

DISPENSA DI OTTIMIZZAZIONE MATEMATICA PER L'ECONOMIA

A.A. 2012/13

FRANCESCO RANIA¹

INDICE

1. Introduzione all'Ottimizzazione	1
1.1. Massimi e minimi di una funzione scalare reale	2
1.2. Il problema di ottimizzazione	4
2. Ottimizzazione libera	5
2.1. Calcolo della soluzione locale	5
2.2. Calcolo della soluzione globale	7
2.3. Esempi	7
3. Ottimizzazione con vincoli di uguaglianza	11
3.1. Metodo di sostituzione	11
3.2. Metodo dei moltiplicatori di Lagrange	12
3.3. Interpretazione geometrica del metodo dei moltiplicatori di Lagrange	15
3.4. La condizione del secondo ordine	16
3.5. Esempi	19
3.6. Significato dei moltiplicatori di Lagrange	21
4. Ottimizzazione con vincoli di disuguaglianza	23
4.1. Problema di massimo con vincoli di non positività	23
4.2. Problema di massimo misto	26
4.3. Problema di minimo	27
Appendice	28
4.4. A proposito delle matrici quadrate	28
4.5. A proposito delle funzioni scalari reali	28

Avvertenze: Per semplicità e brevità di esposizione, in tale dispensa, l'Ottimizzazione sarà sviluppata in $\mathbb{R}^n$ con $n < +\infty$. Per il caso $n = +\infty$ e per tutti quegli argomenti, come da programma, che sono qui omissi, si rimanda lo studente ai cap. §.5, §.6, §.7 di A. VAGLIO, *Matematica per economisti*, Apogeo 2004.

1. INTRODUZIONE ALL'OTTIMIZZAZIONE

Un agente deve operare una scelta tra diversi panieri in base all'utilità che da essi ne trarrebbe; un'azienda vuole riorganizzare capitale e lavoro per il processo di produzione affinché vi sia un incremento di profitto ed un abbattimento dei costi; ecc. Sono soltanto due dei numerosi esempi in Economia, che ci permettono di comprendere l'importanza della ricerca degli stadi privilegiati di un processo economico a cui corrispondono il valore più alto del fenomeno (vedi utilità e profitto) oppure il valore più basso del fenomeno (vedi costi). Tale ricerca è denominata *problema di ottimizzazione*. Pertanto l'*optimum* o più semplicemente ottimo rappresenta la soddisfazione massima che l'operatore raggiunge in merito ad un processo.

Prima di effettuare una rigorosa formalizzazione, risulta doveroso effettuare alcune precisazioni. Implementiamo di informazioni il consumatore alle prese con la scelta tra i panieri e fotografiamo tre diverse situazioni²:

- 1) Il consumatore in modo "edonistico" cerca l'appagamento dei suoi desideri scegliendo il panieriere migliore tra tutti quelli esistenti;
- 2) Il consumatore consapevole di non aver a disposizione credito proprio illimitato e né di poter accedere a forme di credito esogene effettua la scelta solo tra quei panieri che può permettersi di acquistare impiegando interamente la propria ricchezza;

¹Department of Legal, Historical, Economic and Social Sciences, Magna Graecia University of Catanzaro, Campus loc. Germaneto, Viale Europa, 88100 Catanzaro, Italy. Email: raniaf@unicz.it

² Invitiamo lo studente a cimentarsi nella costruzione di tre situazioni aziendali differenti il cui fine è la ricerca dell'ottimo in termini di ricavo o abbattimento dei costi, dimostrando che il livello di soddisfazione pur essendo massimo in ognuna delle circostanze ha però valore differente caso per caso. Le conclusioni a cui perverrà dovranno essere le stesse di quelle ottenute dall'analisi della ricerca della massima utilità di un consumatore.

- 3) Il consumatore oltre ad essere consapevole dei limiti nell'acquisto dei panieri come nel caso 2) effettua la scelta in modo previdente contemplando anche l'eventualità che parte della sua ricchezza non venga spesa.

La semplicità del consumatore del caso 1) porta a considerare il problema in *forma libera* secondo cui il suo ottimo coincide con la massima soddisfazione teorica, con il payoff dell'“Iperuranio”. Il consumatore del caso 2) non può ignorare le regole di acquisto ed unitamente all'impiego della totale ricchezza posseduta considera l'ottimizzazione in *forma vincolata*. Pertanto la soddisfazione massima da raggiungere, deve essere tale che debba valere la condizione: bilancio di pareggio tra ciò che acquista (paniere) e ciò che possiede (reddito). Infine la situazione 3) è una variante del caso 2) ancora più restrittiva, perchè la scelta ricade al più tra i panieri che il consumatore può permettersi, precisando che il disavanzo tra la ricchezza del suo reddito e ciò che ha acquistato è pari al grado di previdenza circa il futuro.

La modellizzazione del problema di ottimizzazione, ad opera della Matematica, scienza preposta all'importante compito, avverrà allora per gradi. Precisato che l'ottimo, a seconda dei casi, coincide con il minimo o il massimo di una funzione scalare reale ad una o più variabili, che rappresenta il fenomeno economico in esame³, dapprima affronteremo l'*ottimizzazione libera* (caso 1). Successivamente, introducendo i vincoli, considereremo l'*ottimizzazione vincolata con vincoli di uguaglianza* (caso 2) e l'*ottimizzazione vincolata con vincoli di disuguaglianza* (caso 3). Inoltre, tutta la teoria sarà sviluppata nella direzione dell'ottimo con l'eccezione di quella con vincoli di disuguaglianza in cui le condizioni di ottimalità non saranno le stesse tra problemi di minimo e problemi di massimo. In tale circostanza preferiremo svolgere la trattazione con riferimento ai problemi di un tipo e poi comprendere come ricondurre i problemi dell'altro tipo sempre ai primi.

1.1. Massimi e minimi di una funzione scalare reale. Siano $f : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ una funzione scalare reale a n variabili reali e $\mathbf{x}_0 \in D$ un punto interno al dominio.

Definizione 1.1. Definiamo $\mathbf{x}_0 \in D$

- punto di *massimo globale* (risp. di *minimo globale*) per f se

$$f(\mathbf{x}_0) \geq f(\mathbf{x}) \quad (\text{risp. } f(\mathbf{x}_0) \leq f(\mathbf{x})) \quad \forall \mathbf{x} \in D$$

- punto di *massimo locale* (risp. di *minimo locale*) per f se

$$\exists \delta > 0 : \forall \mathbf{x} \in B_\delta(\mathbf{x}_0) \cap D \Rightarrow f(\mathbf{x}_0) \geq f(\mathbf{x}) \quad (\text{risp. } f(\mathbf{x}_0) \leq f(\mathbf{x}))$$

Dalla Definizione 1.1 si evince che la differenza sostanziale tra un punto di massimo (risp. minimo) globale e un punto di massimo (risp. di minimo) locale è che il secondo lo è esclusivamente in un suo intorno B_δ mentre il primo lo è su tutto il dominio D della funzione f . Questo fatto determina allora i primi risultati di interesse di cui è omessa la dimostrazione.

Teorema 1. Se $\mathbf{x}_0 \in D$ è punto di massimo globale (risp. minimo globale) per f allora è anche punto di massimo locale (risp. minimo locale) per f .

Osservazione 1. Il viceversa del Teorema 1 non vale. Infatti la funzione $f(x) = x^3 - 3x^2 + 5$ ha 0 come punto di massimo locale e 2 come punto di minimo locale ma non ammette massimo globale e minimo globale in quanto per $x \rightarrow +\infty$ si ha $f(x) = +\infty$ e per $x \rightarrow -\infty$ si ha $f(x) = -\infty$.

Teorema 2. Sia $\mathbf{x}_0 \in D$ è un punto di massimo (risp. di minimo) locale per f .

- Se f è convessa (risp. concava) allora $\mathbf{x}_0$ è un punto di massimo (risp. di minimo) globale per f .
- Se f è strettamente convessa (risp. strettamente concava) allora $\mathbf{x}_0 \in D$ è l'unico punto minimo (risp. di massimo) globale per f .

Esempio 1.1. La funzione $f(x, y) = x^2 + y^2$ è strettamente convessa e pertanto il punto $O = (0, 0)$ di minimo locale è anche un punto di minimo globale e l'unico (vedi Figura 1).

³ Per il consumatore la funzione dei payoff è l'utilità ed il problema consiste nella ricerca del massimo. Se invece il decisore è l'azienda allora la scelta da effettuare ricadrà sul vettore input-output massimo qualora si parli di profitto e sul vettore input-output minimo se in esame è la funzione costo.

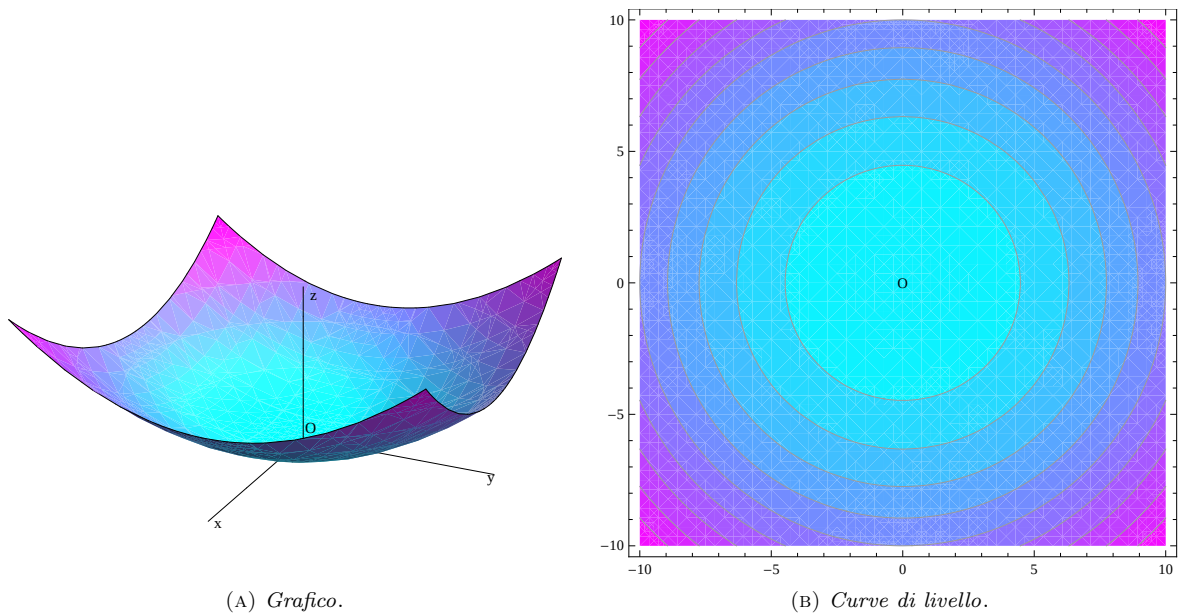


FIGURA 1. Il punto O è un minimo locale ed un minimo globale per la $f(x, y) = x^2 + y^2$

Esempio 1.2. La funzione $f(x, y) = -x^2 - y^2$ è strettamente concava e pertanto il punto $O = (0, 0)$ di massimo locale è anche un punto di massimo globale e l'unico (vedi Figura 2).

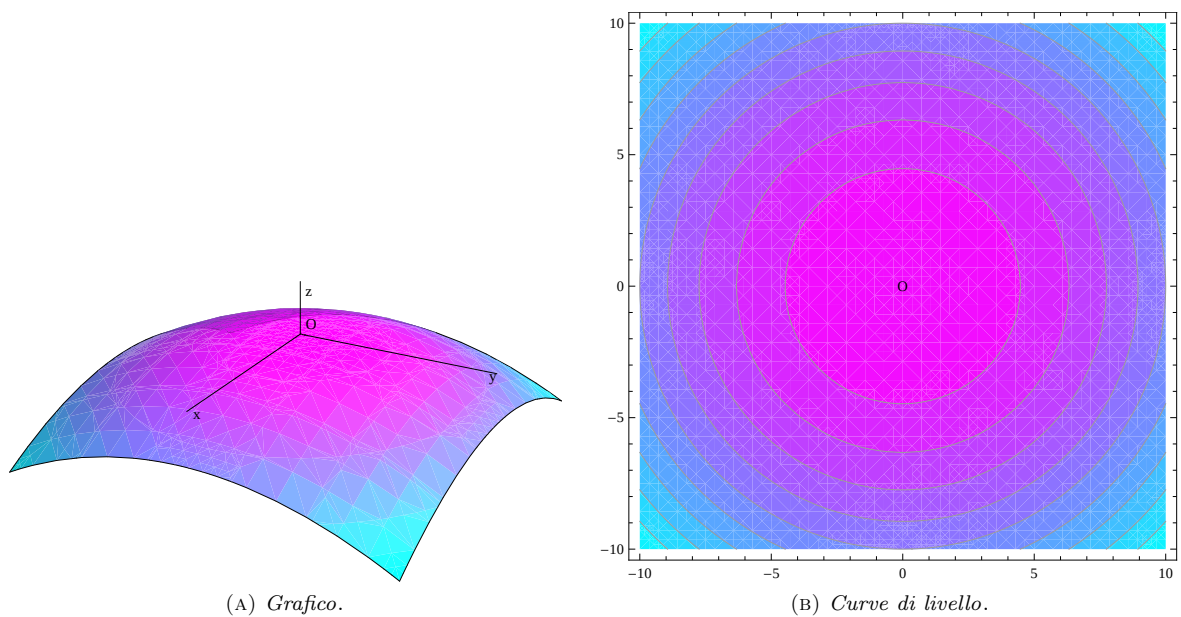


FIGURA 2. Il punto O è un massimo locale ed un massimo globale per la $f(x, y) = -x^2 - y^2$

Teorema 3 (di Weierstrass). Sia $f : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ tale che $f \in \mathcal{C}(\mathbb{R}^n, \mathbb{R})$ e D è un compatto di $\mathbb{R}^n$. Allora f ammette almeno un punto di massimo globale e un punto di minimo globale.

Osservazione 2. Nel Teorema 3 la continuità di f su un dominio compatto garantisce che f sia limitata inferiormente e superiormente e quindi di conseguenza che f ammetta almeno un punto di minimo e un punto di massimo globale (vedi Figura 3).

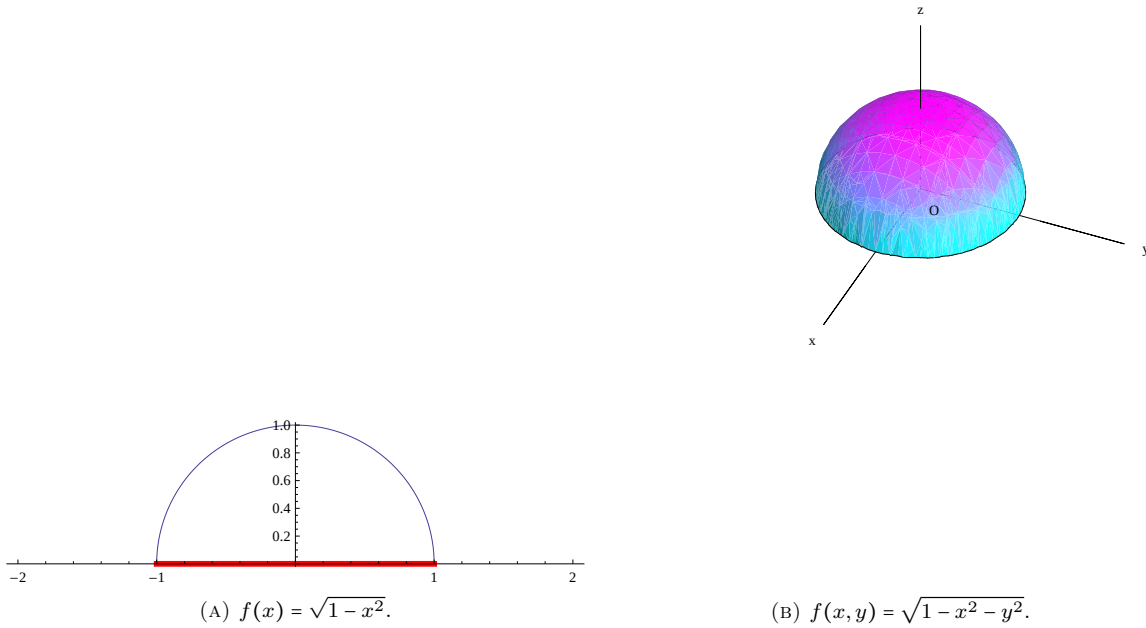


FIGURA 3. Due funzioni limitate perché continue su un compatto

1.2. Il problema di ottimizzazione. La ricerca dell'*optimum* di una situazione economica dovrà senz'altro tener conto di alcune condizioni reali esogene ed endogene che nel modello matematico si tradurrebbero in restrizioni della funzione f . Inoltre la natura delle condizioni limitanti la f detterà la tecnica di indagine da usare.

Siano $f : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ una funzione scalare reale tale che $f \in C^2(\mathbb{R}^n, \mathbb{R})$, $\mathfrak{P}$ una proprietà, $K = \{\mathbf{x} \in \mathbb{R}^n : \mathfrak{P}(\mathbf{x})\}$ l'insieme dei punti di $\mathbb{R}^n$ che verificano la proprietà $\mathfrak{P}$.

La funzione f è detta *funzione obiettivo*, $\mathfrak{P}$ è detta *relazione vincolare*⁴, $\mathfrak{P}(\mathbf{x})$ è detto *vincolo*, K è detto *insieme vincolante* e $V := D \cap K$ è detta *regione ammissibile*.

Problema 1 (di massimizzazione). Definiamo problema di massimizzazione per la funzione f sottoposta al vincolo $\mathfrak{P}(\mathbf{x})$ la ricerca del punto $\mathbf{x}_0$ di massimo globale per la funzione f ristretta alla regione ammissibile. La sua formalizzazione assume una delle seguenti scritture:

$$\text{Trovare} \quad \mathbf{x}_0 \in D \cap K \quad : \quad f(\mathbf{x}_0) \geq f(\mathbf{x}) \quad \forall \mathbf{x} \in D \cap K$$

$$\text{Trovare} \quad \begin{cases} \max_{\mathbf{x}} f(\mathbf{x}) \\ \text{t.c.} \quad \mathbf{x} \in D \cap K \end{cases}$$

$$\text{(P-Max)} \quad \text{Trovare} \quad \max_{\mathbf{x} \in V} f(\mathbf{x})$$

$$\text{Trovare} \quad \max_{\mathbf{x}} f|_V(\mathbf{x})$$

In modo analogo possiamo definire il problema di minimizzazione.

Problema 2 (di minimizzazione). Definiamo problema di minimizzazione per la funzione f sottoposta al vincolo $\mathfrak{P}(\mathbf{x})$ la ricerca del punto $\mathbf{x}_0$ di minimo globale per la funzione f ristretta alla regione ammissibile. La sua formalizzazione assume una delle seguenti scritture:

$$\text{Trovare} \quad \mathbf{x}_0 \in D \cap K \quad : \quad f(\mathbf{x}_0) \leq f(\mathbf{x}) \quad \forall \mathbf{x} \in D \cap K$$

$$\text{Trovare} \quad \begin{cases} \min_{\mathbf{x}} f(\mathbf{x}) \\ \text{t.c.} \quad \mathbf{x} \in D \cap K \end{cases}$$

⁴ In seguito riconoscendo in $\mathfrak{P}$ una funzione vettoriale o un vettore di componenti delle funzioni vettoriali, si dirà più propriamente *funzione vincolare*.

$$(P\text{-min}) \quad \text{Trovare} \quad \min_{\mathbf{x} \in V} f(\mathbf{x})$$

$$\text{Trovare} \quad \min_{\mathbf{x}} f|_V(\mathbf{x})$$

Osservazione 3. Poiché un punto di minimo (*risp.* di massimo) per la funzione f è un punto di massimo (*risp.* di minimo) per la funzione $-f$ allora sussistono le seguenti relazioni tra (P-Max) e (P-min)

$$(1) \quad \max_{\mathbf{x} \in V} f(\mathbf{x}) = \min_{\mathbf{x} \in V} -f(\mathbf{x}) \quad (2) \quad \max_{\mathbf{x} \in V} -f(\mathbf{x}) = \min_{\mathbf{x} \in V} f(\mathbf{x})$$

Se indichiamo con $\tilde{\mathbf{x}} \in V$ l'*ottimo* di f in V e quindi $f(\tilde{\mathbf{x}})$ è il *valore ottimale* di f , diremo che $\tilde{\mathbf{x}} \in V$ è un *ottimo locale* (*risp. globale*) se è soluzione locale⁵ (*risp. globale*) del problema di ottimizzazione nella forma (P-Max) o nella forma (P-min). Inoltre se l'ottimo globale (*risp. locale*) sarà unico allora diremo $\tilde{\mathbf{x}} \in V$ punto di *ottimo forte* (*risp. ottimo debole*).

2. OTTIMIZZAZIONE LIBERA

Nel caso in cui $K = D = V$ ⁶, ovvero nel caso in cui l'insieme ammissibile sia il dominio di f , il problema di ottimizzazione nelle forme (P-Max) e (P-min) diventa la semplice ricerca del punto di massimo e del punto di minimo della funzione f

$$(FP\text{-Max}) \quad \text{Trovare} \quad \max_{\mathbf{x} \in D} f(\mathbf{x})$$

$$(FP\text{-min}) \quad \text{Trovare} \quad \min_{\mathbf{x} \in D} f(\mathbf{x})$$

2.1. Calcolo della soluzione locale. Per determinare le soluzioni locali dei problemi (P-Max) e (P-min) ricorremo al Calcolo Differenziale. Pertanto supporremo che $f : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ sia tale che $f \in \mathcal{C}^2(\mathbb{R}^n, \mathbb{R})$.

Definizione 2.1. Il punto $\mathbf{x}_0 \in \overset{\circ}{D}$ è detto *punto stazionario* per f se il gradiente di f in $\mathbf{x}_0$ è il vettore nullo di $\mathbb{R}^n$ ($\nabla f(\mathbf{x}_0) = \mathbf{0}$)⁷. In caso contrario $\mathbf{x}_0$ è detto *punto regolare*.

Teorema 4 (condizione del primo ordine).⁸

Ogni punto di massimo (risp. minimo) locale è un punto stazionario.

Dimostrazione. Sia $\mathbf{x}_0 = (x_1^0, x_2^0, \dots, x_n^0) \in \overset{\circ}{D}$ un punto di massimo locale di f . Fissato $1 \leq i \leq n$ costruiamo la funzione scalare reale univariabile h definita dalla legge $h(x_i) := (x_1^0, x_2^0, \dots, x_{i-1}^0, x_i, x_{i+1}^0, \dots, x_n^0)$ con x_i la i -esima componente di un qualsiasi $\mathbf{x}$ preso nell'intorno di $\mathbf{x}_0$. Per il Teorema di Fermat la derivata prima in x_i esiste e finita e vale 0 ovvero

$$(3) \quad h'(x_i) = f_{x_i}(\mathbf{x}_0) = 0$$

Poiché la (3) vale per ogni $i = 1, 2, \dots, n$ ne segue che ogni derivata parziale di f è nulla in un $\mathbf{x}$ prossimo a $\mathbf{x}_0$ secondo la direzione della variabile. Pertanto il gradiente di f è nullo ovvero $\mathbf{x}_0$ è punto stazionario. $\square$

Osservazione 4. Il Teorema 4 non è invertibile e pertanto è noto anche come condizione necessaria dell'esistenza dei punti di ottimo. In Figura 4 sono presenti due funzioni $f(x) = x^3$ e $g(x, y) = x^2 + y^2$ che ammettono rispettivamente in $x_0 = 0$ ed in $\mathbf{x}_0 = (0, 0)$ un punto stazionario che non è un ottimo.

⁵ L'aggettivo «locale» è inteso per sottolineare che $\tilde{\mathbf{x}}$ è ottimo di f in un suo intorno.

⁶ È evidente che $\mathfrak{P} = f$.

⁷ In modo del tutto equivalente potremmo affermare che $\mathbf{x}_0 \in \overset{\circ}{D}$ è un punto di stazionarietà per f se in esso si annullano tutte le derivate parziali di f .

⁸ Tale Teorema è la versione generale al caso $n > 1$ del Teorema di Fermat delle funzioni ad una variabile secondo cui:

Sia $f : I \subseteq \mathbb{R} \rightarrow \mathbb{R}$ tale che $f \in \mathcal{C}(\mathbb{R})$. Se $x_0 \in \overset{\circ}{I}$ è un punto di massimo (*risp. di minimo*) locale per f allora $f'(x_0) = 0$. Inoltre tale teorema è detto «del primo ordine» in quanto impiega le derivate parziali prime di f .

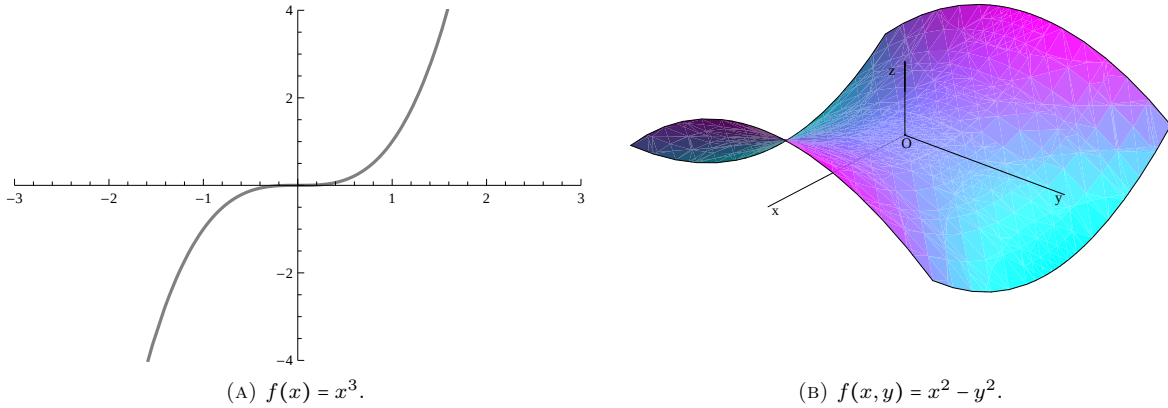


FIGURA 4. Due controesempi del Teorema 4

Dalla Osservazione 4 discende che i punti di stazionarietà per f che non sono punti di ottimo per f sono detti *punti di sella*⁹ per f .

Teorema 5 (condizione del secondo ordine).¹⁰ Siano $f : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ tale che $f \in C^2(\mathbb{R}^n, \mathbb{R}^2)$ e $\mathbf{x}_0 \in \overset{\circ}{D}$ stazionario. Allora

- (i) se $\mathbf{x}_0$ è di minimo (risp. massimo) locale debole per f allora la forma quadratica associata ad $\mathbf{H}_f(\mathbf{x}_0)$ è semidefinita positiva (risp. semidefinita negativa);
- (ii) se la forma quadratica associata a $\mathbf{H}_f(\mathbf{x}_0)$ è definita positiva (risp. negativa) allora $\mathbf{x}_0$ è di minimo (risp. massimo) locale forte per f .

Dimostrazione. La formula di Taylor ci suggerisce che in un intorno di $\mathbf{x}_0$ la funzione f può scriversi per $\mathbf{x} \rightarrow \mathbf{x}_0$ come

$$(4) \quad f(\mathbf{x}) - f(\mathbf{x}_0) = \nabla f(\mathbf{x}_0)\Delta\mathbf{x} + \frac{1}{2}\mathbf{H}_f(\mathbf{x}_0)(\Delta\mathbf{x})^2 + o(\|\mathbf{x} - \mathbf{x}_0\|)$$

con $o(\|\mathbf{x} - \mathbf{x}_0\|)$ infinitesimo di ordine superiore a $\|\mathbf{x} - \mathbf{x}_0\|$. Inoltre se poniamo $\mathbf{x} = \mathbf{x}_0 + \mathbf{h}$ allora la (4) può essere riscritta per $\mathbf{h} \rightarrow \mathbf{0}$ come

$$(5) \quad f(\mathbf{x}_0 + \mathbf{h}) - f(\mathbf{x}_0) = \nabla f(\mathbf{x}_0)\mathbf{h} + \frac{1}{2}\mathbf{h}^T\mathbf{H}_f(\mathbf{x}_0)\mathbf{h} + o(\|\mathbf{x} - \mathbf{x}_0\|)$$

Essendo $\mathbf{x}_0$ un punto stazionario, applicando il risultato del Teorema 4 nella (5) e passando al limite per $\mathbf{h} \rightarrow \mathbf{0}$ otteniamo che la funzione può essere approssimata alla sola forma quadratica associata all'Hessiano calcolato in $\mathbf{x}_0$

$$f(\mathbf{x}_0 + \mathbf{h}) - f(\mathbf{x}_0) \approx \frac{1}{2}\mathbf{h}^T\mathbf{H}_f(\mathbf{x}_0)\mathbf{h}.$$

Quindi se la quantità a secondo membro è positiva si ha che $\mathbf{x}_0$ è di minimo locale altrimenti di massimo locale. $\square$

Combinando il Teorema 5 con il Teorema 17 (vedi § 4.7.2. dell'Appendice) possiamo fornire una condizione del secondo ordine semplificata.

Teorema 6 (condizione del secondo ordine semplificata). Siano $f : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ tale che $f \in C^2(\mathbb{R}^n, \mathbb{R})$ e $\mathbf{x}_0 \in \overset{\circ}{D}$ tale che $\nabla f(\mathbf{x}_0) = \mathbf{0}$. Allora:

- se $(-1)^i \det \mathbf{H}_{i f}(\mathbf{x}_0) > 0$ per ogni $i = 1, 2, \dots, n$ allora $\mathbf{x}_0$ è un massimo locale per f ;
- se $\det \mathbf{H}_{i f}(\mathbf{x}_0) > 0$ per ogni $i = 1, 2, \dots, n$ allora $\mathbf{x}_0$ è un minimo locale per f ;
- se esiste almeno $i = 1, 2, \dots, n$ tale che $\det \mathbf{H}_{i f}(\mathbf{x}_0)$ non segue alcuna delle precedenti regole allora $\mathbf{x}_0$ è un punto di sella per f .

dove $\mathbf{H}_{i f}(\mathbf{x}_0)$ è la sottomatrice principale di nord-ovest di ordine i della Hessiana di f .

Per $n = 2$ il Teorema 6 può essere ancora riscritto come segue:

Teorema 7. Siano $f : D \subseteq \mathbb{R}^2 \rightarrow \mathbb{R}$, $\mathbf{x}_0$ punto interno di D stazionario, $f \in C^2$ in un intorno di $\mathbf{x}_0$. Allora

- (i) se $\det \mathbf{H}_f(\mathbf{x}_0) > 0$ allora $\mathbf{x}_0$ è un punto estemale per f . In dettaglio:
 - se $f_{xx}(\mathbf{x}_0) < 0$ e $f_{yy}(\mathbf{x}_0) < 0$ allora $\mathbf{x}_0$ è punto di massimo locale per f ;

⁹ Nelle funzioni ad una variabile essi sono detti anche punti di flesso orizzontale

¹⁰ Il Teorema 5 è valido solamente nell'implicazione dimostrata e pertanto è detto anche condizione sufficiente dell'esistenza di ottimo. Inoltre la dicitura «condizione del secondo ordine» è riferita all'uso delle derivate parziali seconde mediante l'Hessiano.

- se $f_{xx}(\mathbf{x}_0) > 0$ e $f_{yy}(\mathbf{x}_0) > 0$ allora $\mathbf{x}_0$ è punto di minimo locale per f ;
- (ii) se $\det \mathbf{H}_f(\mathbf{x}_0) < 0$ allora $\mathbf{x}_0$ è punto di sella per f ;
- (iii) se $\det \mathbf{H}_f(\mathbf{x}_0) = 0$ non si sa nulla circa la natura di $\mathbf{x}_0$.

2.2. Calcolo della soluzione globale.

Definizione 2.2. Sia $f : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ tale che $f \in \mathcal{C}^2(\mathbb{R}^n, \mathbb{R})$. Definiamo *punti critici* per f i seguenti:

- tutti i punti stazionari $\mathbf{x}_0 \in \overset{\circ}{D}$;
- tutti i punti $\widehat{\mathbf{x}}_0 \in \partial D$;
- tutti i punti $\widehat{\mathbf{x}}_0 \in \partial D$ tale che in essi f non è differenziabile.

L'ottimo $\tilde{\mathbf{x}} \in D$ dovrà essere individuato tra i punti critici per f ed un metodo per determinarli è offerto dalla regola decisionale seguente.

Regola 1. Nelle ipotesi del Teorema 5 la soluzione globale del problema di ottimizzazione nella forma (P-Max) o nella forma (P-min) si ottiene seguendo i seguenti passi:

1. calcolare tutti i punti critici per f ;
2. consideriamo due possibili situazioni:
 - (a) se vale il Teorema 3 (di Weierstrass) occorre
 - (i) calcolare il valore di f nei punti critici;
 - (ii) scegliere come punto di massimo (*risp.* di minimo) globale il punto $\mathbf{x}_0$ per cui $f(\mathbf{x}_0)$ è il più grande (*risp.* il più piccolo) tra quelli ottenuti nel punto (i).
 - (b) se $D = V$ non è compatto consideriamo due possibili casi:
 - (i) se vale il Teorema 2 (f è (strettamente) convessa (*risp.* concava)) occorre:
 - (A) determinare attraverso il Teorema 4 quali punti stazionari siano ottimi locali;
 - (B) calcolare il valore di f nei punti di ottimo locale e nei punti critici non stazionari;
 - (C) se f è convessa (*risp.* concava) allora esistono solamente punti di minimo (*risp.* di massimo) globale ottenuti dal confronto dei valori di f calcolati nei punti determinati in (B).
 - (ii) se non vale il Teorema 2 occorre:
 - (A) determinare attraverso il Teorema 4 quali punti stazionari siano ottimi locali;
 - (B) calcolare il valore di f nei punti di ottimo locale e nei punti critici non stazionari;
 - (C) se esiste ed finito il più grande (*risp.* il più piccolo) valore di f allora il punto $\mathbf{x}_0$ corrispondente è un punto di massimo (*risp.* di minimo) globale.

Osservazione 5. Sebbene la regola decisionale si sdoppi in due sottoregole ed una di queste in due sottocasi evidenzia procedure e risultati sostanzialmente differenti. Se il dominio è compatto esistono i punti di ottimo globale e la loro individuazione avviene esclusivamente tra i punti critici. Nel caso in cui il dominio non sia compatto siamo solo certi di determinare gli ottimi locali tra i punti di stazionarietà (vedi percorso (b)-(i)-(A) e percorso (b)-(ii)-(A)), e solo nell'eventualità che f sia convessa (*risp.* concava) siamo certi della sola esistenza dei punti di minimo (*risp.* di massimo) globale ma non di massimo (*risp.* di minimo) globale.

2.3. Esempi.

Esempio 2.1. Determinare i punti di ottimo globale delle funzioni:

- I. $f(x) = \sqrt{1-x^2}$;
- II. $g(x) = x^2$;
- III. $h(x, y) = x \ln x + y \ln y$;
- IV. $i(x, y, z) = -3x^2 - 2y^2 - z^2 + \ln(x + y + 3z)$.

SOLUZIONE:

- I. La funzione $f(x) = \sqrt{1-x^2}$ ammette come dominio l'insieme $D = \{x \in \mathbb{R} : 1-x^2 \geq 0\} = [-1, 1]$ che è un compatto. Quindi procediamo ricercando i punti critici. Dal calcolo della derivata prima otteniamo

$$f'(x) = \frac{1}{2\sqrt{1-x^2}} \cdot (-2x) = -\frac{x}{\sqrt{1-x^2}}$$

che i punti $x_1 = -1$ e $x_2 = 1$ sono di non derivabilità e gli unici punti di frontiera di D . Inoltre da

$$\begin{aligned} f'(x) &= 0 \\ -\frac{x}{\sqrt{1-x^2}} &= 0 \end{aligned}$$

otteniamo il punto di stazionarietà $x_3 = 0$. In ossequio alla Regola 1 negli steps 1.-2.-(a)-(i)-(ii) dal calcolo dei valori di f in x_1, x_2 e x_3

$$f(0) = 1 \quad f(-1) = 0 \quad f(1) = 0$$

otteniamo che $x_3 = 0$ è un punto di massimo globale forte ed i punti $x_1 = -1$ e $x_2 = 1$ sono punti di minimo globale deboli.

- II. La funzione $g(x) = x^2$ ammette come dominio l'insieme $D = \mathbb{R}^2$ che è un aperto e quindi non compatto. Il Teorema 3 non vale ma essendo g una funzione convessa ammette per il Teorema 2 il punto x_0 di minimo globale. Per determinare x_0 procediamo secondo gli steps 1.-2.-(b)-(ii)-(A)-(B)-(C) della Regola 1. Dalla derivata prima

$$g'(x) = 2x$$

risulta che non ci sono punti di non derivabilità e il punto stazionario ottenuto da $g'(x) = 0$ è $x_0 = 0$ che è un minimo locale in quanto

$$\begin{aligned} f''(x) &= 2 \\ f''(0) &= 2 > 0 \end{aligned}$$

$x_0 = 0$ è allora il punto di minimo globale di valore $g(0) = 0$.

- III. Il dominio della funzione $h(x, y) = x \ln x + y \ln y$ è rappresentato dall'insieme $D = \{(x, y) \in \mathbb{R}^2 : x > 0 \wedge y > 0\} = \mathbb{R}_{0+}^2$ che non è un compatto.

Risolvendo il sistema

$$\begin{cases} f_x(x, y) = \ln x + 1 = 0 \\ f_y(x, y) = \ln y + 1 = 0 \end{cases}$$

otteniamo l'unico punto di stazionarietà $P = (e^{-1}, e^{-1})$. L'Hessiano di f calcolato in P

$$\mathbf{H}_f(e^{-1}, e^{-1}) = \begin{pmatrix} e & 0 \\ 0 & e \end{pmatrix}$$

è tale da soddisfare alle condizioni del Teorema 7 per cui P è un punto di minimo locale. Inoltre essendo la forma quadratica associata a $\mathbf{H}_f(x, y)$ definita positiva per ogni $(x, y) \in \overset{\circ}{D}$ allora per il Teorema 23 f è strettamente convessa e quindi P è un punto di minimo globale.

- IV. Il dominio della funzione $i(x, y, z) = -3x^2 - 2y^2 - z^2 + \ln(x + y + 3z)$ è rappresentato dall'insieme $D = \{(x, y, z) \in \mathbb{R}^3 : x + 2y + 3z > 0\}$ che non è un compatto¹¹.

L'unico punto di stazionarietà ottenuto risolvendo il sistema

$$\begin{cases} i_x = -6x + \frac{1}{x+2y+3z} = 0 \\ i_y = -4y + \frac{1}{x+2y+3z} \cdot 2 = 0 \\ i_z = -2z + \frac{1}{x+2y+3z} \cdot 3 = 0 \end{cases}$$

$$\text{è } P = \left(\frac{1}{2\sqrt{51}}, \frac{3}{2}\sqrt{51}, \frac{9}{2}\sqrt{51} \right)$$

La forma quadratica associata a

$$(6) \quad \mathbf{H}_i(\mathbf{x}) = \begin{pmatrix} -6 - \frac{1}{(x+2y+3z)^2} & -\frac{2}{(x+2y+3z)^2} & -\frac{3}{(x+2y+3z)^2} \\ -\frac{2}{(x+2y+3z)^2} & -4 - \frac{4}{(x+2y+3z)^2} & -\frac{6}{(x+2y+3z)^2} \\ -\frac{3}{(x+2y+3z)^2} & -\frac{6}{(x+2y+3z)^2} & -2 - \frac{9}{(x+2y+3z)^2} \end{pmatrix}$$

è definita negativa per il Teorema 17 in quanto

$$\det \mathbf{H}_{i_1}(\mathbf{x}) = \left| -6 - \frac{1}{(x+2y+3z)^2} \right| = -6 - \frac{1}{(x+2y+3z)^2} < 0 \quad \forall (x, y, z)$$

$$\det \mathbf{H}_{i_2}(\mathbf{x}) = \begin{vmatrix} -6 - \frac{1}{(x+2y+3z)^2} & -\frac{2}{(x+2y+3z)^2} \\ -\frac{2}{(x+2y+3z)^2} & -4 - \frac{4}{(x+2y+3z)^2} \end{vmatrix} = 24 + \frac{28}{(x+2y+3z)^2} > 0 \quad \forall (x, y, z)$$

$$\det \mathbf{H}_{i_3}(\mathbf{x}) = \begin{vmatrix} -6 - \frac{1}{(x+2y+3z)^2} & -\frac{2}{(x+2y+3z)^2} & -\frac{3}{(x+2y+3z)^2} \\ -\frac{2}{(x+2y+3z)^2} & -4 - \frac{4}{(x+2y+3z)^2} & -\frac{6}{(x+2y+3z)^2} \\ -\frac{3}{(x+2y+3z)^2} & -\frac{6}{(x+2y+3z)^2} & -2 - \frac{9}{(x+2y+3z)^2} \end{vmatrix} = -48 - \frac{272}{(x+2y+3z)^2} < 0 \quad \forall (x, y, z)$$

Per il Teorema 23 allora f è strettamente convessa e quindi per il Teorema 2 il punto P è l'unico punto di minimo globale per f .

Esempio 2.2 (La massimizzazione del profitto). Consideriamo un'impresa con una funzione di produzione Cobb-Douglas $f(x, y) = x^{\frac{1}{2}}y^{\frac{1}{3}}$, nei fattori x (capitale) e y (lavoro). Il prezzo unitario del bene prodotto è 1, mentre i costi unitari dei fattori x e y sono rispettivamente $\frac{1}{2}$ e $\frac{1}{3}$. Massimizzare il profitto dell'impresa.

SOLUZIONE: Ricordiamo che il profitto si ottiene come differenza dei ricavi totali (TR) meno i costi totali (TC)

$$\pi := TR - TC$$

¹¹ Infatti D è un semispazio di $\mathbb{R}^3$ delimitato dal piano $x + 2y + 3z = 0$.

e qui pertanto è rappresentato dalla funzione scalare reale a due variabili $\pi : \mathbb{R}_+^2 \rightarrow \mathbb{R}$ definita dalla legge

$$\pi(x, y) := x^{\frac{1}{2}}y^{\frac{1}{3}} - \frac{1}{2}x - \frac{1}{3}y.$$

Massimizzare il profitto dell'impresa equivale a risolvere il problema di ottimizzazione libera di forma (FP-Max)

$$\max_{(x,y)} x^{\frac{1}{2}}y^{\frac{1}{3}} - \frac{1}{2}x - \frac{1}{3}y$$

dove la soluzione va ricercata, per ragioni pratiche, in $V = D = \mathbb{R}_+^2$ piuttosto che in $\mathbb{R}^2$ come sia legittimo da un punto di vista matematico.

Procediamo con il calcolo dei massimi locali all'interno del dominio. Le derivate parziali del profitto sono:

$$\pi_x(x, y) = \frac{1}{2}x^{-\frac{1}{2}}y^{\frac{1}{3}} - \frac{1}{2}, \quad \pi_y(x, y) = \frac{1}{3}x^{\frac{1}{2}}y^{-\frac{2}{3}} - \frac{1}{3}.$$

Le condizioni del primo ordine diventano:

$$(7) \quad x^{-\frac{1}{2}}y^{\frac{1}{3}} = 1,$$

$$(8) \quad x^{\frac{1}{2}}y^{-\frac{2}{3}} = 1.$$

Per risolvere queste equazioni, dividiamo (8) per (7) termine a termine, ottenendo:

$$(9) \quad \frac{x}{y} = 1, \quad \Leftrightarrow \quad x = y.$$

Inserendo (9) in (7) si ha:

$$x = 1 \quad y = 1.$$

I potenziali estremi locali sono quindi: il punto $(1, 1)$, che soddisfa le condizioni di primo ordine, ed i punti sul bordo, vale a dire i punti con $x = 0$ o $y = 0$, che corrispondono ai punti dove uno dei due fattori, oppure entrambi, non sono impiegati.

Le derivate parziali del secondo ordine sono:

$$\pi_{xx}(x, y) = -\frac{1}{4}x^{-\frac{3}{2}}y^{\frac{1}{3}}, \quad \pi_{xy}(x, y) = \frac{1}{6}x^{-\frac{1}{2}}y^{-\frac{2}{3}}, \quad \pi_{yy}(x, y) = -\frac{2}{9}x^{\frac{1}{2}}y^{-\frac{5}{3}},$$

e nel punto $(1, 1)$ valgono:

$$\pi_{xx}(1, 1) = -\frac{1}{4}, \quad \pi_{xy}(1, 1) = \frac{1}{6}, \quad \pi_{yy}(1, 1) = -\frac{2}{9}.$$

Perciò il risultato delle condizioni del secondo ordine¹²

$$\pi_{xx}(1, 1) = -\frac{1}{4} < 0, \quad \pi_{xx}(1, 1)\pi_{yy}(1, 1) - \pi_{xy}(1, 1)^2 = \frac{1}{36} > 0,$$

stabilisce che $(1, 1)$ è un massimo locale. In questo punto i profitti valgono:

$$\pi(1, 1) = \frac{1}{6} > 0.$$

Per determinare se $(1, 1)$ è anche il massimo globale, dobbiamo considerare i punti sul bordo del dominio. Per i punti con $x = 0$ il profitto vale: $\pi(0, y) = -\frac{y}{3} \leq 0$, che non è positivo dato che $y \geq 0$. Analogamente il profitto è non positivo per i punti con $y = 0$. Questi punti non possono perciò dare massimi globali. Infine consideriamo i punti dove uno dei due fattori, o entrambi sono impiegati in quantità illimitata. Ad esempio, se $x \rightarrow \infty$, con y un valore fisso, vediamo che: $\pi(x, y) = x^{\frac{1}{2}}y^{\frac{1}{3}} - x - y \rightarrow -\infty$. Se invece x e y vanno a ∞ restando uguali, $x = y \rightarrow \infty$, allora $\pi(x, y) = x^{\frac{5}{6}} - 2x \rightarrow -\infty$. Gli altri casi sono simili, e non danno origine a massimi globali.

In conclusione, i profitti raggiungono il massimo globale $\frac{1}{6}$ nel punto $(1, 1)$.

Esercizio 2.1 (Minimizzazione dei costi). Minimizzare la funzione costo medio $AC = \frac{25}{q} + 0.1q^2$.

SOLUZIONE: Calcoliamo i punti di stazionarietà di AC .

$$AC'(q) = -\frac{25}{q^2} + 0.2q$$

$$AC'(q) = 0$$

$$0.2q = \frac{25}{q^2}$$

$$q^3 = 125$$

$$\tilde{q} = 5.$$

¹² Si fa riferimento al Teorema 7.

Il punto $\tilde{q} = 5$ verifica le condizioni del secondo ordine del Teorema 5 applicato al caso in cui $n = 1$ (1 sola variabile) in quanto

$$AC''(q) = \frac{50}{q^3} + 0.2$$

$$AC''(\tilde{q}) = \frac{50}{125} + 0.2 = 0.4 + 0.2 = 0.6 > 0$$

e pertanto concludiamo che AC ammette minimo in $\tilde{q} = 5$ assumendo il valore di $AC(5) = \frac{25}{5} + 0.1 \times 5^2 = 5 + 2.5 = 7.5$.

Esercizio 2.2 (Massimizzare il ricavo). Una ditta vende due beni succedanei. La quantità domandata di un bene è direttamente correlata al prezzo dell'altro. I prezzi dei due beni (espressi in €) sono p_1 e p_2 e le loro funzioni di domanda sono rispettivamente

$$q_1 = 517 - 3.5p_1 + 0.8p_2 \quad q_2 = 770 - 4.4p_2 + 1.4p_1$$

Quale prezzo è necessario attribuire a ciascun bene per massimizzare il suo ricavo totale di vendite?

Soluzione: Per questo tipo di problema è più conveniente esprimere il ricavo totale in funzione del prezzo piuttosto che della quantità. Quindi

$$\begin{aligned} TR &= TR_1 + TR_2 = p_1 q_1 + p_2 q_2 \\ &= p_1(517 - 3.5p_1 + 0.8p_2) + p_2(770 - 4.4p_2 + 1.4p_1) \\ &= 517p_1 - 3.5p_1^2 + 0.8p_1p_2 + 770p_2 - 4.4p_2^2 + 1.4p_1p_2 \\ &= 517p_1 - 3.5p_1^2 + 770p_2 - 4.4p_2^2 + 2.2p_1p_2 \end{aligned}$$

Le condizioni del primo ordine sono

$$(10) \quad \frac{\partial TR}{\partial p_1} = 517 - 7p_1 + 2.2p_2 = 0$$

$$(11) \quad \frac{\partial TR}{\partial p_2} = 770 - 8.8p_2 + 2.2p_1 = 0$$

da cui, moltiplicando la (10) per 4 e sommando la (11) otteniamo

$$\begin{array}{rrrrrr} 2.068 & -28p_1 & +8.8p_2 & = & 0 & + \\ 770 & +2.2p_1 & -8.8p_2 & = & 0 & = \\ \hline 2.838 & -25.8p_1 & & = & 0 & \\ 2.838 & & & = & 25.8p_1 & \\ 110 & & & = & \tilde{p}_1 & \end{array}$$

e poi, sostituendo il valore di p_1 nella (10) ricaviamo

$$\begin{aligned} 517 - 7(110) + 2.2p_2 &= 0 \\ 2.2p_2 &= 253 \\ \tilde{p}_2 &= 115 \end{aligned}$$

Ottenuto il punto di stazionarietà $(\tilde{p}_1, \tilde{p}_2) = (110, 115)$ calcoliamo le derivate seconde

$$\begin{aligned} \frac{\partial^2 TR}{\partial p_1^2}(p_1, p_2) &= \frac{\partial^2 TR}{\partial p_1^2}(\tilde{p}_1, \tilde{p}_2) = -7 < 0 \\ \frac{\partial^2 TR}{\partial p_2^2}(p_1, p_2) &= \frac{\partial^2 TR}{\partial p_2^2}(\tilde{p}_1, \tilde{p}_2) = -8.8 < 0 \\ \frac{\partial^2 TR}{\partial p_1 \partial p_2}(p_1, p_2) &= \frac{\partial^2 TR}{\partial p_1 \partial p_2}(\tilde{p}_1, \tilde{p}_2) = 2.2 \end{aligned}$$

e poiché

$$\begin{aligned} \det \mathbf{H}_{TR}(\tilde{p}_1, \tilde{p}_2) &= \frac{\partial^2 TR}{\partial p_1^2}(\tilde{p}_1, \tilde{p}_2) \cdot \frac{\partial^2 TR}{\partial p_2^2}(\tilde{p}_1, \tilde{p}_2) - \left(\frac{\partial^2 TR}{\partial p_1 \partial p_2}(\tilde{p}_1, \tilde{p}_2) \right)^2 = \\ &= (-7)(-8.8) - (2.2)^2 = 61.6 - 4.84 = 56.76 > 0 \end{aligned}$$

allora, in forza del Teorema 7, concludiamo affermando che i valori dei prezzi richiesti al fine di massimizzare il ricavo totale di vendite sono: $\tilde{p}_1 = \text{€}110$ e $\tilde{p}_2 = \text{€}115$.

3. OTTIMIZZAZIONE CON VINCOLI DI UGUAGLIANZA

Consideriamo adesso il problema di ottimizzazione quando $V = K \cap D \subset D$ e $\mathfrak{P}$ è una relazione di uguaglianza. In tal caso le forme (P-Max) e (P-min) diventano

$$\begin{aligned} \text{(ECP-Max)} \quad & \begin{cases} \max_{\mathbf{x}} f(\mathbf{x}) \\ \text{t.c.} \quad g(\mathbf{x}) = \mathbf{0} \end{cases} & \text{(ECP-min)} \quad & \begin{cases} \min_{\mathbf{x}} f(\mathbf{x}) \\ \text{t.c.} \quad g(\mathbf{x}) = \mathbf{0} \end{cases} \end{aligned}$$

dove la relazione vincolare è data dalla funzione vettoriale reale $g : K \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^h$ tale che $g \in \mathcal{C}^1(\mathbb{R}^n, \mathbb{R}^h)$ con $h < n$ ed il vincolo è il sistema di equazioni $g(\mathbf{x}) = \mathbf{0}$ e pertanto la regione ammissibile entro la quale trovare gli ottimi di f è $V = \{\mathbf{x} \in D \cap K : g(\mathbf{x}) = \mathbf{0}\}$.

Il problema di ottimizzazione, espresso nella forma (ECP-Max) oppure nella forma (ECP-min), sotto specifiche condizioni, può essere risolto mediante il metodo di sostituzione e/o il metodo dei moltiplicatori di Lagrange.

3.1. Metodo di sostituzione. Osservato che

$$g(\mathbf{x}) = \mathbf{0} \quad \Leftrightarrow \quad \begin{cases} g_1(x_1, x_2, \dots, x_n) = 0 \\ g_2(x_1, x_2, \dots, x_n) = 0 \\ \dots \\ g_h(x_1, x_2, \dots, x_n) = 0 \end{cases}$$

se ogni equazione $g_i(x_1, x_2, \dots, x_n) = 0$ (con $i = 1, 2, \dots, n$) nelle variabili $x_1, x_2, \dots, x_n$ è esplicitabile secondo una determinata x_k ossia

$$\begin{cases} g_1(x_1, x_2, \dots, x_n) = 0 \\ g_2(x_1, x_2, \dots, x_n) = 0 \\ \dots \\ g_h(x_1, x_2, \dots, x_n) = 0 \end{cases} \quad \Leftrightarrow \quad \begin{cases} x_1 = \phi_1(x_2, x_3, \dots, x_n) \\ x_2 = \phi_2(x_1, x_3, \dots, x_n) \\ \dots \\ x_h = \phi_h(x_1, x_2, \dots, x_{h-1}, x_{h+1}, \dots, x_n) \end{cases} \quad \Leftrightarrow \quad \begin{cases} x_1 = \Phi_1(x_{h+1}, x_{h+2}, \dots, x_n) \\ x_2 = \Phi_2(x_{h+1}, x_{h+2}, \dots, x_n) \\ \dots \\ x_h = \Phi_h(x_{h+1}, x_{h+2}, \dots, x_n) \end{cases}$$

con $\Phi : \mathbb{R}^{n-h} \rightarrow \mathbb{R}^h$ allora operando la sostituzione degli x_k in f otteniamo

$$f(\mathbf{x}) = f(x_1, x_2, \dots, x_n) = \Phi(x_{h+1}, x_{h+2}, \dots, x_n) = \Phi(\mathbf{x})$$

con $\mathbf{x} \in \mathbb{R}^{n-h}$. Quindi il problema di ottimizzazione vincolata si riduce al seguente problema di ottimizzazione libera nella forma

$$\max_{\mathbf{x}} \Phi(\mathbf{x}) \quad \text{oppure} \quad \min_{\mathbf{x}} \Phi(\mathbf{x})$$

Esempio 3.1 (Massimizzazione dell'utilità). La funzione di utilità di un consumatore è la Cobb-Douglas $U(x, y) = x^{\frac{1}{2}} y^{\frac{1}{3}}$ nei beni $x \geq 0$ e $y \geq 0$ nei beni x e y ; i prezzi unitari dei beni sono rispettivamente $p = \frac{1}{2}$ e $q = \frac{1}{3}$ ed il reddito del consumatore è w . Massimizzare l'utilità sotto il vincolo di bilancio.

SOLUZIONE: Il problema di ottimizzazione richiesto è della forma (ECP-Max) così espresso

$$(12) \quad \begin{cases} \max_{(x,y)} & x^{\frac{1}{2}} y^{\frac{1}{3}} \\ \text{t.c.} & \frac{1}{2}x + \frac{1}{3}y = w \end{cases}$$

avendo inteso con

$$(13) \quad \frac{1}{2}x + \frac{1}{3}y = w$$

il vincolo di bilancio del consumatore.

Osservato che dalla (13) possiamo esplicitare la variabile y come segue

$$y = 3 \left(w - \frac{x}{2} \right)$$

ed operando la sostituzione nella funzione di utilità, per cui questa è adesso è funzione della sola x :

$$\mathcal{U}(x) := U \left(x, 3 \left(w - \frac{x}{2} \right) \right) = 3^{\frac{1}{3}} x^{\frac{1}{2}} \left(w - \frac{x}{2} \right)^{\frac{1}{3}}.$$

il problema (12) è ricondotto al problema di ottimizzazione libera della forma (FP-Max) dove la funzione obiettivo è $\mathcal{U} : [0, 2w] \rightarrow \mathbb{R}$ il cui dominio è dettato da ragioni pratiche¹³:

$$(14) \quad \max_x x^{\frac{1}{2}} \left(w - \frac{x}{2} \right)^{\frac{1}{3}}$$

Calcoliamo i punti stazionari di $\mathcal{U}$:

$$\begin{aligned} \mathcal{U}'(x) &= 0 \\ \frac{1}{2}x^{-\frac{1}{2}} \left(w - \frac{x}{2} \right)^{\frac{1}{3}} + \frac{1}{3}x^{\frac{1}{2}} \left(w - \frac{x}{2} \right)^{-\frac{2}{3}} \left(-\frac{1}{2} \right) &= 0 \\ \frac{1}{2}x^{-\frac{1}{2}} \left(w - \frac{x}{2} \right)^{-\frac{2}{3}} \left[\left(w - \frac{x}{2} \right) - \frac{1}{3}x \right] &= 0 \quad \frac{1}{2}x^{-\frac{1}{2}} \left(w - \frac{x}{2} \right)^{-\frac{2}{3}} \left(w - \frac{5}{6}x \right) = 0 \\ x &= \frac{6}{5}w. \end{aligned}$$

Applicando la condizione del secondo ordine per $\tilde{x} = \frac{6}{5}w$

$$\begin{aligned} \mathcal{U}''(x) &= \frac{1}{2} \left\{ -\frac{1}{2}x^{-\frac{3}{2}} \left(w - \frac{x}{2} \right)^{-\frac{2}{3}} \left(w - \frac{5}{6}x \right) - \frac{2}{3} \left(-\frac{1}{2} \right) x^{-\frac{1}{2}} \left(w - \frac{x}{2} \right)^{-\frac{5}{3}} \left(w - \frac{5}{6}x \right) - \frac{5}{6}x^{-\frac{1}{2}} \left(w - \frac{x}{2} \right)^{-\frac{2}{3}} \right\} \\ \mathcal{U}''\left(\frac{6w}{5}\right) &= -\frac{5}{12} \left(\frac{6w}{5} \right)^{-\frac{1}{2}} \left(\frac{2w}{5} \right)^{-\frac{1}{3}} > 0 \end{aligned}$$

otteniamo che $\tilde{x} = \frac{6w}{5}$ è un massimo locale e l'utilità in esso assume il valore

$$\mathcal{U}\left(\frac{6}{5}w\right) = 3 \left(w - \frac{16}{25}w \right) = \frac{6}{5}w.$$

Poichè $[0, 2w]$ è un chiuso e limitato e quindi compatto, per il Teorema di Weierstrass $\mathcal{U}$ ammette minimo e massimo globale. Calcolando il valore di $\mathcal{U}$ negli estremi $x = 0$ e $x = 2w$

$$\mathcal{U}(0) = w < \frac{6}{5}w \quad \text{e} \quad \mathcal{U}(2w) = 0 < \frac{6}{5}w$$

concludiamo che $\tilde{x} = \frac{6}{5}w$ è il massimo globale di $\mathcal{U}$ e soluzione del problema (14). Sostituendo $\tilde{x}$ nella (13) otteniamo che

$$\tilde{P} = (\tilde{x}, \tilde{y}) = \left(\frac{6}{5}w, \frac{6}{5}w \right)$$

è il paniere ottimo ovvero la soluzione del problema (12).

3.2. Metodo dei moltiplicatori di Lagrange. La procedura secondo la quale esplicitare una variabile differenziale da ogni equazione del vincolo e sostituirla nella funzione obiettivo, non è sempre applicabile¹⁴. Pertanto sotto certe condizioni introdurremo il Metodo dei moltiplicatori di Lagrange.

Siano $f : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ tale che $f \in \mathcal{C}^2(\mathbb{R}^n, \mathbb{R})$ la funzione obiettivo, $g : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^h$ tale che $g \in \mathcal{C}^1(\mathbb{R}^n, \mathbb{R}^h)$ la funzione vincolare e $V = \{\mathbf{x} \in \mathbb{R}^n : g(\mathbf{x}) = \mathbf{0}\}$ la regione ammissibile. Definiamo *Lagrangiana di f* la funzione ausiliare $\mathcal{L} : D \times \mathbb{R}^h \rightarrow \mathbb{R}$ definita dalla legge

$$\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}) := f(\mathbf{x}) + \boldsymbol{\lambda} \cdot g(\mathbf{x}) \quad 15$$

$$\mathcal{L}(x_1, \dots, x_k, \lambda_1, \dots, \lambda_h) := f(x_1, \dots, x_k) + \sum_{i=1}^h \lambda_i g_i(x_1, \dots, x_k)$$

con $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \dots, \lambda_h)$ vettore opportuno di $\mathbb{R}^h$.

Osservazione 6. La Lagrangiana di f è tale che $\mathcal{L} \in \mathcal{C}^2(\mathbb{R}^n, \mathbb{R}^h)$.

¹³ I beni non possono essere negativi, pertanto della retta di equazione $y = 3\left(w - \frac{x}{2}\right)$ si considera solamente la parte del I quadrante ovvero il segmento di estremi $(0, 3w)$ e $(2w, 0)$. Lavorando con la sola variabile x allora si ha la limitazione $0 \leq x \leq 2w$

¹⁴ Basti pensare ad un vincolo che sia un luogo geometrico come una circonferenza $x^2 + y^2 + ax + by + c = 0$.

¹⁵ Per $\boldsymbol{\lambda} \cdot g(\mathbf{x})$ si intende il prodotto scalare o l'equivalente di $\boldsymbol{\lambda}^T g(\mathbf{x})$ se intendiamo i fattori come vettori di $\mathbb{R}^h$ e ricorriamo al prodotto tra matrici di cui i vettori costituiscono un caso particolare.

Inoltre si ha

$$\begin{aligned}
 \mathcal{L}_{\lambda_r}(\mathbf{x}, \boldsymbol{\lambda}) &= g_r(x) & \forall r = 1, \dots, h \\
 \mathcal{L}_{x_s}(\mathbf{x}, \boldsymbol{\lambda}) &= f_{x_s}(\mathbf{x}) + \boldsymbol{\lambda} \cdot g_{x_s}(\mathbf{x}) = f_{x_s}(\mathbf{x}) + \sum_{i=1}^h \lambda_i g_{i x_s}(\mathbf{x}) & \forall s = 1, \dots, k \\
 \mathcal{L}_{x_s \lambda_r}(\mathbf{x}, \boldsymbol{\lambda}) &= g_{r x_s}(\mathbf{x}) & \forall s = 1, \dots, k \quad \text{e} \quad \forall r = 1, \dots, h \\
 \mathcal{L}_{\lambda_r \lambda_t}(\mathbf{x}, \boldsymbol{\lambda}) &= 0 & \forall r, t = 1, \dots, h \\
 \mathcal{L}_{x_s x_v}(\mathbf{x}, \boldsymbol{\lambda}) &= f_{x_s x_v}(\mathbf{x}) + \boldsymbol{\lambda} \cdot g_{x_s x_v}(\mathbf{x}) & \forall s, v = 1, \dots, k
 \end{aligned}$$

Denotata con H_f la matrice $k \times k$ delle derivate $\mathcal{L}_{x_s x_v}$ con $s, v = 1, 2, \dots, n$

$$H_f := \begin{pmatrix} \mathcal{L}_{x_1 x_1} & \mathcal{L}_{x_1 x_2} & \dots & \mathcal{L}_{x_1 x_k} \\ \mathcal{L}_{x_2 x_1} & \mathcal{L}_{x_2 x_2} & \dots & \mathcal{L}_{x_2 x_k} \\ \vdots & \vdots & \ddots & \vdots \\ \mathcal{L}_{x_k x_1} & \mathcal{L}_{x_k x_2} & \dots & \mathcal{L}_{x_k x_k} \end{pmatrix}$$

e con G la matrice $h \times k$ delle derivate $\mathcal{L}_{x_s \lambda_r} = g_{r x_s}$ con $s = 1, 2, \dots, n$ e $r = 1, 2, \dots, h$

$$G := \begin{pmatrix} \mathcal{L}_{x_1 \lambda_1} & \mathcal{L}_{x_2 \lambda_1} & \dots & \mathcal{L}_{x_k \lambda_1} \\ \mathcal{L}_{x_1 \lambda_2} & \mathcal{L}_{x_2 \lambda_2} & \dots & \mathcal{L}_{x_k \lambda_2} \\ \vdots & \vdots & \ddots & \vdots \\ \mathcal{L}_{x_1 \lambda_h} & \mathcal{L}_{x_2 \lambda_h} & \dots & \mathcal{L}_{x_k \lambda_h} \end{pmatrix} = \begin{pmatrix} g_{1 x_1} & g_{1 x_2} & \dots & g_{1 x_n} \\ g_{2 x_1} & g_{2 x_2} & \dots & g_{2 x_n} \\ \vdots & \vdots & \ddots & \vdots \\ g_{h x_1} & g_{h x_2} & \dots & g_{h x_n} \end{pmatrix} = \begin{pmatrix} \nabla g_1 \\ \nabla g_2 \\ \vdots \\ \nabla g_h \end{pmatrix}$$

allora l'Hessiano di $\mathcal{L}$ è a blocchi

$$\mathbf{H}_{\mathcal{L}}(\mathbf{x}, \boldsymbol{\lambda}) = \left(\begin{array}{c|c} H_f & G^T \\ \hline G & 0 \end{array} \right)$$

Definizione 3.1. Diciamo che il vincolo è qualificato se la funzione vincolare g soddisfa la *condizione di regolarità*

$$(15) \quad \nabla g(\mathbf{x}) \neq \mathbf{0} \quad \forall \mathbf{x} \in V$$

Da ora in poi, a meno di precisazioni opportune, quando parleremo di problema di ottimizzazione con vincolo di uguaglianza ci riferiremo a (ECP-Max).

Proposizione 8. Se g è tale da soddisfare la condizione di regolarità e $(\tilde{\mathbf{x}}, \tilde{\boldsymbol{\lambda}})$ è soluzione locale del problema $\max_{(\mathbf{x}, \boldsymbol{\lambda})} \mathcal{L}(\mathbf{x}, \boldsymbol{\lambda})$ allora $\tilde{\mathbf{x}}$ è soluzione locale del problema (ECP-Max).

Dimostrazione. Per ipotesi $(\tilde{\mathbf{x}}, \tilde{\boldsymbol{\lambda}})$ è un punto di massimo locale per $\mathcal{L}$ per cui

$$\exists \delta > 0, \epsilon > 0 : \mathcal{L}(\tilde{\mathbf{x}}, \tilde{\boldsymbol{\lambda}}) \geq \mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}) \quad \forall (\mathbf{x}, \boldsymbol{\lambda}) \in B_\delta(\tilde{\mathbf{x}}) \times B_\epsilon(\tilde{\boldsymbol{\lambda}})$$

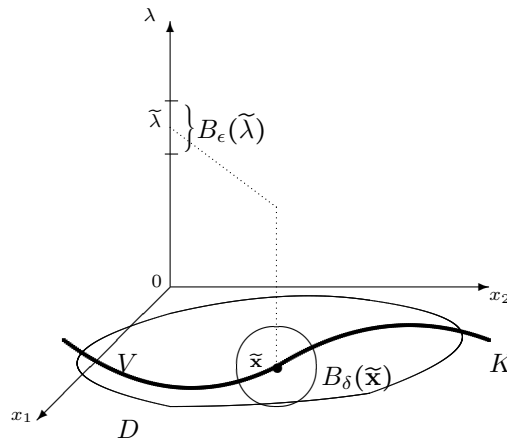


FIGURA 5. Caso di $n = 2$ e $h = 1$ per cui $D \subseteq \mathbb{R}^2$

ed inoltre essendo $\mathcal{L}$ di classe C^1 nel suo dominio accade

$$\mathcal{L}_{\lambda_j}(\tilde{\mathbf{x}}, \tilde{\boldsymbol{\lambda}}) = g_j(\tilde{\mathbf{x}}) = 0 \quad \forall j = 1, \dots, h$$

e quindi $\tilde{\mathbf{x}} \in V$. Infine osserviamo che

$$f(\tilde{\mathbf{x}}) = \mathcal{L}(\tilde{\mathbf{x}}, \tilde{\lambda}) \geq \mathcal{L}(\mathbf{x}, \tilde{\lambda}) = f(\mathbf{x}) \quad \forall \mathbf{x} \in B_\delta(\tilde{\mathbf{x}}) \cap V$$

ossia l'asserto. $\square$

Teorema 9 (condizione del primo ordine).

Siano $\tilde{\mathbf{x}}$ la soluzione locale del problema (ECP-Max) e $G(\tilde{\mathbf{x}})$ la matrice dei $\nabla g_i(\tilde{\mathbf{x}})$ (con $i = 1, 2, \dots, h$) tale che sia di rango massimo¹⁶. Allora esiste un unico $\tilde{\lambda}$ tale che $(\tilde{\mathbf{x}}, \tilde{\lambda})$ sia soluzione locale del problema $\max_{(\mathbf{x}, \lambda)} \mathcal{L}(\mathbf{x}, \lambda)$.

Dimostrazione. Per semplicità di notazione consideriamo il caso $n = 2$ ed $h = 1$. La Lagrangiana in questione sarà

$$\mathcal{L}(x_1, x_2, \lambda) = f(x_1, x_2) + \lambda g(x_1, x_2).$$

Per ipotesi funzione $f|_V$ ammette massimo locale in $\tilde{\mathbf{x}} = (\tilde{x}_1, \tilde{x}_2)$; da $\nabla g(\tilde{\mathbf{x}}) \neq \mathbf{0}$ supponiamo, ad esempio, che $g_{x_1}(\tilde{\mathbf{x}}) \neq 0$; dobbiamo dimostrare che

$$(16) \quad \begin{cases} f_{x_1}(\tilde{\mathbf{x}}) + \tilde{\lambda} g_{x_1}(\tilde{\mathbf{x}}) = 0 \\ f_{x_2}(\tilde{\mathbf{x}}) + \tilde{\lambda} g_{x_2}(\tilde{\mathbf{x}}) = 0 \end{cases}$$

per cui è necessario che

$$\tilde{\lambda} = -\frac{f_{x_1}(\tilde{\mathbf{x}})}{g_{x_1}(\tilde{\mathbf{x}})}$$

e quindi provare che per $\lambda = \tilde{\lambda}$ valga anche la derivata seconda della (16).

Applicando il Teorema 21 al vincolo $g(\mathbf{x}) = g(x_1, x_2) = 0$ si definisce una funzione

$$\phi :]\tilde{x}_2 - \epsilon, \tilde{x}_2 + \epsilon[\rightarrow]\tilde{x}_1 - \delta, \tilde{x}_1 + \delta[\quad \text{per } \epsilon \text{ e } \delta \text{ opportuni}$$

di classe C^1 , ossia esiste un intorno di $\tilde{\mathbf{x}}$ in cui la parte di V che vi cade è il grafico della funzione $x_1 = \phi(x_2)$ (vedi Figura 6).

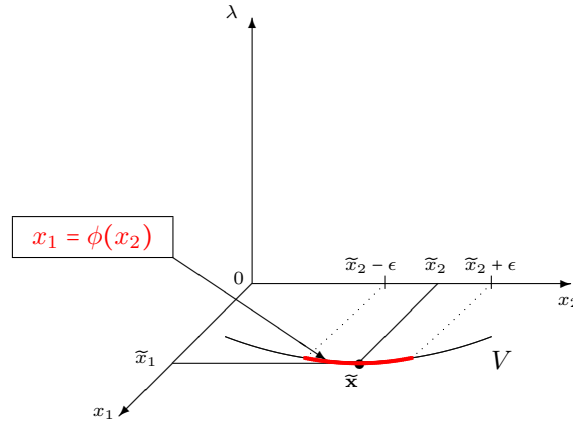


FIGURA 6

Posto $F(t) := f(\phi(t), t)$ dire che $\tilde{\mathbf{x}}$ è di massimo relativo per $f|_V$ equivale a dire che $\tilde{x}_2$ è di massimo relativo per F pertanto abbiamo

$$0 = F'(\tilde{x}_2) = f_{x_1}(\tilde{\mathbf{x}})\phi'(\tilde{x}_2) + f_{x_2}(\tilde{\mathbf{x}}).$$

Applichiamo adesso la regola di derivazione delle funzioni implicite e ricordando il valore di $\tilde{\lambda}$ otteniamo

$$\begin{aligned} 0 &= -\tilde{\lambda}\phi'(\tilde{x}_2)g_{x_1}(\tilde{\mathbf{x}}) + f_{x_2}(\tilde{\mathbf{x}}) = \\ &= \tilde{\lambda}g_{x_2}(\tilde{\mathbf{x}}) + f_{x_2}(\tilde{\mathbf{x}}) \end{aligned}$$

come volevasi dimostrare. $\square$

Corollario 10. Se G sul vincolo V ha sempre rango massimo allora le soluzioni locali del problema (ECP-Max) sono da ricercarsi tra i “primi” $(\mathbf{x}, \lambda) \in D \times \mathbb{R}^h$ tali che $\nabla \mathcal{L} = 0$.

Definizione 3.2. Se è possibile applicare il Corollario 10, allora il vettore $\lambda \in \mathbb{R}^h$ è detto *moltiplicatore di Lagrange*.

¹⁶ $r(G(\tilde{\mathbf{x}})) = h$ è equivalente ad affermare che la funzione vincolare g soddisfa la condizione di regolarità.

3.3. Interpretazione geometrica del metodo dei moltiplicatori di Lagrange. Siano $n = 2$ e $h = 1$ e supponiamo che:

- (1) $\tilde{\mathbf{x}}$ è di massimo (minimo) relativo per $f|_V$;
- (2) $\nabla g(\tilde{\mathbf{x}}) \neq 0$ (per esempio $g_{x_2}(\tilde{\mathbf{x}}) \neq 0$) e G è di rango massimo;
- (3) $\nabla f(\tilde{\mathbf{x}}) \neq 0$.

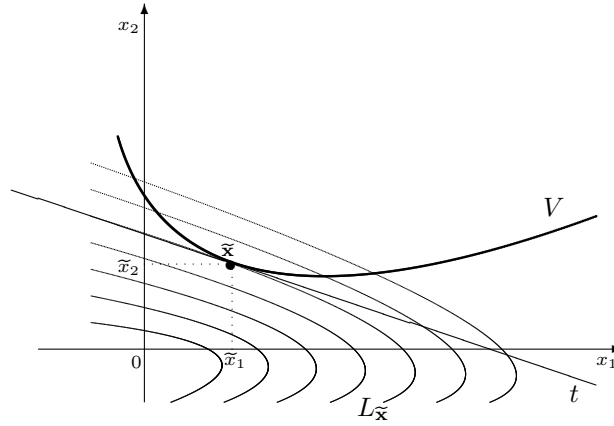


FIGURA 7

Sappiamo che in un intorno di $\tilde{\mathbf{x}}$ il vincolo V è il grafico di una funzione $x_2 = \phi(x_1)$ e la retta tangente a V in $\tilde{\mathbf{x}}$ ha coefficiente angolare

$$m = \phi'(\tilde{x}_1) = -\frac{g_{x_1}(\tilde{\mathbf{x}})}{g_{x_2}(\tilde{\mathbf{x}})}.$$

Indichiamo con $L_{\tilde{\mathbf{x}}}$ la linea di livello di f passante per $\tilde{\mathbf{x}}$ (vedi Figura 7), ossia

$$L_{\tilde{\mathbf{x}}} = \{\mathbf{x} \in D : f(\mathbf{x}) = f(\tilde{\mathbf{x}})\}.$$

Essendo $\nabla f(\tilde{\mathbf{x}}) \neq 0$ esiste in $\tilde{\mathbf{x}}$ la retta tangente alla linea di livello $L_{\tilde{\mathbf{x}}}$. Poichè esiste un unico $\tilde{\lambda}$ per cui $f_{x_s}(\tilde{\mathbf{x}}) + \tilde{\lambda}g_{x_s}(\tilde{\mathbf{x}}) = 0$ per $s = 1, 2$, allora in virtù del Teorema 2.1 si ha $f_{x_2}(\tilde{\mathbf{x}}) \neq 0$.

Anche per $L_{\tilde{\mathbf{x}}}$ possiamo ripetere lo stesso ragionamento fatto per V , ossia nell'intorno di $\tilde{\mathbf{x}}$ la curva di livello $L_{\tilde{\mathbf{x}}}$ è un grafico di una funzione della variabile x_1 . La retta tangente ad $L_{\tilde{\mathbf{x}}}$ nel punto $\tilde{\mathbf{x}}$ ha coefficiente angolare

$$-\frac{f_{x_1}(\tilde{\mathbf{x}})}{f_{x_2}(\tilde{\mathbf{x}})}$$

e questo vale esattamente m .

In conclusione possiamo trarre la seguente:

Regola 2. Nel punto $\tilde{\mathbf{x}}$ il vincolo V e la linea di livello $L_{\tilde{\mathbf{x}}}$ hanno la stessa tangente t . Ossia $\tilde{\mathbf{x}}$ è un punto di tangenza tra la linea di livello passante per esso ed il vincolo V (vedi Figura 8).

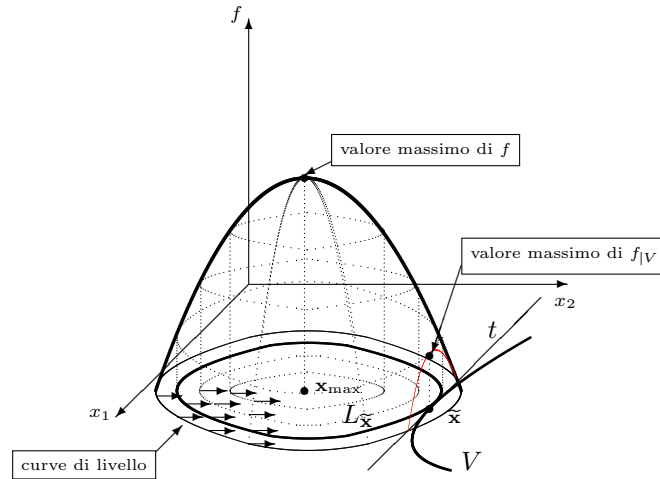


FIGURA 8

3.4. La condizione del secondo ordine. Dimostriamo che il Teorema 9 (e la conseguente Regola 2) è un risultato solo necessario ma non sufficiente per determinare l'ottimo $\tilde{\mathbf{x}}$. A tal proposito analizziamo tre diverse mappe isoquante rappresentate dalle linee di livello I , II e III di una funzione $f(A, B)$ sottoposta al vincolo V anch'esso di forma variabile (vedi Figura 9).

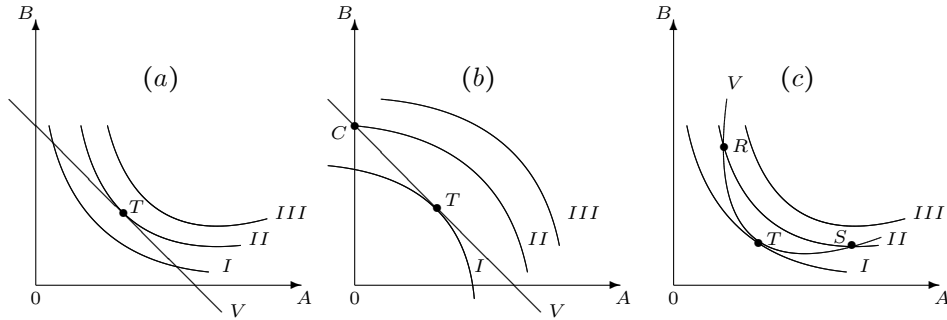


FIGURA 9

In (a) le curve di livello di f sono convesse ed il vincolo V è lineare. Il problema di ottimo vincolato ammette come soluzione il punto T di tangenza. In (b) le curve di livello di f sono concave ed il vincolo V è ancora lineare per cui non è più il punto di tangenza T a renderla massima bensì il punto in alto C . In (c) le curve di livello della funzione obiettivo sono convesse ed il vincolo V non è lineare anzi è rappresentato da una curva convessa “più chiusa” della f . Anche stavolta il punto di tangenza T non è il valore massimo perché possiamo rilevare valori maggiori della f nei punti R ed S per esempio.

Da questi esempi osserviamo che, quando il vincolo è lineare e la funzione obiettivo di due variabili è convessa, il massimo vincolato coincide con il punto di tangenza, negli altri casi questo non accade.

Supponiamo che f e g siano di classe C^2 e sia $\text{Ker}G(\tilde{\mathbf{x}})$ lo spazio vettoriale tangente a V in $\tilde{\mathbf{x}}$ dato da

$$\text{Ker}G(\tilde{\mathbf{x}}) := \{\mathbf{b} \in \mathbb{R}^n : G(\tilde{\mathbf{x}})\mathbf{b} = 0\}$$

allora, sotto le stesse condizioni e le ipotesi del Teorema 9, possiamo enunciare il teorema che assegna condizioni necessarie e sufficienti per determinare l'ottimo locale del problema di ottimizzazione nelle forme (ECP-Max) e (ECP-min).

Teorema 11 (condizione del secondo ordine). *Sia $(\tilde{\mathbf{x}}, \tilde{\lambda}) \in \mathbb{R}^{n+h}$ tale che $\nabla \mathcal{L}(\tilde{\mathbf{x}}, \tilde{\lambda}) = 0$. Allora:*

CN): *Se $\tilde{\mathbf{x}}$ è un punto di massimo (minimo) locale di $f|_V$ e $\tilde{\lambda}$ il relativo moltiplicatore di Lagrange allora risulta che la forma quadratica (hessiana) $\mathbf{H}_{\mathcal{L}}(\tilde{\mathbf{x}}, \tilde{\lambda})$ è semidefinita negativa (positiva).*

CS): Se la forma quadratica (hessiana) $\mathbf{H}_{\mathcal{L}}(\tilde{\mathbf{x}}, \tilde{\boldsymbol{\lambda}})$ è definita negativa (positiva) su $\text{Ker}G(\tilde{\mathbf{x}})$ allora $\tilde{\mathbf{x}}$ è un punto di massimo (minimo) locale proprio per $f|_V$.

Dimostrazione. Dei due risultati certamente il più significativo è il secondo di cui riportiamo la dimostrazione.

CS) Come nell'analogo Teorema 9 proviamo che se H_f è definita positiva esistono un intorno U di $\tilde{\mathbf{x}}$ in $\mathbb{R}^n$ ed un $m \in \mathbb{R}_+$ tali che

$$(17) \quad f(\mathbf{x}) - f(\tilde{\mathbf{x}}) \geq m \|\mathbf{x} - \tilde{\mathbf{x}}\|^2 \quad \text{se } \mathbf{x} \in U \cap V.$$

Supponiamo, per assurdo, che la (17) non sussista per alcun $m \in \mathbb{R}_+$ e per alcun intorno U di $\tilde{\mathbf{x}}$. Allora possiamo costruire una successione $\{\mathbf{x}_i\}_{i \in \mathbb{N}}$ in V convergente verso $\tilde{\mathbf{x}}$ e tale che, per qualsiasi i , si abbia $\mathbf{x}_i \neq \tilde{\mathbf{x}}$ ed anche

$$(18) \quad f(\mathbf{x}_i) - f(\tilde{\mathbf{x}}) < \frac{1}{i} \|\mathbf{x}_i - \tilde{\mathbf{x}}\|^2$$

Passando eventualmente ad una estratta, che per semplicità seguitiamo ad indicare con $\{x_i\}_{i \in \mathbb{N}}$, possiamo supporre che esista $\mathbf{b} \in \mathbb{R}^n$ tale che

$$\lim_{i \rightarrow +\infty} \frac{\mathbf{x}_i - \tilde{\mathbf{x}}}{\|\mathbf{x}_i - \tilde{\mathbf{x}}\|} = \mathbf{b};$$

in modo che risulti $\mathbf{b} \in \text{Ker}G(\tilde{\mathbf{x}})$.

D'altra parte poiché $\mathcal{L}$ è di classe C^2 si ha, dalla (18) per ogni $i \in \mathbb{N}$

$$\frac{1}{i} \|\mathbf{x}_i - \tilde{\mathbf{x}}\|^2 > f(\mathbf{x}_i) - f(\tilde{\mathbf{x}}) = \mathcal{L}(x_i) - \mathcal{L}(\tilde{\mathbf{x}}) = \sum_{s=1}^n \mathcal{L}_{x_s}(\tilde{\mathbf{x}})(x_s^i - x_s) + \frac{1}{2} \sum_{s,v=1}^n \mathcal{L}_{x_s x_v}(\tilde{\mathbf{x}} + \theta_i(\mathbf{x}_i - \tilde{\mathbf{x}}))(x_s^i - \tilde{x}_s)(x_v^i - \tilde{x}_v)$$

essendo $\theta \in [0, 1]$. Ne segue che

$$\frac{1}{i} > \frac{1}{2} \sum_{s,v=1}^n \mathcal{L}_{x_s x_v}(\tilde{\mathbf{x}} + \theta_i(\mathbf{x}_i - \tilde{\mathbf{x}})) \frac{x_s^i - \tilde{x}_s}{\|\mathbf{x}_i - \tilde{\mathbf{x}}\|} \frac{x_v^i - \tilde{x}_v}{\|\mathbf{x}_i - \tilde{\mathbf{x}}\|} \quad \forall i \in \mathbb{N}$$

e passando al limite per $i \rightarrow +\infty$

$$0 \geq \sum_{s,v=1}^n \mathcal{L}_{x_s x_v}(\tilde{\mathbf{x}}) b_s b_v = H_f(\mathbf{b}).$$

Poiché quest'ultima contrasta con l'ipotesi che $H_f(\mathbf{b})$ sia definita positiva su $\text{Ker}G(\tilde{\mathbf{x}})$ la condizione sufficiente è provata. □

Sia $(\tilde{\mathbf{x}}, \tilde{\boldsymbol{\lambda}})$ soluzione locale del problema di ottimizzazione nella forma (ECP-Max) o (ECP-min). La CS) del Teorema 11 può essere così riscritta

$$(19) \quad \mathbf{b}^T H_f \mathbf{b} < 0 \quad (\text{risp. } > 0) \quad \forall \mathbf{b} \in \mathbb{R}^n - \{0\} \text{ tale che } G \cdot \mathbf{b} = 0$$

dove H_f va valutata in $(\tilde{\mathbf{x}}, \tilde{\boldsymbol{\lambda}})$ e G in $\tilde{\mathbf{x}}$.

Come possiamo valutare in modo agevole la condizione (19) ?

Regola 3. Se $G(\tilde{\mathbf{x}})$ ha rango massimo e se $\nabla \mathcal{L}(\tilde{\mathbf{x}}, \tilde{\boldsymbol{\lambda}}) = 0$, per il Teorema 17, la condizione sufficiente del secondo ordine per il massimo relativo è:

$$(-1)^r \left(\begin{array}{ccc|ccc} \mathcal{L}_{x_1 x_1} & \cdots & \mathcal{L}_{x_1 x_r} & g_{1x_1} & \cdots & g_{1x_h} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \mathcal{L}_{x_1 x_r} & \cdots & \mathcal{L}_{x_r x_r} & g_{1x_r} & \cdots & g_{hx_r} \\ \hline g_{1x_1} & \cdots & g_{1x_r} & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ g_{1x_h} & \cdots & g_{hx_r} & 0 & \cdots & 0 \end{array} \right) (\tilde{\mathbf{x}}, \tilde{\boldsymbol{\lambda}}) > 0 \quad \forall r = h+1, \dots, n$$

mentre la condizione sufficiente per il minimo relativo prevede che il determinante suindicato abbia $\forall r = h+1, \dots, n$ segno:

- sempre positivo nel caso di un numero pari di vincoli;
- sempre negativo nel caso di un numero dispari di vincoli.

Dalla Regola 3 sviluppiamo alcuni casi utili per la soluzione di problemi reali.

caso 1) Nel caso di un solo vincolo e due variabili dobbiamo controllare una sola condizione ovvero valutando in $(\tilde{\mathbf{x}}, \tilde{\lambda})$ si deve avere

$$\begin{vmatrix} \mathcal{L}_{x_1 x_1} & \mathcal{L}_{x_1 x_2} & g_{x_1} \\ \mathcal{L}_{x_1 x_2} & \mathcal{L}_{x_2 x_2} & g_{x_2} \\ g_{x_1} & g_{x_2} & 0 \end{vmatrix}(\tilde{\mathbf{x}}, \tilde{\lambda}) > 0 \quad (\text{risp. } < 0)$$

per essere certi che $f|_V$ in $\tilde{\mathbf{x}}$ abbia massimo (risp. minimo) relativo proprio.

caso 2) Nel caso di n variabili e due vincoli si dovranno “orlare” i minori principali di H_f a partire dal terzo ordine

$$\begin{vmatrix} \mathcal{L}_{x_1 x_1} & \mathcal{L}_{x_1 x_2} & \mathcal{L}_{x_1 x_3} & g_{1x_1} & g_{2x_1} \\ \mathcal{L}_{x_1 x_2} & \mathcal{L}_{x_2 x_2} & \mathcal{L}_{x_2 x_3} & g_{1x_2} & g_{2x_2} \\ \mathcal{L}_{x_1 x_3} & \mathcal{L}_{x_2 x_3} & \mathcal{L}_{x_3 x_3} & g_{1x_3} & g_{2x_3} \\ g_{1x_1} & g_{1x_2} & g_{1x_3} & 0 & 0 \\ g_{2x_1} & g_{2x_2} & g_{2x_3} & 0 & 0 \end{vmatrix},$$

$$\begin{vmatrix} \mathcal{L}_{x_1 x_1} & \mathcal{L}_{x_1 x_2} & \mathcal{L}_{x_1 x_3} & \mathcal{L}_{x_1 x_4} & g_{1x_1} & g_{2x_1} \\ \mathcal{L}_{x_1 x_2} & \mathcal{L}_{x_2 x_2} & \mathcal{L}_{x_2 x_3} & \mathcal{L}_{x_2 x_4} & g_{1x_2} & g_{2x_2} \\ \mathcal{L}_{x_1 x_3} & \mathcal{L}_{x_2 x_3} & \mathcal{L}_{x_3 x_3} & \mathcal{L}_{x_3 x_4} & g_{1x_3} & g_{2x_3} \\ \mathcal{L}_{x_1 x_4} & \mathcal{L}_{x_2 x_4} & \mathcal{L}_{x_3 x_4} & \mathcal{L}_{x_4 x_4} & g_{1x_4} & g_{2x_4} \\ g_{1x_1} & g_{1x_2} & g_{1x_3} & g_{1x_4} & 0 & 0 \\ g_{2x_1} & g_{2x_2} & g_{2x_3} & g_{2x_4} & 0 & 0 \end{vmatrix},$$

.....

$$\left| \begin{array}{c|c} H_f & G^T \\ \hline G & 0 \end{array} \right|$$

se i segni dei determinanti suddetti si alternano a partire da quello negativo si ha un massimo relativo; se sono invece, tutti positivi si ha un minimo relativo.

Per le funzioni a due variabili ad un solo vincolo il Teorema 11 si modifica come segue.

Proposizione 12. Siano $f : D \subseteq \mathbb{R}^2 \rightarrow \mathbb{R}$ la funzione obiettivo tale che $f \in \mathcal{C}^2(\mathbb{R}^2, \mathbb{R})$ e $g : D \subseteq \mathbb{R}^2 \rightarrow \mathbb{R}$ la funzione vincolare tale che $g \in \mathcal{C}^2(\mathbb{R}^2, \mathbb{R})$. Se $\nabla \mathcal{L}(\tilde{\mathbf{x}}, \tilde{\lambda}) = 0$ e $\det \mathbf{H}_{\mathcal{L}}(\tilde{\mathbf{x}}, \tilde{\lambda}) > 0$ allora $f|_V$ ammette massimo relativo in $\tilde{\mathbf{x}}$.

Dimostrazione. Certamente si ha $\nabla g(\tilde{\mathbf{x}}) \neq 0$ per cui possiamo ipotizzare che $g_{x_1}(\tilde{\mathbf{x}}) \neq 0$. Intorno a $\tilde{\mathbf{x}}$ la funzione $f|_V$ coincide con la funzione $F(t) = f(\phi(t), t)$ essendo ϕ definita implicitamente da $g(\phi(t), t) = 0$ per qualsiasi t .

Utilizzando $\phi' = -\frac{g_{x_2}}{g_{x_1}}$ e

$$\phi'' = - \left[\frac{(g_{x_2 x_1} \phi' + g_{x_2 x_2})g_{x_1} - (g_{x_1 x_1} \phi' + g_{x_1 x_2})g_{x_2}}{(g_{x_1})^2} \right]$$

calcoliamo

$$\begin{aligned} F' &= f_{x_1} \phi' + f_{x_2}; \\ F'' &= (f_{x_1 x_1} \phi' + f_{x_1 x_2}) \phi' + f_{x_1} \phi'' + (f_{x_2 x_1} \phi' + f_{x_2 x_2}) = \\ &= (\phi')^2 f_{x_1 x_1} + 2f_{x_1 x_2} \phi' - \frac{f_{x_1}}{(g_{x_1})^2} [(g_{x_2 x_1} \phi' + g_{x_2 x_2})g_{x_1} - (g_{x_1 x_1} \phi' + g_{x_1 x_2})g_{x_2}] + f_{x_2 x_2} = \\ &= \frac{(g_{x_2})^2}{(g_{x_1})^2} f_{x_1 x_1} - 2f_{x_1 x_2} \frac{g_{x_2}}{g_{x_1}} - \frac{f_{x_1}}{(g_{x_1})^2} [-2g_{x_1 x_2} g_{x_2} + g_{x_2 x_2} g_{x_1} + g_{x_1 x_1} g_{x_2} \frac{g_{x_2}}{g_{x_1}}] + f_{x_2 x_2} \end{aligned}$$

allora $F''(\tilde{x}_2) < 0$ é equivalente ad affermare

$$(g_{x_2})^2 [f_{x_1 x_1} + \tilde{\lambda} g_{x_1 x_1}] + (g_{x_1})^2 [f_{x_2 x_2} + \tilde{\lambda} g_{x_2 x_2}] - 2[f_{x_1 x_2} + \tilde{\lambda} g_{x_1 x_2}] g_{x_1} g_{x_2} < 0.$$

□

3.5. Esempi.

Esercizio 3.1. Determinare tutti i minimi ed i massimi locali della funzione $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ così definita: $z = f(x, y) = x^2 + 4y^2 - 3$ sottoposta al vincolo $g(x, y) = 4x^2 - 16x + y^2 + 12 = 0$.

SOLUZIONE: Costruiamo la funzione Lagrangiana

$$\mathcal{L}(x, y, \lambda) = x^2 + 4y^2 - 3 + \lambda(4x^2 - 16x + y^2 + 12).$$

Le condizioni necessarie per determinare un ottimo locale in accordo al Teorema 9 sono:

$$(20) \quad \mathcal{L}_x(x, y, \lambda) = 2x + 8\lambda x - 16\lambda = 0$$

$$(21) \quad \mathcal{L}_y(x, y, \lambda) = 8y + 2\lambda y = 0$$

$$(22) \quad \mathcal{L}_\lambda(x, y, \lambda) = 4x^2 - 16x + y^2 + 12 = 0.$$

Mettendo in evidenza $2y$ nell'equazione (21), otteniamo

$$2y(4 + \lambda) = 0.$$

Così, dobbiamo distinguere due casi.

$y = 0$): Sostituendo $y = 0$ nell'equazione (22), otteniamo $4x^2 - 16x + 12 = 0$ equivalente a $x^2 - 4x + 3 = 0$, di soluzioni

$$x_1 = 2 + \sqrt{4 - 3} = 3 \quad \text{e} \quad x_2 = 2 - \sqrt{4 - 3} = 1,$$

e dall'equazione (20) otteniamo i corrispondenti λ -valori:

$$\lambda_1 = -\frac{3}{4} \quad \text{e} \quad \lambda_2 = \frac{1}{4}.$$

Quindi i punti

$$P_1 : (x_1, y_1) = (3, 0) \text{ per } \lambda_1 = -\frac{3}{4} \quad P_2 : (x_2, y_2) = (1, 0) \text{ per } \lambda_2 = \frac{1}{4}$$

sono candidati ad essere i punti locali estremi in questo caso.

$\lambda = 4$): Allora dalla (20) otteniamo l'equazione $2x - 32x + 64 = 0$, che ammette come soluzione

$$x_3 = \frac{64}{30} = \frac{32}{15}.$$

Sostituendo il risultato nell'equazione (22), otteniamo l'equazione

$$4\left(\frac{32}{15}\right)^2 - 16 \cdot \frac{32}{15} + y^2 + 12 = 0$$

equivalente a

$$y^2 = \frac{884}{225}.$$

In definitiva otteniamo i punti

$$P_3 : (x_3, y_3) = \left(\frac{32}{15}, \frac{\sqrt{884}}{15} \approx 1.98\right) \text{ per } \lambda_3 = -4$$

e

$$P_4 : (x_3, y_4) = \left(\frac{32}{15}, -\frac{\sqrt{884}}{15} \approx -1.98\right) \text{ per } \lambda_4 = -4$$

sono candidati ad essere i punti locali estremi in questo caso.

Adesso dobbiamo verificare le condizioni sufficienti del Teorema 11. Determiniamo prima

$$\mathcal{L}_{xx}(x, y; \lambda) = 2 + 8\lambda$$

$$\mathcal{L}_{xy}(x, y; \lambda) = 0 = \mathcal{L}_{yx}(x, y; \lambda)$$

$$\mathcal{L}_{yy}(x, y; \lambda) = 8 + 2\lambda$$

$$\mathcal{L}_{x\lambda}(x, y; \lambda) = 8x - 16 = \mathcal{L}_{\lambda x}(x, y; \lambda)$$

$$\mathcal{L}_{y\lambda}(x, y; \lambda) = 2y = \mathcal{L}_{\lambda y}(x, y; \lambda).$$

Così, siamo in grado di costruire l'Hessiano completo di $\mathcal{L}$

$$H_{\mathcal{L}}(x, y; \lambda) = \left(\begin{array}{cc|c} 2 + 8\lambda & 0 & 8x - 16 \\ 0 & 8 + 2\lambda & 2y \\ \hline 8x - 16 & 2y & 0 \end{array} \right)$$

Siamo nel *caso 1*) ossia in presenza di due variabili in f e di un solo vincolo per cui dobbiamo soltanto calcolare l'Hessiano nei quattro punti di stazionarietà precedentemente trovati e verificarne il segno. In calcoli abbiamo:

$$\left| H_{\mathcal{L}} \left(3, 0; -\frac{3}{4} \right) \right| = \begin{pmatrix} -4 & 0 & 8 \\ 0 & 6.5 & 0 \\ 8 & 0 & 0 \end{pmatrix} = 6.5 \cdot (-64) < 0$$

$$\left| H_{\mathcal{L}} \left(1, 0; \frac{1}{4} \right) \right| = \begin{pmatrix} 4 & 0 & -8 \\ 0 & 8.5 & 0 \\ -8 & 0 & 0 \end{pmatrix} = 8.5 \cdot (-64) < 0$$

$$\left| H_{\mathcal{L}} \left(\frac{32}{15}, \frac{\sqrt{884}}{15}; -4 \right) \right| = \begin{pmatrix} -30 & 0 & \frac{16}{15} \\ 0 & 0 & 2\frac{\sqrt{884}}{15} \\ \frac{16}{15} & 2\frac{\sqrt{884}}{15} & 0 \end{pmatrix} > 0$$

$$\left| H_{\mathcal{L}} \left(\frac{32}{15}, -\frac{\sqrt{884}}{15}; -4 \right) \right| = \begin{pmatrix} -30 & 0 & \frac{16}{15} \\ 0 & 0 & -2\frac{\sqrt{884}}{15} \\ \frac{16}{15} & -2\frac{\sqrt{884}}{15} & 0 \end{pmatrix} > 0$$

A questo punto siamo in grado di trarre le dovute conclusioni per cui per il Teorema 11 riformulato nella Proposizione 12 o che si voglia per semplicità seguendo il *caso 1*), i punti $(3, 0)$ e $(1, 0)$ sono punti di minimo locale, mentre i punti $(\frac{32}{15}, \frac{\sqrt{884}}{15})$ e $(\frac{32}{15}, -\frac{\sqrt{884}}{15})$ sono punti di massimo locale.

Esercizio 3.2. Determinare tutti i minimi ed i massimi della funzione $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ così definita: $f(x, y, z) = x + y + 2z$ sottoposta ai vincoli $g_1(x, y) = x^2 + z^2 - 8 = 0$ e $g_2(x, y) = x - y - 1 = 0$.

SOLUZIONE: La funzione $f(x, y, z)$ è definita su $D = \mathbb{R}^3$ che non è un compatto. La Lagrangiana associata è

$$\mathcal{L}(x, y, z, \lambda, \mu) = x + y + 2z + \lambda(x^2 + z^2 - 8) + \mu(x - y - 1)$$

e la condizione di regolarità è sempre verificata

$$\nabla g(\mathbf{x}) \neq \mathbf{0} \quad \Rightarrow \quad \begin{pmatrix} 2x & 0 & 2z \\ 1 & -1 & 0 \end{pmatrix} \neq \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Le condizioni del primo ordine dettate dal Teorema 9

$$\begin{cases} \mathcal{L}_x(x, y, z, \lambda, \mu) = 1 + 2x\lambda + \mu = 0 \\ \mathcal{L}_y(x, y, z, \lambda, \mu) = 1 - \mu = 0 \\ \mathcal{L}_z(x, y, z, \lambda, \mu) = 2 + 2z\lambda = 0 \\ \mathcal{L}_\lambda(x, y, z, \lambda, \mu) = -8 + x^2 + z^2 = 0 \\ \mathcal{L}_\mu(x, y, z, \lambda, \mu) = -1 + x - y = 0 \end{cases}$$

sono soddisfatte dai punti

$$A = \left(-2, -3, -2, \frac{1}{2}, 1 \right); \quad B = \left(2, 1, 2, -\frac{1}{2}, 1 \right)$$

L'Hessiano della Lagrangiana è

$$\mathbf{H}_{\mathcal{L}}(x, y, z, \lambda, \mu) = \begin{pmatrix} 2\lambda & 0 & 0 & 0 & 2x-1 \\ 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 2\lambda & 2z & 0 \\ 2x & 0 & 2z & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 \end{pmatrix}$$

da cui analizzando i determinanti delle sottomatrici principali di nord-ovest calcolati nei punti A e B e rispettando la Regola 3 nell'accezione del caso 2) otteniamo

$$\bullet \mathbf{H}_{\mathcal{L}}(A) = \begin{pmatrix} 1 & 0 & 0 & -4 & 1 \\ 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & -4 & 0 \\ -4 & 0 & -4 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 \end{pmatrix} \text{ da cui}$$

$$\det \mathbf{H}_{\mathcal{L}_3}(A) = \begin{vmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{vmatrix} = 0 (\geq 0); \quad \det \mathbf{H}_{\mathcal{L}_4}(A) = \begin{vmatrix} 1 & 0 & 0 & -4 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -4 \\ -4 & 0 & -4 & 0 \end{vmatrix} = 0 (\geq 0);$$

$$\det \mathbf{H}_{\mathcal{L}_5}(A) = \begin{vmatrix} 1 & 0 & 0 & -4 & 1 \\ 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & -4 & 0 \\ -4 & 0 & -4 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 \end{vmatrix} = 32 > 0$$

e quindi $A = (-2, -3, -2)$ è punto di minimo locale per $f(x, y, z)$.

$$\bullet \mathbf{H}_{\mathcal{L}}(B) = \begin{pmatrix} -1 & 0 & 0 & 4 & 1 \\ 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 4 & 0 \\ 4 & 0 & 4 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 \end{pmatrix} \text{ da cui}$$

$$\det \mathbf{H}_{\mathcal{L}_3}(B) = \begin{vmatrix} -1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{vmatrix} = 0 (\leq 0); \quad \det \mathbf{H}_{\mathcal{L}_4}(B) = \begin{vmatrix} -1 & 0 & 0 & 4 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 4 \\ 4 & 0 & 4 & 0 \end{vmatrix} = 0 (\geq 0);$$

$$\det \mathbf{H}_{\mathcal{L}_5}(B) = \begin{vmatrix} -1 & 0 & 0 & 4 & 1 \\ 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 4 & 0 \\ 4 & 0 & 4 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 \end{vmatrix} = -32 < 0$$

e quindi $B = (2, 1, 2)$ è punto di massimo locale per $f(x, y, z)$.

Inoltre dal calcolo dei minori principali di nord-ovest dell'Hessiano

$$\det \mathbf{H}_{\mathcal{L}_1}(x, y, z, \lambda, \mu) = 2\lambda$$

$$\det \mathbf{H}_{\mathcal{L}_2}(x, y, z, \lambda, \mu) = 0$$

$$\det \mathbf{H}_{\mathcal{L}_3}(x, y, z, \lambda, \mu) = 0$$

$$\det \mathbf{H}_{\mathcal{L}_4}(x, y, z, \lambda, \mu) =$$

$$\det \mathbf{H}_{\mathcal{L}_5}(x, y, z, \lambda, \mu) = 8x^2\lambda + 8z^2\lambda$$

risulta che

- per $\lambda > 0$ la forma quadratica $Q(x, y, z, \lambda, \mu)$ associata all'Hessiano $\mathbf{H}_{\mathcal{L}}(x, y, z, \lambda, \mu)$ è semidefinita positiva e pertanto per il Teorema 23 f è debolmente convessa e quindi A (ottenuto per $\lambda = \frac{1}{2}$) è anche punto di minimo globale per f ;
- per $\lambda < 0$ la forma quadratica $Q(x, y, z, \lambda, \mu)$ associata all'Hessiano $\mathbf{H}_{\mathcal{L}}(x, y, z, \lambda, \mu)$ è semidefinita negativa e pertanto per il Teorema 23 f è debolmente concava e quindi B (ottenuto per $\lambda = -\frac{1}{2}$) è anche punto di massimo globale per f .

3.6. Significato dei moltiplicatori di Lagrange. Siano f la funzione obiettivo e $g(\mathbf{x}) = \mathbf{b}_0$ l'equazione del vincolo V dove $\mathbf{b}_0 \in \mathbb{R}^h$. La Lagrangiana scritta nella forma

$$\mathcal{L}_{\mathbf{b}_0}(\mathbf{x}, \boldsymbol{\lambda}) = f(\mathbf{x}) + \boldsymbol{\lambda}(\mathbf{b}_0 - g(\mathbf{x}))$$

dà risalto al vettore $\mathbf{b}_0$.

Supponiamo che $(\tilde{\mathbf{x}}, \tilde{\boldsymbol{\lambda}})$ sia un punto tale che é verificata la condizione $\nabla \mathcal{L}_{\mathbf{b}_0}(\tilde{\mathbf{x}}, \tilde{\boldsymbol{\lambda}}) = 0$ e che siano soddisfatte le condizioni del Teorema 11 per cui $\tilde{\mathbf{x}}$ é un punto di massimo relativo per $f|_V$. Osserviamo innanzitutto che le condizioni sufficienti valgono in un opportuno intorno di $(\tilde{\mathbf{x}}, \tilde{\boldsymbol{\lambda}})$ indipendentemente da $\mathbf{b}_0$.

Introduciamo la funzione $F : D \times \mathbb{R}^h \times \mathbb{R}^h \rightarrow \mathbb{R}^{n+h}$ così definita

$$F(\mathbf{x}, \boldsymbol{\lambda}, \mathbf{b}) := \nabla \mathcal{L}_{\mathbf{b}}(\mathbf{x}, \boldsymbol{\lambda})$$

precisando che il gradiente é svolto rispetto alle sole variabili x_i e λ_j .

Consideriamo l'equazione

$$(23) \quad F(\mathbf{x}, \boldsymbol{\lambda}, \mathbf{b}) = 0 \quad \Leftrightarrow \quad \nabla \mathcal{L}_{\mathbf{b}}(\mathbf{x}, \boldsymbol{\lambda}) = 0.$$

Chiaramente $(\tilde{\mathbf{x}}, \tilde{\boldsymbol{\lambda}}, \mathbf{b}_0)$ è una soluzione della (23); inoltre ∇F rispetto alle $\mathbf{x}$ ed alle $\boldsymbol{\lambda}$ è diverso da zero in $(\tilde{\mathbf{x}}, \tilde{\boldsymbol{\lambda}})$. Per il Teorema 21 ne segue che la (23) definisce implicitamente una funzione $(\mathbf{x}, \boldsymbol{\lambda}) = (\phi(\mathbf{b}), \psi(\mathbf{b}))$ da un intorno di $\mathbf{b}_0$ verso un intorno di $(\tilde{\mathbf{x}}, \tilde{\boldsymbol{\lambda}})$.

Avendosi per ogni $\mathbf{b}$

$$(24) \quad \nabla \mathcal{L}_{\mathbf{b}}(\phi(\mathbf{b}), \psi(\mathbf{b})) = 0$$

e per le condizioni del Teorema 11 $x = \phi(\mathbf{b})$ è massimo relativo per f ristretta al vincolo $g(\mathbf{x}) = \mathbf{b}$ e $\boldsymbol{\lambda} = \psi(\mathbf{b})$ è il relativo moltiplicatore di Lagrange.

Abbiamo così ottenuto, almeno localmente, che sia il punto di massimo relativo che il relativo moltiplicatore dipendono dal valore $\mathbf{b}$ del vincolo.

Se l'intento è di massimizzare f ristretta a $g(\mathbf{x}) = \mathbf{b}_0$ allora le funzioni $\mathbf{x} = \phi(\mathbf{b})$ e $\boldsymbol{\lambda} = \psi(\mathbf{b})$ producono rispettivamente il punto di massimo ed il relativo moltiplicatore del problema "perturbato" che prevede come vincolo $g(\mathbf{x}) = \mathbf{b}_0$. Ciò, naturalmente, a patto che $\mathbf{b}$ sia abbastanza vicino a $\mathbf{b}_0$. Con questa limitazione, concludiamo affermando che per la famiglia dei problemi

$$(ECP\text{-}Max\text{-}b) \quad \begin{cases} \max_{\mathbf{x}} f(\mathbf{x}) \\ \text{t.c.} \quad g(\mathbf{x}) = \mathbf{b} \end{cases}$$

le soluzioni sono completamente descritte dalla funzione $\phi(\mathbf{b})$. Evidentemente $v(\mathbf{b}) := f(\phi(\mathbf{b}))$ rappresenta il valore ottimo. Derivando otteniamo

$$(25) \quad v'(\mathbf{b}) = f'(\phi(\mathbf{b}))\phi'(\mathbf{b})$$

e tenendo conto della (24) disponiamo delle due relazioni

$$f'(\phi(\mathbf{b})) = \psi(\mathbf{b})g'(\phi(\mathbf{b})) \quad \text{e} \quad g(\phi(\mathbf{b})) = \mathbf{b}.$$

Derivando la seconda e sostituendo entrambe nella (25) otteniamo

$$v'(\mathbf{b}) = \psi(\mathbf{b})$$

che nel punto $\mathbf{b}_0$ diventa

$$v_{b_s}(\mathbf{b}_0) = \tilde{\lambda}_s \quad \text{con } s = 1, \dots, h.$$

Questo è un risultato "significativo" che può essere riassunto come segue:

nella soluzione del problema (ECP-Max-b) $\tilde{\boldsymbol{\lambda}}$ non è solo il prodotto di un artificio che permette di pervenire al punto di ottimo $\tilde{\mathbf{x}}$; $\tilde{\boldsymbol{\lambda}}$ misura la sensibilità del valore ottimo dell'obiettivo $v^* = f(\tilde{\mathbf{x}})$ alle variazioni del valore b_0 del vincolo.

In particolari contesti ciò porta a significative interpretazioni dei moltiplicatori.

Esercizio 3.3. Una compagnia alloca €600000 da investire in pubblicità e ricerca. La compagnia stima che per investire x centinaia di € in pubblicità e y centinaia di € in ricerca, venderà approssimativamente un totale di $30x^{4/5}y^{1/3}$ unità del suo prodotto. Quanto dovrebbe spendere la società in ricerca ed in pubblicità per massimizzare il totale vendite? Che cosa accade se il budget subisce un incremento dell'1% ed se subisce un decremento dell'1.5%?

SOLUZIONE: Lavoriamo in unità di €1000. In tal modo la funzione obiettivo è

$$f(x, y) = 30x^{4/5}y^{1/3}$$

e l'equazione del vincolo è

$$x + y = 600 \quad \text{oppure} \quad 600 - x - y = 0$$

per cui il problema assume la forma (ECP-Max)

$$(26) \quad \begin{cases} \max_{(x,y)} & 30x^{4/5}y^{1/3} \\ \text{t.c.} & 600 - x - y = 0 \end{cases}$$

La Lagrangiana è

$$\mathcal{L}(x, y, \lambda) = 30x^{4/5}y^{1/3} + \lambda(600 - x - y).$$

Per i punti di stazionarietà occorrono che sia verificate:

$$(27) \quad 0 = \mathcal{L}_x = 24x^{-1/5}y^{1/3} - \lambda$$

$$(28) \quad 0 = \mathcal{L}_y = 10x^{4/5}y^{-2/3} - \lambda$$

$$(29) \quad 0 = \mathcal{L}_\lambda = 600 - x - y$$

Le equazioni (27) e (28) danno

$$(30) \quad \lambda = 24x^{-\frac{1}{5}}y^{\frac{1}{3}} = 10x^{\frac{4}{5}}y^{-\frac{2}{3}}.$$

Perciò $x = 2.4y$. Sostituendo nella (29) otteniamo $600 - 2.4y - y = 0$ equivalente a $y = \frac{600}{3.4} \approx 176.47$. Allora $x = 423.53$ e utilizzando la (30) otteniamo $\lambda = 40.15$. La condizione del secondo ordine attesta che $(\tilde{x}, \tilde{y}) = (423.53, 176.47)$ è soluzione del problema (26) ed il valore ottimale, ossia il totale massimo di vendite, è $f(423.53, 176.47) = 21257.83$.

Per rispondere alla seconda domanda dovremmo ripetere i calcoli utilizzando un nuovo budget ottenuto dal precedente mediante l'incremento dell'1% o il decremento dell'1.5%. La procedura sarebbe senz'altro corretta ma risulterebbe comunque lunga e tediosa! Proviamo a ragionare in modo diverso sfruttando proprio il significato di λ , inteso come rilevatore della sensibilità dell'ottimo dell'obiettivo al variare della costante $\mathbf{b}_0$ del vincolo.

Le variazioni a cui il budget è sottoposto sono molto piccole per cui è intuibile pensare che anche i nuovi problemi di ottimo da risolvere avranno come soluzioni valori che saranno prossimi a quello già rilevato. Occorrerà quindi soltanto calcolare quale sia l'incremento o il decremento che subirà il massimo precedente per determinare i nuovi massimi. Le soluzioni ottenute, anche se risulteranno approssimate rispetto alle reali, rivestiranno comunque un ruolo significativo per il problema da affrontare.

Se il bilancio è aumentato dell'1% a €606000 otteniamo un aumento di 6 nella costante del vincolo, dato che stiamo lavorando in unità di €1000. Il totale massimo delle vendite, di conseguenza, aumenterà approssimativamente di $6\lambda = 6 \cdot 40.15 = 240.90$. Questo è un aumento di circa dell'1.13% sul massimo di vendite precedente. (Se effettuassimo il calcolo con il nuovo budget di €606000, l'incremento attuale sarebbe 241.08.)

Supponiamo adesso che il bilancio decresce dell'1.5% a €591000, il totale massimo delle vendite decrescerebbe approssimativamente di $9\lambda = 9 \cdot 40.15 = 361.35$ perché il vincolo subisce un calo continuo di 9. Questo si ripercuote sul massimo delle vendite determinando un calo del 1.7%. (Il decremento attuale nel massimo delle vendite è dai calcoli pari a 361.02 non troppo dissimile dal valore trovato!)

4. OTTIMIZZAZIONE CON VINCOLI DI DISUGLIANZA

Consideriamo il problema di ottimizzazione nel caso in cui $V = K \cap D \subset D$ e $\mathfrak{P}(\mathbf{x})$ sia una relazione di disuguaglianza.

Per i problemi di ottimizzazione con vincoli di disuguaglianza non è possibile eseguire una formulazione generale in quanto le condizioni di ottimalità (anche quelle necessarie o dette del primo ordine) non sono le stesse tra i problemi del tipo (P-Max) ed i problemi di tipo (P-min). Pertanto svolgeremo la trattazione con riferimento ai problemi di massimo e successivamente definiremo come trasformare i problemi di minimo riconducendoli sempre ai primi.

La complessità della ricerca dei massimi vincolati ci porterà ad esaminare prima il problema con un solo vincolo di non positività, successivamente con più vincoli di non positività ed infine il problema generale in cui, accanto a vincoli di non positività, ci sono anche vincoli di non negatività.

4.1. Problema di massimo con vincoli di non positività. Siano $f : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ la funzione obiettivo tale che $f \in \mathcal{C}^2(\mathbb{R}^n, \mathbb{R})$, $g : K \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^h$ la funzione vincolare tale che $g \in \mathcal{C}^1(\mathbb{R}^n, \mathbb{R}^h)$, $\mathbf{b} \in \mathbb{R}^h$ un vettore reale di h componenti e $g_i(\mathbf{x}) \leq b_i$ (con $i = 1, 2, \dots, h$) gli h vincoli di non negatività per cui $V = \{\mathbf{x} \in D \cap K : g(\mathbf{x}) \leq \mathbf{b}\}$ è la regione ammissibile. In particolare i vincoli di uguaglianza $g_i(\mathbf{x}) = \mathbf{b}$ sono detti *vincoli attivi* o *saturi* o *stringenti*, mentre quelli di disuguaglianza $g_j(\mathbf{x}) < \mathbf{b}$ sono detti *vincoli non attivi* o *non saturi* o *non stringenti*.

Sotto tali ipotesi il problema di massimizzazione della funzione f con h vincoli di non positività assume la forma

$$(DCP\text{-Max}) \quad \begin{cases} \max_{\mathbf{x}} f(\mathbf{x}) \\ t.c. \quad g(\mathbf{x}) \leq \mathbf{b} \end{cases}$$

Osservazione 7. Laddove la relazione $\mathfrak{P}$ è di disuguaglianza $\geq$ basterà moltiplicare ciascun vincolo per -1 è ottenere nuovamente la forma (DCP-Max).

Con analoghe considerazioni fatte nel caso dell'ottimizzazione con uguaglianza introduciamo h moltiplicatori (uno per ciascun vincolo di disuguaglianza).

Sia $\mathcal{L} : D \times \mathbb{R}^n \rightarrow \mathbb{R}^h$ definita dalla legge

$$\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}) = f(\mathbf{x}) - \boldsymbol{\lambda} \cdot (g(\mathbf{x}) - \mathbf{b})$$

per un opportuno $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \dots, \lambda_h) \in \mathbb{R}^h$ la Lagrangiana del problema (DCP-Max).

Come per il Teorema 9 e contemplando l'eventualità dei vincoli non attivi si ha un primo risultato utile:

Teorema 13 (condizione del primo ordine). *Sia $\tilde{\mathbf{x}}$ soluzione locale del problema (DCP-Max) e g tale che sia linearmente indipendenti i gradienti dei vincoli attivi in $\tilde{\mathbf{x}}$ ¹⁷. Allora esiste $\tilde{\lambda} \in \mathbb{R}^h$ che soddisfa contemporaneamente alle seguenti condizioni:*

$$\begin{aligned} (k) \quad & \mathcal{L}_{\mathbf{x}}(\tilde{\mathbf{x}}, \tilde{\lambda}) = 0 \\ (kk) \quad & \mathcal{L}_{\lambda}(\tilde{\mathbf{x}}, \tilde{\lambda}) \geq 0 \quad \wedge \quad \tilde{\lambda} \cdot \mathcal{L}_{\lambda}(\tilde{\mathbf{x}}, \tilde{\lambda}) = 0 \\ (kkk) \quad & \tilde{\lambda} \geq 0 \end{aligned}$$

Dimostrazione. Dimostriamo il Teorema nel caso più semplice di un solo vincolo di non positività

$$(D1CP-Max) \quad \begin{cases} \max_{\mathbf{x}} f(\mathbf{x}) \\ t.c. \quad g(\mathbf{x}) \leq b \end{cases}$$

per cui la condizione di regolarità della funzione vincolare è $\nabla g(\tilde{\mathbf{x}}) \neq 0$ se e solo se $g(\tilde{\mathbf{x}}) = b$ ¹⁸ e quindi occorre provare l'esistenza di $\tilde{\lambda} \in \mathbb{R}$ che soddisfa contemporaneamente alle seguenti condizioni:

$$\begin{aligned} (j) \quad & \mathcal{L}_{\mathbf{x}}(\tilde{\mathbf{x}}, \tilde{\lambda}) = 0 \\ (jj) \quad & \mathcal{L}_{\lambda}(\tilde{\mathbf{x}}, \tilde{\lambda}) \geq 0 \quad \wedge \quad \tilde{\lambda} \mathcal{L}_{\lambda}(\tilde{\mathbf{x}}, \tilde{\lambda}) = 0 \\ (jjj) \quad & \tilde{\lambda} \geq 0 \end{aligned}$$

La condizione (j) è la condizione di uguaglianza del vincolo. Infatti se $g(\tilde{\mathbf{x}}) = b$ allora per il Teorema 11 esiste un $\tilde{\lambda} \geq 0$ tale che $\nabla f(\tilde{\mathbf{x}}) = \tilde{\lambda} \nabla g(\tilde{\mathbf{x}})$ e quindi $\mathcal{L}_{\mathbf{x}}(\tilde{\mathbf{x}}, \tilde{\lambda}) = 0$.

Mentre da $g(\tilde{\mathbf{x}}) < b$ segue che $\nabla f(\tilde{\mathbf{x}}) = 0$ e quindi $\tilde{\lambda} = 0$ dato che $\nabla g(\mathbf{x}) \neq 0$. Pertanto si ha $\mathcal{L}_{\lambda}(\tilde{\mathbf{x}}, \tilde{\lambda}) \geq 0$. Inoltre da $\tilde{\lambda} = 0$ si ha che $\tilde{\lambda} \cdot (-g(\tilde{\mathbf{x}}) + b) = 0$ e quindi $\tilde{\lambda} \mathcal{L}_{\lambda}(\tilde{\mathbf{x}}, \tilde{\lambda}) = 0$. In conclusione la condizione (jj) contiene una condizione di disuguaglianza ed una condizione di uguaglianza detta *condizione di complementarità*.

Infine la condizione (jjj) esclude la possibilità che le altre condizioni possano non valere l'uno rispetto all'altra. □

Come visto nei precedenti casi di ottimizzazione l'ipotesi di convavità e convessità della funzione obiettivo porta a riconoscere tra i punti di stazionarietà i massimi e minimi globali del problema. Pertanto enunciamo i seguenti risultati di cui ometteremo la dimostrazione e che verranno intesi come condizione del secondo ordine per la massimizzazione con vincoli di disuguaglianza.

Teorema 14 (condizione del secondo ordine in forma ristretta). *Siano $f \in \mathcal{C}^1(\mathbb{R}^n, \mathbb{R})$ concava e $g_i \in \mathcal{C}^1(\mathbb{R}^h, \mathbb{R})$ e debolmente convessa per ogni $i = 1, \dots, h$. Se esistono $\tilde{\lambda} = (\tilde{\lambda}_1, \dots, \tilde{\lambda}_h) \in \mathbb{R}^h$ ed $\tilde{\mathbf{x}}$ tali che $(\tilde{\mathbf{x}}, \tilde{\lambda})$ soddisfa le condizioni (k), (kk) e (kkk) allora $\tilde{\mathbf{x}}$ è soluzione del problema (DCP-Max).*

Poiché la condizione che la funzione obiettivo sia concava è troppo restrittiva per risultare utile in alcune applicazioni economiche, allora, l'assunto sarà "generalizzato" supponendo che f sia debolmente concava.

Teorema 15 (condizione del secondo ordine). *Siano $f \in \mathcal{C}^2(\mathbb{R}^n, \mathbb{R})$ e debolmente concava e $g_i \in \mathcal{C}^1(\mathbb{R}^h, \mathbb{R})$ e debolmente convessa per ogni $i = 1, \dots, h$. Se esistono $\tilde{\lambda} = (\tilde{\lambda}_1, \dots, \tilde{\lambda}_h) \in \mathbb{R}^h$ ed $\tilde{\mathbf{x}}$ tali che $(\tilde{\mathbf{x}}, \tilde{\lambda})$ soddisfa le condizioni (k), (kk) e (kkk) e $\nabla f(\tilde{\mathbf{x}}) \neq 0$ allora $\tilde{\mathbf{x}}$ è soluzione del problema (DCP-Max).*

Esempio 4.1. Risolvere il problema

$$\begin{cases} \max_x - (x-2)^2 \\ t.c. \quad 1-x \leq 0. \end{cases}$$

SOLUZIONE: La condizione di regolarità è soddisfatta in quanto g è lineare. La Lagrangiana associata a f è

$$\mathcal{L}(x, \lambda) = -(x-2)^2 - \lambda(1-x)$$

Le condizioni (j), (jj) e (jjj) sono verificate dalle soluzioni del sistema

$$\begin{cases} -2(x-2) + \lambda = 0 \\ x-1 \leq 0 \\ \lambda(1-x) = 0 \\ \lambda \geq 0 \end{cases}$$

Procediamo dalla penultima condizione ottenendo $\lambda = 0$ oppure $x = 1$. Da qui la seguente casistica

$x = 1$: si ottiene $\lambda = -2$ e ciò contrasta con la condizione che $\lambda \geq 0$. Quindi $\tilde{x} = 1$ non è accettabile;

¹⁷ Stavolta la condizione di regolarità può essere espressa come $\nabla g_m(\tilde{\mathbf{x}}) \neq 0_n$ con $m = 1, 2, \dots, k$ se i vincoli attivi investono le prime k delle h componenti della funzione vincolare g .

¹⁸ La condizione di regolarità della funzione vincolare g del problema (D1CP-Max) è del tutto analoga a quella di (ECP-Max) e cioè che ∇g non si annulli. Stavolta però risulta indispensabile che $\nabla g(\tilde{\mathbf{x}}) = 0$ se e solo se $g(\tilde{\mathbf{x}}) = b$ perché se così non fosse ovvero se $g(\tilde{\mathbf{x}}) < b$ si avrebbe $\nabla f = 0$ e quindi potrebbe anche essere $\nabla g = 0$.

$\lambda = 0$: si ottiene $x = 2$.

Concludiamo affermando che il problema ammette soluzione $\tilde{x} = 2$ a cui corrisponde il valore ottimale $f(2) = 0$.

Esempio 4.2. Risolvere il seguente problema

$$\begin{cases} \max_{(x_1, x_2)} & -(x_1 - 4)^2 - (x_2 - 4)^2 \\ t.c. & x_1 + x_2 \leq 4 \\ & x_1 + 3x_2 \leq 9. \end{cases}$$

SOLUZIONE: La funzione obiettivo è concava, ed i vincoli sono entrambi lineari, quindi concavi. La Lagrangiana associata a f è

$$\mathcal{L}(x_1, x_2, \lambda_1, \lambda_2) = -(x_1 - 4)^2 - (x_2 - 4)^2 - \lambda_1(x_1 + x_2 - 4) - \lambda_2(x_1 + 3x_2 - 9)$$

e quindi le soluzioni del problema sono le stesse delle condizioni (k), (kk) e (kkk) ovvero le soluzioni del sistema

$$\begin{cases} -2(x_1 - 4) - \lambda_1 - \lambda_2 = 0 \\ -2(x_2 - 4) - \lambda_1 - 3\lambda_2 = 0 \\ x_1 + x_2 \leq 4, \quad \lambda_1 \geq 0, \quad e \quad \lambda_1(x_1 + x_2 - 4) = 0 \\ x_1 + 3x_2 \leq 9, \quad \lambda_2 \geq 0, \quad e \quad \lambda_2(x_1 + 3x_2 - 9) = 0. \end{cases}$$

Essendo difficile la ricerca immediata delle soluzioni proviamo i vari casi:

$x_1 + x_2 = 4$ e $x_1 + 3x_2 = 9$: In questo caso abbiamo $x_1 = \frac{3}{2}$ e $x_2 = \frac{5}{2}$. Quindi le prime due equazioni sono

$$5 - \lambda_1 - \lambda_2 = 0$$

$$3 - \lambda_1 - 3\lambda_2 = 0$$

vale a dire $\lambda_1 = 6$ e $\lambda_2 = -1$ e ciò viola la condizione $\lambda_2 \geq 0$.

$x_1 + x_2 = 4$ e $x_1 + 3x_2 < 9$: cosicché $\lambda_2 = 0$. Allora le prime due equazioni implicano che $x_1 = x_2 = 2$ e $\lambda_1 = 4$.

Tutte le condizioni sono soddisfatte, dunque $(x_1, x_2, \lambda_1, \lambda_2) = (2, 2, 4, 0)$ è una soluzione.

$x_1 + x_2 < 4$ e $x_1 + 3x_2 = 9$: da cui $\lambda_1 = 0$. Allora le prime due equazioni implicano che $x_1 = \frac{12}{5}$ e $x_2 = \frac{11}{5}$, violando $x_1 + x_2 < 4$.

$x_1 + x_2 < 4$ e $x_1 + 3x_2 < 9$: dunque $\lambda_1 = \lambda_2 = 0$. Allora le prime due equazioni implicano che $x_1 = x_2 = 4$, violando $x_1 + x_2 < 4$.

In conclusione $(\tilde{x}_1, \tilde{x}_2, \tilde{\lambda}_1, \tilde{\lambda}_2) = (2, 2, 4, 0)$ è l'unica soluzione per le condizioni (k), (kk) e (kkk) e quindi per il Teorema 14 l'unica soluzione del problema è $(\tilde{x}_1, \tilde{x}_2) = (2, 2)$.

4.1.1. *Le variabili sono positive.* Nei problemi di ottimizzazione di natura reale sono quasi sempre presenti vincoli di non negatività (ad es. indicano che le quantità sia reali positive) e ciò comporterà una rivisitazione della ricerca come segue.

In particolare prestiamo attenzione al ben noto problema di massimizzazione reale

$$(DCP\text{-}Max\text{-}reale) \quad \begin{cases} \max_{\mathbf{x}} & f(\mathbf{x}) \\ t.c. & g(\mathbf{x}) \leq \mathbf{b} \\ & \mathbf{x} \geq \mathbf{0} \end{cases}$$

a cui è associata la funzione lagrangiana $\mathcal{L} : D \times \mathbb{R}^h \rightarrow \mathbb{R}$ definita dalla legge

$$\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}) = f(\mathbf{x}) - \boldsymbol{\lambda} \cdot (g(\mathbf{x}) - \mathbf{b})$$

Teorema 16. Sia $\tilde{\mathbf{x}}$ soluzione locale del problema (DCP-Max-reale) e g tale che sia linearmente indipendenti i gradienti dei vincoli attivi in $\tilde{\mathbf{x}}$. Allora esiste $\tilde{\boldsymbol{\lambda}} \in \mathbb{R}^h$ che soddisfa contemporaneamente alle seguenti condizioni:

$$(kkt1) \quad \mathcal{L}_{\mathbf{x}}(\tilde{\mathbf{x}}, \tilde{\boldsymbol{\lambda}}) \leq \mathbf{0} \quad \wedge \quad \tilde{\mathbf{x}} \cdot \mathcal{L}_{\mathbf{x}}(\tilde{\mathbf{x}}, \tilde{\boldsymbol{\lambda}}) = 0$$

$$(kkt2) \quad \mathcal{L}_{\boldsymbol{\lambda}}(\tilde{\mathbf{x}}, \tilde{\boldsymbol{\lambda}}) \geq \mathbf{0} \quad \wedge \quad \tilde{\boldsymbol{\lambda}} \cdot \mathcal{L}_{\boldsymbol{\lambda}}(\tilde{\mathbf{x}}, \tilde{\boldsymbol{\lambda}}) = 0$$

$$(kkt3) \quad \tilde{\mathbf{x}}, \tilde{\boldsymbol{\lambda}} \geq \mathbf{0}$$

Le condizioni del primo ordine su riportate sono dette *condizioni di Karush Khun Tucker*.

Osservazione 8. La funzione Lagrangiana non è influenzata dalla presenza di vincoli di non negatività. La condizione di regolarità dei vincoli pur prevedendo che i gradienti dei vincoli attivi siano linearmente indipendenti necessita di considerare sia i vincoli tradizionali sia quelli di non negatività. Infine, le condizioni di Karush Khun Tucker differiscono solamente per la prima diventa ora una condizione con disuguaglianza abbinata ad una nuova condizione di complementarità. Le altre condizioni sono identiche alle precedenti, con l'ovvia condizione aggiuntiva della non negatività del vettore $\mathbf{x}$.

La condizione del secondo ordine è pressocchè simile a quella enunciata nel caso di vincoli di non positività nelle versioni del Teorema 14 e del Teorema 15 con la differenza che occorre considerare le condizioni (kkt1), (kkt2) e (kkt3) al posto delle rispettive (k), (kk) e (kkk).

Esempio 4.3. Risolvere il problema

$$\begin{cases} \max_{(x,y)} xy \\ t.c. \quad x + y \leq 6 \\ x \geq 0 \\ y \geq 0. \end{cases}$$

SOLUZIONE: Le funzioni vincolo sono concave, dunque le condizioni di Karush Kuhn Tucker sono necessarie. Inoltre, la funzione obiettivo è continua e l'insieme vincolare è compatto dunque, per il Teorema 3, il problema ammette soluzione globale. Le soluzioni del problema perciò sono le soluzioni delle condizioni del primo ordine stesse che segnalano i valori più alti della funzione¹⁹.

La Lagrangiana è

$$\mathcal{L}(x, y, \lambda_1, \lambda_2, \lambda_3) = xy - \lambda_1(x + y - 6) + \lambda_2 x + \lambda_3 y.$$

e le condizioni di Karush Kuhn Tucker sono

$$\begin{cases} y - \lambda_1 + \lambda_2 \leq 0 & e \quad x(y - \lambda_1 + \lambda_2) = 0 \\ x - \lambda_1 + \lambda_3 \leq 0 & e \quad y(x - \lambda_1 + \lambda_3) = 0 \\ x + y \leq 6, & e \quad \lambda_1(x + y - 6) = 0 \\ x \geq 0, & e \quad \lambda_2 x = 0 \\ y \geq 0, & e \quad \lambda_3 y = 0 \\ \lambda_1 \geq 0, \lambda_2 \geq 0, \lambda_3 \geq 0 \end{cases}$$

Si ha allora la seguente casistica

- $x > 0$ e $y > 0$: allora $\lambda_2 = \lambda_3 = 0$, quindi $\lambda_1 = x = y$ dalle prime due condizioni. Quindi $x = y = \lambda = 3$ dalla terza condizione. Questi valori soddisfano tutte le condizioni.
- $x = 0$ e $y > 0$: allora $\lambda_3 = 0$ dall'ultima condizione e $\lambda_1 = x = 0$ dalla seconda condizione. Ma ora dalla prima condizione $\lambda_2 = -y < 0$, contraddicendo $\lambda_2 \geq 0$.
- $x > 0$ e $y = 0$: allora $\lambda_2 = 0$, ed un analogo argomento porta ad una contraddizione.
- $x = y = 0$: allora $\lambda_1 = 0$ dal terzo insieme di condizioni, così $\lambda_2 = \lambda_3$ dalla prima e seconda condizione. Questi valori soddisfano tutte le condizioni.

Concludiamo che vi sono due soluzioni che soddisfano alle condizioni di Karush Kuhn Tucker, $(\tilde{x}, \tilde{y}, \tilde{\lambda}_1, \tilde{\lambda}_2, \tilde{\lambda}_3) = (3, 3, 3, 0, 0)$ e $(\hat{x}, \hat{y}, \hat{\lambda}_1, \hat{\lambda}_2, \hat{\lambda}_3) = (0, 0, 0, 0, 0)$. Dato che $f(3, 3) = 9 > 0 = f(0, 0)$ allora la soluzione del problema è $(\tilde{x}, \tilde{y}) = (3, 3)$.

4.2. Problema di massimo misto. Definiamo problema misto quel problema di massimo in cui sono presenti vincoli di uguaglianza, vincoli di non positività e vincoli di non negatività. La sua forma è

$$(DCP\text{-}Max\text{-}misto) \quad \begin{cases} \max_{\mathbf{x}} f(\mathbf{x}) \\ t.c. \quad g(\mathbf{x}) \leq \mathbf{b} \\ h(\mathbf{x}) = \mathbf{0} \\ \mathbf{x} \geq \mathbf{0} \end{cases}$$

La soluzione del problema (DCP-Max-misto) è da ricercare tra le combinazioni dei problemi di tipo (ECP-Max) e del tipo (DCP-Max-reale). Senza dilungarsi in enunciati che pur garantendo le condizioni necessarie e sufficienti della soluzione costituiscono un appesantimento della presente disquisizione presentiamo un esempio dal quale evincere la costruzione delle condizioni.

Esempio 4.4. Risolvere il problema

$$\begin{cases} \max_{(x,y)} f(x, y) = 2x - y \\ t.c. \quad x^2 + y = 2 \\ x + y \geq 0 \\ x \geq 0 \end{cases}$$

¹⁹Alternativamente, possiamo richiamare la sufficienza delle condizioni di Karush Kuhn Tucker: la funzione obiettivo è quasi-concava (ma non concava) nell'insieme vincolare (le sue curve di livello sono iperboli equilateri) e le funzioni vincolo sono lineari, quindi quasiconvesse, per cui se $\tilde{\mathbf{x}}$ risolve le condizioni (kkt1), (kkt2) e (kkt3) e $\nabla f(\tilde{\mathbf{x}}) \neq \mathbf{0}$, allora $\tilde{\mathbf{x}}$ è soluzione del problema. Osserviamo però che questo argomento lascia aperto lo stato dei punti $\tilde{\mathbf{x}}$ che risolvono le condizioni di Karush Kuhn Tucker e soddisfano $\nabla f(\tilde{\mathbf{x}}) = \mathbf{0}$; non si determina quando essi siano soluzione o non lo siano.

SOLUZIONE: Riscriviamo il problema nella forma (DCP-Max-misto) come

$$\begin{cases} \max_{(x,y)} f(x,y) = 2x - y \\ t.c. \quad x^2 + y = 2 \\ \quad -x - y \leq 0 \\ \quad x \geq 0 \end{cases}$$

La Lagrangiana associata è

$$\mathcal{L}(x, y, \lambda_1, \lambda_2) = 2x - y - \lambda_1(x^2 + y - 2) - \lambda_2(-x - y)$$

e le condizioni di ottimalità sono:

$$\begin{cases} 2 - 2\lambda_1 x + \lambda_2 \leq 0 & \wedge & x(2 - 2\lambda_1 x + \lambda_2) = 0 \\ -1 - \lambda_1 + \lambda_2 = 0 \\ x^2 + y = 2 \\ x + y \geq 0 & \wedge & \lambda_2(x + y) = 0 \\ x, \lambda_2 \geq 0. \end{cases}$$

Precisiamo che:

- la prima condizione è data da una disequazione dato che x è vincolata in segno e dalla condizione di complementarità su x ;
- la seconda condizione è data da un'equazione dato che y non è vincolata in segno;
- la terza condizione rappresenta il vincolo di uguaglianza assegnato;
- la quarta condizione è il rispetto del vincolo di disequazione, abbinato alla complementarità con il moltiplicatore del secondo vincolo, cioè λ_2 ;
- la quinta è la non negatività di x e del moltiplicatore λ_2 ²⁰.

Ricerchiamo la soluzione del sistema analizzando i due casi possibili su λ_2 .

Se $\lambda_2 = 0$ si ha subito dalla seconda $\lambda_1 = -1$, per cui la complementarità su x diventa $x(2 + 2x) = 0$, che fornisce $x = 0$ non accettabile in quanto la prima disuguaglianza non è verificata, oppure $x = -1$ che non è accettabile perché $x \geq 0$.

Pertanto deve essere $\lambda_2 > 0$ e quindi il secondo vincolo è attivo. Riscriviamo le condizioni:

$$\begin{cases} 2 - 2\lambda_1 x + \lambda_2 \leq 0 & \wedge & x(2 - 2\lambda_1 x + \lambda_2) = 0 \\ -1 - \lambda_1 + \lambda_2 = 0 \\ x^2 + y = 2 \\ x + y = 0 \\ x \geq 0, \lambda_2 > 0. \end{cases}$$

Dal secondo vincolo si ha $y = -x$ e quindi il primo diventa $x^2 - x - 2 = 0$, che ha per soluzioni $x = 2$ oppure $x = -1$ che scartiamo subito perché $x \geq 0$. Per $x = 2$ si ha $y = -2$ e per la complementarità sulle x le prime due condizioni diventano equazioni entrambe:

$$\begin{cases} 2 - 2\lambda_1 x + \lambda_2 = 0 \\ -1 - \lambda_1 + \lambda_2 = 0 \\ \lambda_2 > 0 \end{cases}$$

che porta a $\lambda_1 = 1$, $\lambda_2 = 2$.

Quindi in conclusione abbiamo soltanto il punto $(\tilde{x}, \tilde{y}) = (2, -2)$ con $\lambda_1 = 1$ e $\lambda_2 = 2$ che risulta di massimo globale vincolato dato che l'insieme dei vincoli è un compatto.

4.3. Problema di minimo. Ricordando la relazione (2), definire i problemi di minimo di tipo (D1CP-min), (DCP-min), (DCP-min-reale) e (DCP-min-misto) è equivalente a trattare i rispettivi problemi di massimo (D1CP-Max), (DCP-Max), (DCP-Max-reale) e (DCP-Max-misto), dove la funzione obiettivo è definita dalla legge $h(\mathbf{x}) = -f(\mathbf{x})$ per ogni $\mathbf{x} \in \mathbb{R}^n$.

Esempio 4.5. Il problema

$$\begin{cases} \min_{(x,y)} f(x,y) = x + y \\ t.c. \quad x^2 + y^2 \geq 2 \\ \quad x + 1 \geq 0 \\ \quad y \geq 0 \end{cases}$$

²⁰ Il moltiplicatore λ_1 è libero poiché il vincolo corrispondente è di equazione.

si può allora trasformare nel problema di massimo in forma (DCP-Max)

$$\begin{cases} \max_{(x,y)} h(x,y) = \max_{(x,y)} -x - y \\ t.c. \quad -x^2 - y^2 \leq -2 \\ \quad -x \leq 1 \\ \quad -y \leq 0, \end{cases}$$

che può essere affrontato con le condizioni di Karush Kuhn Tucker. Si verifica che la soluzione è $\tilde{\mathbf{x}} = (-1, 1)$ a cui corrisponde il valore ottimale $f(\tilde{\mathbf{x}}) = 0$.

APPENDICE

4.4. A proposito delle matrici quadrate. Sia $A = (a_{ij}) \in \mathcal{M}_n$.

4.4.1. *Sottomatrici principali.*

Definizione 4.1. Definiamo *sottomatrice principale* di A di ordine k con $k \leq n$ la matrice A_k ottenuta eliminando $n - k$ colonne ed $n - k$ righe. Il determinante della sottomatrice principale è detto *minore principale*.

Definizione 4.2. La sottomatrice principale A_k ottenuta eliminando le ultime $n - k$ righe e $n - k$ colonne è detta *sottomatrice principale di nord-ovest* ed il suo determinante è detto *minore principale di nord-ovest*.

4.4.2. *Forme quadratiche.* Sia $A = (a_{ij}) \in \mathcal{M}_n$ simmetrica.

Definizione 4.3. La funzione $Q : \mathbb{R}^n \rightarrow \mathbb{R}$ definita dalla legge $\mathbf{x} \in \mathbb{R}^n \mapsto \mathbf{x}^T A \mathbf{x} \in \mathbb{R}$ è detta *forma quadratica* associata ad A .

Definizione 4.4. La forma quadratica $Q(\mathbf{x})$ è detta:

- *semidefinita negativa* se $Q(\mathbf{x}) \leq 0$ per ogni $\mathbf{x} \in \mathbb{R}^n$;
- *definita negativa* se $Q(\mathbf{x}) < 0$ per ogni $\mathbf{x} \in \mathbb{R}^n$ e $\mathbf{x} \neq \mathbf{0}$;
- *semidefinita positiva* se $Q(\mathbf{x}) \geq 0$ per ogni $\mathbf{x} \in \mathbb{R}^n$;
- *definita positiva* se $Q(\mathbf{x}) > 0$ per ogni $\mathbf{x} \in \mathbb{R}^n$ e $\mathbf{x} \neq \mathbf{0}$;
- *indefinita* se cambia di segno al variare di $\mathbf{x} \in \mathbb{R}^n$.

Teorema 17 (di Sylvester-Jacobi). *La forma quadratica $Q(\mathbf{x})$ risulta essere:*

- *definita positiva se e solo se* $\det A_1 > 0, \det A_2 > 0, \dots, \det A_n > 0$;
- *definita negativa se e solo se* $\det A_1 < 0, \det A_2 > 0, \dots, (-1)^n \det A_n > 0$;
- *semidefinita positiva se e solo se esiste almeno A_i tale che* $\det A_i = 0$ *e* $\det A_1 > 0, \det A_2 > 0, \dots, \det A_{i-1} > 0, \det A_{i+1} > 0, \dots, \det A_n > 0$;
- *semidefinita negativa se e solo se esiste almeno A_i tale che* $\det A_i = 0$ *e* $\det A_1 < 0, \det A_2 > 0, \dots, (-1)^n \det A_n > 0$;
- *indefinita se e solo se esiste almeno A_i tale che* $\det A_i \neq 0$ *ma gli altri* A_j *non rispettano alcuna delle precedenti regole.*

dove A_i con $i = 1, 2, \dots, n$ è la sottomatrice principale di nord-ovest di ordine i della matrice A della forma quadratica Q .

4.5. A proposito delle funzioni scalari reali. Siano $f : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ e $\mathbf{x}_0 \in \overset{\circ}{D}$.

4.5.1. *Limiti.*

Definizione 4.5. f ammette:

- limite $l \in \mathbb{R}$ per $\mathbf{x} \rightarrow \mathbf{x}_0$ e scriviamo $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} f(\mathbf{x}) = l$ se

$$\forall \varepsilon > 0 \exists \delta > 0 : \forall \mathbf{x} \in D \text{ con } \|\mathbf{x} - \mathbf{x}_0\| < \delta \Rightarrow \|f(\mathbf{x}) - l\| < \varepsilon$$

- limite ∞ per $\mathbf{x} \rightarrow \mathbf{x}_0$ e scriviamo $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} f(\mathbf{x}) = \infty$ se

$$\forall M > 0 \exists \delta > 0 : \forall \mathbf{x} \in D \text{ con } \|\mathbf{x} - \mathbf{x}_0\| < \delta \Rightarrow \|f(\mathbf{x})\| > M$$

- limite $l \in \mathbb{R}$ per $\mathbf{x} \rightarrow \infty$ e scriviamo $\lim_{\mathbf{x} \rightarrow \infty} f(\mathbf{x}) = l$ se

$$\forall \varepsilon > 0 \exists N > 0 : \forall \mathbf{x} \in D \text{ con } \|\mathbf{x}\| > N \Rightarrow \|f(\mathbf{x}) - l\| < \varepsilon$$

- limite ∞ per $\mathbf{x} \rightarrow \infty$ e scriviamo $\lim_{\mathbf{x} \rightarrow \infty} f(\mathbf{x}) = \infty$ se

$$\forall M > 0 \exists N > 0 : \forall \mathbf{x} \in D \text{ con } \|\mathbf{x}\| > N \Rightarrow \|f(\mathbf{x})\| > M$$

Teorema 18 (Condizione sufficiente per $n = 2$). Posto $\mathbf{x}_0 = (x_0, y_0)$ e

$$x = x_0 + r \cos \theta, \quad y = y_0 + r \sin \theta$$

se esiste ed è finito $l \in \mathbb{R}$ ed esiste la funzione $g(r)$ tale che in un intorno di (x_0, y_0) si abbia

$$|f(x_0 + r \cos \theta, y_0 + r \sin \theta) - l| \leq g(r) \quad \text{con} \quad \lim_{r \rightarrow 0^+} g(r) = 0$$

allora si ha

$$\lim_{(x,y) \rightarrow (x_0,y_0)} f(x,y) = l$$

4.5.2. Gradiente e Hessiano.

Definizione 4.6. Diciamo che f è *derivabile parzialmente* in $\mathbf{x}_0$ rispetto alla variabile x_i se esiste ed è finito il seguente

$$\lim_{\Delta x \rightarrow 0} \frac{f(\mathbf{x}_0 + \Delta x \mathbf{e}_i) - f(\mathbf{x}_0)}{\Delta x}.$$

Il valore del limite è detto *derivata parziale* di f in $\mathbf{x}_0$ rispetto alla variabile x_i ed indicato con $f_{x_i}(\mathbf{x}_0)$.

Diciamo che f è *derivabile* in $\mathbf{x}_0$ se esistono e sono finite le quantità $f_{x_i}(\mathbf{x}_0)$ per ogni $i = 1, 2, \dots, n$.

La funzione $f_{x_i} : D \rightarrow \mathbb{R}$ definita dalla legge $\mathbf{x}_0 \mapsto f_{x_i}(\mathbf{x}_0)$ è detta *funzione derivata parziale* (prima) di f rispetto a x_i .

Sia f derivabile in $\overset{\circ}{D}$.

Definizione 4.7. Definiamo *gradiente* di f in $\mathbf{x}_0$ il vettore riga di $\mathbb{R}^n$

$$\nabla f(\mathbf{x}_0) := (f_{x_1}(\mathbf{x}_0) \quad f_{x_2}(\mathbf{x}_0) \quad \dots \quad f_{x_n}(\mathbf{x}_0))$$

La funzione $\nabla f : D \rightarrow \mathbb{R}^n$ definita dalla legge $\mathbf{x} \mapsto \nabla f(\mathbf{x})$ per ogni $\mathbf{x} \in D$ è detta *funzione gradiente* di f .

Definizione 4.8. Se f_{x_i} è derivabile *definiamo derivata parziale seconda* di f rispetto a x_i ed a x_j la derivata parziale prima di f_{x_i} rispetto a x_j

$$f_{x_i x_j}(\mathbf{x}) := f_{x_j}(f_{x_i}(\mathbf{x}))$$

Definizione 4.9. Definiamo *Hessiano* di f in $\mathbf{x}_0$ la matrice vettoriale di $\mathbb{R}^n \times \mathbb{R}^n = \mathbb{R}^{2n}$

$$\mathbf{H}_f(\mathbf{x}_0) := \begin{pmatrix} f_{x_1 x_1}(\mathbf{x}_0) & f_{x_1 x_2}(\mathbf{x}_0) & \dots & f_{x_1 x_n}(\mathbf{x}_0) \\ f_{x_2 x_1}(\mathbf{x}_0) & f_{x_2 x_2}(\mathbf{x}_0) & \dots & f_{x_2 x_n}(\mathbf{x}_0) \\ \vdots & \vdots & \ddots & \vdots \\ f_{x_n x_1}(\mathbf{x}_0) & f_{x_n x_2}(\mathbf{x}_0) & \dots & f_{x_n x_n}(\mathbf{x}_0) \end{pmatrix}$$

Poniamo

$$\mathcal{C}^n(\mathbb{R}^n, \mathbb{R}) := \{f : \mathbb{R}^n \rightarrow \mathbb{R} : f \text{ e } \underbrace{f_{x_j \dots x_i}}_{n\text{-volte}} \text{ continue}\}$$

Teorema 19 (di Schwartz). Se $f \in \mathcal{C}^2(\mathbb{R}^n, \mathbb{R})$ e $\mathbf{x}_0 \in \overset{\circ}{D}$. Allora

$$f_{x_i x_j}(\mathbf{x}_0) = f_{x_j x_i}(\mathbf{x}_0) \quad \forall i, j = 1, 2, \dots, n$$

Corollario 20. Se valgono le ipotesi del Teorema 19 allora $\mathbf{H}_f(\mathbf{x}_0)$ è una matrice simmetrica.

4.5.3. Funzioni implicite.

Teorema 21 (del Dini). Siano $f \in \mathcal{C}^1(\mathbb{R}^n, \mathbb{R})$, $c \in \mathbb{R}$ e $\mathbf{x}_0 \in \overset{\circ}{D}$ tale che $f(\mathbf{x}_0) = c$. Allora esiste la funzione $\Phi_i : \mathbb{R}^{n-1} \rightarrow \mathbb{R}$ definita dalla legge $(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n) \mapsto x_i$ in un intorno B di $(x_{0_1}, \dots, x_{0_{i-1}}, x_{0_{i+1}}, \dots, x_{0_n})$ tale che:

- $f(x_1, \dots, x_{i-1}, \Phi_i(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n), x_{i+1}, \dots, x_n) = c$ per ogni $(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)$;
- $x_{0_i} = \Phi_i(x_{0_1}, \dots, x_{0_{i-1}}, x_{0_{i+1}}, \dots, x_{0_n})$;
- Φ_i è differenziabile in $(x_{0_1}, \dots, x_{0_{i-1}}, x_{0_{i+1}}, \dots, x_{0_n})$ e

$$\Phi_{i x_j}(x_{0_1}, \dots, x_{0_{i-1}}, x_{0_{i+1}}, \dots, x_{0_n}) = -\frac{f_{x_j}(\mathbf{x}_0)}{f_{x_i}(\mathbf{x}_0)}$$

4.5.4. Funzioni concave e convesse.

Definizione 4.10. La funzione f è *convessa* (risp. *strettamente convessa*) in D se

$$\forall \mathbf{x}, \mathbf{y} \in D \quad \forall t \in [0, 1] \quad f((1-t)\mathbf{x} + t\mathbf{y}) \leq (1-t)f(\mathbf{x}) + tf(\mathbf{y}) \quad (\text{risp. } <)$$

La funzione f è *concava* (risp. *strettamente concava*) in D se

$$\forall \mathbf{x}, \mathbf{y} \in D \quad \forall t \in [0, 1] \quad f((1-t)\mathbf{x} + t\mathbf{y}) \geq (1-t)f(\mathbf{x}) + tf(\mathbf{y}) \quad (\text{risp. } <)$$

Proposizione 22. f è *convava* (risp. *strettamente concava*) in D se $-f$ è *convessa* (risp. *strettamente convessa*) in D .

Teorema 23. Sia $f \in \mathcal{C}^2(\mathbb{R}^2, \mathbb{R})$. Allora

CS: se $\mathbf{H}_f(\mathbf{x})$ è *definita positiva* (risp. *negativa*) per ogni $\mathbf{x} \in \overset{\circ}{D}$ allora f è *strettamente concava* (risp. *strettamente convessa*);

CNeS: f è *debolmente convessa* (risp. *debolmente concava*) se e solo se $\mathbf{H}_f(\mathbf{x})$ è *semidefinita positiva* (risp. *semidefinita negativa*) per ogni $\mathbf{x} \in \overset{\circ}{D}$

4.5.5. Differenziabilità e Polinomio di Taylor.

Definizione 4.11. La funzione f è *differenziabile* in $\mathbf{x}_0$ se

$$(31) \quad \lim_{\mathbf{x} \rightarrow \mathbf{x}_0} \frac{f(\mathbf{x}) - f(\mathbf{x}_0) - \nabla f(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0)}{\|\mathbf{x} - \mathbf{x}_0\|} = 0$$

Definizione 4.12. Data $f : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ tale che $f \in \mathcal{C}^2$ e $\mathbf{x}_0 \in D$ definiamo *polinomio di Taylor* di f in $\mathbf{x}_0$ di ordine 2, la funzione

$$T f_{2, \mathbf{x}_0}(\mathbf{x}) = f(\mathbf{x}) + \nabla f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0) + \frac{1}{2}(\mathbf{x} - \mathbf{x}_0)^T \cdot \mathbf{H}_f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0)$$