+1}} & \cdots & \frac{\partial f_1}{\partial y_m} \\ \frac{\partial f_2}{\partial y_1} & \cdots & \frac{\partial f_2}{\partial y_{i-1}} & \frac{\partial f_2}{\partial y_i} & \frac{\partial f_2}{\partial y_{i+1}} & \cdots & \frac{\partial f_2}{\partial y_m} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \frac{\partial f_m}{\partial y_1} & \cdots & \frac{\partial f_m}{\partial y_{i-1}} & \frac{\partial f_m}{\partial y_i} & \frac{\partial f_m}{\partial y_{i+1}} & \cdots & \frac{\partial f_m}{\partial y_m} \end{vmatrix}} =$$

$$\frac{\partial y_i}{\partial x_k} = -\frac{\left|\frac{\partial(f_1, f_2, .., f_m)}{\partial(y_1, ..., y_{i-1}, x_k, y_{i+1}, ..., y_m)}\right|}{\left|\frac{\partial(f_1, f_2, .., f_m)}{\partial(y_1, ..., y_{i-1}, y_i, y_{i+1}, ..., y_m)}\right|}.$$

Ovvero, la derivata di y_i fatta rispetto a x_k si ottiene dall'opposto di un quoziente fra due determinanti Jacobiani: a denominatore c'è quello delle equazioni rispetto alle variabili dipendenti, a numeratore quello ottenuto sostituendo nel precedente la colonna i-esima, quella della variabile che si deriva, con le derivate delle equazioni fatte rispetto a x_k, la variabile rispetto a cui si vuole derivare y_i.

Funzione implicita $\mathbb{R}^n \to \mathbb{R}^m$: Derivate prime

Vediamo come giustificare quest'ultimo risultato relativo alle derivate della funzione implicita. Dalla composizione di funzioni:

$(x_1, x_2, ..., x_n) \to (x_1, x_2, ..., x_n, y_1, y_2, ..., y_m) \xrightarrow{f_i} w_i = k_i = \text{cost.}$, $1 \le i \le m$ otteniamo:

$$\frac{\partial(w_1, w_2, ..., w_m)}{\partial(x_1, x_2, ..., x_n)} = \frac{\partial(w_1, w_2, ..., w_m)}{\partial(x_1, x_2, ..., x_n, y_1, y_2, ..., y_m)} \cdot \frac{\partial(x_1, x_2, ..., x_n, y_1, y_2, ..., y_m)}{\partial(x_1, x_2, ..., x_n)} = 0.$$

Il secondo termine può essere scritto come:

$$\left\| \frac{\partial(f_1, f_2, .., f_m)}{\partial(x_1, x_2, ..., x_n)} \quad \middle| \quad \frac{\partial(f_1, f_2, .., f_m)}{\partial(y_1, y_2, ..., y_m)} \right\| \cdot \left\| \begin{matrix} \frac{\partial(x_1, x_2, ..., x_n)}{\partial(x_1, x_2, ..., x_n)} \\ - - - - - \\ \frac{\partial(y_1, y_2, ..., y_m)}{\partial(x_1, x_2, ..., x_n)} \end{matrix} \right\| = 0.$$

La matrice di sinistra è una matrice $(m, n+m)$ divisa in due blocchi, il primo (m, n) ed il secondo (m, m); la matrice di destra è una matrice $(n+m, n)$ divisa in due blocchi, il superiore (n, n) e l'inferiore (m, n). Operando per blocchi, dall'uguaglianza otteniamo poi:

$\frac{\partial(f_1, f_2, .., f_m)}{\partial(x_1, x_2, ..., x_n)} \cdot \mathbb{I}_n + \frac{\partial(f_1, f_2, .., f_m)}{\partial(y_1, y_2, ..., y_m)} \cdot \frac{\partial(y_1, y_2, ..., y_m)}{\partial(x_1, x_2, ..., x_n)} = 0$ da cui poi:

$\frac{\partial(f_1, f_2, .., f_m)}{\partial(y_1, y_2, ..., y_m)} \cdot \frac{\partial(y_1, y_2, ..., y_m)}{\partial(x_1, x_2, ..., x_n)} = -\frac{\partial(f_1, f_2, .., f_m)}{\partial(x_1, x_2, ..., x_n)}$ e quindi:

$$\frac{\partial(y_1, y_2, ..., y_m)}{\partial(x_1, x_2, ..., x_n)} = -\left\|\frac{\partial(f_1, f_2, .., f_m)}{\partial(y_1, y_2, ..., y_m)}\right\|^{-1} \cdot \frac{\partial(f_1, f_2, .., f_m)}{\partial(x_1, x_2, ..., x_n)}.$$

In forma compatta le precedenti uguaglianze possono essere espresse come:

$$\frac{\partial(\mathbb{W})}{\partial(\mathbb{X})} = \frac{\partial(\mathbb{W})}{\partial(\mathbb{X}\,|\,\mathbb{Y})}\cdot\frac{\partial(\mathbb{X}\,|\,\mathbb{Y})}{\partial(\mathbb{X})} = \frac{\partial(\mathbb{W})}{\partial(\mathbb{X})}\cdot\mathbb{I}_n + \frac{\partial(\mathbb{W})}{\partial(\mathbb{Y})}\cdot\frac{\partial(\mathbb{Y})}{\partial(\mathbb{X})} = 0\,,$$ da cui:

$$\frac{\partial(\mathbb{Y})}{\partial(\mathbb{X})} = -\left(\frac{\partial(\mathbb{W})}{\partial(\mathbb{Y})}\right)^{-1}\cdot\frac{\partial(\mathbb{W})}{\partial(\mathbb{X})}\,.$$

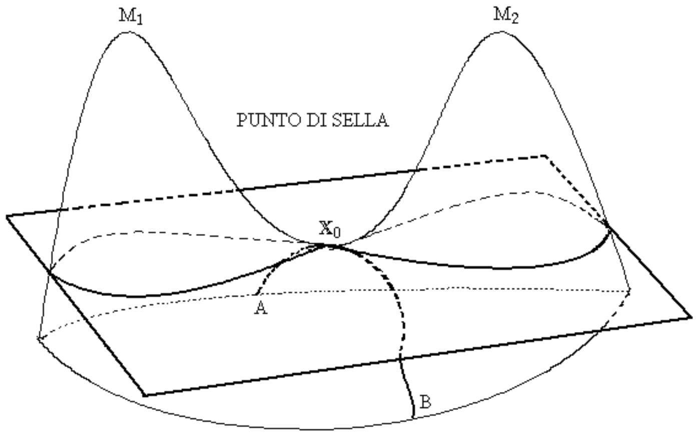
Derivate seconde e Differenziali I° e II di funzioni implicite $\mathbb{R}^n \to \mathbb{R}^m$ in Appendice n.4

MASSIMI E MINIMI PER FUNZIONI $f : \mathbb{R}^n \to \mathbb{R}$

Trattiamo ora il problema della ricerca dei punti di massimo e di minimo (detti anche genericamente estremi) per le funzioni di più variabili.
La definizione di punto di massimo o di minimo relativo per funzioni $f : \mathbb{R}^n \to \mathbb{R}$ è analoga a quella data per le funzioni di una sola variabile, ed è stata enunciata nella Definizione 19.
Per quanto riguarda l'esistenza dei punti di massimo e di minimo assoluti vale il Teorema 6 di Weierstrass. Per le funzioni di una sola variabile vale poi il Teorema, cosiddetto di Fermat, secondo il quale, se una funzione è derivabile in x_0 ed ha in x_0 un punto di massimo o di minimo relativo, allora si ha che $f'(x_0) = 0$.
Per le funzioni $f : \mathbb{R}^n \to \mathbb{R}$, analogamente, vale anzitutto il seguente:
Teorema 22 : Sia $f(\mathbb{X})$ differenziabile in $\mathbb{X}_0$, punto interno al dominio, e sia $\mathbb{X}_0$ punto di massimo o di minimo relativo per f ; allora $\nabla f(\mathbb{X}_0) = \mathbb{O}$, dove $\mathbb{O}$ indica il vettore nullo.
Dimostrazione : Posto $f(\mathbb{X}) = f(\mathbb{X}_0 + t\,v)$, con v versore, abbiamo una funzione composta: $\mathbb{R} \to \mathbb{R}^n \xrightarrow{f} \mathbb{R}$, $t \to (\mathbb{X}_0 + t\,v) \xrightarrow{f} f(\mathbb{X}_0 + t\,v)$, che vediamo come $g(t) = f(\mathbb{X}_0 + t\,v)$, con $g(0) = f(\mathbb{X}_0)$. Se $\mathbb{X}_0$ è punto di massimo (o di minimo) relativo, sarà pure:
$f(\mathbb{X}_0 + t\,v) \le f(\mathbb{X}_0)$ $(f(\mathbb{X}_0 + t\,v) \ge f(\mathbb{X}_0))$, $\forall\, \mathbb{X} \in \mathfrak{I}(\mathbb{X}_0)$, ovvero:
$g(t) \le g(0)$ $(g(t) \ge g(0))$, $\forall\, t \in \mathfrak{I}(0)$. Quindi se $\mathbb{X}_0$ è punto di massimo (o di minimo) relativo per f, altrettanto sarà $t = 0$ per g. Essendo f differenziabile, ed essendo $\mathbb{X}_0 + t\,v$ derivabile rispetto a t, ne segue che $g(t)$ è derivabile in quanto composizione di funzioni differenziabili, e quindi per il Teorema di Fermat dovrà essere $g'(0) = 0$. Ma:

$$g'(0) = \lim_{t\to 0}\frac{g(t) - g(0)}{t} = \lim_{t\to 0}\frac{f(\mathbb{X}_0 + t\,v) - f(\mathbb{X}_0)}{t} = \mathcal{D}_v f(\mathbb{X}_0) = 0\,.$$

Ma $\mathcal{D}_v f(\mathbb{X}_0) = \nabla f(\mathbb{X}_0)\cdot v$, in quanto f è differenziabile, e allora:
$\forall\, v :\ \nabla f(\mathbb{X}_0)\cdot v = 0 \Leftrightarrow \nabla f(\mathbb{X}_0) = \mathbb{O}$.●
La condizione $\nabla f(\mathbb{X}_0) = \mathbb{O}$ è una condizione necessaria, ma non sufficiente, per avere un punto di massimo o di minimo relativo, ed è condizione necessaria solo nell'ipotesi che la funzione sia differenziabile. Viene detta **condizione del I ordine** per la ricerca dei punti di massimo e di minimo.
I punti nei quali $\nabla f(\mathbb{X}_0) = \mathbb{O}$ vengono detti **punti stazionari**, e sono i punti nei quali il piano (o l'iperpiano) tangente al grafico della funzione risulta orizzontale (o parallelo all'iperpiano delle variabili indipendenti).
Per la ricerca dei punti di massimo e minimo relativo occorre quindi anzitutto soddisfare le condizioni del I ordine, ovvero imporre $\nabla f(\mathbb{X}) = \mathbb{O}$, e risolvere il sistema di n equazioni in n incognite che da questa scaturisce. Trovate tutte le soluzioni, ovvero tutti i punti $\mathbb{X}_0$ per i quali $\nabla f(\mathbb{X}_0) = \mathbb{O}$, occorre verificare se detti punti risultano effettivamente di massimo o di minimo relativo, oppure se sono punti di sella (o di colle).



I **punti di sella** sono punti nei quali $\nabla f(\mathbb{X}_0) = \mathbb{O}$, ma non è però soddisfatta nè la definizione di punto di massimo nè quella di punto di minimo, in quanto in ogni intorno di $\mathbb{X}_0$ ci sono punti per i quali $f(\mathbb{X}) > f(\mathbb{X}_0)$ e punti per i quali $f(\mathbb{X}) < f(\mathbb{X}_0)$. Geometricamente parlando, in un punto stazionario che risulti di massimo (o di minimo) il piano (o iperpiano) tangente sta, in un intorno di $\mathbb{X}_0$, tutto al di sopra (al di sotto) del grafico della funzione. Se il punto è invece di sella il piano tangente attraversa il grafico della funzione, e quindi, in ogni intorno di $\mathbb{X}_0$, ci sono punti del grafico al di sopra e punti al di sotto del piano tangente.
Occorrono quindi criteri e metodologie, che vengono detti **condizioni del II ordine**, aventi per scopo quello di stabilire la vera natura di un punto stazionario. Vediamo intanto con un esempio come l'analisi lungo particolari direzioni, al pari di quanto detto per l'operazione di limite, non risulti in generale valida per trarre conclusioni affermative sulla natura di un punto di massimo o di minimo.

Esempio 57 : Data la funzione $f(x,y) = \left(y - x^2\right)\left(y - x^4\right)$, determiniamone anzitutto i punti stazionari; dovrà risultare:

$$\begin{cases} f_x' = -2x\left(y - x^4\right) - 4x^3\left(y - x^2\right) = 6x^5 - 2xy - 4x^3y = 0 \\ f_y' = y - x^4 + y - x^2 = 2y - x^4 - x^2 = 0 \quad \text{ovvero } 2y = x^4 + x^2 \end{cases};$$

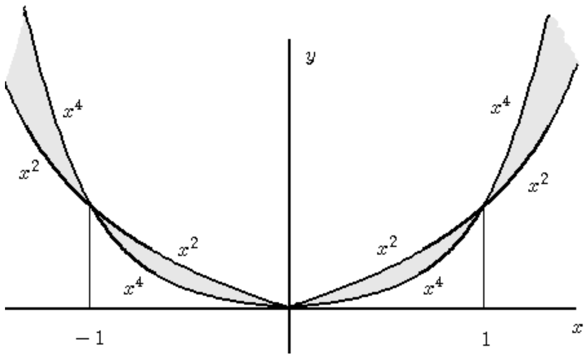
sostituendo nella prima equazione si ha:
$2x^7 - 3x^5 + x^3 = x^3\left(2x^4 - 3x^2 + 1\right) = 0$, da cui otteniamo $x = 0$ e:

$x^2 = \dfrac{3 \pm \sqrt{9-8}}{4}$ e quindi: $\begin{cases} x^2 = 1 & \text{ovvero } x = \pm 1 \\ x^2 = \frac{1}{2} & \text{ovvero } x = \pm\frac{1}{\sqrt{2}} \end{cases}$. Abbiamo cinque punti stazionari:

$$\begin{cases} x = 0 \\ y = 0 \end{cases}; \begin{cases} x = 1 \\ y = 1 \end{cases}; \begin{cases} x = -1 \\ y = 1 \end{cases}; \begin{cases} x = \frac{1}{\sqrt{2}} \\ y = \frac{3}{8} \end{cases}; \begin{cases} x = -\frac{1}{\sqrt{2}} \\ y = \frac{3}{8} \end{cases}.$$

Occupiamoci in questo esempio del solo punto $(0,0)$. Intanto $f(0,0) = 0$. Studiamo il comportamento della funzione $f(x,y)$ lungo una qualunque retta passante per l'origine. Posto $y = mx$, avremo: $f(x,mx) = \left(mx - x^2\right)\left(mx - x^4\right) = x^2\left(x^4 - mx^3 - mx + m^2\right)$.
Studiando il segno di $f(x,mx)$, dato che $x^2 > 0, \forall\, x \neq 0$, per il Teorema della permanenza del segno, dato che $x^4 - mx^3 - mx + m^2$ vale $m^2 > 0$ per $x = 0$, ed essendo la funzione continua, $f(x,mx)$ sarà sicuramente positiva in un intorno del punto $x = 0$.
Quindi il punto $(0,0)$ rappresenta un minimo lungo ogni retta passante per l'origine; sulla retta $x = 0$, che non appartiene alle $y = mx$, abbiamo $f(0,y) = y^2$, e quindi anche su questa retta il punto $(0,0)$ è un punto di minimo. Nonostante la concordanza di tutte le analisi sulle rette passanti per l'origine, il punto $(0,0)$ è un punto di sella. Infatti, se studiamo il segno di $f(x,y)$ avremo che:

$f(x,y) > 0$ per $\begin{cases} y > x^2 \\ y > x^4 \end{cases}$ oppure per $\begin{cases} y < x^2 \\ y < x^4 \end{cases}$.



Nella figura la parte scura rappresenta i punti in cui $f(x,y) < 0$. In ogni intorno di $(0,0)$ ci sono punti in cui $f(x,y) > 0$ e punti in cui $f(x,y) < 0$. Dato che $f(0,0) = 0$, ne segue che $(0,0)$ è un punto di sella, contrariamente a quanto si poteva dedurre analizzando la funzione lungo tutte le rette passanti per l'origine.
Riprenderemo e completeremo l'analisi dei punti stazionari di questa funzione quando disporremo degli strumenti opportuni.

CONDIZIONI DEL II ORDINE

Dato che per le funzioni $f : \mathbb{R}^n \to \mathbb{R}$ non si possono utilizzare, in quanto non definibili, la crescenza o la decrescenza della funzione, non ha alcuna utilità lo studiare il segno delle derivate parziali prime. Per distinguere tra i punti di massimo o minimo e gli eventuali punti di sella possiamo invece utilizzare le condizioni del II ordine, che sono condizioni sufficienti, e sono collegate alla concavità e alla convessità di f in $\mathbb{X}_0$. Vale infatti il seguente:
Teorema 23 : Se $\mathbb{X}_0$ è un punto stazionario per f e se la funzione è differenziabile e concava in un intorno di $\mathbb{X}_0$, allora $\mathbb{X}_0$ è un punto di massimo.
Dimostrazione : Per il Teorema 14, essendo f concava sarà:
$f(\mathbb{X}) \leq f(\mathbb{X}_0) + \nabla f(\mathbb{X}_0)(\mathbb{X} - \mathbb{X}_0)$, in quanto il grafico della funzione si trova al di sotto del (iper)piano tangente in $\mathbb{X}_0$. Ma $\nabla f(\mathbb{X}_0) = \mathbb{O}$, da cui segue $f(\mathbb{X}) \leq f(\mathbb{X}_0)$ in $\mathfrak{I}(\mathbb{X}_0)$, e quindi $\mathbb{X}_0$ è un punto di massimo.●
Analogamente, se f è differenziabile e convessa in un intorno del punto stazionario $\mathbb{X}_0$, allora $\mathbb{X}_0$ sarà un punto di minimo.
Supponendo ora che la funzione sia differenziabile due volte in $\mathbb{X}_0$, avremo, utilizzando il polinomio di Taylor, posto $\mathrm{d}\mathbb{X} = \mathbb{X} - \mathbb{X}_0$:

$$f(\mathbb{X}) = f(\mathbb{X}_0) + \nabla f(\mathbb{X}_0) \cdot \mathrm{d}\mathbb{X} + \frac{1}{2}\,\mathrm{d}\mathbb{X} \cdot \mathbb{H}(\mathbb{X}_0) \cdot \mathrm{d}\mathbb{X}^{\mathrm{T}} + o\big(\|\mathrm{d}\mathbb{X}\|^2\big),$$

ed essendo il punto $\mathbb{X}_0$ un punto stazionario ($\nabla f(\mathbb{X}_0) = \mathbb{O}$) otteniamo:

$$f(\mathbb{X}) - f(\mathbb{X}_0) = \frac{1}{2}\,\mathrm{d}\mathbb{X} \cdot \mathbb{H}(\mathbb{X}_0) \cdot \mathrm{d}\mathbb{X}^{\mathrm{T}} + o\big(\|\mathrm{d}\mathbb{X}\|^2\big) = \frac{1}{2}\,\mathrm{d}^2 f(\mathbb{X}_0) + o\big(\|\mathrm{d}\mathbb{X}\|^2\big).$$

Quindi il segno di $f(\mathbb{X}) - f(\mathbb{X}_0)$ coincide con quello di $\mathrm{d}\mathbb{X} \cdot \mathbb{H}(\mathbb{X}_0) \cdot \mathrm{d}\mathbb{X}^{\mathrm{T}} = \mathrm{d}^2 f(\mathbb{X}_0)$.
Se $\mathrm{d}^2 f(\mathbb{X}_0) < 0$ f è concava in $\mathbb{X}_0$, mentre se $\mathrm{d}^2 f(\mathbb{X}_0) > 0$ f è convessa in $\mathbb{X}_0$.
Per verificare se in un intorno del punto stazionario $\mathbb{X}_0$ risulta soddisfatta la definizione di punto di massimo o quella di punto di minimo, possiamo studiare il segno della differenza $f(\mathbb{X}) - f(\mathbb{X}_0)$, in quanto:
.) $f(\mathbb{X}) < f(\mathbb{X}_0) \Leftrightarrow f(\mathbb{X}) - f(\mathbb{X}_0) < 0 \Leftrightarrow \mathrm{d}^2 f(\mathbb{X}_0) < 0$ e quindi $\mathbb{X}_0$ è punto di massimo;
.) $f(\mathbb{X}) > f(\mathbb{X}_0) \Leftrightarrow f(\mathbb{X}) - f(\mathbb{X}_0) > 0 \Leftrightarrow \mathrm{d}^2 f(\mathbb{X}_0) > 0$ e quindi $\mathbb{X}_0$ è punto di minimo.
Se il segno della differenza $f(\mathbb{X}) - f(\mathbb{X}_0)$, ovvero del $\mathrm{d}^2 f(\mathbb{X}_0)$, non è costante in un intorno di $\mathbb{X}_0$, allora in $\mathbb{X}_0$ non potrà che esserci un punto di sella.

Il differenziale secondo $\mathrm{d}^2 f(\mathbb{X}_0) = \sum_{i,j=1}^{n} \frac{\partial^2 f(\mathbb{X}_0)}{\partial x_i\, \partial x_j}\, \mathrm{d}x_i\, \mathrm{d}x_j$ costituisce una **forma quadratica**, ovvero un polinomio nelle n variabili $\mathrm{d}x_i$ con termini tutti di secondo grado; occorrono quindi ora criteri atti a stabilire il segno di una forma quadratica.

FORME QUADRATICHE

Una forma quadratica è una espressione del tipo: $Q(\mathbb{X}) = Q(x_1, x_2, ..., x_n) = \sum_{i,j=1}^{n} a_{i\,j}\, x_i\, x_j$, ovvero un polinomio in n variabili costituito esclusivamente da termini di secondo grado.
Vediamo come ogni forma quadratica possa scriversi nella forma $Q(\mathbb{X}) = \mathbb{X} \cdot \mathbb{A} \cdot \mathbb{X}^{\mathrm{T}}$, dove $\mathbb{X} \in \mathbb{R}^n$ ed $\mathbb{A}$ è una matrice quadrata di ordine n.

Esempio 58 : $Q(x_1, x_2) = \|x_1\ x_2\| \cdot \begin{Vmatrix} 1 & 2 \\ 4 & 5 \end{Vmatrix} \cdot \begin{Vmatrix} x_1 \\ x_2 \end{Vmatrix} = 1\,x_1^2 + 2x_1\,x_2 + 4x_1\,x_2 + 5x_2^2 =$
$= x_1^2 + 6x_1\,x_2 + 5x_2^2$.

Esempio 59 : $Q(x_1, x_2, x_3) = \|x_1\ x_2\ x_3\| \cdot \begin{Vmatrix} 1 & 2 & 3 \\ 0 & 3 & 1 \\ 1 & 3 & 4 \end{Vmatrix} \cdot \begin{Vmatrix} x_1 \\ x_2 \\ x_3 \end{Vmatrix} =$
$= x_1^2 + 3x_2^2 + 4x_3^2 + 2x_1\,x_2 + 4x_1\,x_3 + 4x_2\,x_3$.

Vediamo però subito che ogni forma quadratica può essere generata da una matrice simmetrica: partendo dalla matrice $\mathbb{A}$, si costruisce la matrice simmetrica $\mathbb{B}$ ponendo $b_{i\,j} = \frac{a_{i\,j} + a_{j\,i}}{2}$.

Esempio 60 : Riprendendo gli Esempi precedenti, si verifica facilmente che:
$Q(x_1, x_2) = \|x_1\ x_2\| \cdot \begin{Vmatrix} 1 & 2 \\ 4 & 5 \end{Vmatrix} \cdot \begin{Vmatrix} x_1 \\ x_2 \end{Vmatrix} = \|x_1\ x_2\| \cdot \begin{Vmatrix} 1 & 3 \\ 3 & 5 \end{Vmatrix} \cdot \begin{Vmatrix} x_1 \\ x_2 \end{Vmatrix}$ e che:
$Q(x_1, x_2, x_3) = \|x_1\ x_2\ x_3\| \cdot \begin{Vmatrix} 1 & 2 & 3 \\ 0 & 3 & 1 \\ 1 & 3 & 4 \end{Vmatrix} \cdot \begin{Vmatrix} x_1 \\ x_2 \\ x_3 \end{Vmatrix} = \|x_1\ x_2\ x_3\| \cdot \begin{Vmatrix} 1 & 1 & 2 \\ 1 & 3 & 2 \\ 2 & 2 & 4 \end{Vmatrix} \cdot \begin{Vmatrix} x_1 \\ x_2 \\ x_3 \end{Vmatrix}$.

Considereremo quindi, d'ora in poi, ogni forma quadratica come generata da una matrice simmetrica, e questa simmetria è garantita quando abbiamo la forma quadratica generata dal differenziale totale secondo di una funzione differenziabile due volte, dato che la matrice di questa forma quadratica è la matrice Hessiana $\mathbb{H}(\mathbb{X}_0)$ di f. Per i nostri scopi, scriveremo una forma quadratica sempre nella forma $Q(\mathrm{d}\mathbb{X}) = \mathrm{d}\mathbb{X} \cdot \mathbb{H} \cdot \mathrm{d}\mathbb{X}^{\mathrm{T}}$, con $\mathrm{d}\mathbb{X} = (\mathrm{d}x_1, \mathrm{d}x_2, .., \mathrm{d}x_n)$, in modo da rappresentarla nella forma di un differenziale secondo.
Lo studio del segno delle forme quadratiche si basa sulle seguenti definizioni:
Definizione 43 : La forma quadratica $Q(\mathrm{d}\mathbb{X}) = \mathrm{d}\mathbb{X} \cdot \mathbb{H} \cdot \mathrm{d}\mathbb{X}^{\mathrm{T}}$ è detta:
definita positiva se $Q(\mathrm{d}\mathbb{X}) > 0, \forall\, \mathrm{d}\mathbb{X} \neq \mathbb{O}$;
definita negativa se $Q(\mathrm{d}\mathbb{X}) < 0, \forall\, \mathrm{d}\mathbb{X} \neq \mathbb{O}$.
Definizione 44 : La forma quadratica $Q(\mathrm{d}\mathbb{X}) = \mathrm{d}\mathbb{X} \cdot \mathbb{H} \cdot \mathrm{d}\mathbb{X}^{\mathrm{T}}$ è detta:
semidefinita positiva se $(Q(\mathrm{d}\mathbb{X}) \geq 0, \forall\, \mathrm{d}\mathbb{X} \neq \mathbb{O})$ **e** $(\exists\, \mathrm{d}\mathbb{X} \neq \mathbb{O} : Q(\mathrm{d}\mathbb{X}) = 0)$;
semidefinita negativa se $(Q(\mathrm{d}\mathbb{X}) \leq 0, \forall\, \mathrm{d}\mathbb{X} \neq \mathbb{O})$ **e** $(\exists\, \mathrm{d}\mathbb{X} \neq \mathbb{O} : Q(\mathrm{d}\mathbb{X}) = 0)$.
Definizione 45 : La forma quadratica $Q(\mathrm{d}\mathbb{X}) = \mathrm{d}\mathbb{X} \cdot \mathbb{H} \cdot \mathrm{d}\mathbb{X}^{\mathrm{T}}$ è detta:
indefinita se $(\exists\, \mathrm{d}\mathbb{X}_1 : Q(\mathrm{d}\mathbb{X}_1) > 0)$ **e** $(\exists\, \mathrm{d}\mathbb{X}_2 : Q(\mathrm{d}\mathbb{X}_2) < 0)$.

Esempio 61 : $Q(x_1,x_2) = \|x_1\ x_2\| \cdot \begin{Vmatrix} 1 & 0 \\ 0 & 3 \end{Vmatrix} \cdot \begin{Vmatrix} x_1 \\ x_2 \end{Vmatrix} = x_1^2 + 3x_2^2$ è una forma definita positiva in quanto $x_1^2 + 3x_2^2 \geq 0\, \forall\, (x_1,x_2)$ mentre $x_1^2 + 3x_2^2 = 0$ se e solo se $x_1 = x_2 = 0$.

Esempio 62 : $Q(x_1,x_2,x_3) = \|x_1\ x_2\ x_3\| \cdot \begin{Vmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 0 \end{Vmatrix} \cdot \begin{Vmatrix} x_1 \\ x_2 \\ x_3 \end{Vmatrix} = x_1^2 + 3x_2^2$ è una forma semidefinita positiva in quanto $x_1^2 + 3x_2^2 \geq 0\, \forall\, (x_1,x_2,x_3)$ mentre $x_1^2 + 3x_2^2 = 0$ quando $x_1 = x_2 = 0, \forall\, x_3 \neq 0$.

Esempio 63 : $Q(x_1,x_2) = \|x_1\ x_2\| \cdot \begin{Vmatrix} 1 & 1 \\ 1 & 1 \end{Vmatrix} \cdot \begin{Vmatrix} x_1 \\ x_2 \end{Vmatrix} = x_1^2 + 2x_1\, x_2 + x_2^2 = (x_1 + x_2)^2$ è una forma semidefinita positiva in quanto $(x_1+x_2)^2 \geq 0\, \forall\, (x_1,x_2)$ mentre $(x_1+x_2)^2 = 0$ ogni volta che $x_1 = -x_2$.

Esempio 64 : $Q(x_1,x_2,x_3) = \|x_1\ x_2\ x_3\| \cdot \begin{Vmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{Vmatrix} \cdot \begin{Vmatrix} x_1 \\ x_2 \\ x_3 \end{Vmatrix} = x_1^2 - x_2^2 + x_3^2$ è una forma indefinita in quanto $x_1^2 - x_2^2 + x_3^2 > 0$ se, ad esempio, $x_1 = x_2 = 0, \forall\, x_3 \neq 0$ mentre risulta invece $x_1^2 - x_2^2 + x_3^2 < 0$ quando, ad esempio, $x_1 = x_3 = 0, \forall\, x_2 \neq 0$.

Vediamo allora i criteri atti a garantire che una forma quadratica risulti definita o semidefinita.
Studiamo il segno del differenziale secondo nel caso di $f : \mathbb{R}^2 \to \mathbb{R}$, e poi generalizzeremo la procedura al caso di $f : \mathbb{R}^n \to \mathbb{R}$. Risulta allora:

$$f(x,y) - f(x_0,y_0) = \frac{1}{2}\,(\mathrm{d}x, \mathrm{d}y) \cdot \mathbb{H}(x_0,y_0) \cdot (\mathrm{d}x, \mathrm{d}y)^{\mathrm{T}} + o\big(\|(x,y) - (x_0,y_0)\|^2\big),$$

e noi dobbiamo studiare il segno di:

$$\mathrm{d}^2 f(x_0,y_0) = f''_{xx}(x_0,y_0)\,(\mathrm{d}x)^2 + 2\, f''_{xy}(x_0,y_0)\,\mathrm{d}x\,\mathrm{d}y + f''_{yy}(x_0,y_0)\,(\mathrm{d}y)^2 .$$

Scrivendo per brevità $\mathrm{d}^2 f = f''_{xx}\,(\mathrm{d}x)^2 + 2\, f''_{xy}\,\mathrm{d}x\,\mathrm{d}y + f''_{yy}\,(\mathrm{d}y)^2$, avremo:

$$\mathrm{d}^2 f = f''_{xx}\,(\mathrm{d}x)^2 + 2\, f''_{xy}\,\mathrm{d}x\,\mathrm{d}y + \frac{\left(f''_{xy}\right)^2}{f''_{xx}}\,(\mathrm{d}y)^2 - \frac{\left(f''_{xy}\right)^2}{f''_{xx}}\,(\mathrm{d}y)^2 + f''_{yy}\,(\mathrm{d}y)^2 =$$

$$\mathrm{d}^2 f = f''_{xx}\left((\mathrm{d}x)^2 + 2\,\frac{f''_{xy}}{f''_{xx}}\,\mathrm{d}x\,\mathrm{d}y + \frac{\left(f''_{xy}\right)^2}{\left(f''_{xx}\right)^2}\,(\mathrm{d}y)^2\right) - \frac{\left(f''_{xy}\right)^2}{f''_{xx}}\,(\mathrm{d}y)^2 + f''_{yy}\,(\mathrm{d}y)^2 =$$

$$\mathrm{d}^2 f = f''_{xx}\left(\mathrm{d}x + \frac{f''_{xy}}{f''_{xx}}\,\mathrm{d}y\right)^2 + \frac{f''_{xx}\, f''_{yy} - \left(f''_{xy}\right)^2}{f''_{xx}}\,(\mathrm{d}y)^2 .$$

In maniera analoga si può ottenere anche una seconda uguaglianza:

$$\mathrm{d}^2 f = f''_{yy}\left(\mathrm{d}y + \frac{f''_{xy}}{f''_{yy}}\,\mathrm{d}x\right)^2 + \frac{f''_{xx}\, f''_{yy} - \left(f''_{xy}\right)^2}{f''_{yy}}\,(\mathrm{d}x)^2 .$$

In ogni caso abbiamo la somma di due termini, ciascuno dei quali è dato dal prodotto di un quadrato (quindi sempre positivo) per un altro termine, di segno invece variabile:

f''_{xx} e $\dfrac{f''_{xx}\, f''_{yy} - \left(f''_{xy}\right)^2}{f''_{xx}}$ oppure f''_{yy} e $\dfrac{f''_{xx}\, f''_{yy} - \left(f''_{xy}\right)^2}{f''_{yy}}$.

Se consideriamo la matrice Hessiana $\left\| \begin{matrix} f''_{xx} & f''_{xy} \\ f''_{yx} & f''_{yy} \end{matrix} \right\|$, vedremo che f''_{xx} e f''_{yy} sono i cosiddetti minori di guida del I ordine della matrice, mentre $f''_{xx}\, f''_{yy} - \left(f''_{xy}\right)^2$ è il determinante della matrice Hessiana, detto anche minore di guida del II ordine.
Diamo allora le seguenti:
Definizione 46 : Si dicono **minori principali** di una matrice i minori aventi come elementi della diagonale principale elementi appartenenti alla diagonale principale della matrice data.
Definizione 47 : Si dicono **minori di guida** di una matrice simmetrica gli n minori principali, di ordine crescente, ottenuti partendo da un qualunque elemento della diagonale principale.

Esempio 65 : Se $\mathbb{H} = \left\| \begin{matrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{matrix} \right\|$, si hanno due minori di guida del I ordine, che sono $|a_{11}|$ e $|a_{22}|$, ed un solo minore di guida del II ordine: $\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}\, a_{22} - a_{12}\, a_{21}$.
Quindi ci sono due sole possibili sequenze di minori di guida:
$|\mathbb{H}_1| = |a_{11}|$ e $|\mathbb{H}_2| = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$ oppure $|\mathbb{H}_1| = |a_{22}|$ e $|\mathbb{H}_2| = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$.
Per una matrice di ordine 2 i minori principali ed i minori di guida coincidono.

Esempio 66 : Se $\mathbb{H} = \left\| \begin{matrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{matrix} \right\|$ abbiamo tre minori di guida del I ordine: $|a_{11}|$, $|a_{22}|$ e $|a_{33}|$; abbiamo due minori di guida del II ordine: $\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$ e $\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix}$, abbiamo un minore di guida del III ordine: $\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$.
Il minore $\begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix}$ è un minore principale ma non è un minore di guida.
Ci sono quindi quattro possibili sequenze di minori di guida:

1) $|\mathbb{H}_1| = |a_{11}|$, $|\mathbb{H}_2| = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$ e $|\mathbb{H}_3| = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$; oppure
2) $|\mathbb{H}_1| = |a_{22}|$, $|\mathbb{H}_2| = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$ e $|\mathbb{H}_3| = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$; oppure
3) $|\mathbb{H}_1| = |a_{22}|$, $|\mathbb{H}_2| = \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix}$ e $|\mathbb{H}_3| = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$; infine
4) $|\mathbb{H}_1| = |a_{33}|$, $|\mathbb{H}_2| = \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix}$ e $|\mathbb{H}_3| = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$.

La prima viene detta sequenza dei Minori di Nord-Ovest, l'ultima dei Minori di Sud-Est.
L'espressione del differenziale secondo per $f : \mathbb{R}^2 \to \mathbb{R}$ precedentemente ottenuta può essere allora riscritta come: $\mathrm{d}^2 f = |\mathbb{H}_1|\, Q_2^2 + \dfrac{|\mathbb{H}_2|}{|\mathbb{H}_1|}\, Q_1^2$, con $|\mathbb{H}_1| = |a_{11}|$ oppure $|\mathbb{H}_1| = |a_{22}|$ men-

tre in ogni caso $|\mathbb{H}_2| = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$. I termini Q_1^2 e Q_2^2 rappresentano, rispettivamente, il quadrato di un monomio e il quadrato di un binomio.
Dire che $\mathrm{d}^2 f < 0$ oppure $\mathrm{d}^2 f > 0$ in un intorno di (x_0, y_0) significa che le due precedenti espressioni sono negative o positive $\forall\,\mathrm{d}x$ e $\forall\,\mathrm{d}y$, ovvero che il loro segno è indipendente dalla scelta del $\mathrm{d}x$ e del $\mathrm{d}y$. Possiamo ottenere tale indipendenza in due soli casi:

M) $\begin{cases} f''_{xx} < 0 \\ f''_{xx}\, f''_{yy} - \left(f''_{xy}\right)^2 > 0 \end{cases}$ o $\begin{cases} f''_{yy} < 0 \\ f''_{xx}\, f''_{yy} - \left(f''_{xy}\right)^2 > 0 \end{cases} \Leftrightarrow \mathrm{d}^2 f < 0\,,\ \forall\,\mathrm{d}x$ e $\forall\,\mathrm{d}y$

e quindi (x_0, y_0) è punto di massimo;

m) $\begin{cases} f''_{xx} > 0 \\ f''_{xx}\, f''_{yy} - \left(f''_{xy}\right)^2 > 0 \end{cases}$ o $\begin{cases} f''_{yy} > 0 \\ f''_{xx}\, f''_{yy} - \left(f''_{xy}\right)^2 > 0 \end{cases} \Leftrightarrow \mathrm{d}^2 f > 0\,,\ \forall\,\mathrm{d}x$ e $\forall\,\mathrm{d}y$

e quindi (x_0, y_0) è punto di minimo.
Se risulta $f''_{xx}\, f''_{yy} - \left(f''_{xy}\right)^2 < 0$, avremo che $\mathrm{d}^2 f$ è la somma di due termini di segno opposto, e quindi il suo segno varia al variare del $\mathrm{d}x$ e del $\mathrm{d}y$; quindi $f''_{xx}\, f''_{yy} - \left(f''_{xy}\right)^2 < 0$ è condizione sufficiente a garantire che (x_0, y_0) è un punto di sella. Osserviamo che questo accade sempre quando f''_{xx} e f''_{yy} hanno segno diverso tra loro.
Nulla si può per il momento concludere nel caso che sia $f''_{xx}\, f''_{yy} - \left(f''_{xy}\right)^2 = 0$.
Tale caso verrà trattato successivamente con le forme semidefinite.
Se ora passiamo a studiare il $\mathrm{d}^2 f$ nel caso $f : \mathbb{R}^3 \to \mathbb{R}$, con procedura simile ma con calcoli molto più consistenti, otterremo un'espressione del tipo:
$\mathrm{d}^2 f = |\mathbb{H}_1|\, Q_3^2 + \dfrac{|\mathbb{H}_2|}{|\mathbb{H}_1|}\, Q_2^2 + \dfrac{|\mathbb{H}_3|}{|\mathbb{H}_2|}\, Q_1^2$, dove $|\mathbb{H}_1|$, $|\mathbb{H}_2|$ e $|\mathbb{H}_3|$ sono una qualsiasi delle quattro possibili sequenze di minori di guida viste nell'Esempio 66, e i termini Q_1^2, Q_2^2 e Q_3^2 rappresentano, rispettivamente, il quadrato di un monomio, di un binomio e di un trinomio.
Passando al caso generale del $\mathrm{d}^2 f$ per $f : \mathbb{R}^n \to \mathbb{R}$, avremo l'espressione:
$\mathrm{d}^2 f = |\mathbb{H}_1|\, Q_n^2 + \dfrac{|\mathbb{H}_2|}{|\mathbb{H}_1|}\, Q_{n-1}^2 + \ldots + \dfrac{|\mathbb{H}_{n-1}|}{|\mathbb{H}_{n-2}|}\, Q_2^2 + \dfrac{|\mathbb{H}_n|}{|\mathbb{H}_{n-1}|}\, Q_1^2$, dove $|\mathbb{H}_1|$, $|\mathbb{H}_2|$,...,$|\mathbb{H}_n|$ è una qualsiasi sequenza di minori di guida e Q_1^2, Q_2^2,...,Q_n^2 sono i quadrati di un monomio, di un binomio, ..., di un n-omio.
Possiamo allora formulare, in analogia a quanto visto nel caso di funzione di due variabili, il seguente criterio atto a stabilire se una forma quadratica in n variabili risulta definita, positiva o negativa. Vale il seguente:
Teorema 24 : La forma quadratica $Q(\mathrm{d}\mathbb{X}) = \mathrm{d}\mathbb{X} \cdot \mathbb{H} \cdot \mathrm{d}\mathbb{X}^{\mathrm{T}}$ risulta:
- definita positiva se e solo se $|\mathbb{H}_i| > 0, \forall\, i : 1 \le i \le n$;
- definita negativa se e solo se $(-1)^i \cdot |\mathbb{H}_i| > 0, \forall\, i : 1 \le i \le n$.

Se la forma $Q(\mathrm{d}\mathbb{X}) = \mathrm{d}\mathbb{X} \cdot \mathbb{H} \cdot \mathrm{d}\mathbb{X}^{\mathrm{T}}$ risulta definita positiva in $\mathbb{X}_0$, in quanto i minori di guida hanno tutti segno positivo, questa è condizione sufficiente (non necessaria) a garantire che in $\mathbb{X}_0$ c'è un punto di minimo. Se la forma $Q(\mathrm{d}\mathbb{X}) = \mathrm{d}\mathbb{X} \cdot \mathbb{H} \cdot \mathrm{d}\mathbb{X}^{\mathrm{T}}$ risulta definita negativa in $\mathbb{X}_0$, in quanto i minori di guida hanno segno alterno, quelli di posto pari positivo e quelli di posto dispari negativo, questa è condizione sufficiente (non necessaria) a garantire che in $\mathbb{X}_0$ c'è un punto di massimo.
Dato che il Teorema è espresso nella forma di condizione necessaria e sufficiente, deduciamo che qualsiasi sequenza di minori di guida, qualunque sia l'elemento sulla diagonale principale da cui si parte, conduce sempre alla stessa conclusione: non c'è quindi da ritenere più utile di altre una scelta nel minore di guida di ordine uno da cui partire.

Qualunque sequenza che non sia la $(+\ +\ ...\ +\ +)$ o la $(-\ +\ -\ +\ .....)$ indica che il punto $\mathbb{X}_0$ è un punto di sella. Quando anche un solo minore di guida fosse uguale a zero le precedenti considerazioni cessano di valere, e siamo nel campo delle forme quadratiche semidefinite.
Per la determinazione delle forme quadratiche semidefinite vale un criterio simile a quello dato per le forme definite. Non è però più sufficiente analizzare una qualunque sequenza di minori di guida, ma occorre esaminare tutti i minori principali della matrice. I minori principali sono quei minori che hanno per elementi della diagonale principale elementi che appartengono alla diagonale principale della matrice data.

Esempio 67 : Se $\mathbb{A} = \left\| \begin{matrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{matrix} \right\|$ abbiamo tre minori principali del I ordine: $|a_{11}|$, $|a_{22}|$ e $|a_{33}|$, e un minore principale del III ordine: $\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$, che coincidono con i minori di guida; ci sono invece tre minori principali del II ordine: $\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$, $\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix}$ e $\begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix}$. L'ultimo di questi non è un minore di guida.

Per vedere se una forma quadratica è semidefinita vale il seguente criterio, nel quale indichiamo con $|MP_i|$ un qualunque minore principale di ordine i:
Teorema 25 : La forma quadratica $Q(\mathrm{d}\mathbb{X}) = \mathrm{d}\mathbb{X} \cdot \mathbb{H} \cdot \mathrm{d}\mathbb{X}^{\mathrm{T}}$ risulta:
- semidefinita positiva se e solo se $|MP_i| \geq 0, \forall\, i : 1 \leq i \leq n$;
- semidefinita negativa se e solo se $(-1)^i \cdot |MP_i| \geq 0, \forall\, i : 1 \leq i \leq n$.
Le sequenze dei segni sono le stesse di quelle relative alle forme definite, ma ora è compresa la possibilità della presenza di zeri. Ogni sequenza di segni che non sia una delle due precedenti porta a concludere che la forma quadratica è indefinita.
Si faccia bene attenzione al fatto che mentre l'essere forma quadratica definita in $\mathbb{X}_0$ permette subito di dedurre la natura del punto stazionario, l'essere nel punto forma semidefinita non permette di concludere, ma solo di escludere: se in $\mathbb{X}_0$ la forma $Q(\mathrm{d}\mathbb{X}) = \mathrm{d}\mathbb{X} \cdot \mathbb{H} \cdot \mathrm{d}\mathbb{X}^{\mathrm{T}}$ è semidefinita positiva, allora in $\mathbb{X}_0$ non ci sarà un punto di massimo, e quindi ci sarà o un punto di minimo o un punto di sella; se invece in $\mathbb{X}_0$ la forma $Q(\mathrm{d}\mathbb{X}) = \mathrm{d}\mathbb{X} \cdot \mathbb{H} \cdot \mathrm{d}\mathbb{X}^{\mathrm{T}}$ è semidefinita negativa, allora in $\mathbb{X}_0$ non ci sarà un punto di minimo, e quindi ci sarà o un punto di massimo o un punto di sella.
Come decidere tra le due possibilità rimanenti dipende dalla funzione in oggetto, a seconda della quale si possono usare varie metodologie. La più frequente consiste nello studiare, con vari artifici, il segno della differenza $f(\mathbb{X}) - f(\mathbb{X}_0)$.
Vediamo attraverso alcuni esempi quanto finora illustrato.

Esempio 68 : Sia $\mathbb{H} = \left\| \begin{matrix} 4 & 2 & 1 \\ 2 & 1 & 2 \\ 1 & 2 & 1 \end{matrix} \right\|$ la matrice della forma quadratica $\mathrm{d}^2 f$. I minori principali del I ordine sono $4 > 0, 1 > 0$ e $1 > 0$; i minori principali del II ordine sono $\begin{vmatrix} 4 & 2 \\ 2 & 1 \end{vmatrix} = 0$, $\begin{vmatrix} 4 & 1 \\ 1 & 1 \end{vmatrix} = 3 > 0$; $\begin{vmatrix} 1 & 2 \\ 2 & 1 \end{vmatrix} = -3 < 0$. La presenza di un minore negativo di ordine pari porta immediatamente a concludere che la forma quadratica è indefinita.

Esempio 69 : Sia $\mathbb{H} = \left\| \begin{matrix} -4 & 2 & 3 \\ 2 & -2 & -1 \\ 3 & -1 & -3 \end{matrix} \right\|$ la matrice della forma quadratica $\mathrm{d}^2 f$. I minori principali del I ordine sono $-4 < 0,\ -2 < 0$ e $-3 < 0$; i minori principali del II ordine sono $\begin{vmatrix} -4 & 2 \\ 2 & -2 \end{vmatrix} = 4 > 0$, $\begin{vmatrix} -4 & 3 \\ 3 & -3 \end{vmatrix} = 3 > 0$; $\begin{vmatrix} -2 & -1 \\ -1 & -3 \end{vmatrix} = 5 > 0$. Il minore del III ordine è $\begin{vmatrix} -4 & 2 & 3 \\ 2 & -2 & -1 \\ 3 & -1 & -3 \end{vmatrix} = -2 < 0$. La forma quadratica è quindi definita negativa.

Per arrivare a questa conclusione bastava però esaminare la sequenza dei soli minori di guida: $|\mathbb{H}_1| = -4 < 0$; $|\mathbb{H}_2| = \begin{vmatrix} -4 & 2 \\ 2 & -2 \end{vmatrix} = 4 > 0$, $|\mathbb{H}_3| = \begin{vmatrix} -4 & 2 & 3 \\ 2 & -2 & -1 \\ 3 & -1 & -3 \end{vmatrix} = -2 < 0$.

Esempio 70 : Sia $\mathbb{H} = \left\| \begin{matrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 3 & 6 & 9 \end{matrix} \right\|$ la matrice della forma quadratica $\mathrm{d}^2 f$.

I minori principali del I ordine sono $1 > 0, 4 > 0$ e $9 > 0$; i minori principali del II ordine sono $\begin{vmatrix} 1 & 2 \\ 2 & 4 \end{vmatrix} = 0$, $\begin{vmatrix} 4 & 6 \\ 6 & 9 \end{vmatrix} = 0$; $\begin{vmatrix} 1 & 3 \\ 3 & 9 \end{vmatrix} = 0$. Il minore principale del III ordine è il determinante di $\mathbb{H}$, palesemente uguale a zero. Quindi la forma quadratica è semidefinita positiva.

Vediamo ora con degli esempi come applicare queste metodologie per lo studio dei punti stazionari di una funzione di più variabili.

Esempio 71 : Consideriamo la funzione $f(x,y) = x^2 + y^4$. Essendo $\nabla f(\mathbb{X}) = \left(2x, 4y^3\right)$, il solo punto stazionario è il punto $(0,0)$. Risulta $\mathbb{H} = \left\| \begin{matrix} 2 & 0 \\ 0 & 12y^2 \end{matrix} \right\|$, da cui, sostituendo, si ha $\mathbb{H}(0,0) = \left\| \begin{matrix} 2 & 0 \\ 0 & 0 \end{matrix} \right\|$. La forma quadratica $\mathrm{d}^2 f(0,0)$ è semidefinita positiva, per cui $(0,0)$ non è punto di massimo. Essendo $f(0,0) = 0$ ed essendo $x^2 + y^4 > 0$, $\forall\,(x,y) \neq (0,0)$, si vede subito che $(0,0)$ è un punto di minimo, oltretutto assoluto.

Esempio 72 : Data la funzione $f(x,y) = x^2 - y^4$, essendo $\nabla f(\mathbb{X}) = \left(2x, -4y^3\right)$, il solo punto stazionario è il punto $(0,0)$. Risulta $\mathbb{H} = \left\| \begin{matrix} 2 & 0 \\ 0 & -12y^2 \end{matrix} \right\|$, da cui, sostituendo, si ha $\mathbb{H}(0,0) = \left\| \begin{matrix} 2 & 0 \\ 0 & 0 \end{matrix} \right\|$.

La forma quadratica $\mathrm{d}^2 f(0,0)$ è semidefinita positiva, per cui $(0,0)$ non può essere punto di massimo. Essendo $f(0,0) = 0, f(x,0) = x^2 > 0$ e $f(0,y) = -y^4 < 0$, si conclude che $(0,0)$ è un punto di sella.

Notiamo quindi che, a parità di matrice Hessiana $\mathbb{H}(0,0) = \left\| \begin{matrix} 2 & 0 \\ 0 & 0 \end{matrix} \right\|$, varia la conclusione a seconda della funzione che stiamo esaminando.

Esempio 73 : Sia $f(x,y) = 3x^3 - y^3 - 3x^2y + 3xy^2 - 3x^2 - 3x + 3y$. Determiniamone gli eventuali punti di massimo e di minimo. Imponendo la condizione del I ordine, $\nabla f(\mathbb{X}) = 0$, otteniamo il sistema:

$\begin{cases} f'_x = 9x^2 - 6xy + 3y^2 - 6x - 3 = 0 \\ f'_y = -3y^2 - 3x^2 + 6xy + 3 = 0 \end{cases}$. Sommando le due equazioni otteniamo il sistema:

$\begin{cases} 6x^2 - 6x = 0 \\ 2xy - y^2 - x^2 + 1 = 0 \end{cases}$, ovvero: $\begin{cases} 6x(x-1) = 0 \\ 2xy - y^2 - x^2 + 1 = 0 \end{cases}$ da cui si ha:

$\begin{cases} x = 0 \\ 1 - y^2 = 0 \end{cases} \Rightarrow \begin{cases} x = 0 \\ y = 1 \end{cases}$ e $\begin{cases} x = 0 \\ y = -1 \end{cases}$, oppure:

$\begin{cases} x = 1 \\ 2y - y^2 = 0 \end{cases} \Rightarrow \begin{cases} x = 1 \\ y(2-y) = 0 \end{cases} \Rightarrow \begin{cases} x = 1 \\ y = 0 \end{cases}$ e $\begin{cases} x = 1 \\ y = 2 \end{cases}$.

Ci sono 4 punti stazionari: $\mathrm{P}_1 = (0,1)$, $\mathrm{P}_2 = (0,-1)$, $\mathrm{P}_3 = (1,0)$, $\mathrm{P}_4 = (1,2)$.
Passiamo quindi alle condizioni del II ordine. Avremo anzitutto:

$\mathbb{H} = \begin{Vmatrix} 18x - 6y - 6 & -6x + 6y \\ -6x + 6y & -6y + 6x \end{Vmatrix}$. Studiando l'Hessiano in ciascuno dei 4 punti si ha:

$\mathbb{H}(\mathrm{P}_1) = \begin{Vmatrix} -12 & 6 \\ 6 & -6 \end{Vmatrix} \Rightarrow \begin{cases} f''_{xx} = -12 < 0 \\ f''_{xx} f''_{yy} - \left(f''_{xy}\right)^2 = 72 - 36 = 36 > 0 \end{cases}$,

e quindi $(0,1)$ è un punto di massimo;

$\mathbb{H}(\mathrm{P}_2) = \begin{Vmatrix} 0 & -6 \\ -6 & 6 \end{Vmatrix} \Rightarrow f''_{xx} f''_{yy} - \left(f''_{xy}\right)^2 = 0 - 36 = -36 < 0$,

e quindi $(0,-1)$ è un punto di sella;

$\mathbb{H}(\mathrm{P}_3) = \begin{Vmatrix} 12 & -6 \\ -6 & 6 \end{Vmatrix} \Rightarrow \begin{cases} f''_{xx} = 12 > 0 \\ f''_{xx} f''_{yy} - \left(f''_{xy}\right)^2 = 72 - 36 = 36 > 0 \end{cases}$,

e quindi $(1,0)$ è un punto di minimo;

$\mathbb{H}(\mathrm{P}_4) = \begin{Vmatrix} 0 & 6 \\ 6 & -6 \end{Vmatrix} \Rightarrow f''_{xx} f''_{yy} - \left(f''_{xy}\right)^2 = 0 - 36 = -36 < 0$,

e quindi $(1,2)$ è un punto di sella.

Esempio 74 : Concludiamo lo studio della funzione $f(x,y) = \left(y - x^2\right)\left(y - x^4\right)$, i cui punti stazionari sono $\begin{cases} x = 0 \\ y = 0 \end{cases}$; $\begin{cases} x = 1 \\ y = 1 \end{cases}$; $\begin{cases} x = -1 \\ y = 1 \end{cases}$; $\begin{cases} x = \frac{1}{\sqrt{2}} \\ y = \frac{3}{8} \end{cases}$; $\begin{cases} x = -\frac{1}{\sqrt{2}} \\ y = \frac{3}{8} \end{cases}$. Passando alle condizioni del II ordine, abbiamo $\mathbb{H}(x,y) = \begin{Vmatrix} 30x^4 - 2y - 12x^2y & -2x - 4x^3 \\ -2x - 4x^3 & 2 \end{Vmatrix}$, da cui:

$\mathbb{H}(0,0) = \begin{Vmatrix} 0 & 0 \\ 0 & 2 \end{Vmatrix}$ per cui la forma quadratica è semidefinita positiva, ma abbiamo già visto che il punto $(0,0)$ è un punto di sella;

$\mathbb{H}(1,1) = \begin{Vmatrix} 16 & -6 \\ -6 & 2 \end{Vmatrix}$ da cui $\begin{cases} |\mathbb{H}_1| = 16 > 0 \\ |\mathbb{H}_2| = 32 - 36 = -4 < 0 \end{cases}$,

e quindi anche $(1,1)$ è un punto di sella;

$\mathbb{H}(-1,1) = \begin{Vmatrix} 16 & 6 \\ 6 & 2 \end{Vmatrix}$ da cui $\begin{cases} |\mathbb{H}_1| = 16 > 0 \\ |\mathbb{H}_2| = 32 - 36 = -4 < 0 \end{cases}$,

e quindi anche $(-1,1)$ è un punto di sella;

$\mathbb{H}\left(\frac{1}{\sqrt{2}}, \frac{3}{8}\right) = \begin{Vmatrix} \frac{9}{2} & -\frac{4}{\sqrt{2}} \\ -\frac{4}{\sqrt{2}} & 2 \end{Vmatrix}$ da cui $\begin{cases} |\mathbb{H}_1| = \frac{9}{2} > 0 \\ |\mathbb{H}_2| = 9 - 8 = 1 > 0 \end{cases}$, e quindi $\left(\dfrac{1}{\sqrt{2}}, \dfrac{3}{8}\right)$ è un punto di minimo;

$$\mathbb{H}\left(-\frac{1}{\sqrt{2}}, \frac{3}{8}\right) = \begin{Vmatrix} \frac{9}{2} & \frac{4}{\sqrt{2}} \\ \frac{4}{\sqrt{2}} & 2 \end{Vmatrix} \text{ da cui } \begin{cases} |\mathbb{H}_1| = \frac{9}{2} > 0 \\ |\mathbb{H}_2| = 9 - 8 = 1 > 0 \end{cases},$$

e quindi $\left(-\dfrac{1}{\sqrt{2}}, \dfrac{3}{8}\right)$ è un punto di minimo.

Esempio 75 : Data $f(x, y, z) = x^2 + y^2 + z^2 - x\,y^2$, determiniamone gli eventuali punti di massimo o minimo relativo. Le condizioni del I ordine danno luogo al sistema:

$$\begin{cases} f'_x = 2x - y^2 = 0 \\ f'_y = 2y - 2x\,y = 2y\,(1 - x) = 0 \\ f'_z = 2z = 0 \end{cases} \text{ da cui otteniamo le soluzioni:}$$

$$\begin{cases} x = 0 \\ y = 0 \\ z = 0 \end{cases} \text{ oppure } \begin{cases} x = 1 \\ y^2 = 2 \\ z = 0 \end{cases} \text{ da cui } \begin{cases} x = 1 \\ y = \sqrt{2} \\ z = 0 \end{cases} \text{ e } \begin{cases} x = 1 \\ y = -\sqrt{2} \\ z = 0 \end{cases}.$$

Risulta poi $\mathbb{H}(x, y, z) = \begin{Vmatrix} 2 & -2y & 0 \\ -2y & 2 - 2x & 0 \\ 0 & 0 & 2 \end{Vmatrix}$ da cui otteniamo:

$$\mathbb{H}(0, 0, 0) = \begin{Vmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{Vmatrix} \text{ e quindi: } \begin{cases} |\mathbb{H}_1| = 2 > 0 \\ |\mathbb{H}_2| = 4 > 0 \\ |\mathbb{H}_3| = 8 > 0 \end{cases}$$

per cui $(0, 0, 0)$ è un punto di minimo;

$$\mathbb{H}\left(1, \sqrt{2}, 0\right) = \begin{Vmatrix} 2 & -2\sqrt{2} & 0 \\ -2\sqrt{2} & 0 & 0 \\ 0 & 0 & 2 \end{Vmatrix} \text{ da cui: } \begin{cases} |\mathbb{H}_1| = 2 > 0 \\ |\mathbb{H}_2| = -8 < 0 \\ |\mathbb{H}_3| = -16 < 0 \end{cases}$$

per cui $\left(1, \sqrt{2}, 0\right)$ è un punto di sella;

$$\mathbb{H}\left(1, -\sqrt{2}, 0\right) = \begin{Vmatrix} 2 & 2\sqrt{2} & 0 \\ 2\sqrt{2} & 0 & 0 \\ 0 & 0 & 2 \end{Vmatrix} \text{ da cui: } \begin{cases} |\mathbb{H}_1| = 2 > 0 \\ |\mathbb{H}_2| = -8 < 0 \\ |\mathbb{H}_3| = -16 < 0 \end{cases}$$

per cui $\left(1, -\sqrt{2}, 0\right)$ è un punto di sella.

Esempio 76 : Data $f(x, y) = (x - 1)^4 - y\,(x - 1)^2 + y^2$, determiniamone gli eventuali punti di massimo o minimo relativo. Le condizioni del I ordine danno luogo al sistema:

$$\begin{cases} f'_x = 4(x - 1)^3 - 2y(x - 1) = 0 \\ f'_y = -(x - 1)^2 + 2y = 0 \end{cases} \Rightarrow \begin{cases} 4(x - 1)^3 - (x - 1)^3 = 0 \\ 2y = (x - 1)^2 \end{cases} \Rightarrow \begin{cases} 3(x - 1)^3 = 0 \\ 2y = (x - 1)^2 \end{cases},$$

per cui $\begin{cases} x = 1 \\ y = 0 \end{cases}$ è l'unico punto stazionario. Inoltre:

$$\mathbb{H}(x, y) = \begin{Vmatrix} 12(x - 1)^2 - 2y & -2(x - 1) \\ -2(x - 1) & 2 \end{Vmatrix} \text{ da cui } \mathbb{H}(1, 0) = \begin{Vmatrix} 0 & 0 \\ 0 & 2 \end{Vmatrix}.$$

Quindi $\mathrm{d}^2 f(1, 0)$ è una forma quadratica semidefinita positiva, per cui $(1, 0)$ non può essere un punto di massimo. Essendo $\mathrm{d}^2 f(1, 0) = 2\,(\mathrm{d}y)^2$, indaghiamo nella direzione $\mathrm{d}y = 0$, ovvero $y = 0$. Avremo $f(x, 0) = (x - 1)^4$, che indicherebbe $x = 1$ (e $y = 0$) come un punto di minimo. Una indagine come questa, cioè di tipo monodimensionale, però non permette di

concludere, in maniera affermativa, che $(1,0)$ è un punto di minimo; potrebbe servire casomai, anche se qui non accade, per escludere che $(1,0)$ sia punto di minimo.
Nel punto $(1,0)$ c'è comunque veramente un punto di minimo; basta scrivere:

$$f(x,y)=(x-1)^4-y\,(x-1)^2+y^2=\frac{1}{4}(x-1)^4-2y\,\frac{(x-1)^2}{2}+y^2+\frac{3}{4}(x-1)^4=$$

$$f(x,y)=\left(\frac{1}{2}(x-1)^2-y\right)^2+\frac{3}{4}(x-1)^4>0=f(1,0)\,,\;\forall\,(x,y)\neq(1,0)\,.$$

Esempio 77 : Data $f(x,y)=3y^2+6xy-x^3-9x-6y$, determiniamone gli eventuali punti di massimo o minimo relativo. Le condizioni del I ordine danno luogo al sistema:

$$\begin{cases} f'_x=6y-3x^2-9=0 \\ f'_y=6y+6x-6=0 \end{cases} \Rightarrow \begin{cases} 2-2x-x^2-3=0 \\ y=1-x \end{cases} \Rightarrow \begin{cases} -(x+1)^2=0 \\ y=1-x \end{cases} \Rightarrow \begin{cases} x=-1 \\ y=2 \end{cases}$$

che è l'unica soluzione. Inoltre $\mathbb{H}(x,y)=\left\|\begin{matrix} -6x & 6 \\ 6 & 6 \end{matrix}\right\|$ da cui $\mathbb{H}(-1,2)=\left\|\begin{matrix} 6 & 6 \\ 6 & 6 \end{matrix}\right\|$, per cui $\mathrm{d}^2 f(-1,2)$ è una forma semidefinita positiva. Quindi $(-1,2)$ non può essere punto di massimo. Essendo $\mathrm{d}^2 f(-1,2)=6\,(\mathrm{d}x+\mathrm{d}y)^2$, risulta $\mathrm{d}^2 f(-1,2)=0$ se $\mathrm{d}x=-\mathrm{d}y$.
Esaminando la funzione sulla retta $y=-x+1$, passante per $(-1,2)$ e parallela alla $y=-x$, si ha: $f(x,1-x)=-x^3-3x^2-3x-3=-3\,(x+1)^3$ ed anche:
$f'_x(x,1-x)=-9\,(x+1)^2$, che risulta negativa $\forall\,x\neq-1$.
Lungo tale retta quindi la funzione è sempre decrescente, quindi il punto $(-1,2)$ non è neppure di minimo, e quindi è un punto di sella. Questa volta l'analisi in una particolare direzione porta a conclusioni negative, cioè ad escludere il punto di minimo e ci da quindi la certezza del punto di sella.

METODO DEGLI AUTOVALORI DELLA MATRICE HESSIANA
Esiste anche un'altra metodologia per lo studio delle forme quadratiche, basata sugli autovalori della matrice simmetrica che genera la forma quadratica. Sappiamo che una matrice simmetrica ammette autovalori tutti reali, e che può essere sempre diagonalizzata mediante una matrice ortogonale. Ovvero, se $\mathbb{H}$ è la matrice simmetrica, esiste una matrice ortogonale $\mathbb{P}$ tale che: $\mathbb{H}\cdot\mathbb{P}=\mathbb{P}\cdot\mathbb{D}$, da cui $\mathbb{P}^{-1}\cdot\mathbb{H}\cdot\mathbb{P}=\mathbb{P}^{\mathrm{T}}\cdot\mathbb{H}\cdot\mathbb{P}=\mathbb{D}$, dove $\mathbb{D}$ è la matrice diagonale avente come elementi della diagonale principale gli autovalori $\lambda_1,\lambda_2,...,\lambda_n$ di $\mathbb{H}$. La matrice $\mathbb{P}$ è la matrice modale di $\mathbb{H}$, ovvero la matrice avente per colonne gli autovettori normalizzati di $\mathbb{H}$, che per le proprietà delle matrici simmetriche sappiamo essere a due a due perpendicolari tra loro. Affinchè la forma quadratica $\mathrm{d}^2 f=\mathrm{d}\mathbb{X}\cdot\mathbb{H}\cdot\mathrm{d}\mathbb{X}^{\mathrm{T}}$ risulti definita dovrà risultare $\mathrm{d}^2 f>0$ oppure $\mathrm{d}^2 f<0$, $\forall\,\mathrm{d}\mathbb{X}\neq\mathbb{O}$. Poniamo $\mathrm{d}\mathbb{X}=\mathrm{d}\mathbb{Y}\cdot\mathbb{P}^{\mathrm{T}}$, dove $\mathbb{P}$ è la matrice modale di $\mathbb{H}$. Essendo $|\mathbb{P}|\neq0$, l'applicazione lineare $\mathrm{d}\mathbb{Y}\to\mathrm{d}\mathbb{Y}\cdot\mathbb{P}^{\mathrm{T}}=\mathrm{d}\mathbb{X}$ è una corrispondenza biunivoca $\mathbb{R}^n\to\mathbb{R}^n$, ed avremo:

$$\mathrm{d}\mathbb{X}\cdot\mathbb{H}\cdot\mathrm{d}\mathbb{X}^{\mathrm{T}}=\mathrm{d}\mathbb{Y}\cdot\mathbb{P}^{\mathrm{T}}\cdot\mathbb{H}\cdot\left(\mathrm{d}\mathbb{Y}\cdot\mathbb{P}^{\mathrm{T}}\right)^{\mathrm{T}}=\mathrm{d}\mathbb{Y}\cdot\mathbb{P}^{\mathrm{T}}\cdot\mathbb{H}\cdot\mathbb{P}\cdot(\mathrm{d}\mathbb{Y})^{\mathrm{T}}=\mathrm{d}\mathbb{Y}\cdot\mathbb{D}\cdot(\mathrm{d}\mathbb{Y})^{\mathrm{T}}.$$

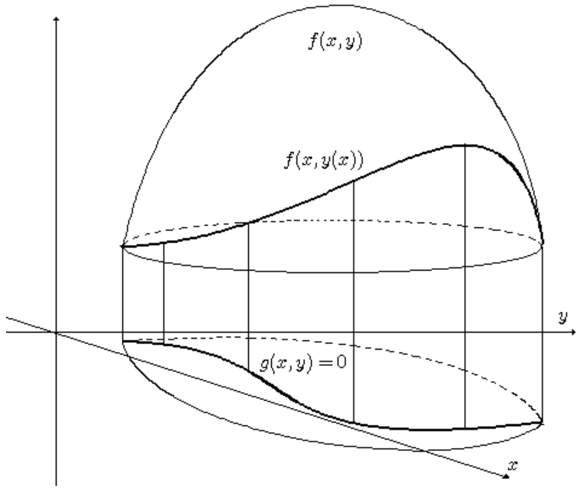
Se $\mathrm{d}\mathbb{Y}\cdot\mathbb{D}\cdot(\mathrm{d}\mathbb{Y})^{\mathrm{T}}>0$ o se $\mathrm{d}\mathbb{Y}\cdot\mathbb{D}\cdot(\mathrm{d}\mathbb{Y})^{\mathrm{T}}<0$, $\forall\,\mathrm{d}\mathbb{Y}\neq\mathbb{O}$, sarà anche $\mathrm{d}\mathbb{X}\cdot\mathbb{H}\cdot\mathrm{d}\mathbb{X}^{\mathrm{T}}>0$ oppure $\mathrm{d}\mathbb{X}\cdot\mathbb{H}\cdot\mathrm{d}\mathbb{X}^{\mathrm{T}}<0$, $\forall\,\mathrm{d}\mathbb{X}\neq\mathbb{O}$. Ma
$\mathrm{d}\mathbb{Y}\cdot\mathbb{D}\cdot(\mathrm{d}\mathbb{Y})^{\mathrm{T}}=\lambda_1\,(\mathrm{d}y_1)^2+\lambda_2\,(\mathrm{d}y_2)^2+...+\lambda_n\,(\mathrm{d}y_n)^2$, dalla quale segue subito il
Teorema 26 : La forma quadratica $\mathrm{d}\mathbb{X}\cdot\mathbb{H}\cdot\mathrm{d}\mathbb{X}^{\mathrm{T}}$ risulta:
- definita positiva se e solo se $\lambda_i>0,\forall\,i:1\leq i\leq n$;
- definita negativa se e solo se $\lambda_i<0,\forall\,i:1\leq i\leq n$;
- semidefinita positiva se e solo se $\lambda_i\geq0,\forall\,i:1\leq i\leq n$ e $\exists\,\lambda_k=0$;
- semidefinita negativa se e solo se $\lambda_i\leq0,\forall\,i:1\leq i\leq n$ e $\exists\,\lambda_k=0$;
- indefinita se $\exists\,\lambda_i>0$ e $\exists\,\lambda_j<0$.

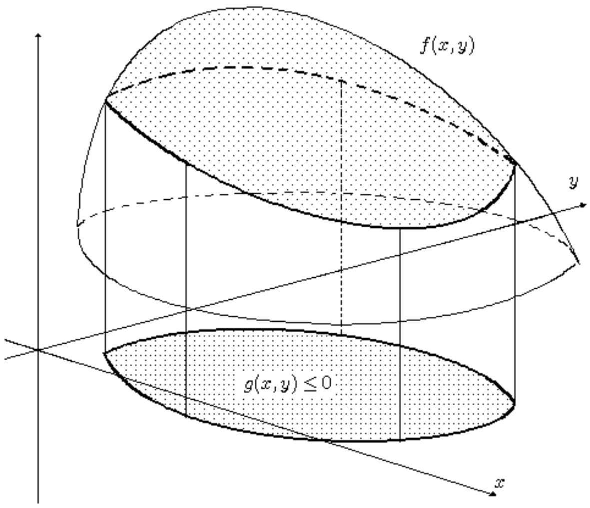
Sia il Teorema 24 che il Teorema 26 esprimono condizioni necessarie e sufficienti affinchè una forma quadratica risulti definita; quindi le due metodologie non sono alternative ma conducono sempre alle stesse conclusioni.
Quanto visto finora riguardo alle condizioni del II ordine per l'analisi dei punti stazionari è basato su di una analisi puntuale, ovvero sulla verifica che $\mathrm{d}^2 f(\mathbb{X}_0)$ risulti una forma definita positiva oppure negativa oppure indefinita; se la forma $\mathrm{d}^2 f(\mathbb{X}_0)$ risulta semidefinita possiamo solo escludere una delle tre possibilità, per poi, con ulteriori analisi, decidere tra le due rimanenti, una delle quali sarà sempre quella del punto di sella.
Diversa è la situazione se si può condurre un'analisi di tipo globale. Se $\mathrm{d}^2 f(\mathbb{X})$ è semidefinita, positiva o negativa, sia in $\mathbb{X}_0$ che in tutto il dominio, allora questo è sufficiente a garantire che $\mathbb{X}_0$ è punto, rispettivamente, di minimo o di massimo.

Esempio 78 : Riprendiamo la funzione $f(x,y) = x^2 + y^4$, con il punto stazionario $(0,0)$ e con matrice Hessiana $\mathbb{H}(x,y) = \left\| \begin{matrix} 2 & 0 \\ 0 & 12y^2 \end{matrix} \right\|$.
Risulta $\mathrm{d}^2 f(x,y) = 2\,(\mathrm{d}x)^2 + 12\,y^2(\mathrm{d}y)^2 \geq 0$, $\forall\,(\mathrm{d}x, \mathrm{d}y) \in \mathbb{R}^2$. Quindi la forma $\mathrm{d}^2 f(x,y)$ è semidefinita positiva non solo in $(0,0)$ ma in tutto $\mathbb{R}^2$, e quindi, come già visto, $(0,0)$ è un punto di minimo.

MASSIMI E MINIMI VINCOLATI

La ricerca dei punti di massimo e di minimo, sia relativi che assoluti, alla luce di quanto visto finora, può essere scomposta in tre diversi problemi. Il primo è quello finora affrontato, che riguarda la ricerca e l'analisi dei punti stazionari di una funzione in tutto il suo campo d'esistenza, e che ha per soluzioni punti che risultano generalmente interni al dominio. Per questa ricerca occorre che la funzione risulti differenziabile. Il secondo problema, del quale ci limitiamo a far notare l'esistenza, senza però scendere nello specifico, riguarda l'analisi dei punti nei quali una funzione è definita ma non risulta differenziabile. Anche questi punti potrebbero essere di massimo o di minimo. Il terzo tipo di problema riguarda la ricerca di punti di massimo e di minimo relativamente ad un certo opportuno sottoinsieme del campo d'esistenza. Il Teorema di Weierstrass ci fornisce una condizione sufficiente a garantire l'esistenza di massimo e minimo per funzioni continue in un insieme compatto. Due sono i modi nei quali si può presentare questo terzo tipo di problema. Il primo è quello dei cosiddetti vincoli di uguaglianza, che nel caso più semplice si presenta nella forma $\begin{cases} \text{Max/min } f(x,y) \\ \text{s.v. : } g(x,y) = 0 \end{cases}$.
La s.v. significa "sotto al vincolo". Non si cercano gli estremi per $f(x,y)$ in tutto il dominio (punti di massimo o minimo che chiameremo da ora in poi "liberi"), ma solo tra i punti che soddisfano l'equazione $g(x,y) = 0$. Come $g(x,y) = 0$ prende il nome di **vincolo**, così la funzione $f(x,y)$ viene detta **funzione obiettivo**. Se la funzione $g(x,y)$ soddisfa ad opportune ipotesi, riconducibili a quelle del Teorema del Dini sulle funzioni implicite, possiamo considerare una delle due variabili come funzione, esplicita o implicita, dell'altra, ad esempio $y = y(x)$, e quindi ricavare una curva nel piano: $(x, y(x))$; quindi il problema è ricondotto alla determinazione dei punti di massimo e di minimo della funzione $f(x, y(x))$, ovvero della curva, proiezione sulla superficie $f(x,y)$ dei punti della curva definita mediante la $g(x,y) = 0$.
All'aumentare delle variabili e del numero delle equazioni avremo rappresentazioni geometriche del problema di dimensioni maggiori di quello ora descritto.





Consideriamo invece, quale secondo tipo di problema, un vincolo espresso in forma di disequazione: $g(x,y) \leq 0$. Normalmente (ma non necessariamente) i punti nei quali è soddisfatta la $g(x,y) \leq 0$ costituiscono l'interno e la frontiera di una regione contenuta nel dominio, ed il problema allora consiste nel trovare i punti di massimo e di minimo della funzione $f(x,y)$ all'interno o sulla frontiera della regione scelta: $\begin{cases} \text{Max/min } f(x,y) \\ \text{s.v. : } g(x,y) \leq 0 \end{cases}$. La ricerca degli estremi all'interno della regione data ricalca quella dei massimi e minimi liberi, quella nei punti di frontiera è analoga a quella dei massimi e minimi con vincoli di uguaglianza. I due problemi che abbiamo descritto passano sotto il nome di massimizzazione o minimizzazione con vincoli di uguaglianza o con vincoli di disuguaglianza.

ESTREMI CON VINCOLI DI UGUAGLIANZA

Iniziamo a trattare il caso descritto per primo: $\begin{cases} \text{Max/min } z = f(x,y) \\ \text{s.v. : } g(x,y) = 0 \end{cases}$. Sia $\mathcal{E}$ l'insieme dei punti che soddisfano il vincolo $g(x,y) = 0$; supponiamo che $f(x,y)$ e $g(x,y)$ siano funzioni differenziabili e che sia $\nabla g(x,y) \neq (0,0)$, $\forall\,(x,y) \in \mathcal{E}$. Quest'ultima condizione permette di applicare il Teorema del Dini in ogni punto di $\mathcal{E}$, e quindi di determinare, esplicitamente o implicitamente, una variabile in funzione dell'altra. Questa condizione può essere formulata anche come $\text{Car}\left(\dfrac{\partial g}{\partial(x,y)}\right) = 1 = \text{Max}$. Se la caratteristica dello Jacobiano, cioè del gradiente di g, è massima, ovvero pari a 1, le due derivate g'_x e g'_y non possono essere contemporaneamente nulle. Supponiamo tra le due possibilità che esista la $y = y(x)$, ed avremo allora due composizioni di funzioni :

$\mathbb{R} \to \mathbb{R}^2 \xrightarrow{f} \mathbb{R}, x \to (x, y(x)) \to f(x,y) = z$, e la

$\mathbb{R} \to \mathbb{R}^2 \xrightarrow{g} \mathbb{R}, x \to (x, y(x)) \to g(x,y) = 0$.

Derivando, nella prima, la variabile z rispetto a x otteniamo:

$$\frac{\mathrm{d}z}{\mathrm{d}x} = \frac{\partial f}{\partial x} \cdot \frac{\mathrm{d}x}{\mathrm{d}x} + \frac{\partial f}{\partial y} \cdot \frac{\mathrm{d}y}{\mathrm{d}x} = f'_x + f'_y \cdot y'(x).$$

Dall'equazione del vincolo ricaviamo, come derivata di funzione implicita, $y'(x)$ ed avremo:

$y'(x) = -\dfrac{g'_x}{g'_y}$ con la quale, sostituendo, otteniamo: $\dfrac{\mathrm{d}z}{\mathrm{d}x} = f'_x - f'_y \cdot \dfrac{g'_x}{g'_y}$.

Supponiamo, per ipotesi, che il punto (x_0, y_0) sia una soluzione, ovvero un punto di massimo o di minimo, per il nostro problema. Essendo la funzione una composizione di funzioni differenziabili, dovrà risultare, per il Teorema di Fermat:

$\dfrac{\mathrm{d}z}{\mathrm{d}x} = f'_x - f'_y \cdot \dfrac{g'_x}{g'_y} = 0$, da cui $f'_x = f'_y \cdot \dfrac{g'_x}{g'_y}$ ovvero $f'_x \cdot g'_y = f'_y \cdot g'_x$ ovvero:
$\begin{vmatrix} f'_x & f'_y \\ g'_x & g'_y \end{vmatrix} = \left|\dfrac{\partial(f,g)}{\partial(x,y)}\right| = 0$. Essendo $\left|\dfrac{\partial(f,g)}{\partial(x,y)}\right| = \begin{vmatrix} \nabla f \\ \nabla g \end{vmatrix} = 0$, risulta che ∇f e ∇g sono vettori linearmente dipendenti, e quindi $\nabla f = \lambda\,\nabla g$, con $\lambda \in \mathbb{R}$.
Ma allora $\nabla f - \lambda\,\nabla g = \nabla(f - \lambda\,g) = \mathbb{O}$.
Abbiamo così ottenuto una condizione necessaria affinchè il punto (x_0, y_0) sia soluzione del problema $\begin{cases} \text{Max/min } f(x,y) \\ \text{s.v. : } g(x,y) = 0 \end{cases}$. Anche ora si deve annullare un gradiente, che però non è quello della funzione obiettivo, come nel caso dei massimi e minimi liberi, ma quello della funzione $f(x,y) - \lambda\,g(x,y)$.
La funzione $\Lambda(x,y,\lambda) = f(x,y) - \lambda\,g(x,y)$ viene detta "**funzione Lagrangiana**" mentre λ prende il nome di "**moltiplicatore di Lagrange**". Vale quindi il
Teorema 27 : Siano f e g funzioni $\mathbb{R}^2 \to \mathbb{R}$ differenziabili con $\nabla g(x,y) \neq (0,0)$, sia (x_0, y_0) soluzione del problema $\begin{cases} \text{Max/min } f(x,y) \\ \text{s.v. : } g(x,y) = 0 \end{cases}$. Se $\Lambda(x,y,\lambda) = f(x,y) - \lambda\,g(x,y)$, esiste allora un valore λ_0 per il quale: $\nabla\Lambda(x_0, y_0, \lambda_0) = \mathbb{O}$.
Imporre le condizioni del I ordine significa quindi imporre $\nabla\Lambda(x,y,\lambda) = \mathbb{O}$, tenendo però presente che le eventuali soluzioni devono comunque soddisfare il vincolo, e quindi si dovrà risolvere il sistema: $\begin{cases} \Lambda'_x = f'_x - \lambda\,g'_x = 0 \\ \Lambda'_y = f'_y - \lambda\,g'_y = 0 \\ g(x,y) = 0 \end{cases}$. Vediamo però che $\Lambda'_\lambda = -\,g(x,y)$, e quindi il precedente sistema può essere sì visto come $\nabla\Lambda(x,y,\lambda) = \mathbb{O}$, intendendo però $\Lambda(x,y,\lambda)$ come funzione delle variabili x, y e λ, e quindi si impone il sistema: $\begin{cases} \Lambda'_x = f'_x - \lambda\,g'_x = 0 \\ \Lambda'_y = f'_y - \lambda\,g'_y = 0 \\ \Lambda'_\lambda = -\,g(x,y) = 0 \end{cases}$.

Supponiamo che il problema sia ora $\begin{cases} \text{Max/min } w = f(x,y,z) \\ \text{s.v. : } g(x,y,z) = 0 \end{cases}$. Abbiamo ancora un solo vincolo, ma le variabili indipendenti sono tre. Nell'ipotesi che $\dfrac{\partial(g)}{\partial(x,y,z)}$ abbia caratteristica pari ad 1 in ogni punto di $\mathcal{E}$, per il Teorema del Dini supponiamo di ricavare (esplicitamente o implicitamente) $z = z(x,y)$, ottenendo la seguente composizione di funzioni:
$\mathbb{R}^2 \to \mathbb{R}^3 \xrightarrow{f} \mathbb{R}$, $(x,y) \to (x,y,z(x,y)) \to f(x,y,z) = w$, e quindi ci proponiamo di trovare massimi e minimi di una funzione non di una ma di due variabili. Supponendo f e g differenziabili, dovremo ora imporre:

$$\begin{cases} \dfrac{\partial w}{\partial x} = \dfrac{\partial f}{\partial x} \cdot \dfrac{\partial x}{\partial x} + \dfrac{\partial f}{\partial y} \cdot \dfrac{\partial y}{\partial x} + \dfrac{\partial f}{\partial z} \cdot \dfrac{\partial z}{\partial x} = f'_x + f'_y \cdot 0 + f'_z \cdot z'_x \\ \dfrac{\partial w}{\partial y} = \dfrac{\partial f}{\partial x} \cdot \dfrac{\partial x}{\partial y} + \dfrac{\partial f}{\partial y} \cdot \dfrac{\partial y}{\partial y} + \dfrac{\partial f}{\partial z} \cdot \dfrac{\partial z}{\partial y} = f'_x \cdot 0 + f'_y + f'_z \cdot z'_y \end{cases}.$$

Ma, per il Teorema del Dini, dall'equazione del vincolo ricaviamo, come derivata di funzione implicita, z'_x e z'_y ed avremo: $z'_x = -\dfrac{g'_x}{g'_z}$ e $z'_y = -\dfrac{g'_y}{g'_z}$, dalle quali, sostituendo, si ha:

$\begin{cases} \dfrac{\partial w}{\partial x} = f'_x - f'_z \cdot \dfrac{g'_x}{g'_z} \\ \dfrac{\partial w}{\partial y} = f'_y - f'_z \cdot \dfrac{g'_y}{g'_z} \end{cases}$. Se, per ipotesi, il punto (x_0, y_0, z_0) è soluzione del problema, dovrà risultare $\dfrac{\partial w}{\partial x} = \dfrac{\partial w}{\partial y} = 0$, ovvero $\begin{cases} \dfrac{\partial w}{\partial x} = f'_x \cdot g'_z - f'_z \cdot g'_x = 0 \\ \dfrac{\partial w}{\partial y} = f'_y \cdot g'_z - f'_z \cdot g'_y = 0 \end{cases}$, da cui segue:
$\begin{vmatrix} f'_x & f'_z \\ g'_x & g'_z \end{vmatrix} = \begin{vmatrix} f'_y & f'_z \\ g'_y & g'_z \end{vmatrix} = 0$. Ma allora $\left\| \begin{matrix} f'_x & f'_y & f'_z \\ g'_x & g'_y & g'_z \end{matrix} \right\| = \dfrac{\partial(f,g)}{\partial(x,y,z)} = \left\| \begin{matrix} \nabla f \\ \nabla g \end{matrix} \right\|$ ha caratteristica pari a 1, e quindi $\nabla f = \lambda \nabla g$, da cui otteniamo ancora $\nabla(f - \lambda g) = \nabla \Lambda = \mathbb{O}$.
L'aumento nel numero delle variabili indipendenti non porta cambiamenti dal punto di vista operativo: in un punto di massimo o minimo vincolato è ancora necessario l'annullamento del gradiente della funzione Lagrangiana, che però dà luogo questa volta al sistema:
$\begin{cases} \Lambda'_x = f'_x - \lambda g'_x = 0 \\ \Lambda'_y = f'_y - \lambda g'_y = 0 \\ \Lambda'_z = f'_z - \lambda g'_z = 0 \\ \Lambda'_\lambda = -g(x,y,z) = 0 \end{cases}$, di quattro equazioni nelle quattro incognite x, y, z e λ.
Con le debite modifiche, si può enunciare per questo caso un Teorema analogo al Teorema 27.

La prossima generalizzazione che faremo del problema interviene invece sui vincoli.
Consideriamo infatti il problema: $\begin{cases} \text{Max/min } w = f(x,y,z) \\ \text{s.v.}: \begin{cases} g(x,y,z) = 0 \\ h(x,y,z) = 0 \end{cases} \end{cases}$, che presenta non uno ma due vincoli nelle tre variabili x, y e z. Poniamo per ipotesi che $\dfrac{\partial(g,h)}{\partial(x,y,z)}$ abbia caratteristica pari a 2, ovvero massima, in ogni punto dell'insieme $\mathcal{E}$ in cui risultano contemporaneamente soddisfatti i due vincoli. Per il Teorema del Dini questo garantisce l'esistenza di una funzione $\mathbb{R} \to \mathbb{R}^2$, ad esempio $x \to (y(x), z(x))$, che genera la seguente composizione di funzioni: $\mathbb{R} \to \mathbb{R}^3 \xrightarrow{f} \mathbb{R}, x \to (x, y(x), z(x)) \to f(x,y,z) = w$, e quindi siamo ricondotti alla determinazione degli estremi di una funzione di una sola variabile. In un punto (x_0, y_0, z_0) che sia soluzione del problema dovrà risultare: $\dfrac{\mathrm{d}w}{\mathrm{d}x} = \dfrac{\partial f}{\partial x} \cdot \dfrac{\mathrm{d}x}{\mathrm{d}x} + \dfrac{\partial f}{\partial y} \cdot \dfrac{\mathrm{d}y}{\mathrm{d}x} + \dfrac{\partial f}{\partial z} \cdot \dfrac{\mathrm{d}z}{\mathrm{d}x} = 0$, ovvero:
$f'_x + f'_y \cdot y'(x) + f'_z \cdot z'(x) = 0$. Ma dal sistema $\begin{cases} g(x,y,z) = 0 \\ h(x,y,z) = 0 \end{cases}$ ricaviamo, per il Teorema del Dini, essendo per ipotesi $\begin{vmatrix} g'_y & g'_z \\ h'_y & h'_z \end{vmatrix} \neq 0$:

$y'(x) = -\dfrac{\begin{vmatrix} g'_x & g'_z \\ h'_x & h'_z \end{vmatrix}}{\begin{vmatrix} g'_y & g'_z \\ h'_y & h'_z \end{vmatrix}}$ e $z'(x) = -\dfrac{\begin{vmatrix} g'_y & g'_x \\ h'_y & h'_x \end{vmatrix}}{\begin{vmatrix} g'_y & g'_z \\ h'_y & h'_z \end{vmatrix}}$, dalle quali, sostituendo e scambiando tra loro le colonne del numeratore di $z'(x)$, otteniamo, fatto il denominatore comune:
$f'_x \cdot \begin{vmatrix} g'_y & g'_z \\ h'_y & h'_z \end{vmatrix} - f'_y \cdot \begin{vmatrix} g'_x & g'_z \\ h'_x & h'_z \end{vmatrix} + f'_z \cdot \begin{vmatrix} g'_x & g'_y \\ h'_x & h'_y \end{vmatrix} = 0$, ovvero:

$\begin{vmatrix} f'_x & f'_y & f'_z \\ g'_x & g'_y & g'_z \\ h'_x & h'_y & h'_z \end{vmatrix} = \begin{vmatrix} \nabla f \\ \nabla g \\ \nabla h \end{vmatrix} = 0$. Le tre righe, ovvero i gradienti, risultano quindi linearmente dipendenti, e quindi avremo: $\nabla f = \lambda_1 \nabla g + \lambda_2 \nabla h$, ovvero $\nabla(f - \lambda_1 g - \lambda_2 h) = 0$.

Dal punto di vista operativo l'aumento dei vincoli comporta che la funzione Lagrangiana venga espressa come differenza tra la funzione obiettivo ed una combinazione lineare dei vincoli, ciascuno quindi con il suo moltiplicatore.

L'aumento nel numero dei vincoli non porta cambiamenti dal punto di vista operativo: in un punto di massimo o minimo vincolato è ancora necessario l'annullamento del gradiente della funzione Lagrangiana, che però dà luogo questa volta al sistema:

$$\begin{cases} \Lambda'_x = f'_x - \lambda_1 g'_x - \lambda_2 h'_x = 0 \\ \Lambda'_y = f'_y - \lambda_1 g'_y - \lambda_2 h'_y = 0 \\ \Lambda'_z = f'_z - \lambda_1 g'_z - \lambda_2 h'_z = 0 \\ \Lambda'_{\lambda_1} = -g(x,y,z) = 0 \\ \Lambda'_{\lambda_2} = -h(x,y,z) = 0 \end{cases}$$

, di cinque equazioni nelle cinque incognite x, y, z, λ_1 e λ_2.

Prendendo spunto dai tre casi introduttivi presentati, passiamo alla formulazione delle condizioni del I ordine nel caso generale di un problema di estremi con vincoli di uguaglianza.

Sia $\mathcal{E}$ l'insieme dei punti che soddisfano il sistema $\begin{cases} g_1(x_1, x_2, ..., x_n) = 0 \\ g_2(x_1, x_2, ..., x_n) = 0 \\ \dots\dots\dots \\ g_m(x_1, x_2, ..., x_n) = 0 \end{cases}$.

Vale il seguente:

Teorema 28 : Siano f e g_i, $1 \le i \le m < n$, funzioni $\mathbb{R}^n \to \mathbb{R}$, differenziabili in $\mathcal{E} \subset \mathbb{R}^n$, sia poi $\text{Car}\left(\dfrac{\partial(g_1, g_2, ..., g_m)}{\partial(x_1, x_2, ..., x_n)}\right) = m$ in ogni punto di $\mathcal{E}$, e sia $(x_1^0, x_2^0, ..., x_n^0)$ un punto soluzione del problema $\begin{cases} \text{Max/min } f(x_1, x_2, ..., x_n) \\ \text{s.v.} \begin{cases} g_1(x_1, x_2, ..., x_n) = 0 \\ g_2(x_1, x_2, ..., x_n) = 0 \\ \dots\dots\dots \\ g_m(x_1, x_2, ..., x_n) = 0 \end{cases} \end{cases}$.

Allora, posto $\Lambda(x_1, x_2, ..., x_n, \lambda_1, \lambda_2, ..., \lambda_m) = f(x_1, x_2, ..., x_n) - \sum_{i=1}^{m} \lambda_i \cdot g_i(x_1, x_2, ..., x_n)$, esiste un vettore $(\lambda_1^0, \lambda_2^0, ..., \lambda_m^0)$ per il quale $\nabla\Lambda\left(x_1^0, x_2^0, ..., x_n^0, \lambda_1^0, \lambda_2^0, ..., \lambda_m^0\right) = \mathbb{O}$.

Il punto $\left(x_1^0, x_2^0, ..., x_n^0, \lambda_1^0, \lambda_2^0, ..., \lambda_m^0\right) = (\mathbb{X}_0, \lambda_0)$ deve cioè essere una soluzione del sistema di $n + m$ equazioni in $n + m$ incognite:

$$\begin{cases} \dfrac{\partial\Lambda}{\partial x_i} = \dfrac{\partial f}{\partial x_i} - \displaystyle\sum_{j=1}^{m} \lambda_j \cdot \dfrac{\partial g_j}{\partial x_i} = 0 & 1 \le i \le n \\ \dfrac{\partial\Lambda}{\partial\lambda_i} = -g_i(x_1, x_2, ..., x_n) = 0 & 1 \le i \le m \end{cases}.$$

E' bene ribadire come queste condizioni del I ordine, al pari di quella, $\nabla f = \mathbb{O}$, per massimi e minimi liberi, siano necessarie, e non sufficienti, per garantire la natura degli estremi. Tra i punti che risultano stazionari per la funzione Lagrangiana ci sono, oltre gli eventuali massimi e minimi, anche punti di flesso o di sella. Avremo punti di flesso quando $m = n - 1$, in quanto i vincoli permettono di esprimere $n - 1$ variabili come funzioni dell'unica rimanente, e quindi il problema consiste nella ricerca di estremi per una funzione di una sola variabile; avremo invece punti di sella se $n - m > 1$, in quanto le variabili che rimangono indipendenti sono in numero maggiore di 1.

CONDIZIONI DEL SECONDO ORDINE PER ESTREMI VINCOLATI

Passiamo ad esaminare le condizioni del II ordine che, come per massimi e minimi liberi, risulteranno condizioni sufficienti e non necessarie per garantire estremi vincolati. Ripercorriamo i casi a dimensione più semplice per poter giustificare (non dimostrare) la formulazione delle condizioni nel caso più generale.

Ritorniamo al problema $\begin{cases} \text{Max/min } z = f(x,y) \\ \text{s.v.} : g(x,y) = 0 \end{cases}$. Vediamo come studiare il segno del differenziale secondo in presenza di un vincolo. Essendo $\mathrm{d}z = f'_x\,\mathrm{d}x + f'_y\,\mathrm{d}y$, ed ipotizzando che sia $y = y(x)$ per il vincolo $g(x,y) = 0$, il differenziale $\mathrm{d}y$ risulta funzione di x e di y, per cui avremo:

$$\mathrm{d}^2 z = \left(f''_{xx}\,\mathrm{d}x + f''_{yx}\,\mathrm{d}y + \frac{\partial(\mathrm{d}y)}{\partial x} \cdot f'_y \right) \cdot \mathrm{d}x + \left(f''_{yx}\,\mathrm{d}x + f''_{yy}\,\mathrm{d}y + \frac{\partial(\mathrm{d}y)}{\partial y} \cdot f'_y \right) \cdot \mathrm{d}y =$$

$$\mathrm{d}^2 z = f''_{xx}\,(\mathrm{d}x)^2 + 2 f''_{xy}\,\mathrm{d}x\,\mathrm{d}y + f''_{yy}\,(\mathrm{d}y)^2 + f'_y \cdot \left(\frac{\partial(\mathrm{d}y)}{\partial x} \cdot \mathrm{d}x + \frac{\partial(\mathrm{d}y)}{\partial y} \cdot \mathrm{d}y \right) =$$

$$\mathrm{d}^2 z = \mathrm{d}^2 f + f'_y \cdot \mathrm{d}(\mathrm{d}y) = \mathrm{d}^2 f + f'_y \cdot \mathrm{d}^2 y\,.$$

Dalla $z = g(x,y) = 0$, supponendo g differenziabile due volte, in maniera analoga si ha: $\mathrm{d}^2 z = \mathrm{d}^2 g + g'_y \cdot \mathrm{d}^2 y = 0$, in quanto sul vincolo z è costante, da cui ricaviamo:

$\mathrm{d}^2 y = -\dfrac{1}{g'_y} \cdot \mathrm{d}^2 g$. Ricordando che $\nabla f = \lambda \nabla g$, da cui $f'_y = \lambda\, g'_y$ e quindi $\dfrac{f'_y}{g'_y} = \lambda$, sostituendo nella precedente, otteniamo:

$$\mathrm{d}^2 z = \mathrm{d}^2 f + f'_y \cdot \mathrm{d}^2 y = \mathrm{d}^2 f - \frac{f'_y}{g'_y} \cdot \mathrm{d}^2 g = \mathrm{d}^2 f - \lambda \cdot \mathrm{d}^2 g = \mathrm{d}^2 (f - \lambda \cdot g) = \mathrm{d}^2 \Lambda\,.$$

Così come le condizioni del I ordine sono esprimibili mediante l'annullamento di un gradiente (di Λ e non di f, però), così anche quelle del II ordine sono esprimibili mediante lo studio del segno di un differenziale secondo, anche questa volta non di f ma della Lagrangiana Λ.

Non tutto però ricalca le metodologie già svolte per i massimi e minimi liberi. Infatti:

$\mathrm{d}^2 \Lambda(x,y) = \Lambda''_{xx}\,(\mathrm{d}x)^2 + 2\,\Lambda''_{xy}\,\mathrm{d}x\,\mathrm{d}y + \Lambda''_{yy}\,(\mathrm{d}y)^2$. Dal vincolo $g(x,y) = 0$ possiamo ricavare $\mathrm{d}y = -\dfrac{g'_x}{g'_y} \cdot \mathrm{d}x$, da cui, sostituendo, si ha:

$$\mathrm{d}^2 \Lambda(x,y) = \Lambda''_{xx}\,(\mathrm{d}x)^2 - 2\,\Lambda''_{xy}\,\frac{g'_x}{g'_y} \cdot (\mathrm{d}x)^2 + \Lambda''_{yy} \left(\frac{g'_x}{g'_y} \right)^2 \cdot (\mathrm{d}x)^2 =$$

$\mathrm{d}^2 \Lambda(x,y) = \left[\Lambda''_{xx}\,(g'_y)^2 - 2\,\Lambda''_{xy}\,g'_x\,g'_y + \Lambda''_{yy}\,(g'_x)^2 \right] \cdot \dfrac{(\mathrm{d}x)^2}{(g'_y)^2}$. Ma si vede che:

$\Lambda''_{xx}\,(g'_y)^2 - 2\,\Lambda''_{xy}\,g'_x\,g'_y + \Lambda''_{yy}\,(g'_x)^2 = -\begin{vmatrix} 0 & g'_x & g'_y \\ g'_x & \Lambda''_{xx} & \Lambda''_{xy} \\ g'_y & \Lambda''_{xy} & \Lambda''_{yy} \end{vmatrix}$ ed inoltre:

$\begin{vmatrix} 0 & g'_x & g'_y \\ g'_x & \Lambda''_{xx} & \Lambda''_{xy} \\ g'_y & \Lambda''_{xy} & \Lambda''_{yy} \end{vmatrix} = \begin{vmatrix} \Lambda''_{\lambda\lambda} & \Lambda''_{\lambda x} & \Lambda''_{\lambda y} \\ \Lambda''_{\lambda x} & \Lambda''_{xx} & \Lambda''_{xy} \\ \Lambda''_{\lambda y} & \Lambda''_{xy} & \Lambda''_{yy} \end{vmatrix} = \left| \overline{\mathbb{H}}(\Lambda(x,y,\lambda)) \right|$. La matrice $\overline{\mathbb{H}}$ prende il nome di "matrice Hessiana orlata"; il cosiddetto orlo è costituito dalla prima riga e dalla prima colonna che sono, tolto lo zero iniziale, il gradiente del vincolo $g(x,y)$. L'elemento nullo iniziale corrisponde a $\Lambda''_{\lambda\lambda}$, così come i restanti elementi dell'orlo sono l'opposto delle derivate seconde della Lagrangiana fatte rispetto a λ e poi rispetto a x oppure a y; per le proprietà del determinante, cambiando segno a due linee il determinante resta invariato.

Il segno del differenziale $\mathrm{d}^2\Lambda$ non viene studiato al variare dei due incrementi indipendenti $\mathrm{d}x$ e $\mathrm{d}y$; la sostituzione $\mathrm{d}y = -\dfrac{g'_x}{g'_y}\cdot \mathrm{d}x$ lascia indipendente il solo incremento $\mathrm{d}x$, coerentemente con il fatto che la presenza di un vincolo lascia indipendente una sola tra le variabili x e y.
Questo fatto può essere interpretato geometricamente dicendo che si deve indagare il $\mathrm{d}^2\Lambda$ con incrementi nelle sole direzioni tangenziali rispetto al vincolo.

Dato che $\mathrm{d}^2\Lambda = -\left|\overline{\mathbb{H}}(\Lambda(x,y,\lambda))\right|$, vale il seguente:
Teorema 29 : Sia (x_0, y_0, λ_0) soluzione del sistema $\nabla\Lambda(x,y,\lambda) = 0$. Allora
- $\left(\left|\overline{\mathbb{H}}(\Lambda(x_0,y_0,\lambda_0))\right| < 0 \Leftrightarrow \mathrm{d}^2\Lambda > 0\right) \Rightarrow (x_0,y_0)$ è punto di minimo vincolato;
- $\left(\left|\overline{\mathbb{H}}(\Lambda(x_0,y_0,\lambda_0))\right| > 0 \Leftrightarrow \mathrm{d}^2\Lambda < 0\right) \Rightarrow (x_0,y_0)$ è punto di massimo vincolato.

Nulla si può concludere se risulta $\left|\overline{\mathbb{H}}(\Lambda(x_0,y_0,\lambda_0))\right| = 0$, in quanto (x_0,y_0) può essere punto di massimo, di minimo oppure di flesso.
Occorrono in questo caso analisi di tipo diverso per stabilire la natura del punto.

Osservazione 1) Anche se la matrice Hessiana orlata è di ordine 3, solo il minore di guida di ordine 3, cioè il determinante della matrice stessa, va preso in considerazione: inutile esaminare il minore di ordine 1, che è sempre nullo, ed inutile esaminare quello di ordine 2, in quanto $\begin{vmatrix} 0 & g'_x \\ g'_x & \Lambda''_{xx} \end{vmatrix} = -\left(g'_x\right)^2 < 0$, quindi sempre negativo.
Osservazione 2) Si sono esaminati solo i minori di guida di Nord-Ovest; non ci sono condizioni esprimibili mediante altre sequenze di minori di guida nè esistono condizioni basate sugli autovalori della Matrice Hessiana orlata.

Se passiamo ora a studiare le condizioni del II ordine per il problema $\begin{cases} \text{Max/min } f(x,y,z) \\ \text{s.v. : } g(x,y,z) = 0 \end{cases}$
saremo ancora ricondotti allo studio del $\mathrm{d}^2\Lambda$. Dal vincolo $g(x,y,z) = 0$ possiamo, per ipotesi, ricavare $z = z(x,y)$; quindi stiamo cercando gli estremi di una funzione di due variabili, e lo studio del segno del $\mathrm{d}^2\Lambda$ sarà basato su due incrementi, $\mathrm{d}x$ e $\mathrm{d}y$.
Lo sviluppo dei calcoli, che qui per brevità vengono omessi, conduce allo studio del segno di una quantità il cui opposto è collegabile alla matrice Hessiana orlata:
$\left\|\begin{matrix} 0 & g'_x & g'_y & g'_z \\ g'_x & \Lambda''_{xx} & \Lambda''_{xy} & \Lambda''_{xz} \\ g'_y & \Lambda''_{xy} & \Lambda''_{yy} & \Lambda''_{yz} \\ g'_z & \Lambda''_{xz} & \Lambda''_{yz} & \Lambda''_{zz} \end{matrix}\right\| = \overline{\mathbb{H}}(\Lambda(x,y,z,\lambda))$. Analogamente a quanto visto nel caso di estremi liberi per una funzione di due variabili, occorre studiare il segno di due minori di guida della sequenza di Nord-Ovest. Come prima, il minore di ordine 1 e quello di ordine 2 non servono. Va studiato il segno dei soli minori $\left|\overline{\mathbb{H}}_3\right|$ e $\left|\overline{\mathbb{H}}_4\right|$ e vale il
Teorema 30 : Sia $(x_0, y_0, z_0, \lambda_0)$ soluzione del sistema $\nabla\Lambda(x,y,z,\lambda) = 0$. Allora
- $\left(\left(\left|\overline{\mathbb{H}}_3\right| < 0 \text{ e } \left|\overline{\mathbb{H}}_4\right| < 0\right) \Leftrightarrow \mathrm{d}^2\Lambda > 0\right) \Rightarrow (x_0,y_0,z_0)$ è punto di minimo vincolato;
- $\left(\left(\left|\overline{\mathbb{H}}_3\right| > 0 \text{ e } \left|\overline{\mathbb{H}}_4\right| < 0\right) \Leftrightarrow \mathrm{d}^2\Lambda < 0\right) \Rightarrow (x_0,y_0,z_0)$ è punto di massimo vincolato.

Se $\left(\left|\overline{\mathbb{H}}_3\right| < 0 \text{ e } \left|\overline{\mathbb{H}}_4\right| > 0\right)$ oppure se $\left(\left|\overline{\mathbb{H}}_3\right| > 0 \text{ e } \left|\overline{\mathbb{H}}_4\right| > 0\right)$ il punto è sicuramente di sella.
Nulla si può concludere se risulta $\left|\overline{\mathbb{H}}_3\right| = 0$ oppure $\left|\overline{\mathbb{H}}_4\right| = 0$, sempre comunque rispettando le precedenti sequenze di segni, in quanto (x_0,y_0,z_0) può essere punto di massimo e/o di minimo, oppure di sella.
Occorrono in questo caso analisi di tipo diverso per stabilire la natura del punto.

Infine, nel problema: $\begin{cases} \text{Max/min } f(x,y,z) \\ \text{s.v. : } \begin{cases} g(x,y,z)=0 \\ h(x,y,z)=0 \end{cases} \end{cases}$ siamo ancora ricondotti allo studio del $d^2\Lambda$, ma dai vincoli $\begin{cases} g(x,y,z)=0 \\ h(x,y,z)=0 \end{cases}$ possiamo, per ipotesi, ricavare $\begin{cases} y=y(x) \\ z=z(x) \end{cases}$; quindi cerchiamo gli estremi di una funzione di una variabile, e lo studio del segno del $d^2\Lambda$ sarà basato su di un solo incremento, dx. Con la presenza di due vincoli non ci sono cambiamenti di segno, ed il segno del $d^2\Lambda$ corrisponde a quello del determinante della matrice Hessiana orlata:
$\left\|\begin{matrix} 0 & 0 & g'_x & g'_y & g'_z \\ 0 & 0 & h'_x & h'_y & h'_z \\ g'_x & h'_x & \Lambda''_{xx} & \Lambda''_{xy} & \Lambda''_{xz} \\ g'_y & h'_y & \Lambda''_{xy} & \Lambda''_{yy} & \Lambda''_{yz} \\ g'_z & h'_z & \Lambda''_{xz} & \Lambda''_{yz} & \Lambda''_{zz} \end{matrix}\right\| = \overline{\mathbb{H}}(\Lambda(x,y,z,\lambda_1,\lambda_2))$. Il minore di ordine 2 ad elementi nulli dipende dal fatto che $\Lambda = f(x,y,z) - \lambda_1\, g(x,y,z) - \lambda_2\, h(x,y,z)$, e quindi sono nulle le quattro derivate seconde di Λ fatte rispetto ai moltiplicatori λ_1 e λ_2.
Si può ora vedere che nei minori di guida (sempre e solo di Nord-Ovest) i primi quattro (e quattro è il doppio di due, il numero dei vincoli) o sono nulli $\left(|\overline{\mathbb{H}}_1| = 0,\ |\overline{\mathbb{H}}_2| = 0,\ |\overline{\mathbb{H}}_3| = 0\right)$ o hanno segno costante $\left(|\overline{\mathbb{H}}_4| < 0\right)$, per cui il segno di $d^2\Lambda$ coincide con quello di $|\overline{\mathbb{H}}_5|$, il determinante della matrice Hessiana orlata. Vale quindi il
Teorema 31 : Sia $(x_0, y_0, z_0, \lambda_1^0, \lambda_2^0)$ soluzione del sistema $\nabla\Lambda(x,y,z,\lambda_1,\lambda_2) = 0$. Allora
- $\left(|\overline{\mathbb{H}}_5(\Lambda(x_0,y_0,z_0,\lambda_1^0,\lambda_2^0))| > 0 \Leftrightarrow d^2\Lambda > 0\right) \Rightarrow (x_0,y_0,z_0)$ è punto di minimo vincolato;
- $\left(|\overline{\mathbb{H}}_5(\Lambda(x_0,y_0,z_0,\lambda_1^0,\lambda_2^0))| < 0 \Leftrightarrow d^2\Lambda < 0\right) \Rightarrow (x_0,y_0,z_0)$ è punto di massimo vincolato.

Nulla si può concludere se risulta $|\overline{\mathbb{H}}_5(\Lambda(x_0,y_0,z_0,\lambda_1^0,\lambda_2^0))| = 0$, in quanto (x_0,y_0,z_0) può essere punto di massimo, di minimo oppure di flesso.
Occorrono al solito analisi di tipo diverso per stabilire la natura del punto.

Da quanto visto nei tre casi trattati, emergono le seguenti considerazioni:
- occorre costruire la matrice Hessiana orlata, costituita dalle derivate seconde della Lagrangiana fatte sia rispetto ai moltiplicatori che rispetto alle variabili;
- si devono esaminare i segni solo dei minori di guida di Nord-Ovest, scartandone un numero di iniziali pari al doppio del numero dei vincoli; il numero dei minori di guida di cui studiare il segno coincide con il numero delle variabili che rimangono indipendenti, sottratte quelle dipendenti implicitamente dai vincoli;
- per i minori di guida rimanenti valgono, per individuare un minimo o un massimo, due sequenze di segni che coincidono con quelle degli estremi liberi se il numero dei vincoli è pari, oppure, nel caso di un numero di vincoli dispari, hanno segni opposti alle precedenti;
- ogni sequenza che non sia una delle due descritte porta a concludere che il punto è di flesso o di sella;
- se tra i minori di guida significativi ve ne fosse anche uno solo uguale a zero, nulla si può concludere sulla natura del punto.

Passiamo quindi al caso generale, cioè al problema $\begin{cases} \text{Max/min } f(x_1,x_2,...,x_n) \\ \text{s.v. } \begin{cases} g_1(x_1,x_2,...,x_n)=0 \\ g_2(x_1,x_2,...,x_n)=0 \\ \dots\dots\dots\dots\dots\dots \\ g_m(x_1,x_2,...,x_n)=0 \end{cases} \end{cases}$.
La matrice Hessiana orlata, quadrata di ordine $m+n$, risulta uguale a:

$$\overline{\mathbb{H}}(\Lambda(x_1, x_2, ..., x_n, \lambda_1, \lambda_2, ..., \lambda_m)) = \left\| \begin{matrix} 0 & ... & 0 & g'_{11} & ... & g'_{1n} \\ ... & ... & ... & ... & ... & ... \\ 0 & ... & 0 & g'_{m1} & ... & g'_{mn} \\ g'_{11} & ... & g'_{m1} & \Lambda''_{11} & ... & \Lambda''_{1n} \\ ... & ... & ... & ... & ... & ... \\ g'_{1n} & ... & g'_{mn} & \Lambda''_{n1} & ... & \Lambda''_{nn} \end{matrix} \right\|,$$

dove, per brevità, si è posto $g'_{ij} = \dfrac{\partial g_i}{\partial x_j}$ e $\Lambda''_{ij} = \dfrac{\partial^2 \Lambda}{\partial x_i \, \partial x_j}$, e può essere così rappresentata:

$$\overline{\mathbb{H}}(\Lambda(x_1, x_2, ..., x_n, \lambda_1, \lambda_2, ..., \lambda_m)) = \left\| \begin{matrix} \mathbb{O} & \dfrac{\partial(g_1, g_2, ..., g_m)}{\partial(x_1, x_2, ..., x_n)} \\ \left(\dfrac{\partial(g_1, g_2, ..., g_m)}{\partial(x_1, x_2, ..., x_n)} \right)^{\mathrm{T}} & \mathbb{H}(\Lambda(x_1, x_2, ..., x_n)) \end{matrix} \right\|.$$

Come si vede, la matrice Hessiana orlata può essere scomposta in quattro blocchi. Il blocco in alto a sinistra è una matrice nulla $m \times m$, costituita dalle derivate della Lagrangiana fatte ambedue le volte rispetto ai moltiplicatori e quindi nulle. In alto a destra e in basso a sinistra ci sono le derivate della Lagrangiana ottenute derivando una volta rispetto ad un moltiplicatore ed una volta rispetto ad una variabile; anche queste costituiscono l'orlo, completato con lo Jacobiano dei vincoli, matrice $m \times n$, in alto come righe e sulla sinistra, trasposto, come colonne. Il blocco rimanente $n \times n$, in basso a destra, è l'Hessiana della Lagrangiana fatta rispetto alle variabili x_i. Vale allora il seguente

Teorema 32 : Sia $(x_1^0, x_2^0, ..., x_n^0, \lambda_1^0, \lambda_2^0, ..., \lambda_m^0)$ soluzione del sistema :
$\nabla\Lambda(x_1, x_2, ..., x_n, \lambda_1, \lambda_2, ..., \lambda_m) = 0$. Allora

- $\left(\begin{cases} (-1)^m \cdot \left|\overline{\mathbb{H}}_i(\Lambda)\right| > 0 \\ 2m+1 \le i \le m+n \end{cases} \Leftrightarrow \mathrm{d}^2\Lambda > 0 \right) \Rightarrow (x_1^0, x_2^0, ..., x_n^0)$ è punto di minimo vincolato;
- $\left(\begin{cases} (-1)^{m+i} \cdot \left|\overline{\mathbb{H}}_i(\Lambda)\right| > 0 \\ 2m+1 \le i \le m+n \end{cases} \Leftrightarrow \mathrm{d}^2\Lambda < 0 \right) \Rightarrow (x_1^0, x_2^0, ..., x_n^0)$ è punto di massimo vincolato.

Nulla si può concludere se anche uno solo dei minori di guida significativi fosse nullo.

Se i vincoli sono in numero pari, zero compreso, una sequenza di segni tutti positivi indica un punto di minimo, mentre una sequenza di segni alterni, a partire dal negativo, indica un punto di massimo.
Se i vincoli sono in numero dispari una sequenza di segni tutti negativi indica un punto di minimo, mentre una sequenza di segni alterni, a partire dal positivo, indica un punto di massimo.
Ogni sequenza che non ricada in quelle descritte indica un punto di flesso o un punto di sella.
La condizione $2m+1 \le i \le m+n$ esprime il fatto che si devono considerare solo i minori di Nord-Ovest significativi, ottenuti scartando i primi $2m$, un numero pari al doppio dei vincoli, iniziando quindi da quello di ordine $2m+1$. Cambiare o non cambiare i segni della sequenza dei minori di guida significativi è il ruolo del fattore $(-1)^m$.
Dato che $m < n$, ci saranno $(m+n) - 2m = n - m$ minori di guida significativi; se $m = 1$, ovvero se c'è un solo vincolo, ci sarà da esaminare il segno di $n-1$ minori di guida; se invece $m = n-1$, che è il massimo numero possibile di vincoli, ci sarà da esaminare il segno di un solo minore, che sarà il determinante della matrice Hessiana orlata.
Dato che ogni vincolo, esplicitamente o implicitamente, rende una variabile dipendente, il numero di minori di guida di cui studiare il segno coincide sempre con il numero delle variabili che rimangono indipendenti.
E' opportuno rimarcare infine che non esiste una condizione del II ordine basata sugli autovalori della matrice Hessiana orlata.

Esempio 79 : Studiamo il problema $\begin{cases} \text{Max/min } f(x,y) = 2x^3 + 3y \\ \text{s.v. : } g(x,y) = x^2 + y^2 = 1 \end{cases}$.

Sia $f(x,y)$ che $g(x,y)$ sono polinomi, e quindi funzioni differenziabili infinite volte. Risulta poi $\nabla g(x,y) = (2x, 2y) = \mathbb{O} \Leftrightarrow (x,y) = (0,0)$; ma $(0,0)$ non soddisfa l'equazione del vincolo, e quindi sono soddisfatte le ipotesi del Teorema 27. Formiamo allora la Lagrangiana ed avremo: $\Lambda(x,y,\lambda) = 2x^3 + 3y - \lambda\left(x^2 + y^2 - 1\right)$ e imponiamo $\nabla\Lambda = 0$, ottenendo:

$$\begin{cases} \Lambda'_x = 0 \Rightarrow 6x^2 - 2\lambda x = 2x\,(3x - \lambda) = 0 \\ \Lambda'_y = 0 \Rightarrow 3 - 2\lambda y = 0 \\ \Lambda'_\lambda = 0 \Rightarrow x^2 + y^2 = 1 \end{cases} \Rightarrow \begin{cases} x = 0 \\ \lambda = \dfrac{3}{2y} \\ y = \pm 1 \end{cases} \Rightarrow \begin{cases} x = 0 \\ \lambda = \frac{3}{2} \\ y = 1 \end{cases} \text{ e } \begin{cases} x = 0 \\ \lambda = -\frac{3}{2} \\ y = -1 \end{cases}$$

oppure $\begin{cases} x = \dfrac{\lambda}{3} \\ y = \dfrac{3}{2\lambda} \\ \dfrac{\lambda^2}{9} + \dfrac{9}{4\lambda^2} = 1 \end{cases} \Rightarrow \begin{cases} x = \dfrac{\lambda}{3} \\ y = \dfrac{3}{2\lambda} \\ \dfrac{4\lambda^4 + 81 - 36\lambda^2}{36\lambda^2} = 0 \end{cases} \Rightarrow \begin{cases} x = \dfrac{\lambda}{3} \\ y = \dfrac{3}{2\lambda} \\ \left(2\lambda^2 - 9\right)^2 = 0 \end{cases}$ da cui si ha:

$\begin{cases} \lambda = \frac{3}{\sqrt{2}} \\ x = \frac{1}{\sqrt{2}} \\ y = \frac{1}{\sqrt{2}} \end{cases}$ e $\begin{cases} \lambda = -\frac{3}{\sqrt{2}} \\ x = -\frac{1}{\sqrt{2}} \\ y = -\frac{1}{\sqrt{2}} \end{cases}$. Si sono quindi ottenuti quattro punti: $\mathrm{P}_1 = \left(0, 1, \dfrac{3}{2}\right)$, $\mathrm{P}_2 = \left(0, -1, -\dfrac{3}{2}\right)$, $\mathrm{P}_3 = \left(\dfrac{1}{\sqrt{2}}, \dfrac{1}{\sqrt{2}}, \dfrac{3}{\sqrt{2}}\right)$ e $\mathrm{P}_4 = \left(-\dfrac{1}{\sqrt{2}}, -\dfrac{1}{\sqrt{2}}, -\dfrac{3}{\sqrt{2}}\right)$.

Passiamo alle condizioni del II ordine. Formiamo l'Hessiana orlata ed avremo:

$\overline{\mathbb{H}}(\Lambda(x,y,\lambda)) = \left\|\begin{matrix} 0 & 2x & 2y \\ 2x & 12x - 2\lambda & 0 \\ 2y & 0 & -2\lambda \end{matrix}\right\|$ e studiamola nei punti trovati. Sarà:

$\left|\overline{\mathbb{H}}(\Lambda(\mathrm{P}_1))\right| = \begin{vmatrix} 0 & 0 & 2 \\ 0 & -3 & 0 \\ 2 & 0 & -3 \end{vmatrix} = 2 \cdot 6 = 12 > 0$, per cui P_1 è un punto di massimo;

$\left|\overline{\mathbb{H}}(\Lambda(\mathrm{P}_2))\right| = \begin{vmatrix} 0 & 0 & -2 \\ 0 & 3 & 0 \\ -2 & 0 & 3 \end{vmatrix} = -2 \cdot 6 = -12 < 0$, per cui P_2 è un punto di minimo;

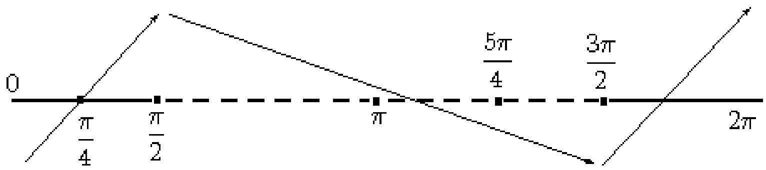
$\left|\overline{\mathbb{H}}(\Lambda(\mathrm{P}_3))\right| = \begin{vmatrix} 0 & \sqrt{2} & \sqrt{2} \\ \sqrt{2} & 3\sqrt{2} & 0 \\ \sqrt{2} & 0 & -3\sqrt{2} \end{vmatrix} = \begin{vmatrix} 0 & \sqrt{2} & \sqrt{2} \\ 0 & 3\sqrt{2} & 3\sqrt{2} \\ \sqrt{2} & 0 & -3\sqrt{2} \end{vmatrix} = 0$, per cui non si può decidere nulla;

$\left|\overline{\mathbb{H}}(\Lambda(\mathrm{P}_4))\right| = \begin{vmatrix} 0 & -\sqrt{2} & -\sqrt{2} \\ -\sqrt{2} & -3\sqrt{2} & 0 \\ -\sqrt{2} & 0 & 3\sqrt{2} \end{vmatrix} = \begin{vmatrix} 0 & -\sqrt{2} & -\sqrt{2} \\ 0 & -3\sqrt{2} & -3\sqrt{2} \\ -\sqrt{2} & 0 & 3\sqrt{2} \end{vmatrix} = 0$, per cui non si può decidere nulla.

Visto che il vincolo rappresenta la circonferenza trigonometrica, per risolvere i due casi rimanenti proviamo la sostituzione $\begin{cases} x = \cos t \\ y = \operatorname{sen} t \end{cases}$ ed avremo: $f(t) = 2\cos^3 t + 3\operatorname{sen} t$. Risulta:

$f'(t) = 6\cos^2 t \cdot (-\operatorname{sen} t) + 3\cos t = 3\cos t \cdot (1 - 2\operatorname{sen} t\cos t) = 3\cos t \cdot (1 - \operatorname{sen} 2t) \geq 0$ quando $\cos t \geq 0$ dato che $\operatorname{sen} 2t \leq 1$, $\forall t$.

Quindi $f'(t) \geq 0$ per $0 \leq t \leq \dfrac{\pi}{2}$ e per $\dfrac{3\pi}{2} \leq t \leq 2\pi$.



Il punto $\left(\dfrac{1}{\sqrt{2}}, \dfrac{1}{\sqrt{2}}\right)$ corrisponde a $t = \dfrac{\pi}{4}$, mentre il punto $\left(-\dfrac{1}{\sqrt{2}}, -\dfrac{1}{\sqrt{2}}\right)$ corrisponde a $t = \dfrac{5\pi}{4}$ e, come si vede dall'analisi della crescenza, sono ambedue punti di flesso e non di massimo o di minimo. Dato che i punti della circonferenza formano un insieme limitato e chiuso, e dato che la funzione $f(x,y)$ è continua, i punti P_1 e P_2 sono rispettivamente il punto di massimo e quello di minimo, non solo relativi, ma pure assoluti.

Esempio 80 : Studiamo il problema $\begin{cases} \text{Max/min } f(x,y,z) = x + y - 2z \\ \text{s.v. : } g(x,y,z) = x^2 + y^2 - z = 0 \end{cases}$.

Sia $f(x,y,z)$ che $g(x,y,z)$ sono polinomi, e quindi funzioni differenziabili infinite volte. Risulta poi $\nabla g(x,y,z) = (2x, 2y, -1) \neq \mathbb{O} \ \ \forall (x,y,z)$; quindi sono soddisfatte le ipotesi del Teorema 27. Formiamo allora la Lagrangiana ed avremo:

$\Lambda(x,y,z,\lambda) = x + y - 2z - \lambda\left(x^2 + y^2 - z\right)$ e imponiamo $\nabla\Lambda = 0$, ottenendo:

$$\begin{cases} \Lambda'_x = 0 \Rightarrow 1 - 2\lambda x = 0 \\ \Lambda'_y = 0 \Rightarrow 1 - 2\lambda y = 0 \\ \Lambda'_z = 0 \Rightarrow -2 + \lambda = 0 \\ \Lambda'_\lambda = 0 \Rightarrow z - x^2 - y^2 = 0 \end{cases} \Rightarrow \begin{cases} x = \dfrac{1}{2\lambda} = \dfrac{1}{4} \\ y = \dfrac{1}{2\lambda} = \dfrac{1}{4} \\ \lambda = 2 \\ z = x^2 + y^2 = \frac{1}{8} \end{cases}$$

. C'è quindi un solo punto stazionario per la Lagrangiana: $P_0 = \left(\dfrac{1}{4}, \dfrac{1}{4}, \dfrac{1}{8}, 2\right)$. Sarà poi:

$$\overline{\mathbb{H}}(\Lambda) = \begin{Vmatrix} 0 & 2x & 2y & -1 \\ 2x & -2\lambda & 0 & 0 \\ 2y & 0 & -2\lambda & 0 \\ -1 & 0 & 0 & 0 \end{Vmatrix} \text{ e } \overline{\mathbb{H}}(\Lambda(P_0)) = \begin{Vmatrix} 0 & \frac{1}{2} & \frac{1}{2} & -1 \\ \frac{1}{2} & -4 & 0 & 0 \\ \frac{1}{2} & 0 & -4 & 0 \\ -1 & 0 & 0 & 0 \end{Vmatrix}$$

, della quale dobbiamo calcolare due minori: $\left|\overline{\mathbb{H}}_3\right|$ e $\left|\overline{\mathbb{H}}_4\right|$, ed avremo:

$$\left|\overline{\mathbb{H}}_3(\Lambda(P_0))\right| = \begin{vmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -4 & 0 \\ \frac{1}{2} & 0 & -4 \end{vmatrix} = \begin{vmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & -4 & 4 \\ \frac{1}{2} & 0 & -4 \end{vmatrix} = \frac{1}{2} \cdot (2 + 2) = 2 > 0, \text{ e}$$

$$\left|\overline{\mathbb{H}}_4(\Lambda(P_0))\right| = \begin{vmatrix} 0 & \frac{1}{2} & \frac{1}{2} & -1 \\ \frac{1}{2} & -4 & 0 & 0 \\ \frac{1}{2} & 0 & -4 & 0 \\ -1 & 0 & 0 & 0 \end{vmatrix} = 1 \cdot \begin{vmatrix} \frac{1}{2} & \frac{1}{2} & -1 \\ -4 & 0 & 0 \\ 0 & -4 & 0 \end{vmatrix} = -16 < 0.$$

Da $\left|\overline{\mathbb{H}}_3(\Lambda(P_0))\right| > 0$ e $\left|\overline{\mathbb{H}}_4(\Lambda(P_0))\right| < 0$ segue che P_0 è punto di massimo.

Si poteva anche ottenere lo stesso risultato in un modo più rapido, dato che il vincolo permette, da $x^2 + y^2 - z = 0$, di esplicitare z come $z = x^2 + y^2$, per cui, sostituendo, otteniamo: $f\left(x, y, x^2 + y^2\right) = x + y - 2x^2 - 2y^2$. Possiamo cercare allora i massimi e minimi liberi di questa funzione di due variabili. Il vincolo è soddisfatto in quanto riportato nella sostituzione.

Avremo allora: $\begin{cases} f'_x = 1 - 4x = 0 \\ f'_y = 1 - 4y = 0 \end{cases} \Rightarrow \begin{cases} x = \dfrac{1}{4} \\ y = \dfrac{1}{4} \end{cases} \Rightarrow z = x^2 + y^2 = \dfrac{1}{8}$. Sarà poi:

$\mathbb{H}(f) = \begin{Vmatrix} -4 & 0 \\ 0 & -4 \end{Vmatrix} = \mathbb{H}(f(\mathrm{P}))$ da cui $|\mathbb{H}_1| = -4 < 0$ e $|\mathbb{H}_2| = 16 > 0$, per cui il punto $\mathrm{P} = \left(\dfrac{1}{4}, \dfrac{1}{4}\right)$ è un massimo per $f(x, y, x^2 + y^2)$ e quindi $\mathrm{P}_0 = \left(\dfrac{1}{4}, \dfrac{1}{4}, \dfrac{1}{8}, 2\right)$ è un massimo vincolato per $f(x, y, z)$.

Esempio 81 : Studiamo il problema $\begin{cases} \text{Max/min } f(x,y,z) = xy - z \\ \text{s.v. : } \begin{cases} g(x,y,z) = x^2 + y^2 = 1 \\ h(x,y,z) = 2x - 2y - z = 0 \end{cases} \end{cases}$.

Risolviamo questo problema in tre modi diversi, raggiungendo ovviamente gli stessi risultati. Cominciamo utilizzando il metodo dei moltiplicatori di Lagrange e costruiamo la funzione Lagrangiana dopo aver controllato che sono soddisfatte tutte le ipotesi. Sono chiaramente differenziabili sia la funzione obiettivo che i vincoli. Poi $\dfrac{\partial(g,h)}{\partial(x,y,z)} = \begin{Vmatrix} 2x & 2y & 0 \\ 2 & -2 & -1 \end{Vmatrix}$.

La caratteristica dello Jacobiano sarà pari a 1 solo se $x = y = 0$, ma tale punto non soddisfa il primo vincolo, e quindi tutte le ipotesi sono soddisfatte.

Risulta $\Lambda(x, y, z, \lambda_1, \lambda_2) = xy - z - \lambda_1\left(x^2 + y^2 - 1\right) - \lambda_2\left(2x - 2y - z\right)$.

Poniamo $\nabla\Lambda = 0$ ed abbiamo: $\begin{cases} \Lambda'_x = 0 \Rightarrow y - 2\lambda_1 x - 2\lambda_2 = 0 \\ \Lambda'_y = 0 \Rightarrow x - 2\lambda_1 y + 2\lambda_2 = 0 \\ \Lambda'_z = 0 \Rightarrow -1 + \lambda_2 = 0 \\ \Lambda'_{\lambda_1} = 0 \Rightarrow x^2 + y^2 = 1 \\ \Lambda'_{\lambda_2} = 0 \Rightarrow 2x - 2y - z = 0 \end{cases}$. Sommando la prima e la seconda equazione si ottiene il sistema: $\begin{cases} (y + x)(1 - 2\lambda_1) = 0 \\ x - 2\lambda_1 y + 2\lambda_2 = 0 \\ \lambda_2 = 1 \\ x^2 + y^2 = 1 \\ 2x - 2y - z = 0 \end{cases}$ e quindi i due sistemi:

$\begin{cases} (y + x) = 0 \\ x - 2\lambda_1 y + 2\lambda_2 = 0 \\ \lambda_2 = 1 \\ x^2 + y^2 = 1 \\ 2x - 2y - z = 0 \end{cases}$ e $\begin{cases} 1 - 2\lambda_1 = 0 \\ x - 2\lambda_1 y + 2\lambda_2 = 0 \\ \lambda_2 = 1 \\ x^2 + y^2 = 1 \\ 2x - 2y - z = 0 \end{cases}$. Per il primo abbiamo:

$\begin{cases} x = -y \\ \lambda_1 = \dfrac{2 - y}{2y} \\ \lambda_2 = 1 \\ y^2 = \frac{1}{2} \\ z = -4y \end{cases} \Rightarrow \begin{cases} x = -\frac{1}{\sqrt{2}} \\ \lambda_1 = \sqrt{2} - \frac{1}{2} \\ \lambda_2 = 1 \\ y = \frac{1}{\sqrt{2}} \\ z = -2\sqrt{2} \end{cases}$ e $\begin{cases} x = \frac{1}{\sqrt{2}} \\ \lambda_1 = -\sqrt{2} - \frac{1}{2} \\ \lambda_2 = 1 \\ y = -\frac{1}{\sqrt{2}} \\ z = 2\sqrt{2} \end{cases}$. Per il secondo:

$\begin{cases} \lambda_1 = \frac{1}{2} \\ x - y + 2 = 0 \\ \lambda_2 = 1 \\ x^2 + y^2 = 1 \\ z = 2x - 2y \end{cases} \Rightarrow \begin{cases} \lambda_1 = \frac{1}{2} \\ y = x + 2 \\ \lambda_2 = 1 \\ x^2 + (x + 2)^2 = 1 \\ z = 2x - 2y \end{cases}$. La quarta equazione non ha però soluzioni reali.

Ci sono quindi per la Lagrangiana due soli punti stazionari:

$$\mathrm{P}_1 = \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, -2\sqrt{2}, \sqrt{2} - \frac{1}{2}, 1 \right) \text{ e } \mathrm{P}_2 = \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, 2\sqrt{2}, -\sqrt{2} - \frac{1}{2}, 1 \right).$$

Costruiamo l'Hessiana orlata: $\overline{\mathbb{H}}(\Lambda) = \left\| \begin{matrix} 0 & 0 & 2x & 2y & 0 \\ 0 & 0 & 2 & -2 & -1 \\ 2x & 2 & -2\lambda_1 & 1 & 0 \\ 2y & -2 & 1 & -2\lambda_1 & 0 \\ 0 & -1 & 0 & 0 & 0 \end{matrix} \right\|$. Essendoci due vincoli, sono inutili i primi quattro minori di guida di Nord-Ovest, per cui basta calcolare il determinante della matrice: $\left|\overline{\mathbb{H}}(\Lambda)\right|$. Avremo:

$$\left|\overline{\mathbb{H}}(\Lambda(\mathrm{P}_1))\right| = \begin{vmatrix} 0 & 0 & -\sqrt{2} & \sqrt{2} & 0 \\ 0 & 0 & 2 & -2 & -1 \\ -\sqrt{2} & 2 & 1-2\sqrt{2} & 1 & 0 \\ \sqrt{2} & -2 & 1 & 1-2\sqrt{2} & 0 \\ 0 & -1 & 0 & 0 & 0 \end{vmatrix} =$$

$$= 1 \cdot (-1) \cdot \begin{vmatrix} 0 & -\sqrt{2} & \sqrt{2} \\ -\sqrt{2} & 1-2\sqrt{2} & 1 \\ \sqrt{2} & 1 & 1-2\sqrt{2} \end{vmatrix} = - \begin{vmatrix} 0 & -\sqrt{2} & \sqrt{2} \\ 0 & 2-2\sqrt{2} & 2-2\sqrt{2} \\ \sqrt{2} & 1 & 1-2\sqrt{2} \end{vmatrix} =$$

$= -\sqrt{2} \cdot \left[-2\sqrt{2}\left(2 - 2\sqrt{2}\right) \right] = 8\left(1 - \sqrt{2}\right) < 0$, quindi P_1 è un punto di massimo;

$$\left|\overline{\mathbb{H}}(\Lambda(\mathrm{P}_2))\right| = \begin{vmatrix} 0 & 0 & \sqrt{2} & -\sqrt{2} & 0 \\ 0 & 0 & 2 & -2 & -1 \\ \sqrt{2} & 2 & 1+2\sqrt{2} & 1 & 0 \\ -\sqrt{2} & -2 & 1 & 1+2\sqrt{2} & 0 \\ 0 & -1 & 0 & 0 & 0 \end{vmatrix} =$$

$$= 1 \cdot (-1) \cdot \begin{vmatrix} 0 & \sqrt{2} & -\sqrt{2} \\ \sqrt{2} & 1+2\sqrt{2} & 1 \\ -\sqrt{2} & 1 & 1+2\sqrt{2} \end{vmatrix} = - \begin{vmatrix} 0 & \sqrt{2} & -\sqrt{2} \\ 0 & 2+2\sqrt{2} & 2+2\sqrt{2} \\ -\sqrt{2} & 1 & 1+2\sqrt{2} \end{vmatrix} =$$

$= \sqrt{2} \cdot \left[2\sqrt{2}\left(2 + 2\sqrt{2}\right) \right] = 8\left(1 + \sqrt{2}\right) > 0$, quindi P_2 è un punto di minimo.

Questo problema si può risolvere in un secondo modo.
Dal vincolo $2x - 2y - z = 0$ esplicitiamo la variable $z : z = 2x - 2y$ e sostituiamola nell'espressione della funzione ottenendo: $f(x, y, z) = f(x, y, 2x - 2y) = xy - 2x + 2y$.

Possiamo ora risolvere il problema $\begin{cases} \text{Max/min } f(x,y) = xy - 2x + 2y \\ \text{s.v. : } g(x,y) = x^2 + y^2 = 1 \end{cases}$, che rimane comunque una ricerca di estremi per una funzione di una sola variabile.

Formiamo la Lagrangiana: $\Lambda(x, y, \lambda) = xy - 2x + 2y - \lambda\left(x^2 + y^2 - 1\right)$ e posto $\nabla\Lambda = \mathbb{O}$, otteniamo, sommando la prima con la seconda equazione:

$$\begin{cases} \Lambda'_x = 0 \Rightarrow y - 2 - 2\lambda x = 0 \\ \Lambda'_y = 0 \Rightarrow x + 2 - 2\lambda y = 0 \\ \Lambda'_\lambda = 0 \Rightarrow x^2 + y^2 = 1 \end{cases} \Rightarrow \begin{cases} (y + x)(1 - 2\lambda) = 0 \\ x + 2 - 2\lambda y = 0 \\ x^2 + y^2 = 1 \end{cases}$$ che si scinde nei due sistemi:

$$\begin{cases} x = -y \\ \lambda = \dfrac{2 - y}{2y} \\ 2y^2 = 1 \end{cases} \Rightarrow \begin{cases} x = -\frac{1}{\sqrt{2}} \\ \lambda = \sqrt{2} - \frac{1}{2} \\ y = \frac{1}{\sqrt{2}} \end{cases} \text{ e } \begin{cases} x = \frac{1}{\sqrt{2}} \\ \lambda = -\sqrt{2} - \frac{1}{2} \\ y = -\frac{1}{\sqrt{2}} \end{cases}$$ oppure

$\begin{cases} \lambda = \frac{1}{2} \\ x + 2 - y = 0 \\ x^2 + y^2 = 1 \end{cases} \Rightarrow \begin{cases} \lambda = \frac{1}{2} \\ x = y - 2 \\ y^2 + 4 - 4y + y^2 = 1 \end{cases}$ che non ha però soluzioni.

Abbiamo quindi due punti stazionari:

$\mathrm{P}_1 = \left(-\dfrac{1}{\sqrt{2}}, \dfrac{1}{\sqrt{2}}, \sqrt{2} - \dfrac{1}{2} \right)$ e $\mathrm{P}_2 = \left(\dfrac{1}{\sqrt{2}}, -\dfrac{1}{\sqrt{2}}, -\sqrt{2} - \dfrac{1}{2} \right)$.

Calcoliamo l'Hessiana orlata ed avremo:

$\overline{\mathbb{H}}(\Lambda(x, y, \lambda)) = \left\| \begin{matrix} 0 & 2x & 2y \\ 2x & -2\lambda & 1 \\ 2y & 1 & -2\lambda \end{matrix} \right\|$ e studiamola nei punti trovati. Sarà:

$$\left|\overline{\mathbb{H}}(\Lambda(\mathrm{P}_1))\right| = \begin{vmatrix} 0 & -\sqrt{2} & \sqrt{2} \\ -\sqrt{2} & 1 - 2\sqrt{2} & 1 \\ \sqrt{2} & 1 & 1 - 2\sqrt{2} \end{vmatrix} = \begin{vmatrix} 0 & -\sqrt{2} & \sqrt{2} \\ 0 & 2 - 2\sqrt{2} & 2 - 2\sqrt{2} \\ \sqrt{2} & 1 & 1 - 2\sqrt{2} \end{vmatrix} =$$

$= \sqrt{2} \cdot \left(8 - 4\sqrt{2} \right) > 0$, per cui P_1 è un punto di massimo; ricaviamo poi $z = -2\sqrt{2}$;

$$\left|\overline{\mathbb{H}}(\Lambda(\mathrm{P}_2))\right| = \begin{vmatrix} 0 & \sqrt{2} & -\sqrt{2} \\ \sqrt{2} & 1 + 2\sqrt{2} & 1 \\ -\sqrt{2} & 1 & 1 + 2\sqrt{2} \end{vmatrix} = \begin{vmatrix} 0 & \sqrt{2} & -\sqrt{2} \\ 0 & 2 + 2\sqrt{2} & 2 + 2\sqrt{2} \\ -\sqrt{2} & 1 & 1 + 2\sqrt{2} \end{vmatrix} =$$

$= -\sqrt{2} \cdot \left(8 + 4\sqrt{2} \right) < 0$, per cui P_2 è un punto di minimo; ricaviamo poi $z = 2\sqrt{2}$.

Ovviamente i risultati coincidono con quelli trovati precedentemente.

Risolviamo infine il problema in un terzo modo.

Dopo aver esplicitato z, riprendiamo il problema $\begin{cases} \text{Max/min } f(x, y) = xy - 2x + 2y \\ \text{s.v. : } g(x, y) = x^2 + y^2 = 1 \end{cases}$, e dato che il vincolo è costituito dalla circonferenza trigonometrica, poniamo $\begin{cases} x = \cos t \\ y = \operatorname{sen} t \end{cases}$, dalla quale otteniamo: $f(x(t), y(t), z(x(t), y(t))) = F(t) = \operatorname{sen} t \cos t - 2\cos t + 2\operatorname{sen} t$. Quindi: $F'(t) = \cos^2 t - \operatorname{sen}^2 t + 2\operatorname{sen} t + 2\cos t = (\operatorname{sen} t + \cos t)(\cos t - \operatorname{sen} t + 2)$.

Dato che $\cos t - \operatorname{sen} t + 2 > 0, \forall t$, sarà $F'(t) \geq 0$ per $\cos t \geq -\operatorname{sen} t$, che risulta verificata per $0 \leq t \leq \dfrac{3}{4}\pi$ e per $\dfrac{7}{4}\pi \leq t \leq 2\pi$. Quindi in $t = \dfrac{3}{4}\pi$ c'è un punto di massimo, mentre in $t = \dfrac{7}{4}\pi$ c'è un punto di minimo. Per $t = \dfrac{3}{4}\pi$ otteniamo $\begin{cases} x = \cos \dfrac{3}{4}\pi = -\dfrac{1}{\sqrt{2}} \\ y = \operatorname{sen} \dfrac{3}{4}\pi = \dfrac{1}{\sqrt{2}} \end{cases}$, mentre per $t = \dfrac{7}{4}\pi$ si ha $\begin{cases} x = \cos \dfrac{7}{4}\pi = \dfrac{1}{\sqrt{2}} \\ y = \operatorname{sen} \dfrac{7}{4}\pi = -\dfrac{1}{\sqrt{2}} \end{cases}$.

Si ricava poi z come fatto in precedenza.

Ci sarebbe anche una quarta procedura, che prevede di esplicitare la y dalla $x^2 + y^2 = 1$, per portare il problema alla ricerca di estremi liberi di una funzione di una sola variabile; si ottiene però $y = \pm\sqrt{1 - x^2}$, e sia il tipo di espressione che il dover risolvere due volte il problema rendono sicuramente non molto conveniente quest'ultima procedura.

MASSIMI E MINIMI CON VINCOLI DI DISUGUAGLIANZA

Passiamo infine a trattare problemi del tipo $\begin{cases} \text{Max/min } f(x_1, x_2, ..., x_n) \\ \text{s.v. : } g_i(x_1, x_2, ..., x_n) \leq 0, 1 \leq i \leq m \end{cases}$.

Posto $\mathbb{X} = (x_1, x_2, ..., x_n)$, sia $\mathcal{E} \subset \mathbb{R}^n$ l'insieme nel quale sono soddisfatte contemporaneamente le m disequazioni $g_i(\mathbb{X}) \leq 0$; $\mathcal{E}$ viene detto **regione ammissibile** e costituisce un sottoinsieme chiuso (limitato o illimitato) di $\mathbb{R}^n$. Cercare massimi e minimi di $f(\mathbb{X})$ in $\mathcal{E}$ vuol dire allora cercare gli estremi che si trovano all'interno di $\mathcal{E}$ ($\forall i : g_i(\mathbb{X}) < 0$) e quelli che si trovano sulla frontiera di $\mathcal{E}$ ($\exists i : g_i(\mathbb{X}) = 0$). La ricerca nei punti interni corrisponde al problema degli estremi liberi, quella nei punti di frontiera alla ricerca di estremi con vincoli di uguaglianza, ovvero i due problemi che sono già stati trattati.
Per descrivere la regione $\mathcal{E}$ con vincoli di disuguaglianza si possono avere più vincoli di quante sono le variabili della funzione obiettivo; non vale quindi più la regola $m < n$.
Definizione 48 : Nel punto $\mathbb{X}_0 \in \mathcal{E}$ il vincolo $g_i(\mathbb{X})$ si dice:
- **soddisfatto**, se $g_i(\mathbb{X}_0) \leq 0$;
- **attivo**, se $g_i(\mathbb{X}_0) = 0$.

In un punto $\mathbb{X}_0 \in \mathcal{E}$ tutti i vincoli devono risultare comunque soddisfatti, qualcuno (anche tutti oppure nessuno) può risultare attivo.
Cominciamo allora a costruire quelle che vengono solitamente dette condizioni di Kuhn-Tucker, che rappresentano la forma più generale delle condizioni del I ordine, e sono quindi condizioni necessarie, per la ricerca degli estremi. Per arrivare all'enunciato finale occorre fare anzitutto due scelte: per prima cosa si sceglie di ricercare i punti di massimo. Quelli di minimo sono collegabili a quelli di massimo, una volta notato che $\min(f(x)) = -\max(-f(x))$. Si sceglie poi di scrivere i vincoli, come fatto in precedenza, nella forma $g(\mathbb{X}) \leq 0$; scriverli nella forma $h(\mathbb{X}) = -g(\mathbb{X}) \geq 0$ porta alla determinazione dello stesso insieme $\mathcal{E}$, ma ad una diversa formulazione delle condizioni di Kuhn-Tucker.
Per quanto visto in precedenza, determinare estremi con vincoli di uguaglianza porta ad imporre le condizioni del I ordine non sulla funzione obiettivo, ma sulla Lagrangiana:

$$\Lambda(\mathbb{X}, \lambda_1, ..., \lambda_m) = f(\mathbb{X}) - \sum_{i=1}^{m} \lambda_i \, g_i(\mathbb{X}) .$$

Riprendiamo questa stessa funzione ed osserviamo che, se il punto $\mathbb{X}_0 \in \mathcal{E}$, ipotetica soluzione del problema, fosse interno rispetto al vincolo $g_k(\mathbb{X})$, ovvero se fosse $g_k(\mathbb{X}_0) < 0$, rispetto a tale vincolo è come se si cercassero gli estremi liberi, e quindi basta prendere $\lambda_k = 0$. Se invece il vincolo fosse attivo in $\mathbb{X}_0$, ovvero se risultasse $g_k(\mathbb{X}_0) = 0$, allora è come trattare un problema di estremi con vincoli di uguaglianza.
L'espressione della funzione Lagrangiana consente quindi di condurre la ricerca degli estremi sia all'interno che sulla frontiera di $\mathcal{E}$; basta azzerare o non azzerare gli opportuni moltiplicatori, vedendo quali vincoli sono attivi in $\mathbb{X}_0$.
Quindi la prima condizione da imporre sarà che risulti $\nabla\Lambda = \mathbb{O}$, nell'ipotesi che $\mathbb{X}_0$ sia un punto di massimo, avendo posto per ipotesi che le funzioni $f(\mathbb{X})$ e $g_i(\mathbb{X})$, $1 \leq i \leq m$, siano differenziabili in $\mathcal{E}$. Nelle condizioni del I ordine vanno rispettati anche i vincoli, che si possono ottenere come derivate della Lagrangiana rispetto ai moltiplicatori. Essendo $\dfrac{\partial\Lambda}{\partial\lambda_i} = -g_i(\mathbb{X})$, dovremo imporre $\dfrac{\partial\Lambda}{\partial\lambda_i} = -g_i(\mathbb{X}) \geq 0$ affinchè il vincolo sia soddisfatto.
La richiesta della differenziabilità su $f(\mathbb{X})$ e $g_i(\mathbb{X})$ non esaurisce il quadro delle ipotesi, che verranno completate nel seguito. Enunciamo per il momento il
Teorema 33 (di Kuhn-Tucker) : Sia $\mathbb{X}_0$ soluzione del problema :
$\begin{cases} \text{Max } f(x_1, x_2, ..., x_n) \\ \text{s.v. : } g_i(x_1, x_2, ..., x_n) \leq 0 \end{cases}$,

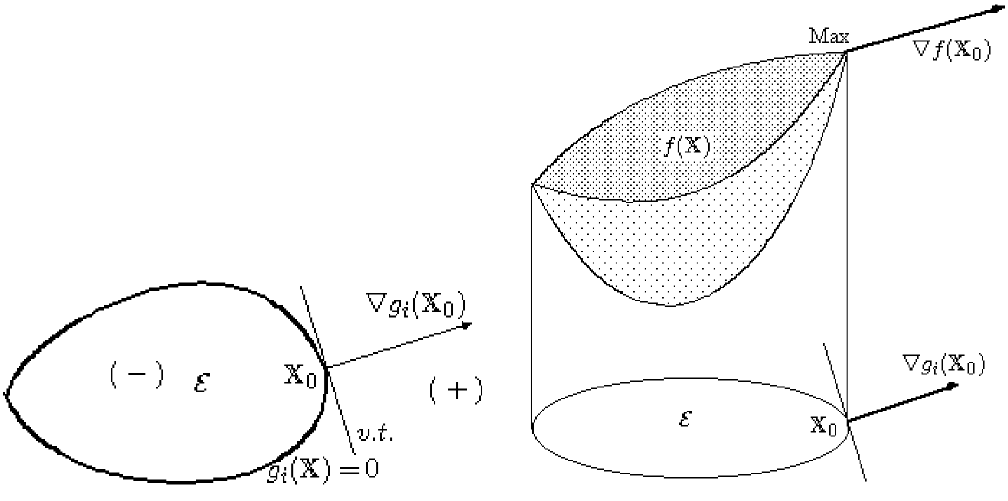
ovvero sia punto di massimo per $f(\mathbb{X})$ sotto ai vincoli $g_i(\mathbb{X}) \le 0$, $1 \le i \le m$. Siano $f(\mathbb{X})$ e $g_i(\mathbb{X})$ differenziabili in $\mathcal{E}$ e siano, nel punto $\mathbb{X}_0$, i vincoli qualificati. Allora esiste un vettore

$$\lambda = (\lambda_1, \lambda_2, ..., \lambda_m) \text{ tale che: } \begin{cases} \dfrac{\partial \Lambda(\mathbb{X}_0)}{\partial x_i} = \dfrac{\partial f(\mathbb{X}_0)}{\partial x_i} - \displaystyle\sum_{j=1}^{m} \lambda_j \dfrac{\partial g_j(\mathbb{X}_0)}{\partial x_i} = 0 & 1 \le i \le n \\ \dfrac{\partial \Lambda}{\partial \lambda_i} = -\, g_i(\mathbb{X}_0) \ge 0 & 1 \le i \le m \cdot \\ \lambda_i \cdot g_i(\mathbb{X}_0) = 0 & 1 \le i \le m \\ \lambda_i \ge 0 & 1 \le i \le m \end{cases}$$

Cosa significhi "vincoli qualificati in $\mathbb{X}_0$" verrà spiegato nel seguito, quando completeremo le ipotesi del Teorema 33.

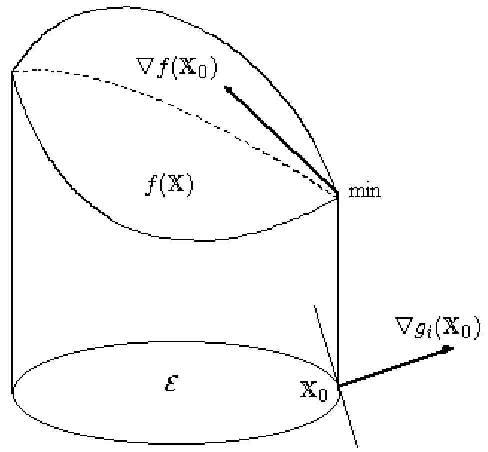
Le prime condizioni: $\dfrac{\partial \Lambda(\mathbb{X}_0)}{\partial x_i} = 0$ ricalcano quanto già visto a proposito dei vincoli di uguaglianza, così come le seconde: $\dfrac{\partial \Lambda}{\partial \lambda_i} = -\, g_i(\mathbb{X}_0) \ge 0$ ribadiscono la necessità che i vincoli vengano soddisfatti. Vediamo la terza condizione: $\lambda_i \cdot g_i(\mathbb{X}_0) = 0$. Dovendosi annullare un prodotto, dovrà essere nullo o il primo o il secondo fattore. Se fosse $\lambda_i \ne 0$, dovrà allora risultare $g_i(\mathbb{X}_0) = 0$, ovvero il vincolo g_i è attivo in $\mathbb{X}_0$, ed è come condurre una ricerca di estremi con vincoli di uguaglianza. Se fosse $g_i(\mathbb{X}_0) \ne 0$ dovrà risultare $\lambda_i = 0$, ovvero il punto $\mathbb{X}_0$ è interno rispetto al vincolo g_i, e rispetto a quel vincolo è come condurre una ricerca di estremi liberi, quindi nella Lagrangiana non compare tale vincolo, e quindi $\lambda_i = 0$. Nel punto $\mathbb{X}_0$ potrebbe anche risultare $\lambda_i = g_i(\mathbb{X}_0) = 0$, e questo dipende dal fatto che un estremo vincolato può coincidere con un estremo libero, circostanza che si verifica quando le coordinate dell'estremo libero soddisfano anche il vincolo. La terza condizione quindi esprime il fatto che la ricerca degli estremi si svolge sia nei punti interni che in quelli di frontiera, e quindi in tutto $\mathcal{E}$.

Dal punto di vista operativo la terza condizione è quella che ci dice come impostare i calcoli, ovvero si devono imporre le equazioni $\dfrac{\partial \Lambda(\mathbb{X})}{\partial x_i} = 0$ e $\dfrac{\partial \Lambda(\mathbb{X})}{\partial \lambda_j} = 0$ per ciascuna delle 2^m possibili combinazioni ottenute imponendo ciascuno dei moltiplicatori λ_i uguale o diverso da zero. Se, ad esempio, avessimo due vincoli e quindi due moltiplicatori, λ_1 e λ_2, dovremo imporre quattro sistemi, quelli corrispondenti ai casi: ($\lambda_1 = 0$ e $\lambda_2 = 0$), ($\lambda_1 \ne 0$ e $\lambda_2 = 0$), ($\lambda_1 = 0$ e $\lambda_2 \ne 0$) ed infine ($\lambda_1 \ne 0$ e $\lambda_2 \ne 0$). Con tre vincoli i casi divengono 8 ecc. ecc..



Veniamo alla quarta ed ultima condizione, che vedremo dipendere strettamente dal segno dei vincoli (abbiamo scelto il negativo) ed al tipo di estremo che cerchiamo (abbiamo scelto di cercare i punti di massimo). Partendo dalla $g_i(\mathbb{X}) = 0$ abbiamo visto che $\nabla g_i(\mathbb{X}_0)$ risulta perpendicolare al vettore tangente al vincolo in $\mathbb{X}_0$. Dato poi che il gradiente esprime la direzione

di massimo accrescimento, essendo $\mathcal{E}$ caratterizzata da valori negativi per il vincolo $g_i(\mathbb{X})$, l'esterno, ovvero il complementare di $\mathcal{E}$, è caratterizzato da valori positivi per $g_i(\mathbb{X})$, da cui ne consegue che il gradiente $\nabla g_i(\mathbb{X}_0)$ non può che essere orientato verso l'esterno di $\mathcal{E}$, come illustrato nella figura, che è basata, come caso particolare, su di un solo vincolo.
Supponiamo che $\mathbb{X}_0$ sia punto di massimo e che in $\mathbb{X}_0$ il vincolo $g_i(\mathbb{X})$ sia attivo: $g_i(\mathbb{X}_0) = 0$. Abbiamo visto che $\nabla g_i(\mathbb{X}_0)$ si dirige verso l'esterno. Il gradiente $\nabla f(\mathbb{X}_0)$ della funzione nel punto, che nella figura, per una migliore lettura, è stato traslato sulla superficie, non può che dirigersi anch'esso verso l'esterno, in quanto indica la direzione di massima crescita; se si dirigesse verso l'interno, partendo dal punto di massimo, indicherebbe invece una decrescita nei valori della funzione. Dalla teoria degli estremi con vincoli di uguaglianza sappiamo che, in un punto che risulti soluzione di un problema di estremo, il gradiente della funzione obiettivo deve risultare una combinazione lineare dei gradienti dei vincoli: $\nabla f(\mathbb{X}_0) = \lambda_i \nabla g_i(\mathbb{X}_0)$. Si è poi visto che $\nabla f(\mathbb{X}_0)$ e $\nabla g_i(\mathbb{X}_0)$ devono dirigersi dalla stessa parte: le due richieste sono entrambe soddisfatte se $\lambda_i > 0$.
Vale un Teorema (Farkas) che estende questa proprietà al caso in cui risultino più vincoli contemporaneamente attivi in $\mathbb{X}_0$: $\nabla f(\mathbb{X}_0) = \sum_{i:1}^{m} \lambda_i \nabla g_i(\mathbb{X}_0)$. Ogni λ_i deve risultare non negativo affinchè $\nabla f(\mathbb{X}_0)$ si diriga verso l'esterno, come fanno i gradienti $\nabla g_i(\mathbb{X}_0)$.

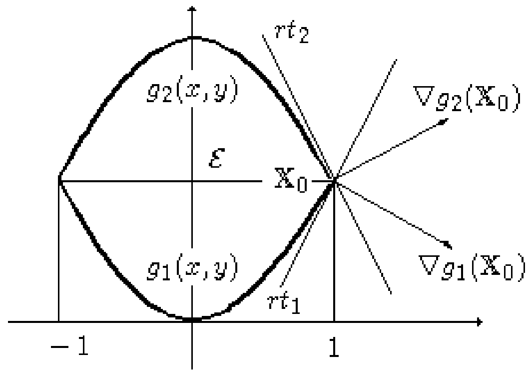


Se, sempre mantenendo i vincoli nella forma $g_i(\mathbb{X}) \leq 0$, si fosse ricercato il punto di minimo, e se questo punto si fosse trovato sul vincolo, $g_i(\mathbb{X}_0) = 0$, come si vede nella figura, il gradiente $\nabla f(\mathbb{X}_0)$ deve ora dirigersi verso l'interno, dove la funzione assume valori maggiori rispetto al minimo. Ma allora dovrà risultare $\lambda_i \leq 0$, dato che $\nabla f(\mathbb{X}_0)$ e $\nabla g_i(\mathbb{X}_0)$ non sono orientati dalla stessa parte, andando il primo verso l'interno ed il secondo verso l'esterno. La ricerca del minimo porta a richiedere moltiplicatori non positivi, lasciando invece inalterate le prime tre condizioni del Teorema 33.
Se i vincoli fossero stati espressi nella forma $g_i(\mathbb{X}) \geq 0$, le precedenti richieste sul segno dei moltiplicatori per un punto di massimo e per un punto di minimo sarebbero risultate invertite. Il segno (non negativo) dei moltiplicatori per un punto di massimo con vincoli $g_i(\mathbb{X}) \leq 0$ è quindi lo stesso di quello per un punto di minimo con vincoli $g_i(\mathbb{X}) \geq 0$.

QUALIFICAZIONE DEI VINCOLI
Completiamo infine la parte mancante riguardo alle ipotesi dove si chiede, per la validità delle condizioni di Kuhn-Tucker, che i vincoli siano qualificati nel punto $\mathbb{X}_0$. Dallo studio del segno dei moltiplicatori λ_i al fine di chiarire la natura del punto che si sta esaminando, emerge una necessità, ovvero che si possa stabilire con chiarezza la direzione del $\nabla f(\mathbb{X}_0)$ in relazione alle direzioni dei gradienti dei vincoli. Questo problema viene detto "**qualificazione**

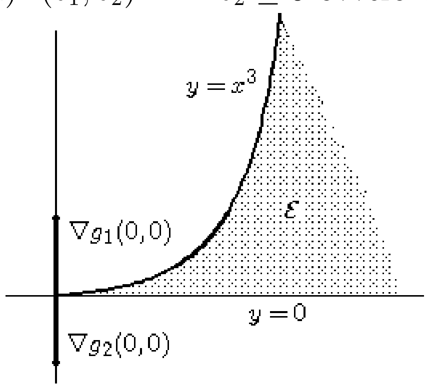
dei vincoli" in $\mathbb{X}_0$. Per arrivare a descrivere questo concetto occorrono alcune definizioni, la prima delle quali è quella di direzione ammissibile:
Definizione 49 : Il vettore $v \in \mathbb{R}^n$ è detto una **direzione ammissibile** in $\mathbb{X}_0$ per $\mathcal{E}$ se esiste una curva, $t \to r(t) \in \mathbb{R}^n$, tale che:
1) $r(0) = \mathbb{X}_0$;
2) $r'(0) = v$;
3) $\exists \varepsilon : r(t) \in \mathcal{E}, \forall t \in [0, \varepsilon[$.
Esiste cioè almeno un tratto, per quanto piccolo, di curva continua che parte da $\mathbb{X}_0$ ed entra in $\mathcal{E}$: la direzione ammissibile v è la tangente in $\mathbb{X}_0$ a tale curva.
Questa definizione riguarda nella pratica i punti che stanno sulla frontiera di $\mathcal{E}$, ovvero i punti nei quali almeno un vincolo è attivo. Se $\mathbb{X}_0$ è interno ad $\mathcal{E}$, ogni direzione risulta ammissibile.
Le direzioni ammissibili in $\mathbb{X}_0$ per $\mathcal{E}$ sono quelle che, partendo da $\mathbb{X}_0$, vanno dentro alla regione ammissibile $\mathcal{E}$, o che sono, come caso limite, tangenziali ad $\mathcal{E}$.
Definizione 50 : Si dice **cono** un insieme $\mathbb{A} \subseteq \mathbb{R}^n$ tale che: $\mathbb{X} \in \mathbb{A} \Rightarrow k \cdot \mathbb{X} \in \mathbb{A}, \forall k \in \mathbb{R}$.
Si dice **cono positivo** un insieme $\mathbb{A} \subseteq \mathbb{R}^n$ tale che: $\mathbb{X} \in \mathbb{A} \Rightarrow k \cdot \mathbb{X} \in \mathbb{A}, \forall k \in \mathbb{R}_+$.
Le direzioni ammissibili in $\mathbb{X}_0$ per $\mathcal{E}$ costituiscono un cono positivo: $\Gamma(\mathbb{X}_0)$.
Per stabilire la natura del punto $\mathbb{X}_0$ abbiamo visto che è necessario, nell'impostazione scelta, che $\nabla f(\mathbb{X}_0)$ si diriga verso l'esterno di $\mathcal{E}$, dalla stessa parte dei $\nabla g_i(\mathbb{X}_0)$. Occorre allora studiare quelle direzioni v che diremo "retroverse" rispetto a $\nabla g_i(\mathbb{X}_0)$, ovvero le direzioni per le quali $\nabla g_i(\mathbb{X}_0) \cdot v \leq 0$.
Essendo $\nabla g_i(\mathbb{X}_0) \cdot v = \|\nabla g_i(\mathbb{X}_0)\| \cdot \|v\| \cdot \cos\alpha$, sarà $\nabla g_i(\mathbb{X}_0) \cdot v \leq 0$ se $\frac{\pi}{2} \leq \alpha \leq \pi$.
Le direzioni ammissibili da $\mathbb{X}_0$ per $\mathcal{E}$ e le direzioni retroverse rispetto ai $\nabla g_i(\mathbb{X}_0)$ dovranno risultare le stesse direzioni affinchè l'analisi del $\nabla f(\mathbb{X}_0)$ risulti fondata.
Si considerino tutti i vincoli attivi in $\mathbb{X}_0$: $g_i(\mathbb{X}_0) = 0$. Con $\widetilde{\Gamma}(\mathbb{X}_0)$ si indica il cono delle **direzioni retroverse** in $\mathbb{X}_0$ rispetto ai gradienti dei vincoli attivi in $\mathbb{X}_0$, ovvero:
$\widetilde{\Gamma}(\mathbb{X}_0) = \{v \in \mathbb{R}^n : \nabla g_i(\mathbb{X}_0) \cdot v \leq 0; g_i(\mathbb{X}_0) = 0\}$.
In generale, come vedremo negli esempi, risulta $\Gamma(\mathbb{X}_0) \subseteq \widetilde{\Gamma}(\mathbb{X}_0)$. Diamo ora la
Definizione 51 : Si dice che i vincoli sono qualificati in $\mathbb{X}_0$ se risulta $\Gamma(\mathbb{X}_0) = \widetilde{\Gamma}(\mathbb{X}_0)$, ovvero direzioni ammissibili e direzioni retroverse rispetto ai gradienti dei vincoli attivi formano lo stesso cono.
E' questa la condizione che va aggiunta alla differenziabilità di f e delle g_i per la validità delle condizioni di Kuhn-Tucker per un punto di massimo o di minimo.

Esempio 82 : Consideriamo come regione ammissibile $\mathcal{E}$ quella data dai due vincoli:
$$\begin{cases} g_1(x,y) = x^2 - y \leq 0 \\ g_2(x,y) = y - 2 + x^2 \leq 0 \end{cases}.$$
I due vincoli sono attivi contemporaneamente nei punti $(1,1)$ e $(-1,1)$. Costruiamo $\Gamma(1,1)$.
Per fare questo troviamo l'equazione della tangente alle curve $g_1 : y = x^2$ e $g_2 : y = 2 - x^2$.
Per la prima, rt_1, avremo $y = 2x - 1$ mentre per la seconda, rt_2, $y = -2x + 3$.



Sia poi $v = (v_1, v_2)$. Per andare da $(1,1)$ verso $\mathcal{E}$ dovrà essere $v_1 \leq 0$, per spostarsi sulla sinistra; l'incremento sulle ordinate dovrà essere compreso tra quelli espressi dalle due tangenti trovate, ovvero dovrà essere: $2v_1 \leq v_2 \leq -2v_1$ ($v_1 \leq 0$!!) e quindi sarà:
$\mathit{\Gamma}(1,1) = \{(v_1, v_2) \in \mathbb{R}^2 : v_1 \leq 0, 2v_1 \leq v_2 \leq -2v_1\}$. Determiniamo poi $\widetilde{\mathit{\Gamma}}(1,1)$.
Da $g_1(x,y) = x^2 - y$ segue $\nabla g_1(x,y) = (2x, -1)$ e quindi $\nabla g_1(1,1) = (2, -1)$.
Da $g_2(x,y) = y - 2 + x^2$ segue $\nabla g_2(x,y) = (2x, 1)$ e quindi $\nabla g_2(1,1) = (2,1)$.
Troviamo le direzioni retroverse ad ambedue i gradienti. Sarà:
$\nabla g_1(1,1) \cdot (v_1, v_2) = (2, -1) \cdot (v_1, v_2) = 2v_1 - v_2 \leq 0$ per $v_2 \geq 2v_1$, mentre
$\nabla g_2(1,1) \cdot (v_1, v_2) = (2,1) \cdot (v_1, v_2) = 2v_1 + v_2 \leq 0$ per $v_2 \leq -2v_1$, e quindi:
$2v_1 \leq v_2 \leq -2v_1$, e questa doppia disequazione implica anche $v_1 \leq 0$.
Quindi $\widetilde{\mathit{\Gamma}}(1,1) = \{(v_1, v_2) \in \mathbb{R}^2 : v_1 \leq 0, 2v_1 \leq v_2 \leq -2v_1\} = \mathit{\Gamma}(1,1)$.
I vincoli sono quindi qualificati nel punto $(1,1)$. Si può vedere in maniera simile che questo accade anche nel punto $(-1,1)$, dove:
$\widetilde{\mathit{\Gamma}}(-1,1) = \{(v_1, v_2) \in \mathbb{R}^2 : v_1 \geq 0, -2v_1 \leq v_2 \leq 2v_1\} = \mathit{\Gamma}(-1,1)$.

Esempio 83 : Sia la regione ammissibile $\mathcal{E}$ data dai due vincoli $\begin{cases} g_1(x,y) = y - x^3 \leq 0 \\ g_2(x,y) = -y \leq 0 \end{cases}$.
Nel punto (0.0) sono attivi ambedue i vincoli. Dato che $g_1 : y = x^3$ ha in $(0,0)$ per tangente l'asse x, l'unica direzione ammissibile da $(0,0)$ per $\mathcal{E}$ è quella del semiasse positivo delle x, per cui: $\mathit{\Gamma}(0,0) = \{(v_1, v_2) \in \mathbb{R}^2 : v_1 \geq 0, v_2 = 0\}$. Determiniamo poi $\widetilde{\mathit{\Gamma}}(0,0)$.
Da $g_1(x,y) = y - x^3$ segue $\nabla g_1(x,y) = (-3x^2, 1)$ e quindi $\nabla g_1(0,0) = (0,1)$.
Da $g_2(x,y) = -y$ segue $\nabla g_2(x,y) = (0, -1)$ e quindi $\nabla g_2(0,0) = (0, -1)$.
Troviamo le direzioni retroverse ad ambedue i gradienti. Sarà:
$\nabla g_1(0,0) \cdot (v_1, v_2) = (0,1) \cdot (v_1, v_2) = v_2 \leq 0$, mentre
$\nabla g_2(0,0) \cdot (v_1, v_2) = (0, -1) \cdot (v_1, v_2) = -v_2 \leq 0$ ovvero $v_2 \geq 0$, e quindi: $v_2 = 0$.



Quindi $\widetilde{\mathit{\Gamma}}(0,0) = \{(v_1, v_2) \in \mathbb{R}^2 : v_2 = 0, \forall v_1\}$, ovvero tutto l'asse x. In questo caso quindi risulta $\mathit{\Gamma}(0,0) \subset \widetilde{\mathit{\Gamma}}(0,0)$ e non $\mathit{\Gamma}(0,0) = \widetilde{\mathit{\Gamma}}(0,0)$, per cui in $(0,0)$ i vincoli non sono qualificati.

Esempio 84 : Sia la regione ammissibile $\mathcal{E}$ data dai tre vincoli $\begin{cases} g_1(x,y) = y - x^3 \leq 0 \\ g_2(x,y) = -y \leq 0 \\ g_3(x,y) = -x \leq 0 \end{cases}$.

La regione $\mathcal{E}$ è esattamente la stessa dell'esempio precedente, anche se caratterizzata da un vincolo ulteriore: $g_3(x,y) = -x$. Essendo anche questo attivo in $(0,0)$, rimanendo validi sia i ragionamenti che i calcoli svolti nell'esempio precedente, rimane inalterato $\mathit{\Gamma}(0,0)$ mentre dobbiamo rideterminare $\widetilde{\mathit{\Gamma}}\,(0,0)$, alla luce della presenza dell'ulteriore vincolo.
Da $g_3(x,y) = -x$ segue $\nabla g_3(x,y) = (-1,0)$ e quindi $\nabla g_3(0,0) = (-1,0)$. Per cui: $\nabla g_3(0,0) \cdot (v_1,v_2) = (-1,0) \cdot (v_1,v_2) = -v_1 \leq 0$ per $v_1 \geq 0$, e questa, assieme alle due già trovate, ci dà: $\mathit{\Gamma}(0,0) = \{(v_1,v_2) \in \mathbb{R}^2 : v_1 \geq 0, v_2 = 0\} = \widetilde{\mathit{\Gamma}}\,(0,0)$.
Con la caratterizzazione operata dal terzo vincolo, abbiamo ora che i vincoli risultano qualificati in $(0,0)$. La qualificazione dei vincoli in un punto dipende non solo dalla forma della regione $\mathcal{E}$ ma anche dai vincoli che la descrivono. La stessa regione descritta da vincoli differenti può avere vincoli che risultano qualificati in un punto mentre prima non lo erano.

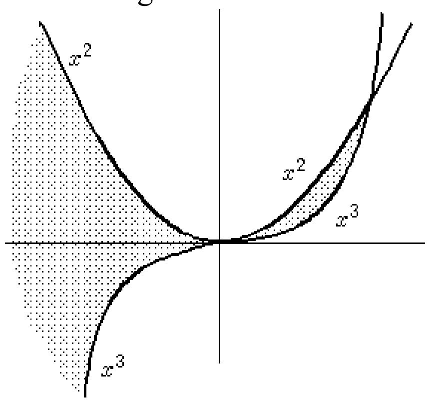
Esistono inoltre alcune condizioni sufficienti a garantire la qualificazione dei vincoli in un dato punto. La più pratica dal punto di vista operativo è sancita dal seguente:
Teorema 34 : Se i vettori $\nabla g_i(\mathbb{X}_0)$ dei vincoli attivi in $\mathbb{X}_0$ sono linearmente indipendenti, allora i vincoli sono qualificati in $\mathbb{X}_0$.
Questa condizione è comunque sufficiente, e non necessaria, per la qualificazione dei vincoli.

Esempio 85 : Riprendiamo la regione $\mathcal{E}$ data dai due vincoli $\begin{cases} g_1(x,y) = x^2 - y \leq 0 \\ g_2(x,y) = y - 2 + x^2 \leq 0 \end{cases}$.

Risulta $\nabla g_1(1,1) = (2,-1)$ e $\nabla g_2(1,1) = (2,1)$, ed essendo $\begin{vmatrix} 2 & -1 \\ 2 & 1 \end{vmatrix} = 4 \neq 0$, i gradienti sono indipendenti ed i vincoli sono qualificati.

Esempio 86 : Consideriamo la regione $\mathcal{E}$ data dai due vincoli $\begin{cases} g_1(x,y) = x^3 - y \leq 0 \\ g_2(x,y) = y - x^2 \leq 0 \end{cases}$.

La regione è rappresentata scurita nella figura:



Le direzioni ammissibili nel punto $(0,0)$ per $\mathcal{E}$, sia sulla sinistra che sulla destra, sono costituite dall'asse x, che risulta retta tangente nel punto $(0,0)$ sia al vicolo $g_1 : y = x^3$ che al vincolo $g_2 : y = x^2$. Sarà quindi $\mathit{\Gamma}(0,0) = \{(v_1,v_2) \in \mathbb{R}^2 : v_2 = 0, \forall\, v_1\}$.
Da $g_1(x,y) = x^3 - y$ segue $\nabla g_1(x,y) = (3x^2, -1)$ e quindi $\nabla g_1(0,0) = (0,-1)$.
Da $g_2(x,y) = y - x^2$ segue $\nabla g_2(x,y) = (-2x, 1)$ e quindi $\nabla g_2(0,0) = (0,1)$.
Troviamo le direzioni retroverse ad ambedue i gradienti. Sarà:
$\nabla g_1(0,0) \cdot (v_1,v_2) = (0,-1) \cdot (v_1,v_2) = -v_2 \leq 0$ per $v_2 \geq 0$, mentre
$\nabla g_2(0,0) \cdot (v_1,v_2) = (0,1) \cdot (v_1,v_2) = v_2 \leq 0$, e quindi: $0 \leq v_2 \leq 0$,

e questa doppia disequazione implica $v_2 = 0$.
Quindi $\widetilde{\Gamma}\,(0,0) = \{(v_1, v_2) \in \mathbb{R}^2 : v_2 = 0, \forall\, v_1\} = \Gamma(0,0)$. I vincoli sono quindi qualificati in $(0,0)$ mentre $\nabla g_1(0,0) = (0, -1)$ e $\nabla g_2(0,0) = (0,1)$ sono vettori linearmente dipendenti, a conferma del fatto che la condizione di indipendenza è sufficiente e non necessaria per la qualificazione.

Esempio 87 : Studiamo il problema $\begin{cases} \text{Max/min } f(x,y) = 1 - x^2 - y^2 \\ \text{s.v. : } \begin{cases} g_1(x,y) = x^3 - y \le 0 \\ g_2(x,y) = y - x^2 \le 0 \end{cases} \end{cases}$.

La regione ammissibile $\mathcal{E}$ e la qualificazione dei vincoli sono stati studiati nell'Esempio 86 per quanto riguarda il punto $(0,0)$. Nel punto $(1,1)$ avremo:
$g_1(x,y) = x^3 - y$ da cui segue $\nabla g_1(x,y) = (3x^2, -1)$ e quindi $\nabla g_1(1,1) = (3, -1)$.
$g_2(x,y) = y - x^2$ da cui segue $\nabla g_2(x,y) = (-2x, 1)$ e quindi $\nabla g_2(1,1) = (-2, 1)$.
Tali vettori sono indipendenti: $\begin{vmatrix} 3 & -1 \\ -2 & 1 \end{vmatrix} = 1 \neq 0$ e quindi i vincoli in $(1,1)$ sono qualificati. Dobbiamo infine esaminare i punti in cui è attivo un solo vincolo e, per usare l'indipendenza lineare, il gradiente del vincolo non dovrà ridursi al vettore nullo. Ma nei due gradienti la seconda componente è sempre costante e diversa da zero, e quindi i vincoli sono sempre qualificati. Infine, notiamo che la regione ammissibile non è un insieme limitato, quindi non è applicabile il Teorema di Weierstrass. Costruiamo allora la Lagrangiana:
$\Lambda(x, y, \lambda_1, \lambda_2) = 1 - x^2 - y^2 - \lambda_1 (x^3 - y) - \lambda_2 (y - x^2)$ e studiamo i quattro casi.

<u>I caso: $\lambda_1 = 0, \lambda_2 = 0$</u> : risolviamo il sistema:

$$\begin{cases} \Lambda'_x = 0 \Rightarrow -2x = 0 \\ \Lambda'_y = 0 \Rightarrow -2y = 0 \\ \Lambda'_{\lambda_1} \ge 0 \Rightarrow x^3 - y \le 0 \\ \Lambda'_{\lambda_2} \ge 0 \Rightarrow y - x^2 \le 0 \end{cases} \Rightarrow \begin{cases} x = 0 \\ y = 0 \\ 0 - 0 \le 0 : \text{ vera} \\ 0 - 0 \le 0 : \text{ vera} \end{cases}$$

. Essendo $\lambda_1 = \lambda_2 = 0$, si tratta di studiare gli estremi liberi della funzione.

Dato che $\mathbb{H}(x,y) = \begin{vmatrix} -2 & 0 \\ 0 & -2 \end{vmatrix} = \mathbb{H}(0,0)$, si ha $\begin{cases} f''_{xx} = -2 < 0 \\ f''_{xx} f''_{yy} - (f''_{xy})^2 = 4 > 0 \end{cases}$ e quindi $(0,0)$ è un punto di massimo, con $f(0,0) = 1$.

<u>II caso: $\lambda_1 \neq 0, \lambda_2 = 0$</u> : risolviamo il sistema:

$$\begin{cases} \Lambda'_x = 0 \Rightarrow -2x - 3\lambda_1 x^2 = 0 \\ \Lambda'_y = 0 \Rightarrow -2y + \lambda_1 = 0 \\ \Lambda'_{\lambda_1} = 0 \Rightarrow y = x^3 \\ \Lambda'_{\lambda_2} \ge 0 \Rightarrow y - x^2 \le 0 \end{cases} \Rightarrow \begin{cases} -x(2 + 3\lambda_1 x) = 0 \\ \lambda_1 = 2y \\ y = x^3 \\ y \le x^2 \end{cases}$$

che dà i due sistemi:

$$\begin{cases} x = 0 \\ \lambda_1 = 0 \\ y = 0 \\ 0 \le 0 : \text{ vera} \end{cases}$$

, ma $(0,0)$ è già stato studiato, e

$$\begin{cases} x = -\dfrac{2}{3\lambda_1} \\ y = \dfrac{\lambda_1}{2} \\ \dfrac{\lambda_1}{2} = -\dfrac{8}{27\lambda_1^3} \Rightarrow \lambda_1^4 = -\dfrac{16}{27} \\ y \le x^2 \end{cases}$$ e che quindi non ammette soluzioni.

III caso: $\lambda_1 = 0, \lambda_2 \neq 0$: risolviamo il sistema:

$$\begin{cases} \Lambda'_x = 0 \Rightarrow -2x + 2\lambda_2 x = 0 \\ \Lambda'_y = 0 \Rightarrow -2y - \lambda_2 = 0 \\ \Lambda'_{\lambda_1} \ge 0 \Rightarrow x^3 \le y \\ \Lambda'_{\lambda_2} = 0 \Rightarrow y = x^2 \end{cases} \Rightarrow \begin{cases} 2x(\lambda_2 - 1) = 0 \\ \lambda_2 = -2y \\ x^3 \le y \\ y = x^2 \end{cases}$$ che dà i due sistemi:

$$\begin{cases} x = 0 \\ \lambda_2 = 0 \\ 0 \le 0 : \text{ vera} \\ y = 0 \end{cases}$$, ma $(0,0)$ è già stato studiato, e

$$\begin{cases} \lambda_2 = 1 \\ y = -\dfrac{1}{2} \\ x^3 \le y \\ x^2 = -\dfrac{1}{2} \end{cases}$$ e che quindi non ammette soluzioni.

IV caso: $\lambda_1 \neq 0, \lambda_2 \neq 0$: risolviamo il sistema:

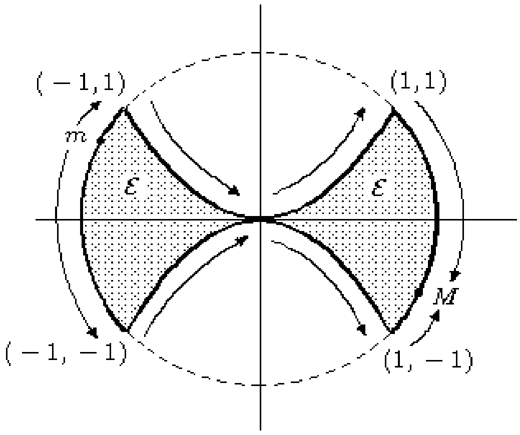
$$\begin{cases} \Lambda'_x = 0 \Rightarrow -2x - 3\lambda_1 x^2 + 2\lambda_2 x = 0 \\ \Lambda'_y = 0 \Rightarrow -2y + \lambda_1 - \lambda_2 = 0 \\ \Lambda'_{\lambda_1} = 0 \Rightarrow x^3 = y \\ \Lambda'_{\lambda_2} = 0 \Rightarrow y = x^2 \end{cases} \Rightarrow \begin{cases} 0 = 0 \\ \lambda_1 = \lambda_2 \\ x = 0 \\ y = 0 \end{cases} \text{ e } \begin{cases} -2 - 3\lambda_1 + 2\lambda_2 = 0 \\ -2 + \lambda_1 - \lambda_2 = 0 \\ x = 1 \\ y = 1 \end{cases} \Rightarrow$$

$$\begin{cases} -2 - 3\lambda_1 + 2\lambda_1 - 4 = 0 \\ \lambda_2 = \lambda_1 - 2 \\ x = 1 \\ y = 1 \end{cases} \Rightarrow \begin{cases} \lambda_1 = -6 \\ \lambda_2 = -8 \\ x = 1 \\ y = 1 \end{cases}$$. Essendo ambedue i moltiplicatori negativi, il punto $(1,1)$ potrebbe essere un punto di minimo. Per risolvere il problema, osserviamo che la funzione $f(x,y) = 1 - x^2 - y^2$, in quanto polinomio, è continua in $\mathbb{R}^2$. La parte destra di $\mathcal{E}$ è un insieme limitato e chiuso, quindi per il Teorema di Weierstrass ci sono minimo e massimo assoluti, che non possono che essere in $(0,0)$ e in $(1,1,)$ per quanto visto in precedenza.

Il massimo $f(0,0) = 1$ è palesemente assoluto, mentre il minimo $f(0,0)$ è solo relativo, in quanto, analizzando ad esempio la funzione sul semiasse negativo delle ascisse, vediamo che $\lim\limits_{x \to -\infty} f(x,0) = -\infty$, e quindi la funzione può assumere in $\mathcal{E}$ valori illimitatamente grandi e negativi, e quindi non può avere un minimo assoluto.

Esempio 88 : Studiamo il problema $\begin{cases} \text{Max/min } f(x,y) = 2x - y \\ \text{s.v. : } \begin{cases} g_1(x,y) = y^2 - x^4 \le 0 \\ g_2(x,y) = x^2 + y^2 - 2 \le 0 \end{cases} \end{cases}$.

La regione $\mathcal{E}$ è rappresentata scurita nella figura:



Da $y^2 \le x^4$ segue $-x^2 \le y \le x^2$, cioè la parte di piano compresa tra le due parabole $y = -x^2$ e $y = x^2$, all'interno della circonferenza $x^2 + y^2 = 2$, di centro $(0,0)$ e raggio pari a $r = \sqrt{2}$.

Da $\begin{cases} y^2 - x^4 = 0 \\ y^2 = 2 - x^2 \end{cases}$ si ha $2 - x^2 - x^4 = 0$, da cui $x^2 = \dfrac{-1 \pm \sqrt{1+8}}{2} \Rightarrow \begin{cases} x^2 = -2 \\ x^2 = 1 \end{cases}$ e quindi $x = \pm 1$, per cui abbiamo quattro intersezioni: $(1,1)$, $(1,-1)$, $(-1,1)$ e $(-1,-1)$.

Da $g_1(x,y) = y^2 - x^4$ segue $\nabla g_1(x,y) = \left(-4x^3, 2y\right)$, da $g_2(x,y) = x^2 + y^2 - 2$ segue $\nabla g_2(x,y) = (2x, 2y)$ e quindi avremo:

$\nabla g_1(1,1) = (-4,2)$ e $\nabla g_2(1,1) = (2,2)$: vettori indipendenti;

$\nabla g_1(1,-1) = (-4,-2)$ e $\nabla g_2(1,-1) = (2,-2)$: vettori indipendenti;

$\nabla g_1(-1,1) = (4,2)$ e $\nabla g_2(-1,1) = (-2,2)$: vettori indipendenti;

$\nabla g_1(-1,-1) = (4,-2)$ e $\nabla g_2(-1,-1) = (-2,-2)$: vettori indipendenti.

Nel punto $(0,0)$ risulta $\nabla g_1(0,0) = (0,0)$, che non consente di garantire la qualificazione dei vincoli. Se vediamo questa parte di $\mathcal{E}$ come $-x^2 \le y \le x^2$, e quindi descritta da due vincoli: $\begin{cases} h_1(x,y) = -y - x^2 \le 0 \\ h_2(x,y) = y - x^2 \le 0 \end{cases}$, essendo $\nabla h_1(0,0) = (0,-1)$ e $\nabla h_2(0,0) = (0,1)$, possiamo ricavare $\widetilde{\Gamma}(0,0) = \left\{(v_1,v_2) \in \mathbb{R}^2 : v_2 = 0, \forall v_1\right\} = \Gamma(0,0)$.

Nel punto $(0,0)$ i vincoli sono quindi qualificati.

Nei punti in cui è attivo un solo vincolo si verifica poi che il vincolo è sempre qualificato.

Costruiamo allora la Lagrangiana:

$\Lambda(x,y,\lambda_1,\lambda_2) = 2x - y - \lambda_1\left(y^2 - x^4\right) - \lambda_2\left(x^2 + y^2 - 2\right)$ e studiamo i quattro casi.

I caso: $\lambda_1 = 0, \lambda_2 = 0$: risolviamo il sistema:

$\begin{cases} \Lambda'_x = 0 \Rightarrow 2 = 0 \\ \Lambda'_y = 0 \Rightarrow -1 = 0 \\ \Lambda'_{\lambda_1} \ge 0 \Rightarrow y^2 - x^4 \le 0 \\ \Lambda'_{\lambda_2} \ge 0 \Rightarrow x^2 + y^2 - 2 \le 0 \end{cases}$. Il sistema non ha soluzioni.

II caso: $\lambda_1 \ne 0, \lambda_2 = 0$: risolviamo il sistema:

$$\begin{cases} \Lambda'_x = 0 \Rightarrow 2 + 4\lambda_1 x^3 = 0 \\ \Lambda'_y = 0 \Rightarrow -1 - 2\lambda_1 y = 0 \\ \Lambda'_{\lambda_1} = 0 \Rightarrow y^2 = x^4 \\ \Lambda'_{\lambda_2} \geq 0 \Rightarrow x^2 + y^2 \leq 2 \end{cases} \Rightarrow \begin{cases} x = -\sqrt[3]{\dfrac{1}{2\lambda_1}} \\ y = -\dfrac{1}{2\lambda_1} \\ \dfrac{1}{4\lambda_1^2} = \sqrt[3]{\dfrac{1}{16\lambda_1^4}} \\ x^2 + y^2 \leq 2 \end{cases} \Rightarrow \begin{cases} x = -\sqrt[3]{\dfrac{1}{2\lambda_1}} \\ y = -\dfrac{1}{2\lambda_1} \\ 16\lambda_1^4 (4\lambda_1^2 - 1) = 0 \\ x^2 + y^2 \leq 2 \end{cases}$$

. La soluzione $\lambda_1 = 0$ non è accettabile, per cui avremo:

$$\begin{cases} x = -\sqrt[3]{\dfrac{1}{2\lambda_1}} \\ y = -\dfrac{1}{2\lambda_1} \\ 4\lambda_1^2 - 1 = 0 \\ x^2 + y^2 \leq 2 \end{cases} \Rightarrow \begin{cases} x = -1 \\ y = -1 \\ \lambda_1 = \frac{1}{2} \\ 1 + 1 \leq 2 : \text{ vera} \end{cases} \text{ e } \begin{cases} x = 1 \\ y = 1 \\ \lambda_1 = -\frac{1}{2} \\ 1 + 1 \leq 2 : \text{ vera} \end{cases} .$$

Quindi il punto $(-1, -1)$ potrebbe essere un punto di massimo $(\lambda_1 > 0)$, mentre il punto $(1, 1)$ potrebbe essere un punto di minimo $(\lambda_1 < 0)$.

<u>III caso: $\lambda_1 = 0, \lambda_2 \neq 0$</u> : risolviamo il sistema:

$$\begin{cases} \Lambda'_x = 0 \Rightarrow 2 - 2\lambda_2 x = 0 \\ \Lambda'_y = 0 \Rightarrow -1 - 2\lambda_2 y = 0 \\ \Lambda'_{\lambda_1} \geq 0 \Rightarrow y^2 \leq x^4 \\ \Lambda'_{\lambda_2} = 0 \Rightarrow x^2 + y^2 = 2 \end{cases} \Rightarrow \begin{cases} x = \dfrac{1}{\lambda_2} \\ y = -\dfrac{1}{2\lambda_2} \\ y^2 \leq x^4 \\ \dfrac{1}{\lambda_2^2} + \dfrac{1}{4\lambda_2^2} = 2 \end{cases} \Rightarrow \begin{cases} x = \dfrac{1}{\lambda_2} \\ y = -\dfrac{1}{2\lambda_2} \\ y^2 \leq x^4 \\ \lambda_2^2 = \dfrac{5}{8} \end{cases}$$

che dà le due soluzioni:

$$\begin{cases} x = 2\sqrt{\frac{2}{5}} \\ y = -\sqrt{\frac{2}{5}} \\ \frac{2}{5} \leq \frac{64}{25} : \text{ vera} \\ \lambda_2 = \frac{1}{2}\sqrt{\frac{5}{2}} \end{cases} \text{ e } \begin{cases} x = -2\sqrt{\frac{2}{5}} \\ y = \sqrt{\frac{2}{5}} \\ \frac{2}{5} \leq \frac{64}{25} : \text{ vera} \\ \lambda_2 = -\frac{1}{2}\sqrt{\frac{5}{2}} \end{cases} .$$

Quindi il punto $\left(2\sqrt{\frac{2}{5}}, -\sqrt{\frac{2}{5}}\right)$ potrebbe essere un punto di massimo $(\lambda_2 > 0)$, mentre il punto $\left(-2\sqrt{\frac{2}{5}}, \sqrt{\frac{2}{5}}\right)$ potrebbe essere un punto di minimo $(\lambda_2 < 0)$.

<u>IV caso: $\lambda_1 \neq 0, \lambda_2 \neq 0$</u> : risolviamo il sistema:

$$\begin{cases} \Lambda'_x = 0 \Rightarrow 2 + 4\lambda_1 x^3 - 2\lambda_2 x = 0 \\ \Lambda'_y = 0 \Rightarrow -1 - 2\lambda_1 y - 2\lambda_2 y = 0 \\ \Lambda'_{\lambda_1} = 0 \Rightarrow y^2 = x^4 \\ \Lambda'_{\lambda_2} = 0 \Rightarrow x^2 + y^2 = 2 \end{cases}$$

dal quale otteniamo quattro sistemi:

$$\begin{cases} 2 + 4\lambda_1 - 2\lambda_2 = 0 \\ -1 - 2\lambda_1 - 2\lambda_2 = 0 \\ x = 1 \\ y = 1 \end{cases} \Rightarrow \begin{cases} \lambda_1 = -\frac{1}{2} \\ \lambda_2 = 0 \\ x = 1 \\ y = 1 \end{cases}$$

il punto $(1, 1)$ è già stato studiato: potrebbe essere un punto di minimo $(\lambda_1 < 0, \lambda_2 \leq 0)$;

$\begin{cases} 2+4\lambda_1-2\lambda_2=0 \\ -1+2\lambda_1+2\lambda_2=0 \\ x=1 \\ y=-1 \end{cases} \Rightarrow \begin{cases} \lambda_1=-\frac{1}{6} \\ \lambda_2=\frac{2}{3} \\ x=1 \\ y=-1 \end{cases}$ il punto non può essere nè di massimo nè di minimo in quanto i moltiplicatori hanno segno diverso;

$\begin{cases} 2-4\lambda_1+2\lambda_2=0 \\ -1-2\lambda_1-2\lambda_2=0 \\ x=-1 \\ y=1 \end{cases} \Rightarrow \begin{cases} \lambda_1=\frac{1}{6} \\ \lambda_2=-\frac{2}{3} \\ x=-1 \\ y=1 \end{cases}$ il punto non può essere nè di massimo nè di minimo in quanto i moltiplicatori hanno segno diverso;

$\begin{cases} 2-4\lambda_1+2\lambda_2=0 \\ -1+2\lambda_1+2\lambda_2=0 \\ x=-1 \\ y=-1 \end{cases} \Rightarrow \begin{cases} \lambda_1=\frac{1}{2} \\ \lambda_2=0 \\ x=-1 \\ y=-1 \end{cases}$ il punto $(-1,-1)$ è già stato studiato: potrebbe essere un punto di massimo $(\lambda_1>0, \lambda_2\geq 0)$.

Dato che $\mathcal{E}$ è un insieme limitato e chiuso e che la funzione $f(x,y)=2x-y$ è continua in tutto $\mathbb{R}^2$, devono esserci un massimo ed un minimo assoluti, e potrebbero essercene anche di relativi. Esaminiamo allora il comportamento di $f(x,y)$ sulla frontiera di $\mathcal{E}$.
Avremo sul vincolo $g_1: y^2=x^4 \Rightarrow y=\pm x^2$:
$f(x,x^2)=2x-x^2 \Rightarrow f'(x)=2-2x\geq 0$ per $x\leq 1$;
$f(x,-x^2)=2x+x^2 \Rightarrow f'(x)=2+2x\geq 0$ per $x\geq -1$.
Indicando con una freccia la direzione nella quale la funzione cresce, abbiamo che sulle due parabole la funzione cresce andando da sinistra verso destra, ovvero cresce sia nel tratto da $(-1,1)$ a $(1,1)$, che nel tratto da $(-1,-1)$ a $(1,-1)$.
Analizziamo infine il comportamento nei tratti relativi alla circonferenza.
Poniamo $\begin{cases} x=\sqrt{2}\cos t \\ y=\sqrt{2}\operatorname{sen} t \end{cases}$ ed avremo $f(x(t),y(t))=2\sqrt{2}\cos t-\sqrt{2}\operatorname{sen} t$ da cui poi:
$f'(t)=\sqrt{2}(-2\operatorname{sen} t-\cos t)\geq 0$ se $\cos t\leq -2\operatorname{sen} t$. Tale disequazione è verificata per $\alpha\leq t\leq\beta$, con $\cos\alpha=\dfrac{x}{\sqrt{2}}=-\dfrac{2}{\sqrt{5}}$ e $\operatorname{sen}\alpha=\dfrac{y}{\sqrt{2}}=\dfrac{1}{\sqrt{5}}$ e con $\cos\beta=\dfrac{x}{\sqrt{2}}=\dfrac{2}{\sqrt{5}}$ e $\operatorname{sen}\beta=\dfrac{y}{\sqrt{2}}=-\dfrac{1}{\sqrt{5}}$. Quindi, partendo dal punto $m=\left(-2\sqrt{\dfrac{2}{5}},\sqrt{\dfrac{2}{5}}\right)$ la funzione cresce in ambedue le direzioni, e quindi questo è un punto di minimo; al contrario, si arriva al punto $M=\left(2\sqrt{\dfrac{2}{5}},-\sqrt{\dfrac{2}{5}}\right)$ sempre crescendo, sia da sopra che da sotto, e questo è allora un punto di massimo. Nel punto $(1,1)$ si arriva crescendo sulla parabola e si continua a crescere ripartendo sulla circonferenza, quindi $(1,1)$ non è punto di minimo. Cosa analoga accade in $(-1,-1)$, che quindi non è un punto di massimo.
Quindi $f\left(2\sqrt{\dfrac{2}{5}},-\sqrt{\dfrac{2}{5}}\right)=5\sqrt{\dfrac{2}{5}}$ e $f\left(-2\sqrt{\dfrac{2}{5}},\sqrt{\dfrac{2}{5}}\right)=-5\sqrt{\dfrac{2}{5}}$ sono, rispettivamente, il massimo ed il minimo di $f(x,y)$ in $\mathcal{E}$.

CONDIZIONI DI KUHN-TUCKER CON VARIABILI NON NEGATIVE

In molti problemi, specialmente di natura economica, la ricerca degli estremi di una funzione si accompagna con la richiesta di non negatività delle variabili indipendenti e si può, in questo contesto, dare una formulazione diversa alle condizioni di Kuhn-Tucker. Sia dato il problema:
$\begin{cases} \text{Max } f(x_1, x_2, ..., x_n) \\ \text{s.v. : } \begin{cases} g_i(x_1, x_2, ..., x_n) \le 0, 1 \le i \le m \\ x_i \ge 0 \Rightarrow -x_i \le 0, 1 \le i \le n \end{cases} \end{cases}$. Posto $\mathbb{X} = (x_1, x_2, ..., x_n)$ e:

$$\widetilde{\Lambda}\,(\mathbb{X}, \lambda_1, .., \lambda_m, \mu_1, .., \mu_n) = f(\mathbb{X}) - \sum_{i:1}^{m} \lambda_i\, g_i(\mathbb{X}) - \sum_{i:1}^{n} \mu_i\,(-x_i) = \Lambda + \sum_{i:1}^{n} \mu_i\, x_i$$

imponendo le condizioni di Kuhn-Tucker per un punto di massimo sia sui moltiplicatori λ_i che sui μ_i, otteniamo:

$$\begin{cases} \dfrac{\partial \widetilde{\Lambda}}{\partial x_i} = \dfrac{\partial \Lambda}{\partial x_i} + \mu_i = 0 \Rightarrow \mu_i = -\dfrac{\partial \Lambda}{\partial x_i} \\ \dfrac{\partial \widetilde{\Lambda}}{\partial \lambda_i} = \dfrac{\partial \Lambda}{\partial \lambda_i} = -g_i(\mathbb{X}) \ge 0 \\ \dfrac{\partial \widetilde{\Lambda}}{\partial \mu_i} = x_i \ge 0 \end{cases}.$$

Dovendo poi imporre $\mu_i \ge 0$, riformuliamo la prima condizione come $\dfrac{\partial \Lambda}{\partial x_i} \le 0$.

Dovremo poi imporre le altre condizioni: $\begin{cases} \lambda_i \cdot g_i(\mathbb{X}) = 0 \\ \mu_i \cdot x_i = 0 \\ \lambda_i \ge 0 \\ \mu_i \ge 0 \end{cases}$. Essendo $\mu_i = -\dfrac{\partial \Lambda}{\partial x_i}$, la seconda condizione viene espressa anche nella forma: $\dfrac{\partial \Lambda}{\partial x_i} \cdot x_i = 0$, così come la quarta, che è equivalente alla $\dfrac{\partial \Lambda}{\partial x_i} \le 0$ già vista in precedenza. Per quanto detto, le condizioni di Kuhn-Tucker per la ricerca del massimo, sotto ai vincoli $g_i(\mathbb{X}) \le 0$ e sotto la condizione di non negatività delle variabili indipendenti, sono esprimibili anche come $\begin{cases} \dfrac{\partial \Lambda}{\partial x_i} \le 0 \\ g_i(\mathbb{X}) \le 0 \\ x_i \ge 0 \\ \lambda_i \cdot g_i(\mathbb{X}) = 0 \\ \dfrac{\partial \Lambda}{\partial x_i} \cdot x_i = 0 \\ \lambda_i \ge 0 \end{cases}$.

APPENDICE 1

Teorema del differenziale totale- Sua dimostrazione e controesempio

Teorema (del differenziale totale) Se la funzione ammette in $\mathbb{X}_0$ tutte le derivate parziali, ovvero se esiste $\nabla f(\mathbb{X}_0)$, e se le funzioni derivate parziali sono continue in $\mathbb{X}_0$, allora f è differenziabile in $\mathbb{X}_0$.
Dimostrazione : Dimostriamo il Teorema nel caso di una funzione di due sole variabili, essendo analoga, anche se con opportuni adattamenti, la dimostrazione nel caso di una funzione di quante si vogliano variabili. Data $f(x, y)$, poniamo:
$f(\mathbb{X}) - f(\mathbb{X}_0) = f(x_0 + h, y_0 + k) - f(x_0, y_0)$.
Sommando e sottraendo $f(x_0, y_0 + k)$ otteniamo:
$f(x_0 + h, y_0 + k) - f(x_0, y_0) =$
$= f(x_0 + h, y_0 + k) - f(x_0, y_0 + k) + f(x_0, y_0 + k) - f(x_0, y_0)$.
Essendo f derivabile rispetto a x, per il Teorema di Lagrange avremo:
$f(x_0 + h, y_0 + k) - f(x_0, y_0 + k) = f'_x(x_0 + \theta_1 h, y_0 + k) \cdot h$, con $0 \leq \theta_1 \leq 1$,
e analogamente, essendo f derivabile rispetto a y:
$f(x_0, y_0 + k) - f(x_0, y_0) = f'_y(x_0, y_0 + \theta_2 k) \cdot k$, con $0 \leq \theta_2 \leq 1$.
Dall'ipotesi di continuità di f'_x e di f'_y in (x_0, y_0) possiamo anche scrivere:
$f'_x(x_0 + \theta_1 h, y_0 + k) \cdot h = [f'_x(x_0, y_0) + \varepsilon_1(h, k)] \cdot h$, e similmente:
$f'_y(x_0, y_0 + \theta_2 k) \cdot k = \left[f'_y(x_0, y_0) + \varepsilon_2(k)\right] \cdot k$,
con $\varepsilon_1(h, k) \to 0$ e con $\varepsilon_2(k) \to 0$ quando $(h, k) \to (0, 0)$. Ma allora:
$f(x_0 + h, y_0 + k) - f(x_0, y_0) = f'_x(x_0, y_0)\, h + f'_y(x_0, y_0)\, k + h\, \varepsilon_1(h, k) + k\, \varepsilon_2(k)$.
Se facciamo vedere che $h\, \varepsilon_1(h, k) + k\, \varepsilon_2(k) = o\left(\sqrt{h^2 + k^2}\right)$ la dimostrazione è conclusa.
Ma $\left|\dfrac{h\, \varepsilon_1(h, k) + k\, \varepsilon_2(k)}{\sqrt{h^2 + k^2}}\right| \leq \dfrac{|h| \cdot |\varepsilon_1(h, k)|}{\sqrt{h^2 + k^2}} + \dfrac{|k| \cdot |\varepsilon_2(k)|}{\sqrt{h^2 + k^2}} \leq |\varepsilon_1(h, k)| + |\varepsilon_2(k)|$, in quanto $\dfrac{|h|}{\sqrt{h^2 + k^2}} \leq 1$ e $\dfrac{|k|}{\sqrt{h^2 + k^2}} \leq 1$.
Dato che $\varepsilon_1(h, k) \to 0$ e $\varepsilon_2(k) \to 0$ quando $(h, k) \to (0, 0)$, la dimostrazione è conclusa.●
Osserviamo comunque come l'avere derivate parziali continue in $\mathbb{X}_0$ sia condizione solo sufficiente, ma non necessaria, per essere funzione differenziabile, come si vede nel seguente:

Esempio 89 : Sia $f(x, y) = \begin{cases} \left(x^2 + y^2\right) \cdot \text{sen}\, \dfrac{1}{\sqrt{x^2 + y^2}} & (x, y) \neq (0, 0) \\ 0 & (x, y) = (0, 0) \end{cases}$.
Vediamo che $f(x, y)$ è continua e differenziabile in $(0, 0)$, ma che le sue funzioni derivate parziali $f'_x(x, y)$ e $f'_y(x, y)$ non sono continue in $(0, 0)$.
Vediamo che $\lim\limits_{(x,y) \to (0,0)} f(x, y) = 0 = f(0, 0)$. Infatti:
$\left|\left(x^2 + y^2\right) \cdot \text{sen}\, \dfrac{1}{\sqrt{x^2 + y^2}} - 0\right| = \left|x^2 + y^2\right| \cdot \left|\text{sen}\, \dfrac{1}{\sqrt{x^2 + y^2}}\right| \leq \left|x^2 + y^2\right|$. Ma
$\left|x^2 + y^2\right| = \|\mathbb{X}\|^2 < \varepsilon$ se $0 < \|\mathbb{X}\| < \sqrt{\varepsilon} = \delta(\varepsilon)$, intorno di $(0, 0)$.
Quindi $f(x, y)$ è continua in $(0, 0)$. Calcoliamo poi $\dfrac{\partial f(0, 0)}{\partial x}$ e $\dfrac{\partial f(0, 0)}{\partial y}$. Avremo:

$$\frac{\partial f(0,0)}{\partial x}=\lim_{h\to 0}\frac{\left((0+h)^2+0^2\right)\cdot \operatorname{sen}\dfrac{1}{\sqrt{(0+h)^2+0^2}}-0}{h}=\lim_{h\to 0}\frac{h^2}{h}\cdot \operatorname{sen}\frac{1}{\sqrt{h^2}}=0\,,$$

in quanto $h\to 0$ mentre $\operatorname{sen}\dfrac{1}{\sqrt{h^2}}$ è una quantità limitata; similmente:

$$\frac{\partial f(0,0)}{\partial y}=\lim_{h\to 0}\frac{\left(0^2+(0+h)^2\right)\cdot \operatorname{sen}\dfrac{1}{\sqrt{0^2+(0+h)^2}}-0}{h}=\lim_{h\to 0}\frac{h^2}{h}\cdot \operatorname{sen}\frac{1}{\sqrt{h^2}}=0\,.$$

Vediamo poi che la funzione è differenziabile in $(0,0)$. Dovrà risultare:

$$f(x,y)-f(0,0)=\frac{\partial f(0,0)}{\partial x}\cdot(x-0)+\frac{\partial f(0,0)}{\partial y}\cdot(y-0)+o\left(\sqrt{(x-0)^2+(y-0)^2}\right)$$

ovvero: $\left(x^2+y^2\right)\cdot \operatorname{sen}\dfrac{1}{\sqrt{x^2+y^2}}-0=0\cdot x+0\cdot y+o\left(\sqrt{x^2+y^2}\right)$, ovvero

$\left(x^2+y^2\right)\cdot \operatorname{sen}\dfrac{1}{\sqrt{x^2+y^2}}=o\left(\sqrt{x^2+y^2}\right)$, e questo è vero in quanto:

$$\lim_{(x,y)\to(0,0)}\frac{\left(x^2+y^2\right)\cdot \operatorname{sen}\dfrac{1}{\sqrt{x^2+y^2}}}{\sqrt{x^2+y^2}}=\lim_{(x,y)\to(0,0)}\sqrt{x^2+y^2}\cdot \operatorname{sen}\frac{1}{\sqrt{x^2+y^2}}=0$$

in quanto $\left|\sqrt{x^2+y^2}\cdot \operatorname{sen}\dfrac{1}{\sqrt{x^2+y^2}}\right|\le\sqrt{x^2+y^2}=\|\mathbb{X}\|$, e preso $\|\mathbb{X}\|<\varepsilon=\delta(\varepsilon)$ ne segue che $|f(x,y)-0|<\varepsilon$, ovvero il limite vale 0. Quindi $f(x,y)$ è differenziabile in $(0,0)$.

Calcoliamo ora la funzione derivata parziale $\dfrac{\partial f(x,y)}{\partial x}$. Risulta:

$$\frac{\partial f(x,y)}{\partial x}=2x\cdot \operatorname{sen}\frac{1}{\sqrt{x^2+y^2}}-\frac{x}{\sqrt{x^2+y^2}}\cdot\cos\frac{1}{\sqrt{x^2+y^2}}\,.$$

Passando al calcolo del $\displaystyle\lim_{(x,y)\to(0,0)}\frac{\partial f(x,y)}{\partial x}$ vediamo che tale limite non esiste, in quanto $2x\cdot \operatorname{sen}\dfrac{1}{\sqrt{x^2+y^2}}\to 0$, mentre sia $\cos\dfrac{1}{\sqrt{x^2+y^2}}$ che $\dfrac{x}{\sqrt{x^2+y^2}}$ non ammettono limite per $(x,y)\to(0,0)$, e quindi la funzione $\dfrac{\partial f(x,y)}{\partial x}$ non è continua in $(0,0)$. Risultato analogo si ottiene esaminando la continuità di $\dfrac{\partial f(x,y)}{\partial y}$ in $(0,0)$.

APPENDICE 2

Derivate seconde e Differenziali I e II di funzioni implicite $\mathbb{R}\to\mathbb{R}^2$

Funzione implicita $\mathbb{R}\to\mathbb{R}^2$: Derivate seconde

Dal sistema $\begin{cases} w_1=f(x,y_1,y_2)=k_1\\ w_2=g(x,y_1,y_2)=k_2\end{cases}$, posto $(x|\,\mathbb{Y})=(x,y_1,y_2)$, abbiamo queste due composizioni di funzioni:

$\mathbb{R}\to\mathbb{R}^3\xrightarrow{f}\mathbb{R},\ x\to(x|\,\mathbb{Y})\xrightarrow{f}w_1=f(x|\,\mathbb{Y})=k_1$, e

$\mathbb{R}\to\mathbb{R}^3\xrightarrow{g}\mathbb{R},\ x\to(x|\,\mathbb{Y})\xrightarrow{g}w_2=g(x|\,\mathbb{Y})=k_2$.

Derivando rispetto a x si ha:
$\dfrac{\mathrm{d}(x|\,\mathbb{Y})}{\mathrm{d}x} = (1, y_1', y_2')$ e $\dfrac{\mathrm{d}^2(x|\,\mathbb{Y})}{\mathrm{d}x^2} = (0, y_1'', y_2'')$ dalle quali otteniamo, per il Teorema 18:

$$\frac{\mathrm{d}^2 w_1}{\mathrm{d}x^2} = \frac{\mathrm{d}(x|\,\mathbb{Y})}{\mathrm{d}x} \cdot \mathbb{H}(f(x|\,\mathbb{Y})) \cdot \left(\frac{\mathrm{d}(x|\,\mathbb{Y})}{\mathrm{d}x}\right)^{\mathrm{T}} + \left(f_x', f_{y_1}', f_{y_2}'\right) \cdot (0, y_1'', y_2'') = 0 \text{ e}$$

$$\frac{\mathrm{d}^2 w_2}{\mathrm{d}x^2} = \frac{\mathrm{d}(x|\,\mathbb{Y})}{\mathrm{d}x} \cdot \mathbb{H}(g(x|\,\mathbb{Y})) \cdot \left(\frac{\mathrm{d}(x|\,\mathbb{Y})}{\mathrm{d}x}\right)^{\mathrm{T}} + \left(g_x', g_{y_1}', g_{y_2}'\right) \cdot (0, y_1'', y_2'') = 0$$

ovvero il sistema, avente per incognite y_1'' e y_2'':

$$\begin{cases} f_{y_1}'\, y_1'' + f_{y_2}'\, y_2'' = -\dfrac{\mathrm{d}(x|\,\mathbb{Y})}{\mathrm{d}x} \cdot \mathbb{H}(f(x|\,\mathbb{Y})) \cdot \left(\dfrac{\mathrm{d}(x|\,\mathbb{Y})}{\mathrm{d}x}\right)^{\mathrm{T}} \\ g_{y_1}'\, y_1'' + g_{y_2}'\, y_2'' = -\dfrac{\mathrm{d}(x|\,\mathbb{Y})}{\mathrm{d}x} \cdot \mathbb{H}(g(x|\,\mathbb{Y})) \cdot \left(\dfrac{\mathrm{d}(x|\,\mathbb{Y})}{\mathrm{d}x}\right)^{\mathrm{T}} \end{cases}$$

la cui soluzione può essere formalmente espressa mediante la:

$$\left\|\begin{matrix} y_1'' \\ y_2'' \end{matrix}\right\| = -\left\|\frac{\partial(f,g)}{\partial(y_1,y_2)}\right\|^{-1} \cdot \left\|\begin{matrix} \dfrac{\mathrm{d}(x|\,\mathbb{Y})}{\mathrm{d}x} \cdot \mathbb{H}(f(x|\,\mathbb{Y})) \cdot \left(\dfrac{\mathrm{d}(x|\,\mathbb{Y})}{\mathrm{d}x}\right)^{\mathrm{T}} \\ \dfrac{\mathrm{d}(x|\,\mathbb{Y})}{\mathrm{d}x} \cdot \mathbb{H}(g(x|\,\mathbb{Y})) \cdot \left(\dfrac{\mathrm{d}(x|\,\mathbb{Y})}{\mathrm{d}x}\right)^{\mathrm{T}} \end{matrix}\right\|.$$

Funzione implicita $\mathbb{R} \to \mathbb{R}^2$: Differenziali I e II

Dal sistema $\begin{cases} w_1 = f(x, y_1, y_2) = k_1 \\ w_2 = g(x, y_1, y_2) = k_2 \end{cases}$, differenziando una prima volta otteniamo:
$\begin{cases} \mathrm{d}w_1 = f_x'\,\mathrm{d}x + f_{y_1}'\mathrm{d}y_1 + f_{y_2}'\mathrm{d}y_2 = 0 \\ \mathrm{d}w_2 = g_x'\,\mathrm{d}x + g_{y_1}'\mathrm{d}y_1 + g_{y_2}'\mathrm{d}y_2 = 0 \end{cases}$, ovvero un sistema lineare di due equazioni nelle due incognite $\mathrm{d}y_1$ e $\mathrm{d}y_2$ che ha per soluzione: $\left\|\begin{matrix} \mathrm{d}y_1 \\ \mathrm{d}y_2 \end{matrix}\right\| = -\left\|\dfrac{\partial(f,g)}{\partial(y_1,y_2)}\right\|^{-1} \cdot \left\|\begin{matrix} f_x'\,\mathrm{d}x \\ g_x'\,\mathrm{d}x \end{matrix}\right\|$.

Differenziando una seconda volta, generalizzando quanto visto nel caso di una sola equazione, avremo: $\begin{cases} \mathrm{d}^2 w_1 = \mathrm{d}^2 f(x, y_1, y_2) + f_{y_1}'\mathrm{d}^2 y_1 + f_{y_2}'\mathrm{d}^2 y_2 = 0 \\ \mathrm{d}^2 w_2 = \mathrm{d}^2 g(x, y_1, y_2) + g_{y_1}'\mathrm{d}^2 y_1 + g_{y_2}'\mathrm{d}^2 y_2 = 0 \end{cases}$, ovvero un sistema lineare di due equazioni nelle due incognite $\mathrm{d}^2 y_1$ e $\mathrm{d}^2 y_2$ che ha per soluzione:

$$\left\|\begin{matrix} \mathrm{d}^2 y_1 \\ \mathrm{d}^2 y_2 \end{matrix}\right\| = \left\|\begin{matrix} y_1''(\mathrm{d}x)^2 \\ y_2''(\mathrm{d}x)^2 \end{matrix}\right\| = -\left\|\frac{\partial(f,g)}{\partial(y_1,y_2)}\right\|^{-1} \cdot \left\|\begin{matrix} \mathrm{d}^2 f(x, y_1, y_2) \\ \mathrm{d}^2 g(x, y_1, y_2) \end{matrix}\right\|.$$

APPENDICE 3

Derivate seconde e Differenziali I e II di funzioni implicite $\mathbb{R}^2 \to \mathbb{R}^2$

Funzione implicita $\mathbb{R}^2 \to \mathbb{R}^2$: Derivate seconde

Da $\begin{cases} w_1 = f(x_1, x_2, y_1, y_2) = k_1 \\ w_2 = g(x_1, x_2, y_1, y_2) = k_2 \end{cases}$, dalle due composizioni di funzioni:

$\mathbb{R}^2 \to \mathbb{R}^4 \overset{f}{\to} \mathbb{R},\ \mathbb{X} \to (\mathbb{X}|\,\mathbb{Y}) \overset{f}{\to} w_1 = f(\mathbb{X}|\,\mathbb{Y}) = k_1$, e

$\mathbb{R}^2 \to \mathbb{R}^4 \overset{g}{\to} \mathbb{R},\ \mathbb{X} \to (\mathbb{X}|\,\mathbb{Y}) \overset{g}{\to} w_2 = g(\mathbb{X}|\,\mathbb{Y}) = k_2$,

essendo $\dfrac{\partial(\mathbb{X}|\,\mathbb{Y})}{\partial x_1} = \left(1, 0, \dfrac{\partial y_1}{\partial x_1}, \dfrac{\partial y_2}{\partial x_1}\right)$ e $\dfrac{\partial(\mathbb{X}|\,\mathbb{Y})}{\partial x_2} = \left(0, 1, \dfrac{\partial y_1}{\partial x_2}, \dfrac{\partial y_2}{\partial x_2}\right)$,

derivando nuovamente si ha:

$\dfrac{\partial^2(\mathbb{X}|\,\mathbb{Y})}{\partial x_1^2} = \left(0,0,\dfrac{\partial^2 y_1}{\partial x_1^2},\dfrac{\partial^2 y_2}{\partial x_1^2}\right)$, $\dfrac{\partial^2(\mathbb{X}|\,\mathbb{Y})}{\partial x_2^2} = \left(0,0,\dfrac{\partial^2 y_1}{\partial x_2^2},\dfrac{\partial^2 y_2}{\partial x_2^2}\right)$,
$\dfrac{\partial^2(\mathbb{X}|\,\mathbb{Y})}{\partial x_1 \partial x_2} = \dfrac{\partial^2(\mathbb{X}|\,\mathbb{Y})}{\partial x_2 \partial x_1} = \left(0,0,\dfrac{\partial^2 y_1}{\partial x_1 \partial x_2},\dfrac{\partial^2 y_2}{\partial x_1 \partial x_2}\right)$, per le quali otteniamo, applicando il Teorema 18, 4 sistemi di 2 equazioni in 2 incognite, che si riducono a 3 sistemi di 2 equazioni in 2 incognite dato che $\dfrac{\partial^2 y_k}{\partial x_i \partial x_j} = \dfrac{\partial^2 y_k}{\partial x_j \partial x_i}$:

$$\begin{cases} \dfrac{\partial^2 w_1}{\partial x_i \partial x_j} = \dfrac{\partial(\mathbb{X}|\,\mathbb{Y})}{\partial x_i} \cdot \mathbb{H}(f(\mathbb{X}|\,\mathbb{Y})) \cdot \left(\dfrac{\partial(\mathbb{X}|\,\mathbb{Y})}{\partial x_j}\right)^{\mathrm{T}} + \nabla f(\mathbb{X}|\,\mathbb{Y}) \cdot \dfrac{\partial^2(\mathbb{X}|\,\mathbb{Y})}{\partial x_i \partial x_j} = 0 \\ \dfrac{\partial^2 w_2}{\partial x_i \partial x_j} = \dfrac{\partial(\mathbb{X}|\,\mathbb{Y})}{\partial x_i} \cdot \mathbb{H}(g(\mathbb{X}|\,\mathbb{Y})) \cdot \left(\dfrac{\partial(\mathbb{X}|\,\mathbb{Y})}{\partial x_j}\right)^{\mathrm{T}} + \nabla g(\mathbb{X}|\,\mathbb{Y}) \cdot \dfrac{\partial^2(\mathbb{X}|\,\mathbb{Y})}{\partial x_i \partial x_j} = 0 \end{cases},$$

$1 \le i,j \le 2$, le cui soluzioni sono esprimibili in forma compatta come:

$$\left\| \begin{matrix} \dfrac{\partial^2 y_1}{\partial x_i \partial x_j} \\ \dfrac{\partial^2 y_2}{\partial x_i \partial x_j} \end{matrix} \right\| = -\left\| \dfrac{\partial(f,g)}{\partial(y_1,y_2)} \right\|^{-1} \cdot \left\| \begin{matrix} \dfrac{\partial(\mathbb{X}|\,\mathbb{Y})}{\partial x_i} \cdot \mathbb{H}(f(\mathbb{X}|\,\mathbb{Y})) \cdot \left(\dfrac{\partial(\mathbb{X}|\,\mathbb{Y})}{\partial x_j}\right)^{\mathrm{T}} \\ \dfrac{\partial(\mathbb{X}|\,\mathbb{Y})}{\partial x_i} \cdot \mathbb{H}(g(\mathbb{X}|\,\mathbb{Y})) \cdot \left(\dfrac{\partial(\mathbb{X}|\,\mathbb{Y})}{\partial x_j}\right)^{\mathrm{T}} \end{matrix} \right\|,$$

$1 \le i,j \le 2\,.$

Funzione implicita $\mathbb{R}^2 \to \mathbb{R}^2$: Differenziali I e II

Da $\begin{cases} w_1 = f(x_1,x_2,y_1,y_2) = k_1 \\ w_2 = g(x_1,x_2,y_1,y_2) = k_2 \end{cases}$, dalle due composizioni di funzioni:

$\mathbb{R}^2 \to \mathbb{R}^4 \xrightarrow{f} \mathbb{R},\ \mathbb{X} \to (\mathbb{X}|\,\mathbb{Y}) \xrightarrow{f} w_1 = f(\mathbb{X}|\,\mathbb{Y}) = k_1$, e
$\mathbb{R}^2 \to \mathbb{R}^4 \xrightarrow{g} \mathbb{R},\ \mathbb{X} \to (\mathbb{X}|\,\mathbb{Y}) \xrightarrow{g} w_2 = g(\mathbb{X}|\,\mathbb{Y}) = k_2$, differenziando una prima volta si ha:
$\begin{cases} \mathrm{d}w_1 = f'_{x_1}\,\mathrm{d}x_1 + f'_{x_2}\,\mathrm{d}x_2 + f'_{y_1}\mathrm{d}y_1 + f'_{y_2}\mathrm{d}y_2 = 0 \\ \mathrm{d}w_2 = g'_{x_1}\,\mathrm{d}x_1 + g'_{x_2}\,\mathrm{d}x_2 + g'_{y_1}\mathrm{d}y_1 + g'_{y_2}\mathrm{d}y_2 = 0 \end{cases}$, ovvero un sistema lineare di due equazioni nelle due incognite $\mathrm{d}y_1$ e $\mathrm{d}y_2$ che ha per soluzione:

$$\left\| \begin{matrix} \mathrm{d}y_1 \\ \mathrm{d}y_2 \end{matrix} \right\| = -\left\| \dfrac{\partial(f,g)}{\partial(y_1,y_2)} \right\|^{-1} \cdot \left\| \begin{matrix} f'_{x_1}\,\mathrm{d}x_1 + f'_{x_2}\,\mathrm{d}x_2 \\ g'_{x_1}\,\mathrm{d}x_1 + g'_{x_2}\,\mathrm{d}x_2 \end{matrix} \right\|.$$

Differenziando una seconda volta, generalizzando quanto visto in precedenza nel caso di funzione implicita $\mathbb{R}^2 \to \mathbb{R}$, avremo:
$\begin{cases} \mathrm{d}^2 w_1 = \mathrm{d}^2 f(\mathbb{X}|\,\mathbb{Y}) + f'_{y_1}\mathrm{d}^2 y_1 + f'_{y_2}\mathrm{d}^2 y_2 = 0 \\ \mathrm{d}^2 w_2 = \mathrm{d}^2 g(\mathbb{X}|\,\mathbb{Y}) + g'_{y_1}\mathrm{d}^2 y_1 + g'_{y_2}\mathrm{d}^2 y_2 = 0 \end{cases}$, ovvero un sistema lineare di due equazioni nelle due incognite $\mathrm{d}^2 y_1$ e $\mathrm{d}^2 y_2$ che ha per soluzione:

$$\left\| \begin{matrix} \mathrm{d}^2 y_1 \\ \mathrm{d}^2 y_2 \end{matrix} \right\| = -\left\| \dfrac{\partial(f,g)}{\partial(y_1,y_2)} \right\|^{-1} \cdot \left\| \begin{matrix} \mathrm{d}^2 f(\mathbb{X}|\,\mathbb{Y}) \\ \mathrm{d}^2 g(\mathbb{X}|\,\mathbb{Y}) \end{matrix} \right\|.$$

APPENDICE 4

Derivate seconde e Differenziali I e II di funzioni implicite $\mathbb{R}^n \longrightarrow \mathbb{R}^m$

Funzione implicita $\mathbb{R}^n \to \mathbb{R}^m$: Derivate seconde

Generalizzando quanto visto nei casi precedenti, dalla composizione di funzioni:
$\mathbb{R}^n \to \mathbb{R}^{n+m} \xrightarrow{f} \mathbb{R}^m,\ \mathbb{X} \to (\mathbb{X}|\,\mathbb{Y}) \xrightarrow{f} \mathbb{W} = f(\mathbb{X}|\,\mathbb{Y}) = \text{cost.}$,

applicando il Teorema 18, otteniamo $\dfrac{n^2+n}{2}$ sistemi di m equazioni in m incognite, le cui soluzioni sono esprimibili in forma compatta mediante le:

$$\frac{\partial^2(y_1,y_2,...,y_m)}{\partial x_i\,\partial x_j} = -\left\|\frac{\partial(f_1,f_2,..,f_m)}{\partial(y_1,y_2,...,y_m)}\right\|^{-1}\cdot\left\|\begin{array}{c}\dfrac{\partial(\mathbb{X}|\,\mathbb{Y})}{\partial x_i}\cdot\mathbb{H}(f_1(\mathbb{X}|\,\mathbb{Y}))\cdot\left(\dfrac{\partial(\mathbb{X}|\,\mathbb{Y})}{\partial x_j}\right)^{\mathrm{T}}\\ \cdots\cdots\cdots\cdots\cdots\cdots\\ \dfrac{\partial(\mathbb{X}|\,\mathbb{Y})}{\partial x_i}\cdot\mathbb{H}(f_m(\mathbb{X}|\,\mathbb{Y}))\cdot\left(\dfrac{\partial(\mathbb{X}|\,\mathbb{Y})}{\partial x_j}\right)^{\mathrm{T}}\end{array}\right\|,$$

$1\le i,j\le n$, dove $\dfrac{\partial^2(y_1,y_2,...,y_m)}{\partial x_i\,\partial x_j}$ e $\left\|\begin{array}{c}\dfrac{\partial(\mathbb{X}|\,\mathbb{Y})}{\partial x_i}\cdot\mathbb{H}(f_1(\mathbb{X}|\,\mathbb{Y}))\cdot\left(\dfrac{\partial(\mathbb{X}|\,\mathbb{Y})}{\partial x_j}\right)^{\mathrm{T}}\\ \cdots\cdots\cdots\cdots\cdots\cdots\\ \dfrac{\partial(\mathbb{X}|\,\mathbb{Y})}{\partial x_i}\cdot\mathbb{H}(f_m(\mathbb{X}|\,\mathbb{Y}))\cdot\left(\dfrac{\partial(\mathbb{X}|\,\mathbb{Y})}{\partial x_j}\right)^{\mathrm{T}}\end{array}\right\|$ sono vettori colonna $(m\cdot 1)$, mentre $\left\|\dfrac{\partial(f_1,f_2,..,f_m)}{\partial(y_1,y_2,...,y_m)}\right\|^{-1}$ è una matrice $(m\cdot m)$.

Funzione implicita $\mathbb{R}^n\to\mathbb{R}^m$: Differenziali I e II

Da $\begin{cases} w_1 = f_1(x_1,x_2,...,x_n,y_1,y_2,...,y_m) = k_1\\ w_2 = f_2(x_1,x_2,...,x_n,y_1,y_2,...,y_m) = k_2\\ \cdots\cdots\cdots\cdots\cdots\cdots\\ w_m = f_m(x_1,x_2,...,x_n,y_1,y_2,...,y_m) = k_m\end{cases}$ dalla composizione di funzioni:

$\mathbb{R}^n\to\mathbb{R}^{n+m}\xrightarrow{f}\mathbb{R}^m$, $\mathbb{X}\to(\mathbb{X}|\,\mathbb{Y})\xrightarrow{f}\mathbb{W}=f(\mathbb{X}|\,\mathbb{Y})=$ cost., differenziando una volta si ottiene un sistema di m equazioni nelle m incognite $\mathrm{d}y_1,...,\mathrm{d}y_m$ esprimibile come:

$$\left\|\begin{array}{c}\mathrm{d}w_1\\ \cdots\\ \mathrm{d}w_m\end{array}\right\| = \left\|\frac{\partial(f_1,f_2,..,f_m)}{\partial(x_1,x_2,...,x_n)}\right\|\cdot\left\|\begin{array}{c}\mathrm{d}x_1\\ \cdots\\ \mathrm{d}x_n\end{array}\right\| + \left\|\frac{\partial(f_1,f_2,..,f_m)}{\partial(y_1,y_2,...,y_m)}\right\|\cdot\left\|\begin{array}{c}\mathrm{d}y_1\\ \cdots\\ \mathrm{d}y_m\end{array}\right\| = \mathbb{O}$$

che ha soluzione: $\left\|\begin{array}{c}\mathrm{d}y_1\\ \cdots\\ \mathrm{d}y_m\end{array}\right\| = -\left\|\dfrac{\partial(f_1,f_2,..,f_m)}{\partial(y_1,y_2,...,y_m)}\right\|^{-1}\cdot\left\|\dfrac{\partial(f_1,f_2,..,f_m)}{\partial(x_1,x_2,...,x_n)}\right\|\cdot\left\|\begin{array}{c}\mathrm{d}x_1\\ \cdots\\ \mathrm{d}x_n\end{array}\right\|.$

Differenziando una seconda volta si ottiene un sistema, di m equazioni nelle m incognite $\mathrm{d}^2y_1,...,\mathrm{d}^2y_m$, esprimibile come:

$$\left\|\begin{array}{c}\mathrm{d}^2w_1\\ \cdots\\ \mathrm{d}^2w_m\end{array}\right\| = \left\|\begin{array}{c}\mathrm{d}^2f_1(\mathbb{X}|\,\mathbb{Y})\\ \cdots\\ \mathrm{d}^2f_m(\mathbb{X}|\,\mathbb{Y})\end{array}\right\| + \left\|\frac{\partial(f_1,f_2,..,f_m)}{\partial(y_1,y_2,...,y_m)}\right\|\cdot\left\|\begin{array}{c}\mathrm{d}^2y_1\\ \cdots\\ \mathrm{d}^2y_m\end{array}\right\| = \mathbb{O}$$

che ha per soluzione:

$$\left\|\begin{array}{c}\mathrm{d}^2y_1\\ \cdots\\ \mathrm{d}^2y_m\end{array}\right\| = -\left\|\frac{\partial(f_1,f_2,..,f_m)}{\partial(y_1,y_2,...,y_m)}\right\|^{-1}\cdot\left\|\begin{array}{c}\mathrm{d}^2f_1(\mathbb{X}|\,\mathbb{Y})\\ \cdots\\ \mathrm{d}^2f_m(\mathbb{X}|\,\mathbb{Y})\end{array}\right\|.$$