# Funzioni di più variabili 3

March 18, 2021

# Derivazione delle funzioni composte

Se $f : \mathbb{R}^n \to \mathbb{R}$ è *differenziabile* in $\mathbf{x}_0$ e $g : \mathbb{R} \to \mathbb{R}$ è derivabile in $f(\mathbf{x}_0)$, la funzione $g \circ f : \mathbb{R}^n \to \mathbb{R}$ è differenziabile in $\mathbf{x}_0$ e vale

$$\nabla(g \circ f)(\mathbf{x}_0) = g'\big(f(\mathbf{x}_0)\big)\,\nabla f(\mathbf{x}_0)\,.$$

Se $\mathbf{r} : \mathbb{R} \to \mathbb{R}^m$ è derivabile in $t_0$ e $f : \mathbb{R}^m \to \mathbb{R}$ è *differenziabile* in $\mathbf{r}(t_0)$, la funzione $f \circ \mathbf{r} : \mathbb{R} \to \mathbb{R}$ è derivabile in $t_0$ e vale

$$(f \circ \mathbf{r})'(t_0) = \nabla f(\mathbf{r}(t_0)) \cdot \mathbf{r}'(t_0)\,.$$

Gradiente di funzioni radiali.

La funzione $\mathbf{x} \mapsto |\mathbf{x}|$ è differenziabile in $\mathbb{R}^n \setminus \{\mathbf{0}\}$ (le derivate parziali sono continue al di fuori dell'origine).

Se $g$ è derivabile in $\mathbb{R}_+$, la funzione radiale $g(|\mathbf{x}|)$ è differenziabile in $\mathbb{R}^n \setminus \{\mathbf{0}\}$ e vale

$$\nabla g(|\mathbf{x}|) = g'(|\mathbf{x}|)\, \nabla |\mathbf{x}| = g'(|\mathbf{x}|)\, \frac{\mathbf{x}}{|\mathbf{x}|}\,.$$

In particolare, se $n = 3$, $\; r = \sqrt{x^2 + y^2 + z^2}$, $\; g(r) = \dfrac{k}{r}\;$ :

$$\nabla g(r) = -\frac{k}{r^2}\Big(\frac{x}{r}\,\mathbf{i} + \frac{y}{r}\,\mathbf{j} + \frac{z}{r}\,\mathbf{k}\Big) \; = -k\,\frac{\mathbf{r}}{r^3} \qquad (\mathbf{r} = x\,\mathbf{i} + y\,\mathbf{j} + z\,\mathbf{k}).$$

Gradiente e curve di livello.

Sia $f(x, y)$ differenziabile (in un aperto di $\mathbb{R}^2$) e supponiamo che l'insieme di livello

$$\{(x, y) \mid f(x, y) = c\}$$

sia (sostegno di) una curva regolare $\mathbf{r} = x(t)\,\mathbf{i} + y(t)\,\mathbf{j}$, $t \in I$.

Allora:

$$g(t) := f\big(x(t), y(t)\big) = c, \qquad \forall\, t \in I\,.$$

Derivando l'equazione rispetto a $t$:

$$g'(t) = \nabla f\big(x(t), y(t)\big) \cdot \mathbf{r}'(t) = 0\,.$$

Dunque, in ogni punto regolare di una curva di livello il gradiente di $f$ (se $\neq \mathbf{0}$) è un vettore *perpendicolare* alla tangente alla curva.

(Si dice in breve che il gradiente è ortogonale alle curve di livello).

# Derivate (e differenziali) di ordine superiore.

Sia $f$ una funzione definita in un aperto di $\mathbb{R}^n$ e supponiamo che una derivata parziale $\dfrac{\partial f}{\partial x_i}$ esista in un *intorno* di un punto $\mathbf{x} \in \mathbb{R}^n$.

Se la funzione $\dfrac{\partial f}{\partial x_i}$ è a sua volta derivabile rispetto a $x_j$ in $\mathbf{x}$, è definita la **derivata parziale seconda** rispetto a $x_i$ e a $x_j$

$$\frac{\partial^2 f}{\partial x_j \partial x_i}(\mathbf{x}) := \frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_i}\right)(\mathbf{x}) \, .$$

Notazioni alternative:

$$\partial_{x_j x_i} f \, , \quad f_{x_j x_i} \, , \quad D^2_{ji} f \, , \ldots$$

Se $f$ è derivabile in un aperto di $R^n$ e le derivate parziali sono a loro volta derivabili, sono definite $n^2$ derivate parziali seconde.

Esempio (n=2).

Sia $f(x,y) = x^2 \cos y\,,\ \ (x,y) \in \mathbb{R}^2\,;$ le derivate parziali prime sono:

$$\partial_x f(x,y) = 2x \cos y, \quad \partial_y f(x,y) = -x^2 \sin y\,, \quad \text{da cui:}$$

$$\partial_{xx} f(x,y) = 2 \cos y\,, \quad \partial_{xy} f(x,y) = -2x \sin y\,,$$

$$\partial_{yx} f(x,y) = -2x \sin y\,, \quad \partial_{yy} f(x,y) = -x^2 \cos y\,.$$

in ogni punto $(x,y) \in \mathbb{R}^2$.

L'uguaglianza delle derivate miste non è casuale:

Teorema (Schwarz).

Sia $f$ definita in un aperto di $\mathbb{R}^n$ e assumiamo che le derivate parziali seconde $\partial_{x_i x_j} f,\ \partial_{x_j x_i} f\ \ (i \neq j)$ esistano in un *intorno* di un punto $\mathbf{x}$ e siano *continue in* $\mathbf{x}$.
Allora $\partial_{x_i x_j} f(\mathbf{x}) = \partial_{x_j x_i} f(\mathbf{x})$.

Se una funzione ha derivate seconde continue in un aperto $D \subseteq \mathbb{R}^n$, si dice che è *di classe* $\mathcal{C}^2(D)$.

Le definizioni di derivate parziali successive si generalizzano alle derivate di ordine $k$.

Se $f$ ha derivate parziali continue fino all'ordine $k$ in un aperto $D \subseteq \mathbb{R}^n$, si dice che $f \in \mathcal{C}^k(D)$. Anche per queste funzioni, le derivate rispetto a variabili differenti si possono commutare.

Esempio

Le derivate parziali terze di $f(x, y) = x^2 \cos y$, sono:

$$\partial_{xxx} f(x, y) = 0\,, \quad \partial_{yxx} f(x, y) = \partial_{xyx} f(x, y) = \partial_{xxy} f(x, y) = -2 \sin y\,,$$

$$\partial_{xyy} f(x, y) = \partial_{yxy} f(x, y) = \partial_{yyx} f(x, y) = -2x \cos y\,, \quad \partial_{yyy} f(x, y) = x^2 \sin y\,.$$

Esercizio

Calcolare tutte le derivate parziali della funzione $g(x, y, z) = xyz$.

## Differenziale secondo

Se $f$ è differenziabile in un aperto $D \subseteq \mathbb{R}^n$, esistono nello stesso aperto le derivate parziali $\partial_{x_i} f$, $i = 1, 2, ..., n$. Se queste derivate sono a loro volta differenziabili in $\mathbf{x}_0 \in D$, si dice che $f$ è *due volte differenziabile* in $\mathbf{x}_0$.

Si definisce *differenziale secondo* di $f$ in $\mathbf{x}_0$ la funzione (forma quadratica)

$$d^2 f(\mathbf{x}_0) \; : \; \mathbf{h} \mapsto \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j} f(\mathbf{x}_0) h_i \, h_j \, , \qquad \mathbf{h} = (h_1, h_2, ..., h_n) \in \mathbb{R}^n \, .$$

Più in breve scriveremo

$$d^2 f(\mathbf{x}_0) \; = \; \sum_{i=1}^{n} \sum_{j=1}^{n} \partial_{x_i x_j} f(\mathbf{x}_0) h_i \, h_j$$

o anche

$$d^2 f(\mathbf{x}_0) \; = \; \sum_{i=1}^{n} \sum_{j=1}^{n} \partial_{x_i x_j} f(\mathbf{x}_0) dx_i \, dx_j \, .$$

*Osservazioni*:

Se $f \in \mathcal{C}^2(D)$, le derivate prime $\partial_{x_i} f$ sono differenziabili (per la condizione sufficiente applicata alle derivate paziali).

Dunque: $f \in \mathcal{C}^2(D) \;\Rightarrow\; f$ due volte differenziabile.

Possiamo scrivere il differenziale secondo (come qualsiasi forma quadratica) nel seguente modo:

$$d^2 f(\mathbf{x}_0) \,=\, H_f(\mathbf{x}_0)\,\mathbf{h} \cdot \mathbf{h} \quad \big( = \mathbf{h}^T\, H_f(\mathbf{x}_0)\,\mathbf{h}\,\big)\,,$$

dove ora **h** è pensato come *vettore colonna* e $H_f(\mathbf{x}_0)$ è la *matrice dei coefficienti*

$$H_f(\mathbf{x}_0) = \begin{pmatrix} f_{x_1 x_1}(\mathbf{x}_0) & f_{x_1 x_2}(\mathbf{x}_0) \; \dots & f_{x_1 x_n}(\mathbf{x}_0) \\ f_{x_2 x_1}(\mathbf{x}_0) & f_{x_2 x_2}(\mathbf{x}_0) \; \dots & f_{x_2 x_n}(\mathbf{x}_0) \\ \dots & \dots & \dots \\ f_{x_n x_1}(\mathbf{x}_0) & f_{x_n x_2}(\mathbf{x}_0) \; \dots & f_{x_n x_n}(\mathbf{x}_0) \end{pmatrix}$$

detta **matrice Hessiana** di $f$ in $\mathbf{x}_0$.
Per quanto visto sopra, questa matrice è *simmetrica*.

Caso $n = 2$: $f : D \to \mathbb{R}$, $\mathbf{x}_0 = (x_0, y_0) \in D$, $\mathbf{h} = (h, k)$:

$$df(x_0, y_0) = \partial_x f(x_0, y_0)h + \partial_y f(x_0, y_0)k = \nabla f(x_0, y_0) \cdot \mathbf{h}\,;$$

$$d^2 f(x_0, y_0) \,=\, \partial_{xx} f(x_0, y_0)h^2 + \partial_{xy} f(x_0, y_0)hk + \partial_{yx} f(x_0, y_0)kh + \partial_{yy} f(x_0, y_0)k^2$$

$$=\, \partial_{xx} f(x_0, y_0)h^2 + 2\partial_{xy} f(x_0, y_0)hk + \partial_{yy} f(x_0, y_0)k^2 \,=\, H_f(\mathbf{x}_0)\,\mathbf{h}\, \cdot \mathbf{h}\,,$$

dove

$$H_f(\mathbf{x}_0) = \begin{pmatrix} f_{xx}(x_0, y_0) & f_{xy}(x_0, y_0) \\ f_{yx}(x_0, y_0) & f_{yy}(x_0, y_0) \end{pmatrix}.$$

Esempio

Calcoliamo il differenziale secondo di $f(x, y) = x^2 e^{3y}$ nel punto $(1, 0)$.

Derivate prime: $f_x(x, y) = 2xe^{3y}$ , $f_y(x, y) = 3x^2 e^{3y}$ ;

Derivate seconde:

$$f_{xx}(x, y) = 2e^{3y} , \quad f_{xy}(x, y) = f_{yx}(x, y) = 6xe^{3y} , \quad f_{yy}(x, y) = 9x^2 e^{3y} ;$$

Dunque:

$$f_{xx}(1, 0) = 2 , \quad f_{xy}(1, 0) = f_{yx}(1, 0) = 6 , \quad f_{yy}(1, 0) = 9 ;$$

$$d^2 f(1, 0) \, = \, 2h^2 + 12hk + 9k^2 .$$

Matrice Hessiana nel punto:

$$H_f(1, 0) = \begin{pmatrix} 2 & 6 \\ 6 & 9 \end{pmatrix}$$

# Formula di Taylor (2° ordine)

Per una funzione due volte differenziabile, possiamo ottenere approssimazioni locali più accurate. Per semplicità assumeremo $f \in \mathcal{C}^2(D)$, $D \subseteq \mathbb{R}^n$ aperto.

Sia $\mathbf{x}_0 \in D$ e $\mathbf{h} \in \mathbb{R}^n$ tale che $\mathbf{x}_0 + \mathbf{h} \in D$. Valgono allora le formule:

**Taylor, resto secondo Lagrange**

$$f(\mathbf{x}_0 + \mathbf{h}) = f(\mathbf{x}_0) + \sum_{i=1}^{n} \partial_{x_i} f(\mathbf{x}_0) h_i + \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \partial_{x_i x_j} f(\mathbf{x}_0 + \delta\,\mathbf{h}) h_i\, h_j$$

per qualche $\delta \in (0, 1)$.

**Taylor, resto secondo Peano**

$$f(\mathbf{x}_0 + \mathbf{h}) = f(\mathbf{x}_0) + \sum_{i=1}^{n} \partial_{x_i} f(\mathbf{x}_0) h_i + \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \partial_{x_i x_j} f(\mathbf{x}_0) h_i\, h_j + o(|\mathbf{h}|^2)$$

Le due formule si possono scrivere in forma più compatta.

*Lagrange*:

$$\begin{aligned} f(\mathbf{x}_0 + \mathbf{h}) &= f(\mathbf{x}_0) + \nabla f(\mathbf{x}_0) \cdot \mathbf{h} + \frac{1}{2} H_f(\mathbf{x}_0 + \delta\,\mathbf{h})\mathbf{h} \cdot \mathbf{h} \\ &= f(\mathbf{x}_0) + df(\mathbf{x}_0) + \frac{1}{2} d^2 f(\mathbf{x}_0 + \delta\,\mathbf{h}), \end{aligned}$$

*Peano*:

$$\begin{aligned} f(\mathbf{x}_0 + \mathbf{h}) &= f(\mathbf{x}_0) + \nabla f(\mathbf{x}_0) \cdot \mathbf{h} + \frac{1}{2} H_f(\mathbf{x}_0)\mathbf{h} \cdot \mathbf{h} + o(|\mathbf{h}|^2) \\ &= f(\mathbf{x}_0) + df(\mathbf{x}_0) + \frac{1}{2} d^2 f(\mathbf{x}_0) + o(|\mathbf{h}|^2). \end{aligned}$$

*Dimostrazione*.

Nel caso di una variabile, se $g(t) \in \mathcal{C}^2(I)$ e $t_0 \in I$, vale la formula

$$g(t) = g(t_0) + g'(t_0)(t - t_0) + \frac{1}{2}g''(\delta)(t - t_0)^2 \,,$$

per qualche $\delta \in (t_0, t) \subset I$.

In particolare, se $t_0 = 0$ e $t = 1$,

$$g(1) = g(0) + g'(0) + \frac{1}{2}g''(\delta) \,,$$

con $\delta \in (0, 1)$.

Definiamo ora

$$g(t) = f(\mathbf{x}_0 + t\mathbf{h}) \,;$$

osserviamo che la funzione composta è definita in $[0, 1]$ e che vale:

$$g(0) = f(\mathbf{x}_0)\,; \qquad g(1) = f(\mathbf{x}_0 + \mathbf{h})\,.$$

Inoltre $g$ è due volte derivabile con continuità in $[0,1]$ e applicando (due volte) la formula di derivazione delle funzioni composte si ottiene

$$g'(t) = \sum_{i=1}^{n} \partial_{x_i} f(\mathbf{x}_0 + t\,\mathbf{h}) h_i\,,$$

$$g''(t) = \sum_{i=1}^{n} \sum_{j=1}^{n} \partial_{x_i x_j} f(\mathbf{x}_0 + t\,\mathbf{h}) h_i\, h_j$$

Ponendo $t = 0$ nella prima, $t = \delta$ nella seconda e inserendo tutte le espressioni nella formula

$$g(1) = g(0) + g'(0) + \frac{1}{2} g''(\delta)\,,$$

si ottiene Taylor con il resto di Lagrange. $\square$

Traccia della dimostrazione della formula con il resto secondo Peano.
Si parte dalla formula con il resto di Lagrange

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

e la si scrive nella forma:

$$= f(\mathbf{x}_0) + \nabla f(\mathbf{x}_0) \cdot \mathbf{h} + \frac{1}{2} H_f(\mathbf{x}_0)\mathbf{h} \cdot \mathbf{h} + \frac{1}{2}\Big[H_f(\mathbf{x}_0 + \delta\, \mathbf{h}) - H_f(\mathbf{x}_0)\Big]\mathbf{h} \cdot \mathbf{h}\,.$$

Per la continuità delle derivate seconde (coefficienti della matrice Hessiana) si dimostra che

$$\lim_{\mathbf{h} \to \mathbf{0}} \frac{\Big[H_f(\mathbf{x}_0 + \delta\, \mathbf{h}) - H_f(\mathbf{x}_0)\Big]\mathbf{h} \cdot \mathbf{h}}{|\mathbf{h}|^2} = 0\,,$$

cioè che l'ultimo termine è $o(|\mathbf{h}|^2)$, come richiesto.

Esempio.

Sviluppo di Taylor di $f(x, y) = x^2 e^{3y}$ nell'intorno di $(1, 0)$.

Valore di $f$ nel punto: $f(1, 0) = 1$ ; derivate nel punto (calcolate a p.11):

$$f_x(1,0) = 2\,, \quad f_y(1,0) = 3\,, \quad f_{xx}(1,0) = 2\,, \quad f_{xy}(1,0) = 6\,, \quad f_{yy}(1,0) = 9\,.$$

$$\Downarrow$$

$$(1+h)^2 e^{3k} = 1 + 2h + 3k + h^2 + 6hk + \frac{9}{2}k^2 + o(h^2 + k^2)\,.$$

Ponendo $x = 1 + h$, $y = k$, si può anche scrivere:

$$x^2 e^{3y} = 1 + 2(x-1) + 3y + (x-1)^2 + 6(x-1)y + \frac{9}{2}y^2 + o\big((x-1)^2 + y^2\big)\,.$$

# Estremi liberi

Problema.

Costruire la scatola di superficie totale minima che racchiude un volume assegnato $V$.

Detti $x$, $y$ e $z$ gli spigoli del parallelepipedo, si cerca il minimo della funzione $2(xy + xz + yz)$ con la condizione $xyz = V$.
Ricavando $z = V/xy$, si pone il problema nella forma:

Trovare il minimo di

$$f(x,y) = xy + \frac{V}{x} + \frac{V}{y} \quad \text{in} \quad D = \{(x,y) \,|\, x > 0\,,\ y > 0\,\}\,.$$

**Problemi di estremi liberi**: trovare massimi e minimi di una funzione in un insieme *aperto* (o nei punti *interni* di un dominio).

*Osservazione*: nella formulazione originale in *tre* variabili, la funzione era definita su un sottoinsieme chiuso di $\mathbb{R}^3$. Quando l'insieme in cui si cercano gli estremi di una funzione non è aperto, si parla di *estremi vincolati*.

Definizione (Estremi locali)

Sia $f : D \to \mathbb{R}$, $D \subseteq \mathbb{R}^n$ aperto. Un punto $\mathbf{x}_0 \in D$ si dice *punto di massimo (minimo) relativo o locale* di $f$ se esiste $B_r(\mathbf{x}_0)$ tale che

$$\mathbf{x} \in B_r(\mathbf{x}_0) \quad \Rightarrow \quad f(\mathbf{x}) \leq f(\mathbf{x}_0) \quad \big(f(\mathbf{x}) \geq f(\mathbf{x}_0)\big)$$

Se le disuguaglianze sono strette, si parla di massimi o minimi (locali) *forti*.

Come trovare i punti di massimo e minimo locale ?

Teorema (Fermat)

Se $\mathbf{x}_0$ è un punto di massimo o minimo locale di $f$ *e se* $f$ è derivabile in $\mathbf{x}_0$, allora $\nabla f(\mathbf{x}_0) = \mathbf{0}$.

*Dimostrazione*:

Nelle ipotesi del teorema, la funzione $g(t) := f(\mathbf{x}_0 + t\,\mathbf{e}_i)$ è definita in un intorno di $t = 0$ ed ha un estremo locale in 0 per ogni $i = 1, 2, ..., n$.

Inoltre, $g$ è derivabile e vale: $g'(0) = \partial_{x_i} f(\mathbf{x}_0)$.

Il risultato segue allora dal teorema di Fermat in una variabile.

*Osservazioni*

i) La condizione $\nabla f(\mathbf{x}_0) = \mathbf{0}$ è solo *necessaria* perché il punto sia un estremo locale.
ii) Può succedere che in un punto di estremo la funzione *non* sia derivabile. Per esempio, $f(x, y) = \sqrt{x^2 + y^2}$ ha un minimo (globale) nell'origine, dove non è derivabile.
iii) Se $f$ è differenziabile in un punto di estremo $\mathbf{x}_0$, il piano tangente è orizzontale.

I punti $\mathbf{x}$ tali che $\nabla f(\mathbf{x}) = \mathbf{0}$ si dicono **punti critici o stazionari** di $f$.

Se un punto critico non è un punto di estremo, si dice che è un *punto di sella o colle*.

Esempi. Le funzioni:

$$f(x, y) = x^2 + y^2\,, \quad g(x, y) = -x^2 - y^2\,, \quad h(x, y) = x^2 - y^2\,,$$

hanno tutte $(0, 0)$ come punto critico. L'origine è: un punto di minimo per $f$, di massimo per $g$ e un punto di sella per $h$.

## Classificazione dei punti stazionari.

Se ora $f \in \mathcal{C}^2(D)$ e $\mathbf{x}_0 \in D$ è un punto critico di $f$, applicando la formula di Taylor abbiamo

$$f(\mathbf{x}_0 + \mathbf{h}) - f(\mathbf{x}_0) = \frac{1}{2} H_f(\mathbf{x}_0)\mathbf{h} \cdot \mathbf{h} + o(|\mathbf{h}|^2)\,,$$

dove

$$H_f(\mathbf{x}_0)\mathbf{h} \cdot \mathbf{h} = \sum_{i,j=1}^{n} \partial_{x_i x_j} f(\mathbf{x}_0) h_i\, h_j\,.$$

Dunque, la variazione $\Delta f := f(\mathbf{x}_0 + \mathbf{h}) - f(\mathbf{x}_0)$ in un intorno di un punto critico $\mathbf{x}_0$ *dipende dalla forma quadratica* $H_f(\mathbf{x}_0)\mathbf{h} \cdot \mathbf{h}$ (differenziale secondo).

Dobbiamo allora considerare il problema (algebrico) dello *studio del segno di una forma quadratica*

$$q(\mathbf{h}) = A\mathbf{h} \cdot \mathbf{h} = \sum_{i,j=1}^{n} a_{i\,j} h_i\, h_j$$

associata ad una matrice simmetrica $A$.

Una forma quadratica

$$q(\mathbf{h}) = A\mathbf{h} \cdot \mathbf{h} \ \big( = \ \mathbf{h}^T A\mathbf{h} \big) \qquad \mathbf{h} \in \mathbb{R}^n ,$$

si dice

i) definita positiva (negativa) se $\forall\, \mathbf{h} \neq \mathbf{0} \ \ q(\mathbf{h}) > 0$ ($< 0$);

ii) semidefinita positiva (negativa) se $\forall\, \mathbf{h} \neq \mathbf{0} \ \ q(\mathbf{h}) \geq 0$ ($\leq 0$) ed esiste $\mathbf{h} \neq \mathbf{0}$ t.c. $q(\mathbf{h}) = 0$;

iii) indefinita se $\exists \ \mathbf{h}_1, \mathbf{h}_2$ t.c. $q(\mathbf{h}_1) > 0$ e $q(\mathbf{h}_2) < 0$.

Esempi ($n = 3$)

Se $\mathbf{h} = (h_1, h_2, h_3)$, le forme quadratiche

$$q(\mathbf{h}) = \begin{cases} h_1^2 + 3h_2^2 + 2h_3^2, \\ -2h_1^2 - h_2^2 - 4h_3^2, \\ h_1^2 + h_3^2, \\ -h_2^2 - h_3^2, \\ h_1^2 - h_2^2 + h_3^2, \end{cases}$$

sono rispettivamente: definita positiva, definita negativa, semidefinita positiva, semidefinita negativa e indefinita.

## Test degli autovalori

Remind: un numero $\lambda$ si dice *autovalore* di una matrice $A$ ($n$ x $n$) se esiste un vettore non nullo $\mathbf{v}$, detto *autovettore*, tale che $A\mathbf{v} = \lambda\,\mathbf{v}$.
Questo succede se e solo se $\lambda$ è soluzione di

$$\det(A - \lambda\, I) = 0 \qquad \text{(equazione caratteristica)}$$

Se $A$ è *simmetrica*:

i) Gli autovalori $\lambda_1$, $\lambda_2$,...,$\lambda_n$, sono reali;
ii) Esistono $n$ autovettori ortogonali che formano una base in $\mathbb{R}^n$;
iii) Esiste una matrice $S$ che trasforma la base canonica di $\mathbb{R}^n$ nella base di autovettori e diagonalizza $A$:

$$S^T A S = \begin{pmatrix} \lambda_1 & 0 \ldots & 0 \\ 0 & \lambda_2 \ldots & 0 \\ \ldots & \ldots & \ldots \\ 0 & 0 \ldots & \lambda_n \end{pmatrix}$$

Usando queste proprietà si dimostra che la forma $q(\mathbf{h}) = A\mathbf{h} \cdot \mathbf{h}$ è:

i) definita positiva (negativa) se e solo se tutti gli autovalori di $A$ sono $> 0$ $(< 0)$;
ii) semidefinita positiva (negativa) se e solo se tutti gli autovalori di $A$ sono $\geq 0$ $(\leq 0)$ ed esiste almeno un autovalore nullo;
iii) indefinita se esistono autovalori di segno opposto.

Inoltre:

Se $q$ è definita positiva, allora

$$q(\mathbf{h}) \geq \lambda_{min}|\mathbf{h}|^2 ,$$

dove $\lambda_{min} > 0$ è il minimo autovalore di $A$.

Se $q$ è definita negativa, allora

$$q(\mathbf{h}) \leq \lambda_{max}|\mathbf{h}|^2 ,$$

dove $\lambda_{max} < 0$ è il massimo autovalore di $A$.

Esempio

Classificare la forma quadratica in $\mathbb{R}^2$: $q(h,k) = h^2 - hk + k^2$.
La matrice dei coefficienti è

$$A = \begin{pmatrix} 1 & -1/2 \\ -1/2 & 1 \end{pmatrix}$$

Gli autovalori sono $\lambda_1 = 1/2$, $\lambda_2 = 3/2$. Quindi $q$ è definita positiva.

Osserviamo che

$$q(h,k) = \frac{1}{2}(h-k)^2 + \frac{1}{2}(h^2+k^2) \geq \frac{1}{2}(h^2+k^2)\,.$$

Esercizio

Verificare che la forma quadratica in $\mathbb{R}^3$:

$$q(\mathbf{h}) = 6h_1^2 + 3h_2^2 - 8h_1h_3 + 4h_3^2\,,$$

è definita positiva.

*Classificazione senza il calcolo degli autovalori (<u>solo</u> per $n = 2$).*

Se

$$q(h,k) = ah^2 + 2bhk + ck^2,$$

la matrice dei coefficienti è

$$A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$$

e l'equazione agli autovalori ha la forma

$$\det \begin{pmatrix} a-\lambda & b \\ b & c-\lambda \end{pmatrix} = \lambda^2 - (a+c)\lambda + ac - b^2 = 0\,.$$

Detti $\lambda_1$, $\lambda_2$ gli autovalori, abbiamo:

$$\lambda_1 + \lambda_2 = a + c\ = \mathrm{Tr}(A); \quad \lambda_1\lambda_2 = ac - b^2\ = \det(A)\,.$$

Dunque:

$\det(A) > 0$ e $\mathrm{Tr}(A) > 0$ ($< 0$) $\Leftrightarrow$ $q$ è definita positiva (negativa);

$\det(A) = 0$ e $\mathrm{Tr}(A) > 0$ ($< 0$) $\Leftrightarrow$ $q$ è semidefinita positiva (negativa);

$\det(A) < 0$ $\Leftrightarrow$ $q$ è indefinita.

Teorema (test delle derivate seconde).

Sia $f \in \mathcal{C}^2(D)$, $D \subseteq \mathbb{R}^n$ e sia $\mathbf{x}_0 \in D$ un punto critico di $f$.
Se la forma quadratica $q(\mathbf{h}) = H_f(\mathbf{x}_0)\mathbf{h} \cdot \mathbf{h}$ è

- definita positiva (negativa) $\Rightarrow$ $\mathbf{x}_0$ punto di minimo (massimo) locale forte;
- indefinita $\Rightarrow$ $\mathbf{x}_0$ punto di sella.

*Dimostrazione:*

Nelle ipotesi del teorema, se $q(\mathbf{h})$ è definita positiva:

$$\Delta f := f(\mathbf{x}_0 + \mathbf{h}) - f(\mathbf{x}_0) = \frac{1}{2} H_f(\mathbf{x}_0)\mathbf{h} \cdot \mathbf{h} + o(|\mathbf{h}|^2)$$

$$\geq \frac{\lambda_{min}}{2}|\mathbf{h}|^2 + o(|\mathbf{h}|^2) = \frac{\lambda_{min}}{2}|\mathbf{h}|^2\big(1 + o(1)\big) .$$

L'ultimo termine è $> 0$ per ogni $\mathbf{h}$ di norma *sufficientemente piccola.*

Dunque, in un intorno di $\mathbf{x}_0$ vale $\Delta f > 0$, cioè $\mathbf{x}_0$ è punto di minimo locale forte.

Se $q(\mathbf{h})$ è definita negativa, si dimostra in modo simile che $\Delta f < 0$ in un intorno di $\mathbf{x}_0$ e quindi che abbiamo un massimo locale forte.

Se $q(\mathbf{h})$ è indefinita, esistono due vettori $\mathbf{h}_1$, $\mathbf{h}_2$ tali che $q(\mathbf{h}_1) > 0$ e $q(\mathbf{h}_2) < 0$.
Possiamo supporre che siano di lunghezza unitaria (versori).
Allora, per $t$ sufficientemente piccolo:

$$f(\mathbf{x}_0 + t\mathbf{h}_1) - f(\mathbf{x}_0) = \frac{1}{2} t^2 H_f(\mathbf{x}_0)\mathbf{h}_1 \cdot \mathbf{h}_1 + o(t^2)$$

$$= t^2 \big( q(\mathbf{h}_1)/2 + o(1) \big) > 0\,,$$

$$f(\mathbf{x}_0 + t\mathbf{h}_2) - f(\mathbf{x}_0) = \frac{1}{2} t^2 H_f(\mathbf{x}_0)\mathbf{h}_2 \cdot \mathbf{h}_2 + o(t^2)$$

$$= t^2 \big( q(\mathbf{h}_2)/2 + o(1) \big) < 0\,.$$

Dunque, lungo due direzioni uscenti da $\mathbf{x}_0$ la variazione $\Delta f$ ha segno *opposto*.
Si conclude che $\mathbf{x}_0$ è un punto di sella.

*Osservazione*.

Se la forma quadratica $q(\mathbf{h}) = H_f(\mathbf{x}_0)\mathbf{h} \cdot \mathbf{h}$ è *semidefinita* (positiva o negativa) il test non decide e occorre un'ulteriore analisi.

Per esempio, le funzioni $f(x, y) = x^2 + y^4$ e $g(x, y) = x^2 - y^4$, hanno entrambe un unico punto critico nell'origine e per entrambe la matrice Hessiana nell'origine è

$$H_f(0,0) = H_g(0,0) = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$$

Quindi, *in entrambi casi* la forma quadratica associata è $q(h, k) = 2h^2$, semidefinita positiva.
Ma l'origine è punto di minimo per $f$ e un colle per $g$.

La situazione cambia se la forma quadratica è semidefinita positiva (negativa) in un *intorno* di un punto critico $\mathbf{x}_0$

In questo caso, usando la formula di Taylor con il resto di Lagrange:

$$\Delta f := f(\mathbf{x}_0 + \mathbf{h}) - f(\mathbf{x}_0) = \frac{1}{2} H_f(\mathbf{x}_0 + \delta\mathbf{h})\mathbf{h} \cdot \mathbf{h} \ \geq 0 \ \ (\leq 0)$$

per ogni $\mathbf{h}$ sufficientemente piccolo.

Questo significa che $\mathbf{x}_0$ è un punto di minimo (massimo) locale debole.