Ottimizzazione in più variabili - Sommario

Tutto sull'ottimizzazione in più variabili: generalità, ottimizzazione libera e vincolata.

A. NOZIONI PRELIMINARI

A1. Weierstraß generalizzato

Teorema di Weierstraß Generalizzato

Condizione sufficiente per l'esistenza di un minimo e massimo di una funzione in più variabili: il teorema generalizzato di Weierstraß.

---

0. Voci correlate

- Teoremi sulle funzioni continue

1. Teorema di Weierstraß Generalizzato

Generalizziamo il *teorema di Weierstraß* (1) su più variabili.

#Teorema

> Teorema (di Weierstraß, generalizzato).
>
> Se $f : K \subseteq \mathbb{R}^N \longrightarrow \mathbb{R}$ è *continua* su K con K *compatto* (ovvero chiuso e limitato) per $\mathbb{R}^N$, allora
>
> $$\exists x_*, x_\bullet \in K : x_* = \min_E f, x_\bullet = \max_E f$$
>
> ovvero esistono sia *il minimo* che il *massimo* della funzione.

Tuttavia, la compattezza del dominio potrebbe sembrare troppo *restrittiva* (ed effettivamente potrebbe esserlo: potrei avere insiemi illimitati). Quindi troviamo una *"versione più debole"* che ci garantisca comunque l'esistenza di uno dei due punti $x_*, x_\bullet$.

Parleremo in particolare di *coercitività* (o *anticoercitività*) (1) e la sua proprietà fondamentale (2).

A2. Funzione Coerciva

Funzione Coerciva

Definizione di funzione coerciva, proprietà fondamentale delle funzioni coerciva (forma debole del teorema di Weierstraß)

---

0. Voci correlate

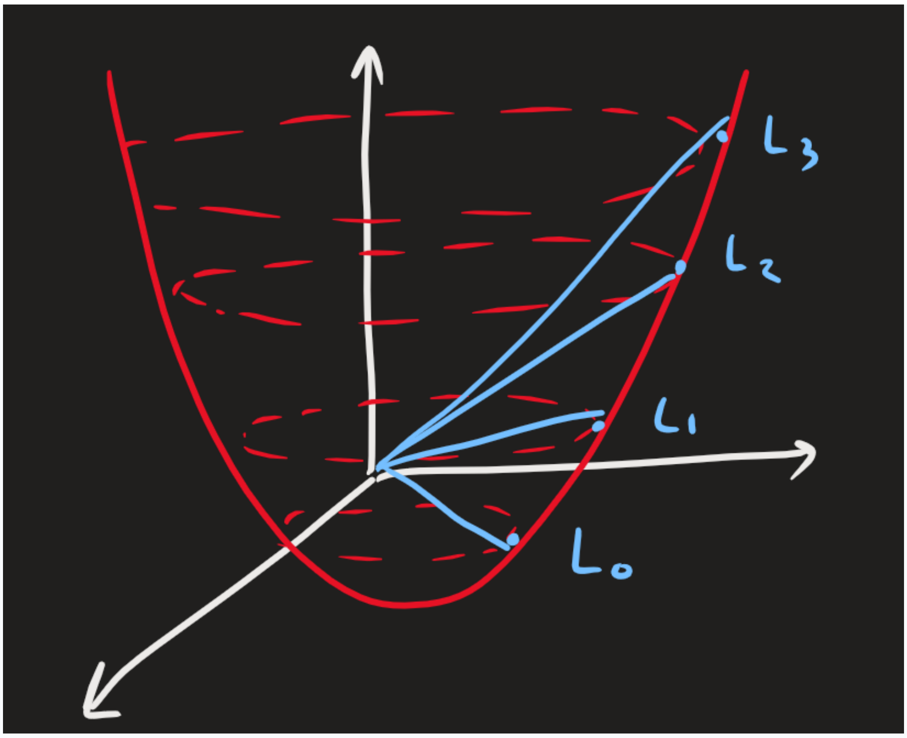
- Definizione di Spazio Metrico

1. Definizione di Funzione Coerciva

#Definizione

> Definizione (funzione coerciva).
>
> Si dice che una funzione in più variabili definita come
>
> $$f: E = \mathbb{R}^N \longrightarrow \mathbb{R}$$
>
> è *coerciva* (o *anticoerciva, in rosso*) se vale il limite
>
> $$\lim_{\|\underline{x}\| \to +\infty} f(\underline{x}) = +\infty \; ({\color{red}-\infty})$$

L'idea della *coercitività* è di avere un *campo scalare* che si *"distanzia dall'origine all'infinito"* al salire di livello.



2. Proprietà delle Funzioni Coercive

#Teorema

> Teorema (proprietà delle funzioni coercive, forma debole di Weierstraß).
>
> Se $f : E = \mathbb{R}^N \longrightarrow \mathbb{R}$ è *continua* e *coerciva* (*anticoerciva, in rosso*) allora vale che
>
> $$\exists \min_{E=\mathbb{R}^2} f \left(\exists \min_{E=\mathbb{R}^2} f \right)$$

CASO D'USO (Idea). Questa proprietà è utile nella *prassi*, in particolare per la *minimizzazione di funzioni*.
Supponiamo di avere una funzione costo f a variare su N parametri. In particolare, questo *fitta dei dati*. Per minimizzare questa funzione in riferimento di un valore y_{data}, posso impostare la funzione ϕ definita come

$$\phi(\underline{x}) = \|f(\underline{x}) - y_{data}\|$$

e ricavarne dunque il *minimo* mediante operazione di ottimizzazioni. Tuttavia, il dato y_{data} potrebbe essere *"sporco"* dato che è *prono ad errori*. Quindi il minimo trovato in ϕ non potrebbe coincidere col minimo effettivo.
Per risolvere questo problema, impostiamo un'altra funzione ϕ_α definita come

$$\phi_\alpha(\underline{x}) = \|f(\underline{x}) - y_{data}\| + \alpha\|\underline{x}\|^2$$

dove α è il *"termine di correzione"*. Notiamo in particolare che il termine $\|\underline{x}\|^2$ tende a infinito, rendendo coercitiva la funzione: possiamo dunque applicare il macchinario appena visto per trovare il *minimo effettivo* della funzione, avendo

$$\min \phi \sim \min \phi_\alpha$$

A3. Estremo Relativo e Assoluto

Definizione di Estremo per una Funzione in più variabili

Definizione di estremo relativo e assoluto per una funzione. Osservazione: la ricerca degli estremi in una sfera.

---

0. Voci correlate

- Funzioni
- Topologia in RN

1. Definizione di Estremo Relativo e Assoluto

#Definizione

> Definizione (estremo relativo e assoluto).
>
> Sia $f: E \subseteq \mathbb{R}^N \longrightarrow \mathbb{R}$ e $\underline{x_0} \in E$.
> Si dice che $\underline{x_0}$ è di *massimo* (o *minimo, in rosso*) *relativo* per f se vale che
>
> $$\exists U(\underline{x_0}) : \forall \underline{x} \in U(\underline{x_0}) \cap \mathcal{C}_E(\underline{x_0}), \begin{matrix} f(\underline{x}) < f(\underline{x_0}) \\ f(\underline{x}) > f(\underline{x_0}) \end{matrix}$$
>
> (come si definisce nel caso unidimensionale). Si nota che U è l'intorno di un punto (1).
>
> Si dice invece che è di *massimo* (o *minimo, in rosso*) *assoluto* per f se vale che
>
> $$M = f(\underline{x_0}) = \max_E f \ \left(m = f(\underline{x_0}) = \min_E f\right)$$

A4. Generalità sull'Ottimizzazione

Problemi di Ottimizzazione

Generalità sui problemi di ottimizzazione: schema generale in ambito interdisciplinare, schema specifico per il nostro corso (integrazione tra appunti e libro di Pagani-Salsa)

---

0. Voci correlate

- Definizione di Estremo per una Funzione in più variabili

1. Problemi di Ottimizzazione

Significato. Per *"ottimizzazione"* si intende un'ampia categoria di problemi; di solito si tratta di *massimizzare* o *minimizzare* un obbiettivo, una funzione (1). Per affrontare un problema del genere, bisogna porsi *una serie di domande* (e questo fungerà da schema generale per la risoluzione di un problema di ottimizzazione).

Schema. Poniamo la seguente serie di domande.

1. *Esiste* il punto di minimo o massimo?
   - Per garantirci di questo, dovremmo applicare il macchinario appena sviluppato su *Weierstraß* (1) o sulle *funzioni coercive* (2).
2. Il punto estremo è *unico*?
   - Ovviamo l'*estremo* in sé (ovvero il $\max$ o $\min$) è unico, ma non è detto che il *punto di estremo* sia unico.
   - Questo è importante per le eventuali proprietà di *convessità* del dominio di una funzione (di cui non occuperemo).
3. *Come si comportano questi punti?*
   - Sapendo che *esiste* l'*estremo* della funzione, dobbiamo sapere *come si comporta* questo punto; ovvero vogliamo sviluppare delle *condizioni necessarie* per questi punti di estremo, così per sapere come trovarli. Ci occuperemo di questa parte.
4. *Come si calcolano questi punti?*
   - Questo rientra nell'ambito dell'*analisi numerica* (non parte del nostro programma).

2. La Ricerca degli Estremi

Adesso svisceriamo lo schema appena sviluppato, soffermandoci sulla domanda *"Come si comportano questi punti?"*; ovvero vogliamo trovare *dove* si trovano questi punti (una domanda quasi analoga).

Problema. Supponiamo di avere una funzione del tipo

$$f: \underbrace{B(0;1)(\mathbb{R}^N)}_{\{x \in \mathbb{R}^N : \|x\| \leq 1\}} \longrightarrow \mathbb{R}$$

Per Weierstraß so che la funzione ammette *almeno* un minimo e un massimo (1). Ma dove stanno? Possiamo cercali in *"due luoghi"*

- Li cerchiamo nei punti $\|x\| < 1$; qui parliamo di *minimi e massimi liberi* e potrebbe essere che il gradiente di questi punti sono nulli; nei prossimi capitoli definiremo degli strumenti per capire la loro natura
- Li cerchiamo nei punto $\|x\| = 1$; qui parliamo di *minimi e massimi vincolati*: non possiamo più utilizzare il gradiente della funzione, in quanto non vi è più un legame. Dobbiamo dunque sviluppare degli strumenti sulle *curve e superfici*, che rappresentano un *vincolo* per la nostra funzione. Infatti troveremo che gli estremi vincolati hanno delle proprietà particolari (in particolare i gradienti della restrizione e della funzione sono paralleli).
  Vedremo di risolvere questi problemi in questo capitolo.

B. OTTIMIZZAZIONE LIBERA

A1. Test del Gradiente (Fermat)

Test del Gradiente

Condizione necessaria per punti di estremo relativo (o noto come test del gradiente, oppure teorema di Fermat generalizzato). Dimostrazione.

---

0. Voci correlate

- Definizione di Estremo per una Funzione in più variabili
- Teorema di Fermat

1. Test del Gradiente

Come *primo approccio* ai problemi di *ottimizzazione*, consideriamo una *condizione necessaria* per i punti di *estremo relativo*. Questo teorema sarà noto come il *test del gradiente*, oppure il *teorema di Fermat generalizzato* (infatti questo teorema sarà analoga alla sua controparte unidimensionale, 1).

#Teorema

> Teorema (test del gradiente o teorema di Fermat).
>
> Sia $f: E \subseteq \mathbb{R}^N \longrightarrow \mathbb{R}$ una funzione *differenziabile* (1).
> Sia $\underline{x_0}$ un punto *interno* del dominio E (ovvero in matematichese si dice $\underline{x_0} \in E^\circ$) (2).

Vale che *se* il punto $\underline{x}_0$ è un *estremo relativo*, allora il suo *gradiente* è nullo.

$$\nabla f(\underline{x}_0) = \underline{0}$$

#Dimostrazione

DIMOSTRAZIONE del Teorema 1 (test del gradiente o teorema di Fermat)

Introduco la funzione

$$g(t) = f(\underline{x}_0 + t\underline{e}_1)$$

con $\underline{e}_1$ il membro della base canonica $\mathcal{E}$.

Notiamo subito che $g(t = 0)$ è un *estremo relativo*. Infatti, per $t > 0$ ho che la funzione deve *"distanziarsi"* dall'estremo, in un modo o l'altro. Quindi per il teorema di Fermat ho

$$g'(0) = 0$$

Per la *differenziazione della composta* (1) possiamo scrivere

$$g'(\underline{x}_0) = \langle \nabla f(\underline{x}_0 + t\underline{e}_1), \underline{e}_1 \rangle = \frac{\partial f}{\partial x_1}(\underline{x}_0 + t\underline{e}_1)$$

possiamo scrivere pure

$$g'(0) = \langle \nabla f(\underline{x}_0), \underline{e}_1 \rangle = \frac{\partial f}{\partial x_1}(\underline{x}_0) = 0$$

Ripetendo il ragionamento per tutti i vettori $\underline{e}_i \in \mathcal{E}$ segue la tesi. ■

I punti critici. Questo teorema ci prepara per dare una definizione *ben posta* dei *punti critici*, come nel caso unidimensionale: il teorema di Fermat ci dà tutti i poteri effettivi per definire i punti stazionari.

A2. Punto Critico

Punto Critico per una Funzione in più variabili

Definizione di punto critico per una funzione in più variabili. Classificazione dei punti critici: punti estremi e punti di sella.

---

0. Voci correlate

1. Definizione di Punto Critico

Dal *test del gradiente* possiamo dare una definizione *ben posta* di un *punto critico*.

#Definizione

Definizione (punto critico per una funzione in più variabili).

Sia $f : E \subseteq \mathbb{R}^N \longrightarrow \mathbb{R}$ una funzione *differenziabile*. Si sceglie un punto interno $\underline{x_0} \in E^\circ$.

Si dice che *"$\underline{x_0}$ è un punto critico per f"* se vale che

$$\nabla f(\underline{x_0}) = \underline{0}$$

Dal *test del gradiente* sappiamo che tutti i punti estremi sono *punti critici*. Ma vale il contrario? Prendiamo il seguente esempio dal *Pagani-Salsa* (esempio 1.5., p. 61)

#Esempio

Esempio (esempio 1.5. del pagani salsa, pagina 61).

Sia $f(x,y) = y^2 - 3x^2y + 2x^4$. Si ha che $(0,0)$ è un *punto critico*.

Tuttavia, è un punto estremo?

Se prendo le *"direzioni uscenti dall'origine"* ovvero del tipo $f(x, mx), m \in \mathbb{R}$ allora trovo che $(0,0)$ è un *punto di minimo locale*. Infatti ho

$$f''(x) = 2m^2 - 18mx + 24x \implies f''(0) = 2m^2 > 0$$

Però il discorso cambia se prendiamo le *"direzioni uscenti dalle parabole"* (ovvero del tipo $f(x, mx^2)$). Infatti ho

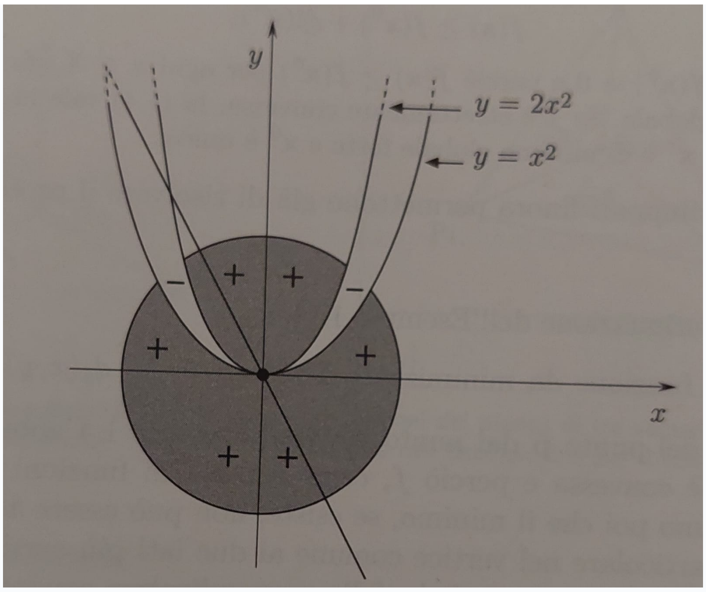
$$f''(x) = 2m^2x - 9mx^2 + 2x^4$$

troviamo che considerando un qualsiasi punto dell'intorno $B(0)$, abbiamo che

$$f''(x \in B(0)) < 0$$

che dimostra $(0,0)$ *non essere* un punto critico.

FIGURA 2.1. (*Esempio 1.5. del Pagani-Salsa, p. 61*)



Conclusione. Da qui c'è la necessità di classificare i *punti critici* in un altro modo, ovvero i *punti di sella*.

2. Definizione di Punto di Sella

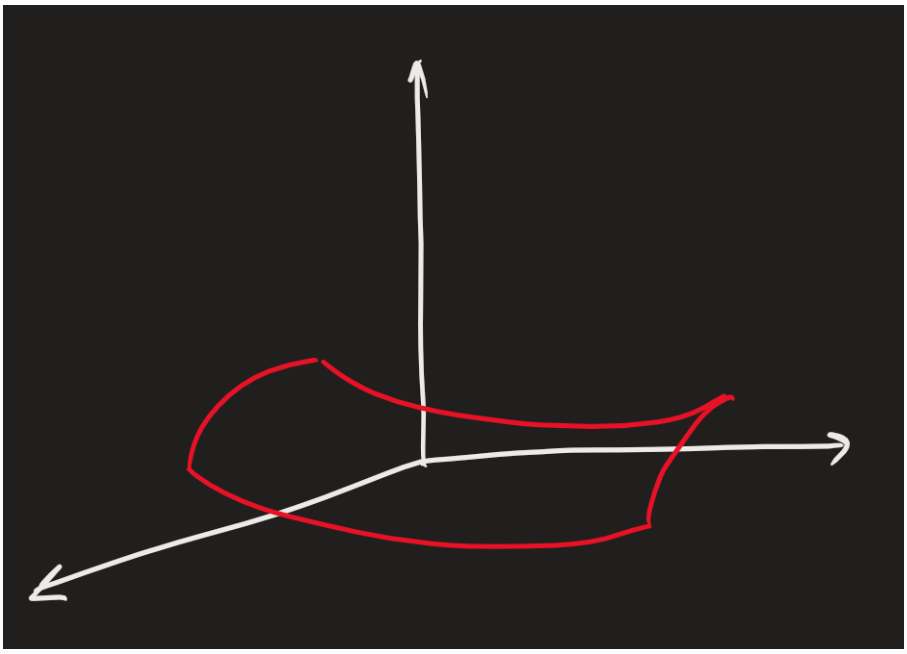
#Definizione

> Definizione (punto di sella).
>
> Sia $f : E \subseteq \mathbb{R}^N \longrightarrow \mathbb{R}$ una funzione *differenziabile*. Sia $\underline{x_0} \in E^\circ$ un *punto critico*.
>
> Si dice che $\underline{x_0}$ *è di sella* se prendendo un qualsiasi intorno $B(\underline{x_0})$ *esistono sia punti* maggiore di $f(\underline{x_0})$ che minore di $f(\underline{x_0})$.
>
> $$\forall B(\underline{x_0}), \exists \underline{x^+}, \underline{x^-} : f(\underline{x^+}) > f(\underline{x}) \land f(\underline{x^-}) < f(\underline{x})$$

#Esempio

> Esempio (le patatine delle Pringles).
>
> Un esempio è $f(x, y) = x^2 - y^2$ con $(0, 0)$ un *punto di sella*. Infatti da un lato ho una *"curvatura positiva"* e dall'altro una *"curvatura negativa"*.
>
> Se la figura della superficie dovesse sembrarvi familiare, molto probabilmente è dovuto al fatto che la funzione *assomiglia* alla forma delle patatine della Pringles. Infatti, la forma di questo cibo non è stato scelto a caso: questa forma conferisce una *robustezza maggiore*.

FIGURA 2.1. (*Pringles*)



3. Classificazione dei Punti Critici

Riassumiamo ciò che abbiamo visto col seguente teorema (nota questo teorema è stato scritto da me solamente per riassumere tutto).

#Teorema

> Teorema (della classificazione dei punti critici).
>
> Sia $f: E \subseteq \mathbb{R}^N \longrightarrow \mathbb{R}$ una funzione *differenziabile*. Si sceglie un punto interno $\underline{x_0} \in E^\circ$. Sia definito il valore del punto scelto come $\xi := f(\underline{x_0})$
>
> Allora vale che:
>
> $$\nabla f(\underline{x_0}) = 0 \implies (\underbrace{\xi \in \{\max f, \min f\}}_{i.} \vee \underbrace{\underline{x_0} \text{ è un punto di sella}}_{ii.})$$
>
> Ovvero o ξ è l'*estremo* della funzione ($i.$) o $\underline{x_0}$ è il *punto di sella* ($ii.$) (ovviamente abbiamo un disgiuntivo *aut-aut*).

#Dimostrazione

DIMOSTRAZIONE del Teorema 5 (della classificazione dei punti critici)
Omessa, basta riferirsi alle definizioni. ■

A3. Segno di una Matrice

Segno di una Matrice

Nozione preliminare per l'ottimizzazione dei liberi: segno delle matrici.

---

0. Voci correlate

- Forme Lineari e Quadratiche
- Matrice
- Determinante

1. Definizione del Segno di una Matrice

Prima di enunciare un *criterio* per *distinguere i punti critici*, definiamo il segno di una matrice (nozioni che useremo poi sulla matrice hessiana).

#Definizione

Definizione (segno di una matrice).

Sia $A \in M_{m,n}(\mathbb{R})$ una matrice. Sia $Q : \mathbb{R}^N \longrightarrow \mathbb{R}$ la sua *forma quadratica* associata (1).

Dato un qualsiasi $\underline{h} \neq \underline{0}$, si dice che Q è:

- *Positiva* se $Q(\underline{h}) > 0$
  - *Semipositiva* se $Q(\underline{h}) \geq 0$
- *Negativa* se $Q(\underline{h}) < 0$
  - *Seminegativa* se $Q(\underline{h}) \leq 0$
- *Indefinita* se $\exists \underline{u}, \underline{v} \in \mathbb{R}^N$ tali che

$$Q(\underline{v}) < 0 < Q(\underline{u})$$

Si definisce il *segno della sua matrice* come il *segno della sua forma quadratica* Q.

Vediamo delle *condizioni equivalenti* per classificare la *positività* e la *negatività* della matrice.

#Proposizione

Proposizione (condizioni equivalenti per la positività e la negatività del segno).

Sia Q una *forma quadratica*. Si ha che

$$Q \text{ positiva} \iff Q(\underline{h}) \geq m||\underline{h}||^2, \forall \underline{h} \in \mathbb{R}^N$$

e

$$Q \text{ negativa} \iff Q(\underline{h}) \leq m||\underline{h}||^2, \forall \underline{h} \in \mathbb{R}^N$$

#Dimostrazione

DIMOSTRAZIONE della Proposizione 2 (condizioni equivalenti per la positività e la negatività del segno).

Omessa. ■

2. Criterio di Sylvester

Vediamo il teorema più utile per poter classificare il segno della matrice.

#Teorema

Teorema (criterio di Sylvester).

Sia $Q : \mathbb{R}^N \longrightarrow \mathbb{R}$ una *forma quadratica* con $Q(\underline{h}) = \langle A \cdot \underline{h}, \underline{h} \rangle$. Sia A una matrice simmetrica ($A = {}^tA$).

Allora si ha che:

$$Q > 0 \iff \begin{cases} \det A_1 = a_{11} > 0 \\ \det A_2 = \det \begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix} > 0 \\ \vdots \\ \det A_N = \det A > 0 \end{cases}$$

Ovvero prendendo *tutte le determinanti di ogni sottomatrice di* A ho solo numeri positivi

Inoltre ho che

$$Q < 0 \iff \begin{cases} \det A_1 = a_{11} > 0 \\ \det A_2 = \det \begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix} < 0 \\ \vdots \\ (-1)^N \det A_N = (-1)^N \det A > 0 \end{cases}$$

Ovvero prendendo *tutte le determinante di ogni sottomatrice di* A ho un'oscillazione tra il negativo-positivo.

Se non vale nessuna delle condizioni equivalenti, si dice che il segno della Q è *indefinita*.

#Esempio

Esempio (caso $N = 2$).

Abbiamo che

$$
\begin{aligned}
&Q > 0 \iff a_{11} > 0 \land \det A > 0 \\
&Q < 0 \iff a_{11} < 0 \land \det A > 0 \\
&Q \gtrless 0 \text{ (indeterminata)} \iff a_{11} \in \mathbb{R} \land \det A < 0
\end{aligned}
$$

Adesso siamo pronti per enunciare il *test della Hessiana*.

A4. Test della Hessiana

Test della Hessiana

Breve descrizione qui

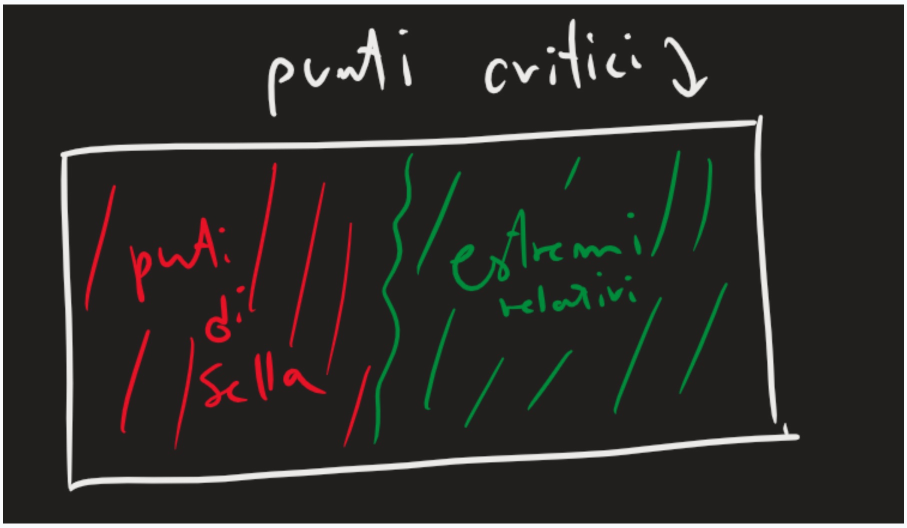
---

0. Voci correlate

- Segno di una Matrice
- Punto Critico per una Funzione in più variabili

1. Preambolo

Recap. Dal *teorema della classificazione dei punti critici* (1) abbiamo che nei *punti critici* abbiamo due *"sottoinsiemi"* di punti: gli *estremi relativi* e i *punti di sella*. Ovvero abbiamo la situazione del tipo



Tuttavia questo non ci basta. Vogliamo sviluppare degli strumenti per *distinguere* i *punti* *critici* a seconda della sua natura: dato un punto critico, voglio sapere se è un *estremo relativo* (in particolare di minimo o di massimo) o se è un *punto di sella*.

L'idea. L'idea è questa: prendiamo lo *sviluppo in serie di Taylor al secondo ordine* di una funzione (1). Sia $\underline{h} := \underline{x} - \underline{x_0}$.

$$f(\underline{x_0} + \underline{h}) - f(\underline{x_0}) = \langle \nabla f(\underline{x_0}), \underline{h} \rangle + \frac{1}{2} \langle Hf(\underline{x_0})(\underline{h}), \underline{h} \rangle + o\left(\|\underline{h}\|^2\right)$$

Possiamo cancellare alcune parti di questa equazione (quelle sottolineate in rosso). Prima di tutto sappiamo che $\underline{x_0}$ è un punto critico, quindi possiamo già cancellare il prodotto scalare $\langle \nabla f(\underline{x_0}), \underline{h} \rangle = 0$. Inoltre, il termine $o\left(\|\underline{h}\|^2\right)$ va a zero, quindi possiamo renderlo *"trascurabile"* nel senso di $o\left(\|\underline{h}\|^2\right) \to 0$. Infine abbiamo che rimane solo la parte

$$f(\underline{x_0} + \underline{h}) - f(\underline{x_0}) \simeq \frac{1}{2} \langle Hf(\underline{x_0})(\underline{h}), \underline{h} \rangle$$

Adesso, nel caso negativo della hessiana, supponiamo

$$\langle Hf_{\underline{x_0}}(\underline{h}), \underline{h} \rangle \leq m \|\underline{h}\|^2, m > 0$$

così possiamo studiare bene la matrice hessiana H. Nel caso unidimensionale avrei

$$f''(x_0)(x - x_0)(x - x_0) = f''(x_0)h^2 \leq mh^2 \implies f''(x_0) \leq m$$

Come studiamo la matrice H? Pensando all'equivalente unidimensionale, ho che se $f''(x)$ è *positivo* allora ho un *minimo locale*; se ho invece $f''(x)$ *negativo* allora ho un *massimo locale*. Vedremo che in caso generalizzato avremo una situazione simile.

2. Test della Hessiana

#Teorema

> Teorema (test della Hessiana).
>
> Sia $f : E \subseteq \mathbb{R}^N \longrightarrow \mathbb{R}$ *due-volte differenziabile* su $\underline{x_0} \in E^\circ$ e $\nabla f(\underline{x_0}) = \underline{0}$.
>
> Si ha che:
>
> $$\begin{gathered} Hf_{\underline{x_0}} > 0 \implies f(\underline{x_0}) \text{ minimo relativo} \\ Hf_{\underline{x_0}} < 0 \implies f(\underline{x_0}) \text{ massimo relativo} \\ Hf_{\underline{x_0}} \text{ indefinita} \implies \underline{x_0} \text{ di sella} \\ Hf_{\underline{x_0}} \underset{\text{o } \geq}{\leq} x_0 \implies \text{non posso dire nulla} \end{gathered}$$

C. OTTIMIZZAZIONE VINCOLATA

C1. Vincolo

Vincolo per una Funzione

Definizione di vincolo per una funzione (campo scalare). Definizione di estremo vincolato di una funzione in un vincolo.

---

0. Voci correlate

- Definizione di Estremo per una Funzione in più variabili

1. Definizione di Vincolo

Prima di poter parlare di *ottimizzazione vincolata*, parliamo (giustamente) di cosa intendiamo per un *"vincolo"*

#Definizione

> Definizione (vincolo per una funzione).
>
> Sia $f: E \subseteq \mathbb{R}^N \longrightarrow \mathbb{R}$ un *campo scalare*. Un insieme $V \subset E$ si dice *"vincolo di E per f* se soddisfa
>
> $$V \neq \emptyset, V \neq E$$
>
> Ovvero il *vincolo* dev'essere una *"parte selezionata"* di E.

#Osservazione

> Osservazione (caso comune).
>
> Di solito come *vincolo* poniamo delle *curve* o *superfici*, che vengono espresse come funzioni γ o superfici σ, poi per considerare i loro sostegni Γ, Σ.

2. Definizione di Estremo Vincolato

#Definizione

> Definizione (estremo vincolato per una funzione).

> Sia $f: E \subseteq \mathbb{R}^N \longrightarrow \mathbb{R}$ un *campo scalare* e V un suo *vincolo*.
>
> Un punto $\underline{x_0} \in V$ si dice *"estremo vincolato per f in V"* se vale che $\underline{x_0}$ è un *punto estremo* (1) di $f_{|V}$ (ovvero *"f ristretta in V"*).

Notiamo che non c'è nessuna relazione tra *estremi vincolati* e *punti estremi* (ovvero di $\max, \min$) per f vista *"globalmente"*. Come obbiettivo di questa sezione ci poniamo quello di *sviluppare* teoremi che ci diano delle *condizioni necessarie* per l'esistenza di punti vincolati, in tal modo da poter capirne la loro natura.

C2. Ottimizzazione in Curve e Superfici Parametriche

Ottimizzazione in Curve e Superfici Parametriche

Ottimizzazione in curve parametriche.

---

0. Voci correlate

- Curve e Superfici Parametriche
- Problemi di Ottimizzazione
- Vincolo per una Funzione

1. Teorema per l'Ottimizzazione in Curve Parametriche

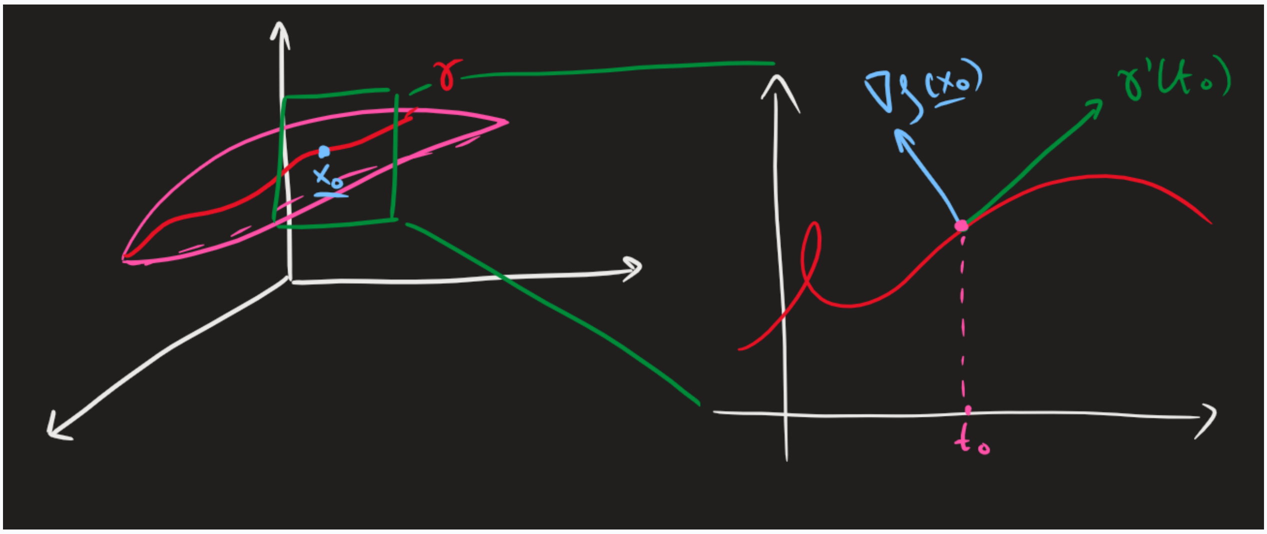
Vediamo una prima *condizione necessaria* per *estremi vincolati* su curve γ.

#Teorema

> Teorema (condizione necessaria per estremi vincolati, caso curvilineo in forma parametrica).
>
> Sia $f: A \subseteq \mathbb{R}^2 \longrightarrow \mathbb{R} \in \mathcal{C}^1(A)$. Sia $\gamma: I \in \mathcal{B}(\mathbb{R}) \longrightarrow A$ una *curva regolare semplice* (1) e sia $\underline{x_0} = \gamma(t_0) \in \Gamma$ un *punto regolare*, con $t_0 \in I^\circ$ e sia posto $\Gamma = V$ come vincolo.
>
> Allora vale che se $\underline{x_0}$ è un *estremo vincolato* per $f_{|\Gamma}$ allora vale che $\gamma'(t_0)$ e $\nabla f(\underline{x_0})$ sono *ortogonali*, ovvero

$$\boxed{\langle \nabla f(\underline{x_0}), \gamma'(t_0) \rangle = 0}$$

FIGURA 1.1. (*L'idea è questa*)



#Dimostrazione

DIMOSTRAZIONE del Teorema 1 (condizione necessaria per estremi vincolati, caso curvilineo in forma parametrica)

Consideriamo la composta $f(\gamma(t)) : I \subseteq \mathbb{R} \to A \to \mathbb{R}$, e la definiamo come $\psi(t)$. Poiché $\underline{x_0} = \gamma(t_0)$ è un *estremo vincolato*, abbiamo che t_0 è un *estremo libero* per ψ (1). Dunque per il *teorema di Fermat* (1), abbiamo che $\psi'(t_0) = 0$, da cui considerando la composizione di funzioni si ricava la tesi

$$\psi'(t_0) = \langle \nabla f(\underline{x_0}), \gamma'(t_0) \rangle = 0$$

che prova il teorema. ■

Notiamo che in questo caso abbiamo usato *curve già note* a priori. Nel caso in cui avessimo *curve implicite*, bisognerà passare al *teorema del Dini* per trovare la curva. Intanto generalizziamo questo teorema sulle *superfici*.

2. Ottimizzazione sulle Curve Parametriche

#Teorema

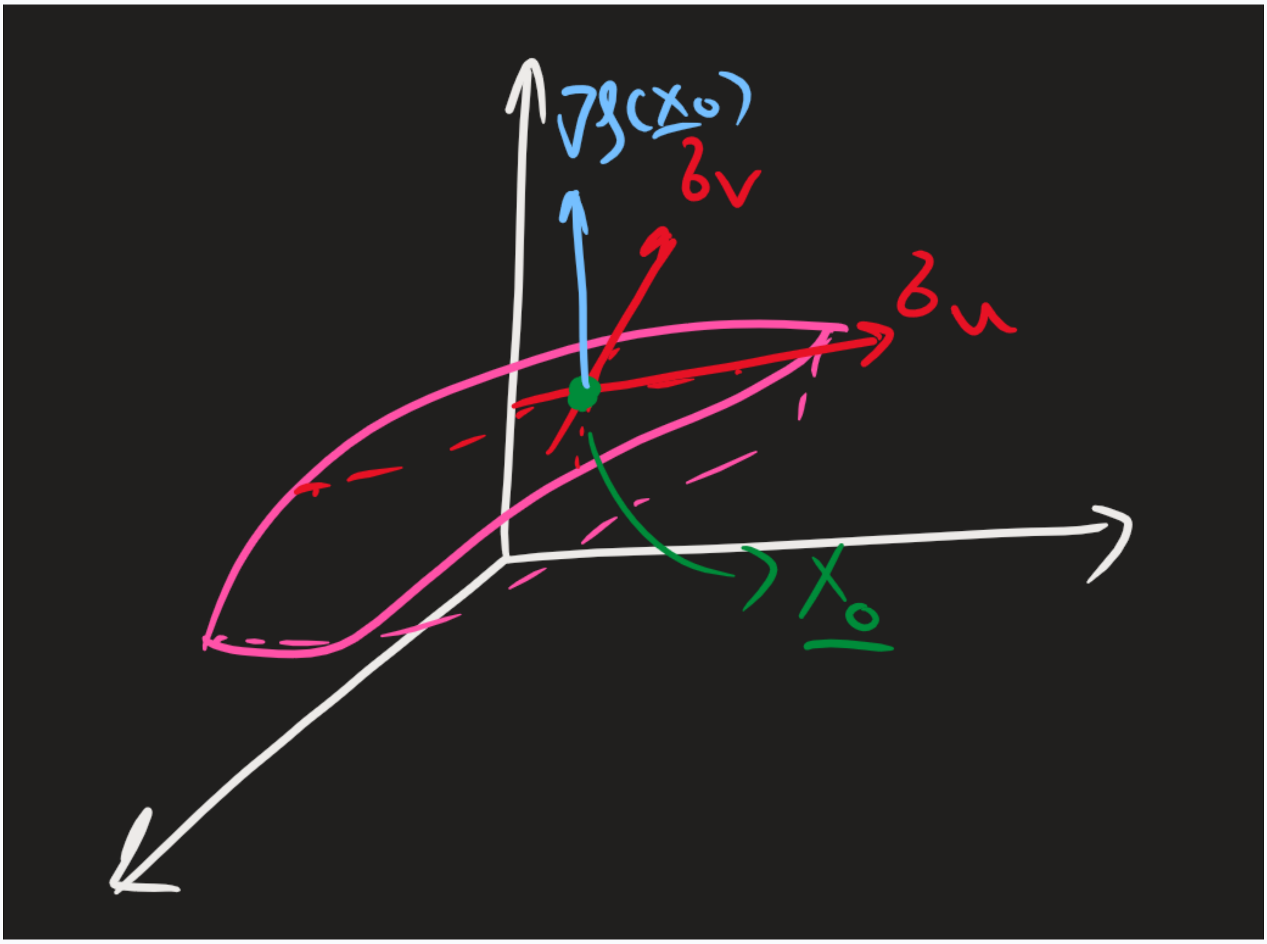
Teorema (condizione necessaria per estremi vincolata, caso superficie).

Sia $f : A \subseteq \mathbb{R}^3 \longrightarrow \mathbb{R} \in \mathcal{C}^1(A)$. Sia $\sigma : K \subseteq \mathbb{R}^2 \longrightarrow \mathbb{R}^3$ con K la *chiusura di un insieme aperto* (ovvero prendiamo $K = \overline{B^\circ}$) e σ una *superficie regolare semplice*.

Allora se $\underline{x_0} = \sigma(u_0, v_0) = (x(u_0, v_0), y(u_0, v_0), z(u_0, v_0)) \in \Sigma := \sigma(K)$ è un *punto di estremo vincolato* per $f_{|\Sigma}$ e se $\underline{u_0} = (u_0, v_0) \in B$, allora vale che le derivate σ_u, σ_v valutate in $\underline{u_0}$ sono ortogonali al gradiente $\nabla f(\underline{x_0})$. Ovvero

$$\boxed{\langle \nabla f(\underline{x_0}), \sigma_u(\underline{u_0}) \rangle = 0 = \langle \nabla f(\underline{x_0}), \sigma_v(\underline{u_0}) \rangle}$$

FIGURA 1.1. (*L'idea è questa*)



#Dimostrazione

DIMOSTRAZIONE del Teorema 2 (condizione necessaria per estremi vincolata, caso superficie)

La dimostrazione è completamente analoga a quella precedente (^71d25d). Ovvero, considerando la composizione $\psi = f \circ \sigma : K \subseteq \mathbb{R}^2 \longrightarrow \mathbb{R}$ abbiamo che ψ ha un *estremo* in $\underline{u_0} = (u_0, v_0)$ che implica, per il test del gradiente (1), $\nabla \psi(\underline{u_0}) = \underline{0}$. Da qui segue l'equazione (1)

$$\nabla \psi(\underline{u_0}) = \underline{0} \implies \psi_u(\underline{u_0}) = \psi_v(\underline{u_0}) = \underline{0} \implies \begin{cases} \langle \nabla f(\psi(\underline{u_0})), \sigma_u(\underline{u_0}) \rangle = 0 \\ \langle \nabla f(\psi(\underline{u_0})), \sigma_v(\underline{u_0}) \rangle = 0 \end{cases}$$

Considerando che $\psi(\underline{u_0})$ non è altro che $\underline{x_0}$ stessa, abbiamo la tesi. ■

C3. Metodo dei Moltiplicatori di Lagrange

Metodo dei Moltiplicatori di Lagrange

Metodo dei moltiplicatori del Lagrange. Teorema: enunciato e dimostrazione. Applicazione pratica del teorema.

---

0. Voci correlate

- Ottimizzazione in Curve e Superfici Parametriche
- Teorema del Dini
- Curva in Forma Implicita

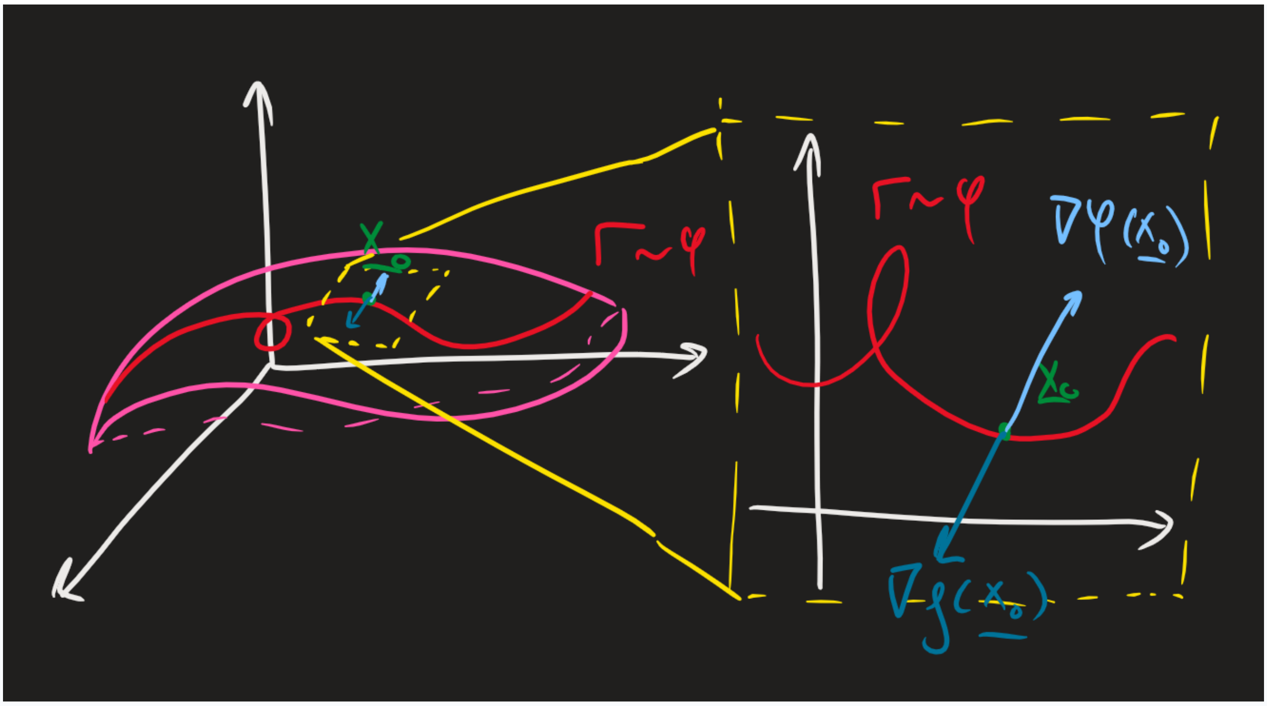
1. Risultato Teorico

Prima di parlare del *metodo dei moltiplicatori di Lagrange*, utilizzato per *ottimizzare* funzioni vincolate su *curve implicite*, enunciamo il risultato teorico.

#Teorema

> Teorema (dei moltiplicatori di Lagrange).
>
> Sia $f : A \subseteq \mathbb{R}^2 \longrightarrow \mathbb{R} \in \mathcal{C}^1(A)$ con A un aperto. Sia φ una funzione definita similmente, che rappresenta un suo *vincolo curvilineo*; infatti sia definita la curva Γ come la sua curva di livello $L_0(\gamma)$ (ovvero $\Gamma := L_0(\gamma)$).
>
> Se vale che $\underline{x_0} = (x_0, y_0) \in \Gamma$ è un *punto regolare* per Γ (ovvero $\nabla\varphi(\underline{x_0}) \neq 0$) e $\underline{x_0}$ è un *punto di estremo vincolato* per $f_{|\Gamma}$, allora
>
> $$\boxed{\exists \lambda \in \mathbb{R} : \nabla f(\underline{x_0}) = \lambda \nabla\varphi(\underline{x_0})}$$
>
> Ovvero il "*gradiente* $\nabla f(\underline{x_0})$ *è parallelo al gradiente della restrizione* $\nabla\varphi(\underline{x_0})$".

FIGURA 1.1. (*L'idea è questa*)



#Dimostrazione

DIMOSTRAZIONE del Teorema 1 (dei moltiplicatori di Lagrange)
Partiamo subito dal *teorema del Dini* (1), che ci dà una curva γ tale che il suo sostegno $\gamma(I)$ sia espressione di Γ per un certo intorno W di $\underline{x}_0$. Ovvero,

$$\nabla\varphi(\underline{x}_0) \neq \underline{0} \implies \exists W(\underline{x}_0) : \Gamma \cap W = \gamma(I)$$

In particolare sappiamo che questa è una *curva cartesiana*, ovvero una curva del tipo $\gamma(t_0) = \underline{x}_0$ e $\gamma(t) = (t, g(t))$ (scegliamo il caso in cui esista la g, l'altro caso si dimostra identicamente). Dato che ho una *curva parametrica*, ho che $\underline{x}_0 = \gamma(t_0)$ è un *estremo* per $f_{|\Gamma}$, il che ci dà (2)

$$(*)\ \langle \nabla f(\underline{x}_0), \gamma'(t_0)\rangle = 0$$

Mi ricordo delle *ipotesi necessarie* per il *teorema del Dini*, ovvero che $\nabla\varphi(\underline{x}_0) \neq \underline{0}$ e $\gamma'(t_0) \neq \underline{0}$. Dopodiché considero un altro risultato dello stesso teorema, che ci dà il risultato (3)

$$(**)\ \nabla\varphi(\underline{x}_0) \neq \underline{0}, \gamma'(t_0) \neq \underline{0}, \langle \nabla\varphi(\underline{x}_0), \gamma'(t_0)\rangle = 0$$

Combinando le $(*)$ e $(**)$, scopro che $\nabla\varphi(\underline{x}_0)$ è ortogonale a $\gamma'(t_0)$ come lo è pure $\nabla f(\underline{x}_0)$. In definita abbiamo il risultato finale

$$\nabla f(\underline{x}_0) = \lambda \nabla\varphi(\underline{x}_0)$$

che prova la tesi. ■

2. Punto di vista pratico

#Osservazione

Osservazione (punto di vista pratico).

Da un punto di vista pratico, possiamo applicare questo teorema per risolvere problemi di *minimizzazione vincolata* su *curve regolari in forma implicita* φ (ovvero che soddisfano sempre il *teorema del Dini*).

Nel caso sfortunato in cui avessimo *curve non regolari*, basta separarli in *due casi*: una in cui possiamo usare ancora i *moltiplicatori di Lagrange* (ovvero il gradiente non è nullo) e l'altra in cui i punti vengono studiati separatamente (dove il gradiente è nullo).

In definita, qualora fossimo nel primo caso, possiamo introdurre la funzione $\mathcal{L}(x, y, \lambda)$ (detta *"lagrangiana"*) definita come

$$\mathcal{L}(x, y, \lambda) := f(x, y) - \lambda\varphi(x, y)$$

dopodiché per risolvere il problema di minimizzazione basta risolvere il sistema

$$\nabla\mathcal{L} = \underline{0} : \begin{cases} \mathcal{L}_x = f_x - \lambda\varphi_x = 0 \\ \mathcal{L}_y = f_y - \lambda\varphi_y = 0 \\ \mathcal{L}_\lambda = \varphi = 0 \end{cases}$$

Le prime due equazioni coincidono col *teorema dei moltiplicatori di Lagrange*, l'ultima è invece la *condizione di vincolo*.

Questo è equivalente alla scrittura

$$\nabla f(x, y) = \lambda\nabla\varphi(x, y) \wedge \varphi(x, y) = 0$$

3. Generalizzazione

Si può generalizzare il risultato appena visto. Lo facciamo in particolare per $N = 3$ e caso superficiale (con superfici).

#Teorema

Teorema (dei moltiplicatori di Lagrange).

Sia $f : A \subseteq \mathbb{R}^3 \longrightarrow \mathbb{R} \in \mathcal{C}^1(A)$ con A un aperto. Sia φ una funzione definita similmente, che rappresenta un suo *vincolo curvilineo*; infatti sia definita la curva Σ come la sua curva di livello $L_0(\gamma)$ (ovvero $\Sigma := L_0(\gamma)$).

Se vale che $\underline{x_0} = (x_0, y_0, z_0) \in \Sigma$ è un *punto regolare* per Σ (ovvero $\nabla\varphi(\underline{x_0}) \neq 0$) e $\underline{x_0}$ è un *punto di estremo vincolato* per $f_{|\Sigma}$, allora

$$\boxed{\exists\lambda \in \mathbb{R} : \nabla f(\underline{x_0}) = \lambda\nabla\varphi(\underline{x_0})}$$

> Ovvero il *"gradiente $\nabla f(\underline{x}_0)$ è parallelo al gradiente della restrizione $\nabla\varphi(\underline{x}_0)$"*.

Le applicazioni pratiche sono identiche, con la lagrangiana definita come

$$\mathcal{L}(x,y,z,\lambda) := f(x,y,z) - \lambda(\varphi(x,y,z))$$

e la risoluzione si applica analogamente, ponendo $\nabla\mathcal{L} = \underline{0}$.

D. ESERCIZI

Esercizi sulla Ottimizzazione in più variabili

Esercizi sull'ottimizzazione in più variabili.

Sezione A.

Keywords: funzione coerciva, proprietà della funzione coerciva.

#Esercizio

> Esercizio.
>
> Stabilire se $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ definita come
>
> $$f(x,y) = x^4 + y^4 + 3xy$$
>
> ammette minimo o massimo.
>
> *Consiglio: usare la trasformazione in coordinate polari, mediante $x = \rho\cos\theta; y = \rho\sin\theta$.*

#Esercizio

> Esercizio.
>
> Sia $f(x) = 3x$ e $g(x,y) = x - 12xy + 8y^3$.
> Stabilire se le funzioni f, g siano coercive o meno.

Sezione B.

Keywords: segno della matrice, ricerca dei minimi e massimi locali liberi, test della Hessiana, test del gradiente.

#Esercizio

Esercizio.

Si cerchino i punti di massimo e minimo locali per la funzione

$$f(x, y) = x^2 - 2x + y^4 + y^2$$

#Esercizio

Esercizio.

Sia la funzione definita come

$$f(x, y) = e^{2x^2+y^2+xy}$$

Cercare i punti critici per f. Dire eventualmente di che tipologia sono.

Sezione C.

#Esercizio

Esercizio.

Sia $f(x, y) = x + y$ ristretta sulla curva $\Gamma : x^4 + y^4 - 4xy = 1$. Dire se esistono gli estremi, se sì determinarli.

#Esercizio

Esercizio.

Determinare gli estremi della funzione $f(x, y) = x$ in $E = \{(x, y) \in \mathbb{R}^2 : x^3 = y^2\}$. Notare che E non è regolare!

#Esercizio

Esercizio.

Data $f : \mathbb{R}^3 \longrightarrow \mathbb{R}$ come

$$f(x, y, z) = x + y + z$$

e la superficie Σ come

$$\Sigma = \{(x, y, z) \in \mathbb{R}^3 : x^2 - xy + y^2 + z^2 = 1\}$$

Trovare estremi su $f_{|\Sigma}$.

#Esercizio

Esercizio.

Trovare i massimi e minimi assoluti di

$$f(x, y) = y^3 + 4x^2y - 4y$$

sulla restrizione $E = \left\{(x, y) \in \mathbb{R}^2 : \frac{1}{2} \leq x^2 + y^2 \leq 1\right\}$.