MASSIMI E MINIMI
LIBERI E VINCOLATI CON 2 (O PIÙ) VARIABILI INDIPENDENTI[1]

DISPENSE SINTETICHE PER STUDENTI DI MICROECONOMIA

Bruno Bosco

Avvertenza. Questi appunti hanno esclusivamente la funzione di sintetizzare, quasi come un pro memoria, le tecniche impiegate per la ricerca di massimi e minimi, liberi o vincolati, in funzioni – generalmente di utilità o di produzione – comunemente utilizzate in microeconomia. Ovviamente questi appunti non possono lontanamente sostituire i testi di analisi o, quando richiamati, di algebra di vettori e matrici, cui costantemente si rimanda. Per Analisi I e II continuo a pensare che resti all'avanguardia il manuale di Luigi Amerio, *Analisi Matematica*, Utet, vari anni. In particolare, ai nostri fini, il Volume I nella nuova edizione ampliata (1990) saltando il Cap. 7 e il Cap. 9. E' scritto in forma piana, ottimo italiano e usando una simbologia intuitiva. Oltre tutto non è stampato su carta lucida o patinata, come usa oggi. Utilissimo resta Boris Pavlovič Demidovič, *Esercizi e Problemi di Analisi Matematica*, Editori Riuniti, 1975, traduzione dall'edizione originale russa del 1968. Per algebra lineare la scelta è più ampia.

1. LA RICERCA DI UN MASSIMO/MINIMO NON VINCOLATO

1.1 Funzioni in *R e Teoremi di* Weierstrass *e di* Fermat

Iniziamo ricordando la definizione di massimo e di minimo per una funzione in una sola variabile. Data una funzione $f: D \subseteq R \to R$ si dice massimo (minimo) di tale funzione il più grande (piccolo) dei valori che essa assume nel dominio *D*. In particolare si dice che x_0 è un punto di massimo (minimo) locale (o relativo) per la funzione *f* se esiste in *D* un intorno di x_0 tale che:

$$\left.\begin{array}{l} f(x) \le f(x_0) \ \text{(massimo)} \\ f(x) \ge f(x_0) \ \text{(minimo)} \end{array}\right\} \forall x \text{ appartenenti all'intorno considerato}$$

. Il grafico seguente fornisce un esempio.

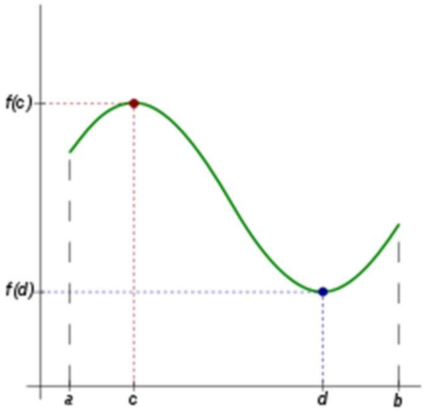


Fig. 1 *Esempio di funzione continua nell'intervallo [a, b] che ammette un massimo e un minimo, rispettivamente in c e in d*

[1] Per esercizi di analisi vai a http://www.slader.com/textbook/9781285057095-calculus-10th-edition/512/exercises/35/#

Per funzioni continue definite in un intervallo chiuso e limitato (compatto), l'esistenza, di minimi e massimi è assicurata dal **Teorema di Weierstrass** in base al quale:

> Sia $[a, b] \subset R$ un intervallo non vuoto, chiuso e limitato e sia
>
> $$f : [a\,,\, b] \to R$$
>
> una funzione continua. Allora $f(x)$ **ammette (almeno) un punto di massimo assoluto e un punto di minimo assoluto** nell'intervallo $[a, b]$.

La dimostrazione del Teorema (insiemi compatti) esula dalle nostre finalità. Soffermiamoci soltanto sulla **necessità** delle ipotesi impiegate.

a) Uso della nozione di compattezza

Poiché $f : [a\,,\, b] \to R$ è una funzione continua, essa trasforma insiemi compatti in insiemi compatti. Dato che $[a, b]$ è un intervallo chiuso e limitato, per il teorema di Heine-Borel esso è un compatto; quindi anche la sua immagine mediante f (codominio) sarà un compatto di R, e dunque f è provvista di massimo e minimo, ovvero assume un valore massimo e uno minimo in essa. Le loro contro immagini in $[a, b]$ sono rispettivamente il punto di massimo e di minimo assoluti.

b) Necessità delle ipotesi?

L'ipotesi di (1) continuità della f su un intervallo, (2) di intervallo chiuso e (3) limitato è sufficiente ma **non necessaria**. Il fatto che una funzione non soddisfi le ipotesi del teorema di Weierstrass, non implica che non esistano massimo o minimo della funzione, ma semplicemente che la loro esistenza non è garantita. Inoltre, come si vedrà nei controesempi, queste sono le ipotesi più larghe possibili per cui vale l'enunciato stesso. Il teorema non vale se cade anche solo una delle tre ipotesi. Analizziamole una per una.

- Caso di f non continua: Si consideri $f : [-1\,,\, 1] \to R$ tale che $f(x) = |\, x \,|$ per $x \neq 0$ e $f(0) = 1$, che non è continua in $x = 0$ (vedi fig. 2). Il teorema non è applicabile; infatti la funzione non ha un minimo ma solo un estremo inferiore uguale a 0.
- L'intervallo non è chiuso: Si consideri. f: $[0\,,\, 1) \to R\,,\, x \to x$. Essa è continua nell'intervallo limitato $[0, 1)$, che però non è chiuso. Il teorema non è applicabile, infatti non ha un massimo ma solo un estremo superiore uguale a 1.
- L'intervallo non è limitato: Si consideri f: $[\,0\,,\, +\infty\,) \to R\,,\, x \to Arctang\,(x)$ con dominio $(-\infty, \infty)$. Essa è continua in $[0, +\infty)$ tuttavia l'intervallo è illimitato. Il teorema non è applicabile; infatti la f non ha un massimo ma solo un estremo superiore uguale a $\pi/2$ per $x \mapsto +\infty$.

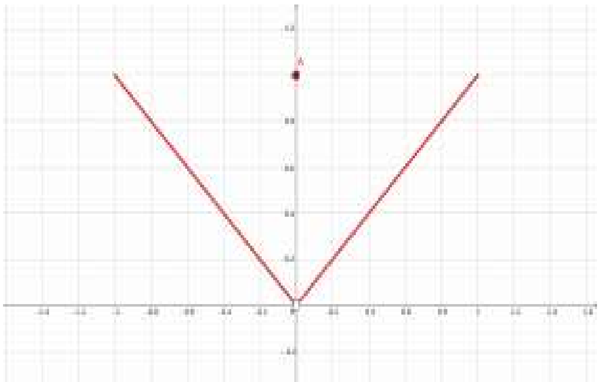


Fig. 2 Controesempio n°1. La funzione f (x) = | x | nell'intervallo [-1,1] ridefinita in x = 0 non è continua. Il teorema di Weierstrass non è quindi valido.

Noi faremo sempre riferimento a funzioni che soddisfano le condizioni del Teorema di Weierstrass.

Come trovare massimi e minimi?

Il **Teorema di Weierstrass** ci dice che la funzione continua $f : [a , b] \rightarrow R$, dove [a, b] $\subset R$ è un intervallo chiuso e limitato non vuoto ammette massimo e minimo. Ma non ci dice, di per sé, come trovarli. **Se la funzione è anche derivabile** ovunque, il **Teorema di Fermat** ci dà la condizione necessaria del primo ordine per un massimo o per un minimo, anche se esso non garantisce di trovare un massimo/minimo **assoluto** se il dominio della *f* è un insieme aperto. Detto che i punti di massimo e minimo locale si dicono punti di estremo locale, o estremali, e analogamente le loro immagini si dicono valori estremi e che il termine locale (a volte sostituito da relativo) contrapposto a globale (o assoluto), si riferisce ad estremi che sono tali su tutto il dominio, il Teorema di Fermat afferma che

> Sia $f:(a , b) \rightarrow R$ una funzione e si supponga che $x_o \in$ (a , b) sia un punto di *estremo* locale di *f.* Se *f* è derivabile nel punto x_o, allora $df(x_0) / dx = 0$.

Il teorema fornisce quindi un metodo per la ricerca dei punti di estremo (ovvero di massimo o minimo) di una funzione **derivabile** definita su un insieme aperto e mostra che ogni punto di estremo locale è un **punto stazionario** della funzione (cioè la derivata prima della funzione si annulla in quel punto). In tal modo, utilizzando il teorema di Fermat, il problema della ricerca dei punti estremi di una funzione è ridotto alla risoluzione di un'equazione legata al seguente problema:

per quale/i valore/i di x vale $df(x)/dx = 0$[2]?

Purtroppo il Teorema di Fermat fornisce solo una condizione **necessaria** per il valore degli estremi della funzione: è vero che tutti i punti estremi sono stazionari, ma esistono anche alcuni punti stazionari (ovvero di annullamento della derivata prima) che non sono punti estremi, ma possono essere punti di **flesso** (o, nel caso di una funzione a più variabili, punti di sella o punti di diversa natura). Per valutare se un punto stazionario è un valore estremo e per distinguere se tale punto è di massimo o di minimo, sarà necessario, in genere, analizzare **anche la derivata seconda** della funzione (se esiste). Procediamo con ordine.

[2] In quanto segue la derivata prima, anche parziale, potrà essere segnata come $df(x)/dx, f', f_x$ (e anche quelle di ordine superiore al primo)

Abbiamo detto che **nei punti di massimo o minimo locali di una funzione derivabile che siano <u>interni</u> al dominio la derivata prima è nulla.** Essendo questa una condizione necessaria ma non sufficiente, per trovare i massimi e minimi dobbiamo seguire una strategia più articolata. Per procedere prendiamo in considerazione:

1) tutti i punti locali dove $f'(x_0) = 0$; Confrontandoli troviamo, se esiste, il **massimo** e il **minimo assoluti.**

Sembra un sistema semplice e immediato. Però: che succede se non possiamo definire f' in qualche punto (ovvero se non ipotizziamo le condizioni di Fermat e ammettiamo che la $f:[a\,,\,b] \rightarrow R$ possa non essere sempre derivabile ? I casi principali (ma non tutti quelli possibili) sono trattati nei puti 2 e 3:

2) eventuali punti angolosi o punti a questi assimilabili (ad esempio, cuspidi) e flessi a tangente verticale;
3) estremi dell'intervallo.

I punti 2 e 3 li trattiamo congiuntamente alla fine del paragrafo. Adesso approfondiamo il punto 1.
Punto 1). Se $f(x)$ è derivabile, la condizione **necessaria** affinché un punto sia di massimo o di minimo locale deve essere: $f'(x) = 0$. Dunque si tratterà di risolvere tale equazione per determinare quel/ei valore/i della x in corrispondenza del/i quale/i si realizza/no il/i possibile/i punto/i di massimo o di minimo locale. I punti in cui si annulla la derivata prima si dicono **punti stazionari o punti critici**. Tuttavia, i valori in questione sono solo *probabili* punti di massimo o minimo locale, in quanto potrebbero anche essere punti di flesso. Per sapere se questi sono punti di massimo o di minimo per la funzione f si può procedere in **2 modi** (apparentemente distinti). Li riassumiamo nella tabella seguente.

<table>
<tr><td>I^ MODO</td><td>Si studia il segno della derivata prima, studiando la disequazione

$$f'(x) > 0$$
nell'intorno del punto, chiamiamolo x_o, in cui la derivata si annulla. Il calcolo della derivata prima serve per determinare gli intervalli in cui la funzione cresce o decresce, facendoci comprendere se i punti trovati sono di massimo o di minimo. In pratica:

I punti di massimo sono quelli tali che $f'(x_o) = 0$ mentre $f'(x) > 0$ a sinistra di x_o (ovvero per ogni x minore di x_o) e $f'(x) < 0$ a destra di x_o (ovvero per ogni x maggiore di x_o)
I punti di minimo sono quelli tali che $f'(x_o) = 0$ con $f'(x) < 0$ a sinistra di x_o e $f'(x) > 0$ a destra di x_o
Se invece la derivata nell'intorno di tali punti non cambia di segno, questi punti non sono né di massimo né di minimo. Es. $f(x)=x^3$ con $x \in [-1, 1]$</td></tr>
<tr><td>II^ MODO</td><td>Si sostituiscono le ascisse dei punti x_o (vedi sopra) nella derivata seconda e si guarda il segno che questa assume.

Il calcolo della derivata seconda serve per determinare gli intervalli in cui la curva è concava o convessa: se $f''(x_o) > 0$ allora la funzione è convessa e perciò il punto è di minimo;

se $f''(x_o) < 0$ allora la funzione è concava e perciò il punto è di massimo;

se invece $f''(x_o) = 0$ allora non possiamo concludere nulla (come sopra).</td></tr>
</table>

Ad esempio, sia *f: (- ∞, ∞)* → *R* la funzione continua (parabola) $f(x) = 2x^2 + x$. Poiché $f'(x) = 0$ per $x_o = -1/4 = -0.25$ la figura seguente rappresenta il grafico della derivata prima su tutto l'intervallo che contiene –1/4: per valori minori la funzione derivata è negativa; per valori superiori la funzione derivata è positiva. In $x = -0.25$ la funzione derivata si annulla. Quindi in $x_o = -1/4$ la $f(x)$ consegue un minimo (alla sua sinistra la funzione *y* "sale" e alla sua destra la funzione *y* "scende").

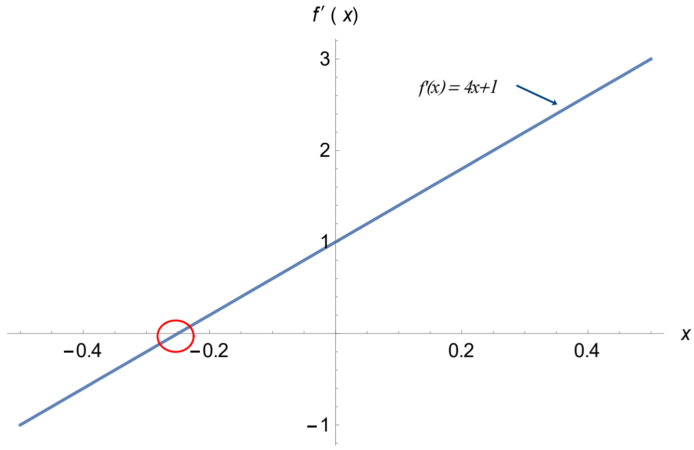
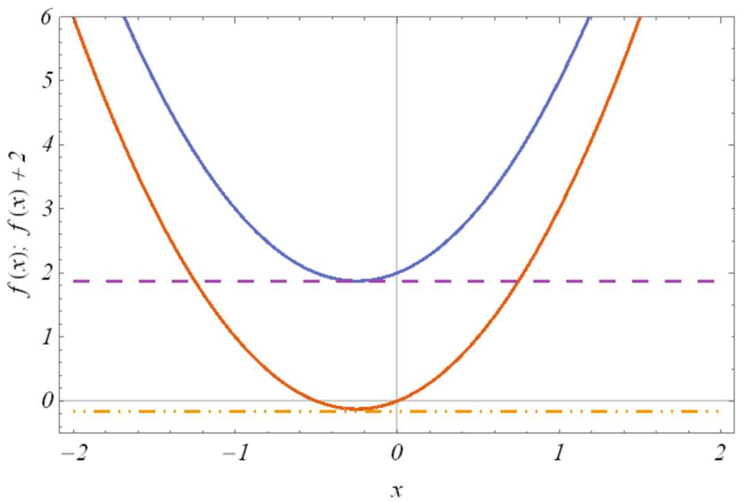


Grafico di df(x)/dx = 4x + 1. Si annulla in x = –1/4 = – 0.25

La derivata seconda è sempre 4 > 0, per qualsiasi *x* nel dominio. Il secondo metodo dà gli stessi risultati del primo. Da ricordare:

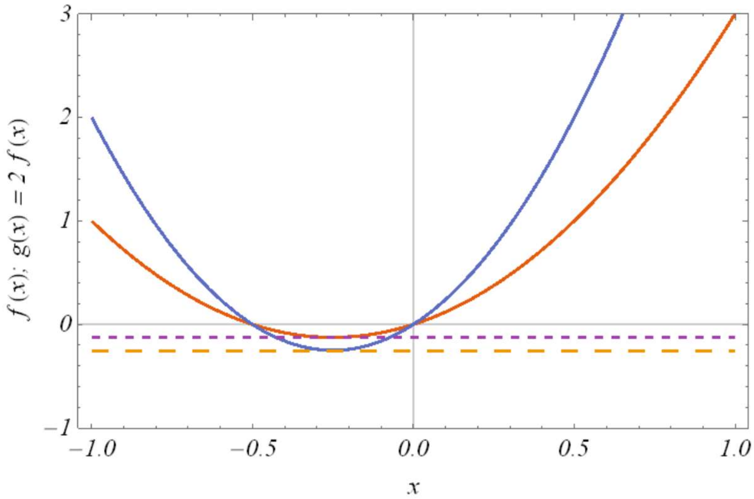
1) Se si somma una costante *c* qualsiasi alla funzione $f(x)$ allora la funzione $f(x) + c$ ha gli stessi punti di massimo e minimo assoluti negli stessi punti della *f(x)* ma (ovviamente) non avrà gli stessi valori di massimo e di minimo della *f(x)* con $c = 0$.

Ad esempio la funzione $f(x) = 2x^2 + x$ conseguiva un minimo con $x_o = -1/4$ e aveva valor minimo della *f* pari a -1/8. La funzione $2+f(x) =2 + 2x^2 + x$ consegue un minimo con $x_o = -1/4$ come la *f(x)* ma ha valor minimo pari a 15/8. Vedi figura.



2) Se si moltiplica la $f(x)$ per una costante positiva *c* allora la funzione $cf(x)$ ha gli stessi punti di massimo e minimo assoluti negli stessi punti della *f(x)* ma (ovviamente) non gli stessi valori di massimo e di minimo.

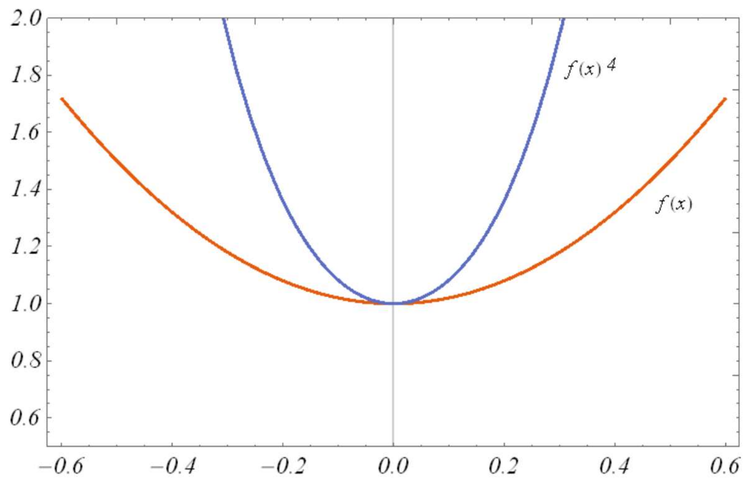
Ad esempio la funzione $f(x) = 2x^2 + x$ conseguiva un minimo con $x_o = -1/4$ e aveva valor minimo pari a -1/8. La funzione $2f(x) = 4x^2 + 2x$ consegue sempre un minimo con $x_o = -1/4$ ma ha valor minimo pari a -1/4. Vedi figura.



3) Se $f(x) > 0$ per qualsiasi x nell'intervallo e consideriamo $1/f(x)$, allora la funzione $f(x)$ ha il massimo assoluto dove $1/f(x)$ ha il minimo assoluto e viceversa.

Ad esempio la funzione $f(x) = 2ln(3x+1)$ conseguiva un minimo con $x_o = \infty$ mentre la funzione $f^{-1}(x) = 1/3\ (e^{y/2} - 1)$ consegue un minimo con $y = -\infty$.

4) Se e solo se $f(x) > 0$ per ogni x si ha che $f(x)$ e $[f(x)]^n$ hanno massimi e i minimi assoluti negli stessi punti. La figura seguente si riferisce a $f(x) = 1 + 2x^2 >> 0$ per ogni x della retta reale e $n = 4$.



Esempio 1

Trovare nell'intervallo [0, 5] il massimo e il minimo assoluto (se esistono) della funzione:
$f(x) = \left(\sqrt{x+1}\right)\left(\ln(x+1)\right)$ il cui CE è $x \in (-1, +\infty)$. Il grafico della funzione è riprodotto di seguito.

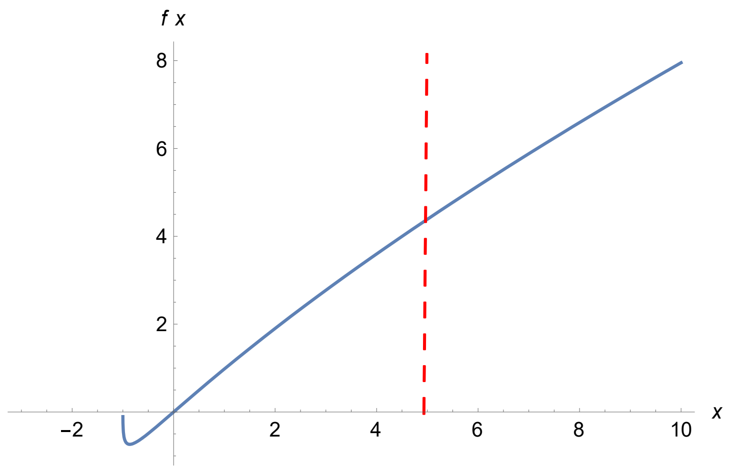


Fig. 3 *Grafico della funzione* $f(x) = \left(\sqrt{x+1}\right)\left(\ln(x+1)\right)$. *Essa è priva di massimo assoluto*

a) Derivata prima:

$$\frac{df(x)}{dx} = f'(x) = \frac{1}{\sqrt{x+1}} + \frac{\ln(x+1)}{2\sqrt{x+1}} = \frac{2+\ln(x+1)}{2\sqrt{x+1}}$$

b) Studio del segno della derivata prima (I^ metodo)

Tale derivata è sempre positiva in [0, 5]. Infatti $x + 1 > 0$ per $x > -1$ e $ln(x+1) > 0$ per ogni x positivo. Allora poiché la funzione è derivabile in tutto l'intervallo deduciamo che il massimo e il minimo nell'intervallo si trovano agli estremi dell'intervallo.

c) Massimi e minimi nell'intervallo

Calcoliamo $f(0) = 0$ e $f(5) = \ln(6)\sqrt{6}$. Quindi il minimo assoluto è 0 e il massimo assoluto è $\ln(6)\sqrt{6}$ in accordo col fatto che la derivata è positiva in [0,5].

d) Minimo assoluto su tutto il C.E.

$$f'(x) = \frac{2+\ln(x+1)}{2\sqrt{x+1}} = 0 \ \text{ per } x = e^{-2} - 1 = -0.8646$$

Allora valutando

$$f''(x) = -\frac{\ln(x+1)}{4(x+1)^{\frac{3}{2}}} \text{ nel punto } x = -0.8646 \text{ otteniamo } -\frac{\ln(-0.8646+1)}{4(-0.8646+1)^{\frac{3}{2}}} = -\frac{\ln(0.1354)}{4(0.1354)^{\frac{3}{2}}} > 0 .$$

Con $x = -0.8646$ la derivata seconda è positiva[3]. Quindi $x = -0.8646$ è un minimo assoluto (vedere grafico) mentre non esiste un massimo assoluto per valori finiti di x.

[3] Equivalentemente, se piace di più, nel punto stazionario la derivata seconda vale $e^3/2 > 0$.

1.2 Casi di non derivabilità

Torniamo al Punto 2 lasciato in sospeso. Si tratta di punti nel dominio in cui non è definita la derivata prima della funzione. **Siamo quindi fuori dall'ipotesi di derivabilità su cui poggia il Teorema di Fermat**. Il primo caso è quello dei **punti angolosi** già intravisto nella Fig. 2 relativa al grafico di $f(x) = |x|$. In generale una funzione presenta un punto angoloso nell'intorno di un punto del dominio se in tale intorno i limiti destro e sinistro del rapporto incrementale esistono e sono **finiti ma non coincidono**. Questa è una caratteristica delle funzioni definite nei valori assoluti ma anche di quelle definite a tratti spesso usate in economia (e per questo motivo che le menzioniamo). Se invece i limiti destro e sinistro del rapporto incrementale sono **infiniti**, e in particolare mostrano **segno opposto**, allora in corrispondenza del punto del dominio in cui ciò accade abbiamo una **cuspide**. Nello zero presenta una cuspide la funzione $f(x) = \sqrt{|x|}$ così come altre funzioni di radici di indice pari. Una cuspide consiste quindi in un punto in cui la funzione cresce con pendenza infinita in uno dei due intorni (sinistro o destro) del punto e decresce con pendenza infinita nell'altro. Il primo grafico seguente illustra il caso di $f(x) = \sqrt[4]{|x|}$

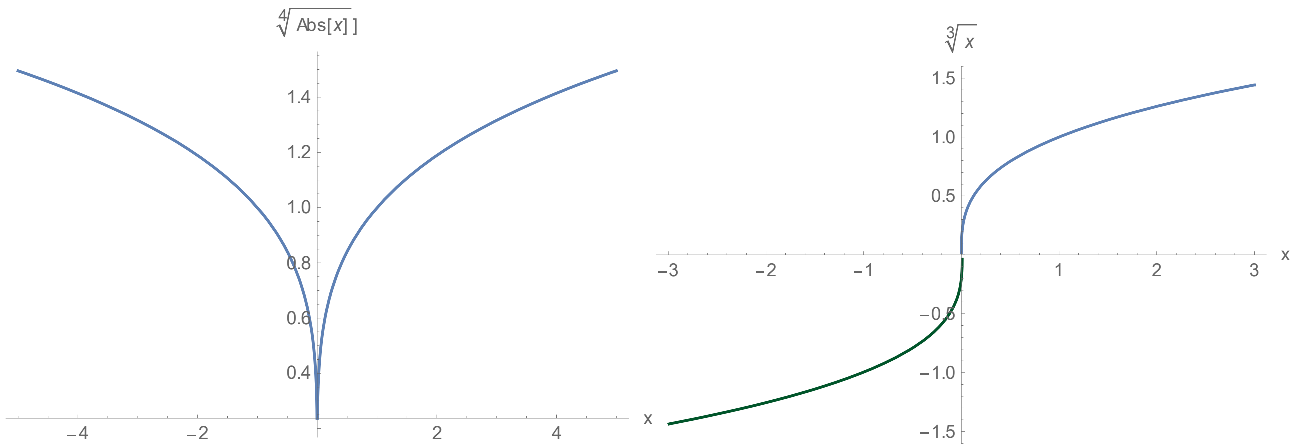


Fig. *4 Casi di non derivabilità in singoli punti*

Il caso dei **flessi a tangente verticale** è quello in cui i limiti sinistro e destro del rapporto incrementale sono **infiniti dello stesso segno**. È il caso delle funzioni di radici di indice dispari (vedi il secondo grafico riprodotto sopra)

Torniamo al Punto 3 lasciato in sospeso. (Cenno) Agli estremi del dominio la funzione potrebbe essere derivabile con la derivata che si annulla proprio agli estremi. In questi casi la derivata potrebbe essere pari a zero, ad esempio, se il punto fosse all'infinito. Ad esempio, con $f(x) = \sqrt{x}$ sul dominio reale positivo, la derivata prima $f_x' = \sqrt{x}/2x$ si annulla per x pari a infinito (prendere il limite di f' per x tendente a infinito e applicare la regola di de l'Hospital).

Esempio 2

Sia la funzione $f(x) = \arcsin(x)$ per ogni $-1 < x < 1$ il cui grafico è

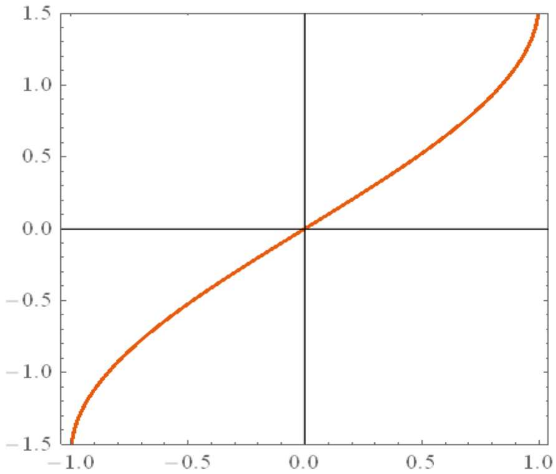


Fig. 5 $f(x) = \arcsin(x)$

Chiaramente $df(x)/dx = 1/\sqrt{[1-x^2]}$. Rappresentare graficamente e valutare se si annulla nel C.E.

Esercizio 1

Usando la funzione di cui all'Esempio 1, verificare **se** trovano applicazione tutti i risultati 1) – 4) ricordati prima.

Esercizio 2

Sia la funzione $f(x) = x[\arcsin(x)]$ per ogni $-1 < x < 1$ il cui grafico è

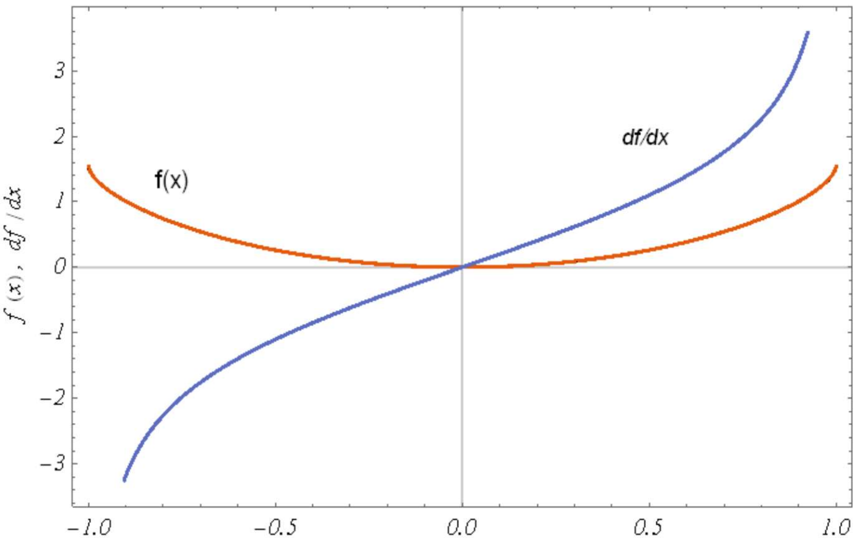


Fig. 5 $f(x) = x(\arcsin(x))$

Mostrare che $df(x)/dx = x/\sqrt{1-x^2} + \arcsin(x)$ si annulla per $x = 0$ e che *f(0)* è un minimo in $x = (-1,1)$. Osservare il passaggio di quadrante della funzione *df/dx*.

1.3 Massimi e minimi per funzioni in R^2

Passiamo adesso al caso di funzioni in due variabili. Sia $f(x, y)$: $R^2 \to R$ di classe C^2. Poiché $f(x, y)$ è derivabile, la condizione **necessaria** affinché almeno un punto $P(x_0, y_0)$ sia di massimo o di minimo locale deve essere: $f_x(x, y) = f_y(x, y) = 0$. Ovvero $\nabla_f(x,y)^T = \left(f_x, f_y\right)^T = 0$. Dunque si tratterà innanzi tutto di risolvere tali equazioni per determinare quel/ei valore/i della x e della y in corrispondenza del/i quale/i si realizza/no il/i possibile/i punto/i di massimo o di minimo locale. Anche in questo caso, i punti in cui si annullano le derivate prime si dicono punti stazionari o punti critici.

Esempio 3

Sia $f(x, y) = 2x^2 + 3xy - 2y$ esistente sull'intero piano reale. Siamo interessati alla ricerca di un suo massimo rispetto a x e a y su tutto il suo dominio che è R^2. Sappiamo che lo troveremo risolvendo inizialmente il sistema non omogeneo in 2 incognite

$$\begin{cases} \dfrac{\partial f}{\partial x} = 4x + 3y = 0 \\ \dfrac{\partial f}{\partial y} = 3x - 2 \quad = 0 \end{cases} \quad \text{ovvero} \quad \underset{(2\times2)}{\begin{bmatrix} 4 & 3 \\ 3 & 0 \end{bmatrix}} \underset{(2\times1)}{\begin{bmatrix} x \\ z \end{bmatrix}} = \underset{(2\times1)}{\begin{bmatrix} 0 \\ 2 \end{bmatrix}}$$

Il determinante della matrice dei coefficienti è $\Delta = (4\times0) - (3\times3) = -9$ che è diverso da zero. Inoltre, nel nostro caso, r(**A**) = r(**A|B**) = 2 (ovvero si verifica l'uguaglianza del rango della matrice dei coefficienti **A** e del rango della matrice c.d. completa o completata al numero delle incognite [**A|B**]) per il Teorema di Rouché-Capelli ci attendiamo un'unica soluzione (x^*, y^*). Questa è data da (con Cramer[4])

$$x^* = \frac{\begin{Vmatrix} 0 & 3 \\ 2 & 0 \end{Vmatrix}}{-9} = -\frac{-6}{9} = \frac{2}{3} \qquad\qquad y^* = \frac{\begin{Vmatrix} 4 & 0 \\ 3 & 2 \end{Vmatrix}}{-9} = -\frac{8}{9}$$

Valutata con $x = x^*$ e $y = y^*$ la funzione f vale 8/9. Quindi il punto $P(x_0, y_0)$ è (2/3, -8/9). Ma abbiamo trovato un massimo o un minimo (o un punto di sella)? Il segno delle derivate seconde (**che però da sole non vogliono dire niente**) è positivo e quindi P_0 = (2/3, -8/9) sembra davvero un minimo. Aspettiamo ancora un momento prima di rispondere. Il grafico seguente mostra la funzione, che nel suo minimo (se tale è per davvero) vale 8/9.

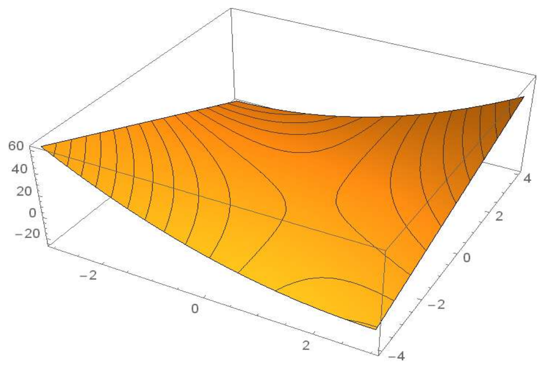


Fig. 6 *Grafico di* $2x^2 + 3xz - 2z$

[4] Con $\begin{Vmatrix} a & b \\ c & d \end{Vmatrix} = (ad) - (cd)$ intendiamo il determinante di $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$.

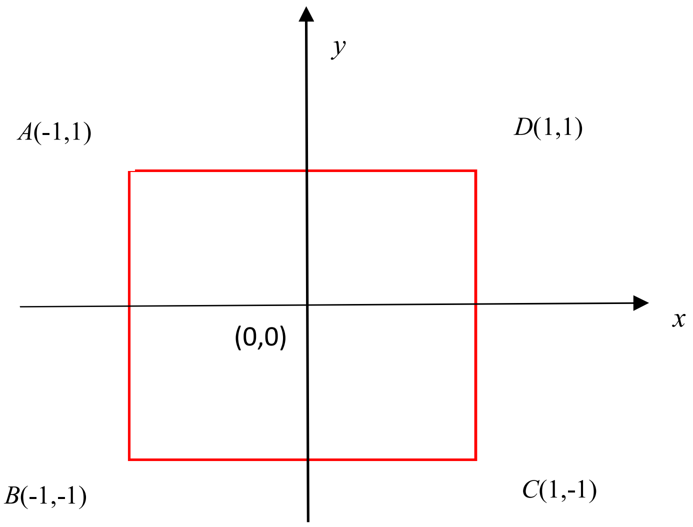
Come facciamo a dire che il punto di singolarità (stazionarietà) ottenuto è un massimo o un minimo o altro ancora? Con riferimento al problema non vincolato (ma vedremo dopo anche nei problemi con vincolo) ci aiutano le condizioni per la concavità o la convessità della funzione. Rinviamo quindi la questione ai paragrafi successivi. Adesso cerchiamo la risposta analizzando solo la natura dei punti di stazionarietà. Prima di procedere facciamo solo un altro esempio.

Esempio 4 (un filo più elaborato)

Sia $f: R^2 \to R$ di classe C^2 sull'intero piano reale

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

di cui cerchiamo i massimi e minimi assoluti nel quadrato $A(-1,1)$, $B(-1,-1)$, $C(1,-1)$ e $D(1,1)$. Effettivamente i massimi e i minimi esistono nel quadrato perché questo è un insieme chiuso e limitato in R^2 e definisce uno spazio metrico dotato della topologia usuale: quindi vale il Teorema di Weirstrass.
Il grafico seguente serve solo a dirci dove siamo richiesti di cercare i massimi e i minimi.



Cerchiamo i punti stazionari interni e di frontiera.

Punti interni. Risolviamo il sistema

$$\begin{matrix} f_x = 4x \\ f_y = 6y \end{matrix} \quad \begin{cases} 4x = 0 \\ 6y = 0 \end{cases} \to (0,0) \to f(0,0) = 0$$

L'origine, che appartiene a R^2, è l'unico punto di stazionarietà interno.

Punti di frontiera. Esaminiamo i 4 segmenti.

AD: $\begin{matrix} y=1 \text{ con } -1 \le x \le 1 \\ f_x = 4x = 0 \end{matrix} \rightarrow \begin{cases} f(x,1) = 2x^2+3 \\ x=0 \end{cases} \rightarrow \begin{cases} f(x,1) \rightarrow f(0,1) = 3 \text{ nel punto } (0,1) \\ \end{cases}$

DC: $\begin{matrix} x=1 \text{ con } -1 \le y \le 1 \\ f_y = 6y = 0 \end{matrix} \rightarrow \begin{cases} f(1,y) = 2+3y^2 \\ y=0 \end{cases} \rightarrow \begin{cases} f(x,1) = 3 \text{ nel punto } (1,0) \\ \end{cases}$

BC: $\begin{matrix} y=-1 \text{ con } -1 \le x \le 1 \\ f_x = 4x = 0 \end{matrix} \rightarrow \begin{cases} f(x,-1) = 2x^2+3 \\ x=0 \end{cases} \rightarrow \begin{cases} f(x,-1) = 3 \text{ nel punto } (0,-1) \\ \end{cases}$

AB: $\begin{matrix} x=-1 \text{ con } -1 \le y \le 1 \\ f_y = 6y = 0 \end{matrix} \rightarrow \begin{cases} f(-1,y) = 2+3y^2 \\ y=0 \end{cases} \rightarrow \begin{cases} f(1,y) = 2 \text{ nel punto } (-1,0) \\ \end{cases}$

Esaminiamo infine i “vertici” del quadrato:

$f(A) = f(-1,1) = 5$; $f(B) = f(-1,-1) = 5$; $f(C) = f(1,-1) = 5$; $f(D) = f(1,1) = 5$

La Tabella seguente riassume i risultati. La penultima colonna riporta le derivate seconde parziali e l’ultima colonna indica la natura del punto. Le informazioni di cui alle altre celle della tabella non bastano a giustificare i risultati dell’ultima colonna. Li riporto per permettere di svolgere successivamente un altro esercizio.

Tab.1 *Punti di stazionarietà e di non stazionarietà della funzione f(x, y) = 2x² +3y²*

f valutata in

Origine	*f(0,0)*	*f = 0*	*(0,0,0)*	$f''_x= 4; f''_y= 6$	*Minimo assoluto*
Segmento AD	*f(0,1)*	*f =3*	*(0,1,3)*	$f''_x= 4; f''_y= 6$	*?*
Segmento DC	*f(1,0)*	*f = 2*	*(1,0,2)*	$f''_x= 4; f''_y= 6$	*?*
Segmento BC	*f(0,-1)*	*f = 3*	*(0,-1,3)*	$f''_x= 4; f''_y= 6$	*?*
Segmento AB	*f(-1,0)*	*f = 2*	*(-1,0,2)*	$f''_x= 4; f''_y= 6$	*?*
Punto A	*f(-1,1)*	*f = 5*	*(-1,1,5)*	$f''_x= 4; f''_y= 6$	*Massimo assoluto*
Punto B	*f(-1,-1)*	*f =5*	*(-1,-1,5)*	$f''_x= 4; f''_y= 6$	*Massimo assoluto*
Punto C	*f(-1,-1)*	*f =5*	*(1,-1,5)*	$f''_x= 4; f''_y= 6$	*Massimo assoluto*
Punto D	*f(1,1)*	*f =5*	*(1,1,5)*	$f''_x= 4; f''_y= 6$	*Massimo assoluto*

Il grafico seguente illustra la funzione nel sotto insieme del dominio dato dal quadrato.

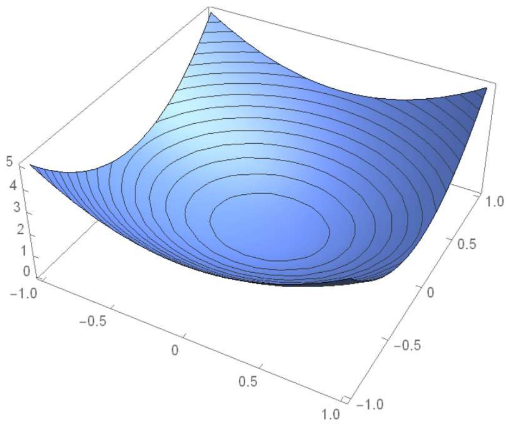


Fig. 7 *$f(x, y) = 2x^2 + 3y^2$ con minimo assoluto in (0,0)*

1.4 Massimi e minimi non vincolati: sintesi operativa (caso R^2) per chi ha fretta

Se il segno delle derivate seconde non basta a chiarire la natura del/i punto/i di stazionarietà trovati ciò vuol dire che la metodologia che cerchiamo è più "impegnativa". In questo paragrafo presentiamo, per chi ha fretta, la pappa già fatta delle condizioni per punti di massimo/minimo non vincolati e rimandiamo la sua ricetta e l'illustrazione degli ingredienti alla discussione delle condizioni di concavità/convessità della funzione studiata.
Sia *f(x, z)* su R^2 una funzione continua e derivabile almeno sino alla derivata seconda su tutto il campo di esistenza. Allora:

1) si calcolano le derivate prime; le si azzerano e si trovano le corrispondenti $x*$ e $z*$ di stazionarietà.
2) Si calcolano le derivate seconde e quelle incrociate (saranno in genere uguali tra loro: come mai?).
3) si montano i pezzi della matrice Hessiana simmetrica (vedi punto 2) valutata in $x*$ e $z*$.
4) Si valutano i segni dei determinanti dei primi due minori (non ce ne sono altri…).

La matrice hessiana (simmetrica) è

$$H(x^*, z^*) = \begin{bmatrix} \frac{\partial^2 f(x,z)}{\partial x^2} & \frac{\partial^2 f(x,z)}{\partial x \partial z} \\ \frac{\partial^2 f(x,z)}{\partial z \partial x} & \frac{\partial^2 f(x,z)}{\partial z^2} \end{bmatrix}$$

Allora, se

$$\|H\|_1(x^*, z^*) = \frac{\partial^2 f(x,z)}{\partial x^2} < 0$$

e

$$\|H\|_2(x^*, z^*) = \left\| \begin{matrix} \frac{\partial^2 f(x,z)}{\partial x^2} & \frac{\partial^2 f(x,z)}{\partial x \partial z} \\ \frac{\partial^2 f(x,z)}{\partial z \partial x} & \frac{\partial^2 f(x,z)}{\partial z^2} \end{matrix} \right\| > 0$$

Siamo in presenza di un massimo in $P(x^*, z^*)$.

Se

$$\|H\|_1(x^*,z^*) = \frac{\partial^2 f(x,z)}{\partial x^2} > 0$$

e

$$\|H\|_2(x^*,z^*) = \begin{Vmatrix} \frac{\partial^2 f(x,z)}{\partial x^2} & \frac{\partial^2 f(x,z)}{\partial x \partial z} \\ \frac{\partial^2 f(x,z)}{\partial z \partial x} & \frac{\partial^2 f(x,z)}{\partial z^2} \end{Vmatrix} > 0$$

Siamo in presenza di un minimo in $P(x^*,z^*)$.

Se invece

$$\|H\|_1(x^*,z^*) = \frac{\partial^2 f(x,z)}{\partial x^2} \gtreqless 0$$

e

$$\|H\|_2(x^*,z^*) = \begin{Vmatrix} \frac{\partial^2 f(x,z)}{\partial x^2} & \frac{\partial^2 f(x,z)}{\partial x \partial z} \\ \frac{\partial^2 f(x,z)}{\partial z \partial x} & \frac{\partial^2 f(x,z)}{\partial z^2} \end{Vmatrix} < 0$$

Siamo in presenza di un punto di sella in $P(x^*,z^*)$.
Infine, se

$$\|H\|_1(x^*,z^*) = \frac{\partial^2 f(x,z)}{\partial x^2} \gtreqless 0$$

e

$$\|H\|_2(x^*,z^*) = \begin{Vmatrix} \frac{\partial^2 f(x,z)}{\partial x^2} & \frac{\partial^2 f(x,z)}{\partial x \partial z} \\ \frac{\partial^2 f(x,z)}{\partial z \partial x} & \frac{\partial^2 f(x,z)}{\partial z^2} \end{Vmatrix} = 0$$

siamo in una condizione di indeterminazione per "risolvere" la quale valutiamo $f(x^*, y^*)$ alla luce del segno delle derivate seconde.
Gli ingredienti della pappa già fatta sono quindi a) le derivate seconde e miste e b) il segno del determinante dell'hessiano da esse composto, valutato in $P(x^*,z^*)$, ovvero sostituendo ad x e ad z i valori ottenuti risolvendo il sistema delle derivate prime parziali azzerate.
Come si vede esiste una fondamentale corrispondenza tra condizioni per un massimo/minimo e condizioni di concavità/convessità della funzione. Nel parleremo nel prossimo paragrafo dove, al fine di generalizzare le condizioni di massimo e minimo al caso R^N, riprenderemo brevemente le condizioni di concavità/convessità. Adesso eseguiamo il seguente

Esercizio 4

Tornare alla funzione $f(x, y) = 2x^2 + 3y^2$ e giustificare la natura dei punti di stazionarietà di cui all'ultima colonna della tabella. Valutare la natura dei punti indicati con il punto di domanda nella Tabella e spiegare perché non esistono punti di sella.

Sintesi

Sempre nella logica della pappa già fatta riassumiamo relativamente al caso in R^2 le condizioni in precedenza ottenute nella seguente tabella.

Tab.2 *Massimi e minimi non vincolati*

Condizione	Massimo di $f(x, y)$	Minimo di $f(x, y)$	Sella di $f(x, y)$
c.d. Primo Ordine	$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial y} = 0$	$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial y} = 0$	$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial y} = 0$
c.d. Secondo ordine	$\frac{\partial^2 f}{\partial x^2} < 0, \frac{\partial^2 f}{\partial y^2} < 0$ e $\left(\frac{\partial^2 f}{\partial x^2}\right)\left(\frac{\partial^2 f}{\partial y^2}\right) > \left[\frac{\partial^2 f}{\partial x \partial y}\right]^2$	$\frac{\partial^2 f}{\partial x^2} > 0, \frac{\partial^2 f}{\partial y^2} > 0$ e $\left(\frac{\partial^2 f}{\partial x^2}\right)\left(\frac{\partial^2 f}{\partial y^2}\right) > \left[\frac{\partial^2 f}{\partial x \partial y}\right]^2$	$\frac{\partial^2 f}{\partial x^2} = ? \quad \frac{\partial^2 f}{\partial y^2} = ?$ e $\left(\frac{\partial^2 f}{\partial x^2}\right)\left(\frac{\partial^2 f}{\partial y^2}\right) < \left[\frac{\partial^2 f}{\partial x \partial y}\right]^2$

N.B: Tutte le derivate seconde sono valutate nel punto di stazionarietà (cioè dove si annullano le derivate prime). Le condizioni sulle derivate seconde sono **sufficienti, ma non necessarie**.

Per tornare all'**Esercizio 3**, e concluderlo, diciamo che effettivamente avevamo trovato un minimo, visto che le derivate seconde erano positive e $\left(\frac{\partial^2 f}{\partial x^2}\right)\left(\frac{\partial^2 f}{\partial z^2}\right) > \left[\frac{\partial^2 f}{\partial x \partial z}\right]^2$ poiché (6×4) > 0.

Chiariamo ancora di più la procedura con un secondo esempio.

Esempio 3

Sia: $f(x,y) = 8x^3 + 2xy - 3x^2 + y^2 + 1$

Valutare se nel suo C.E. la funzione ammette massimi o minimi. Calcoliamo per prima cosa le derivate prime e seconde

$$\frac{\partial f}{\partial x} = 24x^2 + 2y - 6x \qquad \frac{\partial f}{\partial y} = 2x + 2y$$

$$\frac{\partial^2 f}{\partial x^2} = 48x - 6 \qquad \frac{\partial^2 f}{\partial y^2} = 2$$

$$\underbrace{\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}}_{\text{Spiegare perchè}} = 2$$

Le condizioni del primo ordine richiedono la soluzione del sistema

$$\begin{aligned} 24x^2+2y-6x&=0\\ 2y+2x&=0\end{aligned}$$

Dalla seconda equazione otteniamo $y = -x$, che sostituita nella prima ci dà l'equazione di secondo grado $24x^2 - 8x = 0$. Allora le due soluzioni sono

$$\bar{x}_{1,2}=\frac{8\pm\sqrt{64-0}}{48}=\frac{8\pm 8}{48}=\begin{cases}0 \quad \text{che implica } \bar{y}_1=-\bar{x}_1=0\\ 1/3 \text{ che implica } \bar{y}_2=-1/3\end{cases}$$

I punti (0, 0) e (1/3, -1/3) annullano entrambe le derivate prime. Il primo rigo della tabella è a posto! Guardiamo adesso le derivate seconde. Con (0, 0) abbiamo:

$$\frac{\partial^2 f}{\partial x^2}=-6 \text{ e } \frac{\partial^2 f}{\partial y^2}=2. \text{ Di conseguenza : } \left(\frac{\partial^2 f}{\partial x^2}\right)\left(\frac{\partial^2 f}{\partial y^2}\right)=-12<4=\left(\frac{\partial^2 f}{\partial x\partial y}\right)^2$$

Con (0,0) il secondo rigo della tabella è "a posto" MA solo sotto la colonna Sella. Infatti il segno opposto delle due derivate seconde suggerisce che in (0,0) la superficie di *f* "gira" verso l'alto con riferimento alla *y* e verso il basso con riferimento alla *x* nel punto considerato. Siamo quindi in presenza di un punto di sella. Quasi sempre una brutta cosa!

Guardiamo adesso all'altra soluzione dell'equazione di secondo grado risolta in precedenza. Il punto del piano *(x, y)* era (1/3, – 1/3). Quindi

$$\frac{\partial^2 f}{\partial x^2}=10 \text{ e } \frac{\partial^2 f}{\partial y^2}=2 \quad \text{e di conseguenza } \left(\frac{\partial^2 f}{\partial x^2}\right)\left(\frac{\partial^2 f}{\partial y^2}\right)=20>2^2=\left(\frac{\partial^2 f}{\partial x\partial y}\right)^2$$. Stavolta la tabella ci dice che siamo in presenza di un minimo. In tale punto la *f* vale 23/27.

Riassumendo:

Esiste un solo estremo relativo (minimo) della funzione studiata in *x* = 1/3 e *y* = -1/3 con *f* che in tale punto vale 23/27. Il seguente grafico della *f* studiata aiuta a capire (guardare cosa succede nell'origine) …

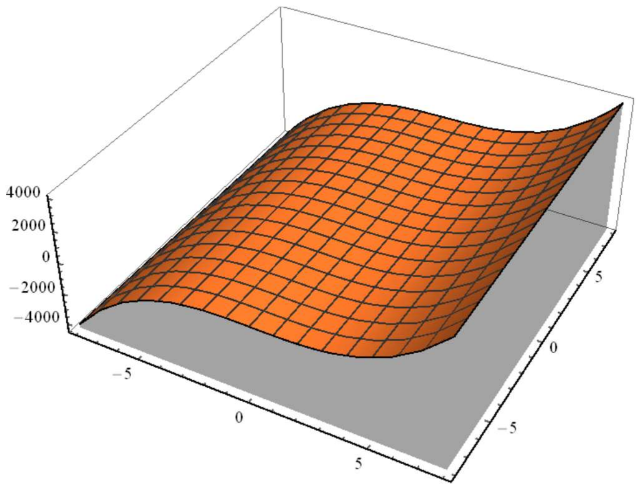


Fig. 8 Grafico di $f(x, y) = 8x^3 + 2xy - 3x^2 + y^2 + 1$

Sempre nella logica della pappa già pronta, diamo adesso una versione più "compatta" delle condizioni richieste per una massimo/minimo non vincolati e discutiamo in modo più esteso anche il caso del punto di sella. Toglieremo in questo modo i punti di domanda dalla Tabella 1 precedente.
Sia, in generale, la funzione $f(x, y)$ definita in un insieme A. Si dice che il punto $P(x_0,y_0) \in A$ è un punto di massimo (minimo) **assoluto** se per ogni punto $P(x, y) \in A$ vale:

$$f(x,y) \leq f(x_0,y_0) \text{ o corrispondentemente } f(x,y) \geq f(x_0,y_0)$$

Il valore $V = f(x_0,y_0)$ che la funzione assume in $P(x_0, y_0)$ si dice massimo (minimo) **assoluto** di f in A. Ci si ricordi che una funzione continua in un insieme chiuso e limitato (compatto) è ivi dotata di massimo e minimo assoluto per via del Teorema di Weierstrass (ripassare).
Detto dell'esistenza del massimo e del minimo assoluto, ripassiamo la nozione di massimo e minimo **relativo**.
Sia $f(x, y)$ definita in A. Si dice che il punto $P_0(x_0, y_0) \in A$ è un punto di massimo (minimo) relativo se esiste un intorno circolare I del punto $P_0(x_0, y_0) \in A$ tale che per ogni punto $P(x, y)$ di A appartenente a I risulta

$$f(x,y) \leq f(x_0,y_0) \text{ o corrispondentemente } f(x,y) \geq f(x_0,y_0)$$

Se il segno di eguaglianza vale solo per $(x, y) = (x_0, y_0)$ si dice che $P_0(x_0, y_0)$ è un punto di massimo (minimo) **relativo proprio**.
Analogamente al caso di funzioni di una sola variabile reale vale il seguente Teorema:

TEOREMA 1

Sia $f(x, y)$ una funzione definita in un insieme A e sia $\boldsymbol{P_0(x_0,y_0)}$ un punto di massimo (minimo) **relativo** interno ad A. Se la funzione $f(x, y)$ è parzialmente derivabile in $P_0(x_0,y_0)$, allora le derivate parziali prime di f si annullano nel punto $P(x, y) \in A$:

$$\frac{\partial f}{\partial x}(x_0,y_0) = 0 \quad \text{e} \quad \frac{\partial f}{\partial y}(x_0,y_0) = 0$$

Osservazione 1: Tali condizioni sono **necessarie** ma non sufficienti affinché il punto $\boldsymbol{P_0(x_0, y_0)}$ sia un punto di massimo (minimo) **relativo.** Osservazione 2: La ricerca dei punti di massimo/minimo va però comunque fatta tra i punti (se più di uno) in cui si annullano simultaneamente le derivate parziali prime. Tutti questi punti vengono denominati punti critici o stazionari o di **singolarità**, e come diceva Sherlock Holmes: *singularities are almost inevitably a clue*! Ovvero: in quei punti deve succedere qualcosa di particolare che offre indizi sul comportamento della funzione. Quindi un punto di massimo o di minimo è sempre un punto di singolarità. Non vale il contrario.
Siamo allora in presenza di condizioni che sono solo necessarie. Per stabilire la natura (massimo o minimo) di un punto di singolarità occorre procedere affrontando il punto 5 del *vade mecum* precedente.
Ricordiamo che $f: R^2 \to R$ è di classe C^2 ovvero che esistono e sono continue e finite le derivate seconde e seconde miste. Usiamole per costruire la matrice Hessiana (2×2 nel nostro caso) seguente

$$H(x_0,y_0) = \begin{bmatrix} \dfrac{\partial^2 f(x_0,y_0)}{\partial x^2} & \dfrac{\partial^2 f(x_0,y_0)}{\partial x \partial y} \\ \dfrac{\partial^2 f(x_0,y_0)}{\partial y \partial x} & \dfrac{\partial^2 f(x_0,y_0)}{\partial y^2} \end{bmatrix}$$

Dove il Teorema di Young (per il quale $\frac{\partial^2 f(x_0,y_0)}{\partial x \partial y} = \frac{\partial^2 f(x_0,y_0)}{\partial y \partial x}$, cercare di ricordarselo) assicura la simmetria.

Vale allora il

TEOREMA 2

Sia $f : R^2 \to R$ di classe C^2 (ovvero f sia continua con le sue derivate parziali prime e seconde) e sia il punto $P_0(x_0,y_0)$ interno al Campo di Esistenza di $f(x, y)$ e si abbia in $P_0(x_0,y_0)$

$$\frac{\partial f}{\partial x}(x_0, y_0) = 0 \quad \text{e} \quad \frac{\partial f}{\partial y}(x_0, y_0) = 0$$

SE:

1) $\frac{\partial^2 f}{\partial x^2}(x_0,y_0) > 0$, $\frac{\partial^2 f}{\partial y^2}(x_0,y_0) > 0$ e $\|H(x_0,y_0)\| > 0$ allora $P_0(x_0,y_0)$ è un punto di **minimo relativo proprio**

2) $\frac{\partial^2 f}{\partial x^2}(x_0,y_0) < 0$, $\frac{\partial^2 f}{\partial y^2}(x_0,y_0) < 0$ e $\|H(x_0,y_0)\| > 0$ allora $P_0(x_0,y_0)$ è un punto di **massimo relativo proprio**

3) $\frac{\partial^2 f}{\partial x^2}(x_0,y_0)$, $\frac{\partial^2 f}{\partial y^2}(x_0,y_0)$ hanno segno qualsiasi e $\|H(x_0,y_0)\| < 0$ allora $P_0(x_0,y_0)$ non è né un massimo né un minimo relativo proprio ma si denomina **punto di sella.**

Il quarto caso è quello in cui $\frac{\partial^2 f}{\partial x^2}(x_0,y_0)$ ha qualsiasi segno e $\|H(x_0,y_0)\| = 0$. Quando il determinante della matrice Hessiana si annulla nel punto $P_0(x_0, y_0)$ nulla si può dire e, come già accennato, bisogna studiare direttamente il comportamento della funzione esaminando il segno della differenza

$$f(x,y) - f(x_0,y_0)$$

<u>Osservazione</u>: nei 2 casi in cui $\|H(x_0,y_0)\| > 0$ la derivata seconda rispetto a x può essere sostituita dalla derivata seconda rispetto a y. Si può infatti mostrare che esse hanno lo stesso segno (dato il determinante positivo dell'Hessiana: provarci).

Che succede se $f: R^N \to R$ è di classe C^2, ovvero se sono N le variabili indipendenti? Niente di speciale. L'Hessiana sarà una matrice N×N e permetterà il calcolo del segno del determinante di N minori c.d. di Nord-Ovest

$$\|H_1(x_0,y_0)\|, \|H_2(x_0,y_0)\|, \ldots, \|H_N(x_0,y_0)\|.$$

Allora, **fermo restando l'annullamento di tutte le derivate prime parziali** nel punto $P_0(x_0, y_0, z_0, \ldots)$, diremo del tale punto di singolarità che:

Il punto di singolarità è un **MASSIMO** relativo se, tutti i determinanti di Nord-Ovest di indice **dispari**

$$\|H_1(x_0,y_0)\|,\|H_3(x_0,y_0)\|,\|H_5(x_0,y_0)\|,\ldots$$

sono **NEGATIVI** e quelli di indice **pari**

$$\|H_2(x_0,y_0)\|,\|H_4(x_0,y_0)\|,\|H_6(x_0,y_0)\|,\ldots$$

sono **POSITIVI** (notare che H_1 coincide con la derivata seconda rispetto alla prima variabile; infatti le derivate seconde sono gli elementi della traccia dell'Hessiana)

Al contrario

Il punto di singolarità è un **MINIMO** relativo se, **tutti** i determinanti di Nord-Ovest

$$\|H_1(x_0,y_0)\|,\|H_2(x_0,y_0)\|,\|H_3(x_0,y_0)\|,\ldots,\|H_N(x_0,y_0)\|$$

sono **POSITIVI**. (Tutti vuol dire tutti, a prescindere dal pari o dispari).

ESEMPIO. Rielaborazione dell'esempio 2. La funzione era $f(x,y) = 8x^3 + 2xy - 3x^2 + y^2 + 1$

Avevamo già calcolato le derivate prime e seconde

$$\frac{\partial f}{\partial x} = 24x^2 + 2y - 6x \qquad \frac{\partial f}{\partial y} = 2x + 2y$$

$$\frac{\partial^2 f}{\partial x^2} = 48x - 6 \qquad \frac{\partial^2 f}{\partial y^2} = 2$$

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x} = 2$$

Le condizioni sulle derivate parziali prime richiedono la soluzione del sistema

$$\begin{aligned} 24x^2 + 2y - 6x &= 0 \\ 2y + 2x &= 0 \end{aligned}$$

Da cui le due soluzioni (vedi sopra)

$$\overline{x}_{1,2} = \frac{8 \pm \sqrt{64-0}}{48} = \frac{8 \pm 8}{48} = \begin{cases} 0 & \text{che implica } \overline{y}_1 = -\overline{x}_1 = 0 \\ 1/3 & \text{che implica } \overline{y}_2 = -1/3 \end{cases}$$

Guardiamo adesso le derivate seconde. Con la prima coppia di soluzioni dell'equazione di secondo grado (0, 0) abbiamo:

$\frac{\partial^2 f}{\partial x^2} = -6$ e $\frac{\partial^2 f}{\partial y^2} = 2$. L'Hessiana nel punto *P(0, 0)* sarà $H(0,0) = \begin{bmatrix} -6 & 2 \\ 2 & 2 \end{bmatrix}$. Pertanto

$$\| H(0,0) \| = \begin{Vmatrix} -6 & 2 \\ 2 & 2 \end{Vmatrix} = -12 - 4 = -16 < 0 .$$

Allora nel punto $P(0, 0)$ avremo

$$\left.\begin{array}{l} \frac{\partial^2 f}{\partial x^2} < 0 \\ \\ \| H(0,0) \| < 0 \end{array}\right\} \text{Quindi: } P(0,0) \text{ è un Punto di Sella}$$

Proviamo la seconda coppia di valori (1/3, -1/3) ovvero il punto *P(1/3, -1/3)*. Avremo che L'Hessiana nel punto P(1/3,-1/3) sarà $H(\frac{1}{3}, -\frac{1}{3}) = \begin{bmatrix} 10 & 2 \\ 2 & 2 \end{bmatrix}$. Pertanto $\left\| H(\frac{1}{3}, -\frac{1}{3}) \right\| = \begin{Vmatrix} 10 & 2 \\ 2 & 2 \end{Vmatrix} = 20 - 4 = 16 > 0$.

Allora nel punto $P(1/3,-1/3)$:

$$\left.\begin{array}{l} \frac{\partial^2 f}{\partial x^2} > 0 \\ \\ \left\| H(\frac{1}{3}, -\frac{1}{3}) \right\| > 0 \end{array}\right\} \text{Quindi: } P(\tfrac{1}{3}, -\tfrac{1}{3}) \text{ è un Minimo relativo proprio}$$

Guardare la figura.

Qui finisce la pappa già fatta, utile a risolvere esercizi, più che a comprendere la logica del problema. Spero però che nella/nel lettrice/re sia soprattutto maturata la "curiosità" di capire **come mai** quelle che abbiamo esposto sono le condizioni che cerchiamo. Spero anche che non sia subentrata la fretta (tipica degli economisti che devono "superare" gli esami) di passare subito alla ricerca delle condizioni di massimo e di minimo vincolati. La cosa più importante da capire e imparare riguarda la convessità/concavità della funzione oggetto di studio.

2. FUNZIONI CONCAVE/CONVESSE IN R, R^2 E IN R^N E CONDIZIONI PER MASSIMI E MINIMI

Trattiamo adesso la questione più importante: in che modo le condizioni per la concavità o la convessità della funzione ci aiutano nella interpretazione della natura dei punti di stazionarietà. In questo modo affronteremo la questione del come mai quelle che abbiamo trovato sono proprio le condizioni che cercavamo. Occorre iniziare dalla definizione di insiemi convessi e insiemi non convessi[5].

[5] Non esistono gli insiemi concavi.

2.1 Insiemi convessi e insiemi non convessi

Premessa: Un vettore y in R^n si dice combinazione convessa di $k \geq 1$ vettori $x_1, \ldots, x_k$ se risulta

$$y = \sum_{i=1}^{k} \lambda_i x_i \quad \text{con} \quad \sum_{i=1}^{k} \lambda_i = 1 \quad \lambda_i \geq 0 \quad i = 1, \ldots, k$$

Un insieme $X \subseteq R^n$ è convesso se per ogni coppia di punti appartenenti all'insieme, appartengono all'insieme anche tutti i vettori ottenibili come loro combinazione convessa. Utilizzando il concetto di segmento chiuso[6], la definizione di insieme convesso può essere riformulata nel modo seguente: Un insieme X è convesso se per ogni coppia di vettori $x, y \in X$ si ha $[x, y] \subseteq X$. Dalla definizione segue che l'insieme vuoto e l'insieme costituito da un solo vettore sono insiemi convessi (banali). Il più semplice insieme convesso non banale è il segmento di estremi.

Esempio

Sia l'insieme

$$S = \{(x, y, z) | x + 3y - 2z = 5\}$$

Esso è convesso. Infatti siano (x_o, y_o, z_o) e (x_1, y_1, z_1) punti appartenenti a S tali che

$$x_0 + 3y_0 - 2z_0 = 5 \qquad x_1 + 3y_1 - 2z_1 = 5$$

Definiamo un terzo punto x, y e z quale combinazione convessa degli altri due come segue

$$\begin{gathered} x = \lambda x_0 + (1 - \lambda) x_1 \\ y = \lambda y_0 + (1 - \lambda) y_1 \\ z = \lambda z_0 + (1 - \lambda) z_1 \end{gathered}$$

con $0 \leq \lambda \leq 1$. Allora sostituendo

$$x + 3y - 2z = \lambda(x_0 + 3y_0 - 2z_0) + (1 - \lambda)(x_0 + 3y_1 - 2z_1) = \lambda(5) + (1 - \lambda)(5) = 5$$

Pertanto (x, y, z) appartiene ad S ed S è convesso, come da definizione.

Un'applicazione per economisti: il vincolo di bilancio quale iperpiano

L'equazione x + 3y -2 z = 5 rappresenta un iperpiano in R^3. Ogni insieme di forma

$$S = \{\boldsymbol{x} | \boldsymbol{a}^T \boldsymbol{x} = b\}$$

dove b è uno scalare è chiamato iperpiano in R^N dove $\boldsymbol{a}$ è un vettore colonna ad N elementi non nulli. L'iperpiano è quindi un insieme convesso. Nell'esempio di cui sopra abbiamo

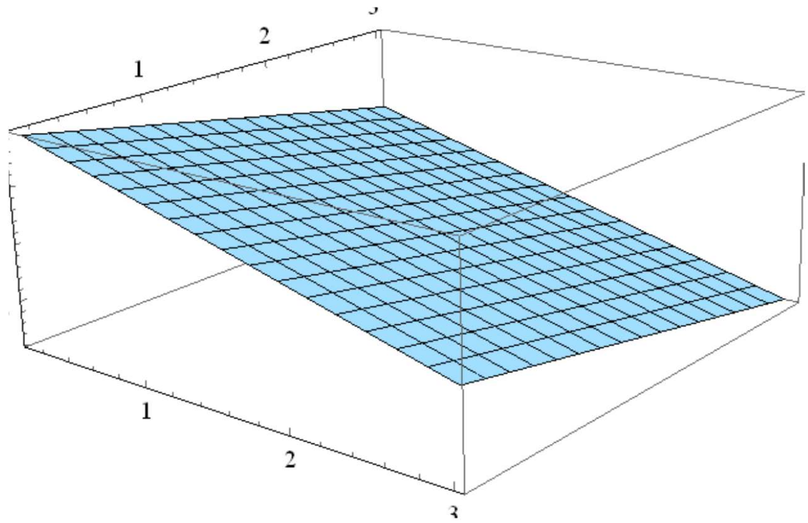
$$\begin{bmatrix} 1 & 3 & -2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = 5$$

[6] Un segmento chiuso di estremi [a, b] è definibile come segue: dati due punti a e b in R^n, l'insieme dei punti di R ottenuti con $z = (1-\beta)a + \beta b$ al variare di $\beta \in [0, 1]$ è detto segmento chiuso perché comprende entrambi i suoi estremi.

Supponiamo adesso di essere in R^3 e che gli elementi di ***a*** siano $p_x > 0$, $p_y > 0$ e $p_z > 0$ e che $b = M > 0$ dove M è la somma spendibile sul mercato (alcuni dicono il reddito del consumatore) mentre p_x , p_y e p_z sono i prezzi delle merci x, y e z. Allora

$$[p_x \quad p_y \quad p_x]\begin{bmatrix} x \\ y \\ z \end{bmatrix} = M$$

è l'iperpiano che indica il vincolo di bilancio del consumatore rispetto a due merci, che è un insieme convesso (anche se le merci fossero N). Ponendo $p_x = 3$, $p_y = 4$ e $p_z = 2$ e che $b = 100$ otteniamo



I tre assi misurano x, y e z. Eliminando un asse si ottiene, ad esempio, il grafico del vincolo di bilancio del consumatore relativo a due beni (retta).

1.3.2 Funzioni convesse e funzioni concave

Iniziamo con il semplice caso di funzioni in R.

Una funzione definita su un insieme convesso S della retta reale è detta concava (convessa) se per una *qualsiasi* coppia di punti x_0 e x_1 appartenenti a S e per $0 < \lambda < 1$ abbiamo

$$f(\lambda x_0 + (1-\lambda)x_1) \geq \lambda f(x_0) + (1-\lambda)f(x_1) \quad concavità$$

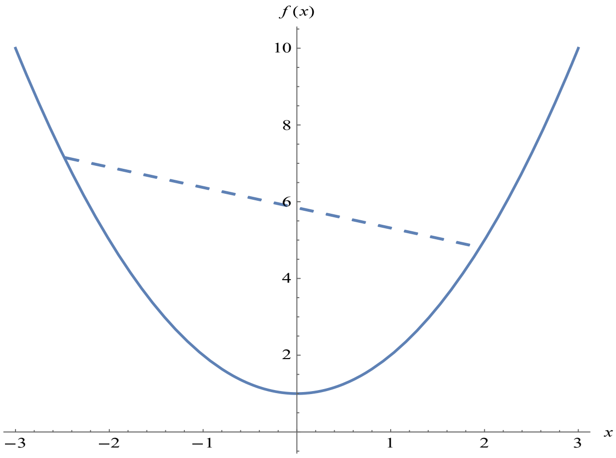
$$f(\lambda x_0 + (1-\lambda)x_1) \leq \lambda f(x_0) + (1-\lambda)f(x_1) \quad convessità$$

$$f(\lambda x_0 + (1-\lambda)x_1) < \lambda f(x_0) + (1-\lambda)f(x_1) \quad stretta\ concavità \quad dato\ x_0 \neq x_1$$

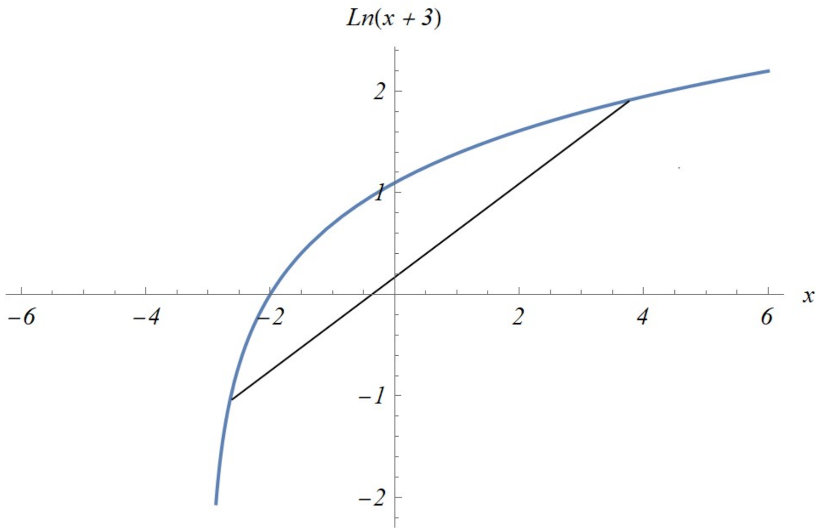
$$f(\lambda x_0 + (1-\lambda)x_1) < \lambda f(x_0) + (1-\lambda)f(x_1) \quad stretta\ convessità\ dato\ x_0 \neq x_1$$

La funzione lineare non è né concava né convessa.

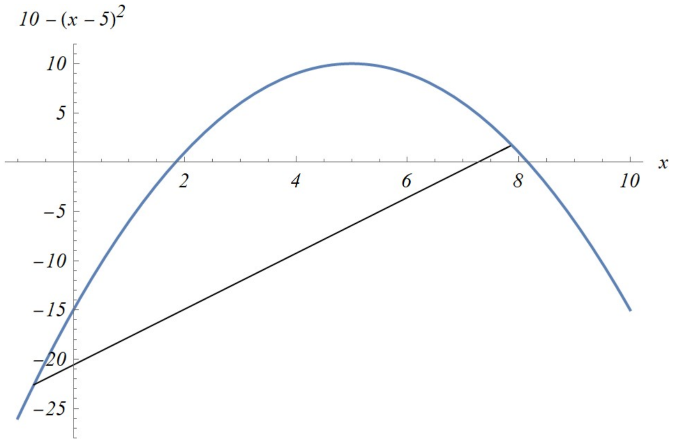
Esempio $f(x)= 1 + x^2$



La corda (funzione della combinazione convessa dei due punti) giace sempre al di sopra della funzione. La funzione è convessa. Le derivate prima e seconda sono 2x > 0 e 2 per x > 0 e -2x > 0 e -2 per x < 0.
Che succede con funzioni concave? Lo mostra l'esempio seguente per funzioni monotone crescenti.
Esercizio. Mostrare che $f(x) = \ln(3 + x)$ è strettamente concava nel dominio (-3 < x < ∞).



La corda (funzione della combinazione convessa dei due punti) giace sempre al di sotto della funzione. La derivata prima è 1/(x+3) e la derivata seconda è $-1/(x+3)^2$.
Con funzioni non monotone, quali $f(x) = 10 - (x - 5)^2$nel dominio (-∞ < x < ∞) avremo:



Valutare le derivate prima e seconda nel tratto crescente e decrescente.
Esercizio. Studiare la concavità e convessità della funzione $f(x) = 10 - (x-5)^3$

Funzioni in $\mathbb{R}^2$

In R^2 le cose sono un po' più complesse.
Sia *f*: $A \rightarrow R$ con $A \subseteq R^2$ insieme aperto e convesso, con *f* di classe C^2. Si dice che *f* è convessa (concava) se e solo se vale la condizione

$$f(x,z) \geq f(x^0,z^0) + \nabla_f(x^0,z^0)^T \begin{bmatrix} x - x^0 \\ z - z^0 \end{bmatrix} \quad \forall (x,z) \in A, \forall (x^0,z^0) \in A$$

$$f(x,z) \leq f(x^0,z^0) + \nabla_f(x^0,z^0)^T \begin{bmatrix} x - x^0 \\ z - z^0 \end{bmatrix} \quad \forall (x,z) \in A, \forall (x^0,z^0) \in A$$

Nel caso di una funzione convessa (concava) il grafico della funzione non sta mai al di sotto (al di sopra) del piano tangente al punto del grafico avente coordinate

$$((x^0,z^0), f(x^0,z^o)).$$

Valutiamo la funzione di cui all'Esempio 2 di cui alla Fig. 4: *f(x, z) = 2x² +3xz –2z* il cui punto di stazionarietà è (2/3, -8/9). Quindi

$$f(x,z) \geq f\left(\frac{2}{3}, -\frac{8}{9}\right) + [0\ 0] \begin{bmatrix} x - \frac{2}{3} \\ z + \frac{8}{9} \end{bmatrix} = \frac{8}{9} + [0\ 0] \begin{bmatrix} x - \frac{2}{3} \\ z + \frac{8}{9} \end{bmatrix} = \frac{8}{9}$$

La funzione nel punto di stazionarietà trovato sta sempre sopra *f(.)* = 8/9. È sufficiente per dire che *f* è convessa e che -8/9 è un minimo? Però perché *f* sia convessa (concava) occorre che

$$\frac{\partial^2 f(x,z)}{\partial x^2} \geq 0 \quad \frac{\partial^2 f(x,z)}{\partial z^2} \geq 0 \quad H(x,z) \geq 0 \ \ \forall (x,z) \in A$$

$$\frac{\partial^2 f(x,z)}{\partial x^2} \leq 0 \quad \frac{\partial^2 f(x,z)}{\partial z^2} \leq 0 \quad H(x,z) \geq 0 \;\; \forall (x,z) \in A$$

dove $H(x,z) = \left\| \begin{matrix} \frac{\partial^2 f(x,z)}{\partial x^2} & \frac{\partial^2 f(x,z)}{\partial x \partial z} \\ \frac{\partial^2 f(x,z)}{\partial z \partial x} & \frac{\partial^2 f(x,z)}{\partial z^2} \end{matrix} \right\|$ con le derivate valutate in (x^0, z^0).

Nel nostro esempio, poiché abbiamo $\frac{\partial^2 f}{\partial x^2} = 4 > 0$; $\frac{\partial^2 f}{\partial z^2} = 0$; $\frac{\partial^2 f}{\partial x \partial z} = \frac{\partial^2 f}{\partial z \partial x} = 3$, otteniamo

$\|H(x,z)\| = \left\| \begin{matrix} 4 & 3 \\ 3 & 0 \end{matrix} \right\| = 0 - 9 = -9 < 0$. La funzione non è convessa.

Altro esempio.

Sia $f(x,z) = \alpha \ln(x) + \beta \ln(z)$ definita sul campo di esistenza $A = \left\{ (x,z) \in R^2 : x > 0, z > 0 \right\}$ e valutata per α e β positivi. Applichiamo il criterio delle derivate seconde e della matrice Hessiana ricordando che il gradiente è $\nabla_f = (\alpha / x \quad \beta / z)^T$

$\frac{\partial^2 f}{\partial x^2} = -\frac{\alpha}{x^2} < 0$; $\frac{\partial^2 f}{\partial z^2} = -\frac{\beta}{z^2} < 0$; $\frac{\partial^2 f}{\partial x \partial z} = \frac{\partial^2 f}{\partial z \partial x} = 0$. Le derivate seconde dirette sono negative.

Allora $\|H(x,z)\| = \left\| \begin{matrix} -\frac{\alpha}{x^2} & 0 \\ 0 & -\frac{\beta}{z^2} \end{matrix} \right\| = \frac{\alpha\beta}{(xz)^2} > 0$. La funzione è concava per ogni *x* e *z* appartenenti all'insieme *A*. Le condizioni di cui al riquadro precedente torneranno buone in seguito.

2.2. Forme quadratiche elementari, concavità o convessità della funzione e determinatezza della matrice Hessiana

Diamo una trattazione più generale della questione vista nel paragrafo precedente. Supponiamo che sia *z*: A → R con $A \subseteq R^2$ insieme aperto e convesso, con *f* di classe C^2:

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

Il nostro problema è capire se essa è concava o convessa (o quasi concava, quasi convessa, ecc.).
Se la funzione fosse $z = f(x)$ ed esistessero (finite) la derivata prima e seconda, diremmo che la *f* è concava (convessa) se la derivata seconda è negativa (positiva). Come adattiamo questo test al caso di funzione definita su R^2? Prima di procedere ricordiamo che $dz/dx = f_x$ può essere riproposta in termini di differenziale come $dz = f_x dx$ e che la derivata seconda può essere riproposta come differenziale secondo $d^2z = f_{xx}(dx)^2$. Allora in *R* la funzione è concava (convessa) se d^2z è negativo (positivo).
Ritorniamo adesso alla funzione $z = f(x,y)$. Prendiamone il differenziale totale primo che scriviamo come

$$dz = f_x(x,y)dx + f_y(x,y)dy$$

e successivamente il differenziale totale del secondo ordine

$$\begin{aligned} d^2z = d(dz) &= \frac{\partial(dz)}{\partial x}dx + \frac{\partial(dz)}{\partial y}dy \\ &= f_{xx}(dx)^2 + f_{xy}dxdy + f_{yx}dydx + f_{yy}(dy)^2 \end{aligned}$$

Applicando il teorema di Young alle derivate seconde incrociate

$$d^2z = f_{xx}(dx)^2 + 2f_{xy}dxdy + f_{yy}(dy)^2$$

Mostriamo che d^2x corrisponde ad una forma quadratica molto semplice. Chiamiamo

$$\text{Variabili: } \begin{cases} u = dx \\ v = dy \end{cases}$$

$$\text{Coefficienti: } \begin{cases} a = f_{xx} \\ b = f_{yy} \\ h = f_{xy} \end{cases}$$

Rimontando i pezzi

$$d^2z = q = au^2 + 2huv + bv^2$$

$$= au^2 + huv + huv + bv^2$$

Ovvero

$$d^2z = \underbrace{\begin{bmatrix} u & v \end{bmatrix}}_{1\times 2} \underbrace{\begin{bmatrix} a & h \\ h & b \end{bmatrix}}_{2\times 2} \underbrace{\begin{bmatrix} u \\ v \end{bmatrix}}_{2\times 1}$$

Ovvero

$$d^2z = \underbrace{\begin{bmatrix} dx & dy \end{bmatrix}}_{1\times 2} \underbrace{\begin{bmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{bmatrix}}_{2\times 2} \underbrace{\begin{bmatrix} dx \\ dy \end{bmatrix}}_{2\times 1}$$

$$= \boldsymbol{d'Hd}$$

Allora d^2z o q definiscono in modo semplificato gli oggetti noti come *forme* [espressioni polinomiali con componenti di grado uniforme] *quadratiche* [ogni componente è al più di secondo grado], per le quali esistono criteri ben definiti per stabilire se esse hanno segno positivo, negativo, non positivo, non negativo purché dx e dy non siano entrambi nulli. Nel nostro caso gli elementi rilevanti del polinomio (le variabili u e v) o sono di secondo grado o sono lineari ma moltiplicati tra loro[7]. Prima di procedere presentiamo il seguente esempio.

[7] In R, per la funzione $f(x)$ di classe C^2 avevamo il "test" del segno della derivata seconda. In R^2 lo sostituiamo con un "test" analogo, esteso però attraverso differenziale secondo a tutte le variabili. Se da $\boldsymbol{d'Hd}$ eliminiamo i termini in y supponendo che z dipenda solo da x torniamo al caso trattato in R, ovvero alla derivata seconda.

Esempio

Si determini il segno della forma quadratica

$$q(\boldsymbol{x}) = x_1^2 + 5x_2^2 + 6x_3^2 + 4x_1x_2 - 2x_1x_3$$

Per prima cosa dobbiamo trasformare la $q(\boldsymbol{x})$ nella versione $\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x}$. Riscriviamo la $q(\boldsymbol{x})$ come segue

$$\begin{aligned} q(\boldsymbol{x}) &= x_1^2 + 5x_2^2 + 6x_3^2 + 4x_1x_2 - 2x_1x_3 \\ &= x_1^2 + 5x_2^2 + 6x_3^2 + 2x_1x_2 + 2x_2x_1 - x_1x_3 - x_3x_1 \end{aligned}$$

Poniamo $\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ e costruiamo la matrice simmetrica dei coefficienti $\boldsymbol{A}$

$$\boldsymbol{A} = \begin{bmatrix} \text{Coefficiente di } x_1^2 & \text{Coefficiente di } x_1x_2 & \text{Coefficiente di } x_1x_3 \\ \text{Coefficiente di } x_2x_1 & \text{Coefficiente di } x_2^2 & \text{Coefficiente di } x_2x_2 \\ \text{Coefficiente di } x_3x_1 & \text{Coefficiente di } x_3x_2 & \text{Coefficiente di } x_3^2 \end{bmatrix} = \begin{bmatrix} 1 & 2 & -1 \\ 2 & 5 & 0 \\ -1 & 0 & 6 \end{bmatrix}. \text{ Allora}$$

$$q(\boldsymbol{x}) = \underbrace{\underbrace{\underbrace{\begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix}}_{1\times3} \begin{bmatrix} \underbrace{\begin{matrix} 1 & 2 & -1 \\ 2 & 5 & 0 \\ -1 & 0 & 6 \end{matrix}}_{3\times3} \end{bmatrix}}_{1\times3} \begin{bmatrix} x_1 \\ x_2 \\ \underbrace{x_3}_{3\times1} \end{bmatrix}}_{1\times1}$$

Da cui

$$\|A_1\| = 1 > 0; \quad \|A_2\| = \begin{Vmatrix} 1 & 2 \\ 2 & 5 \end{Vmatrix} = 1 > 0; \quad \|A\| = \begin{Vmatrix} 1 & 2 & -1 \\ 2 & 5 & 0 \\ -1 & 0 & 6 \end{Vmatrix} = 1 > 0$$

La $q(\boldsymbol{x})$ ha tutti i minori principali positivi e quindi è positivamente definita[8]. Lo stesso risultato si dimostra in termini differenziali (vedi sopra). Infatti

[8] Gli autovalori di $\boldsymbol{A}$ sono (approssimando) $\lambda_1 = 6.4$, $\lambda_2 = 5.6$, $\lambda_3 = 0.03$: tutti positivi. Anche il test degli autovalori conferma il segno di q. Vedi pro memoria alla fine della dispensa. Anche gli autovalori della matrice dei coefficienti della versione differenziale sono tutti positivi. Calcolarli e spiegare la relazione con quelli precedenti.

$$d^2q(\boldsymbol{x}) = \underbrace{\underbrace{\underbrace{\begin{bmatrix} dx_1 & dx_2 & dx_3 \end{bmatrix}}_{1\times3} \begin{bmatrix} \underbrace{\begin{matrix} 2 & 4 & -2 \\ 4 & 10 & 0 \\ -2 & 0 & 12 \end{matrix}}_{3\times3} \end{bmatrix}}_{1\times3} \begin{bmatrix} \underbrace{\begin{matrix} dx_1 \\ dx_2 \\ dx_3 \end{matrix}}_{3\times1} \end{bmatrix}}_{1\times1}$$

Da cui

$$\|A_1\| = 2 > 0; \ \|A_2\| = \begin{Vmatrix} 2 & 4 \\ 4 & 10 \end{Vmatrix} = 4 > 0; \ \|A\| = \begin{Vmatrix} 2 & 4 & -2 \\ 4 & 10 & 0 \\ -2 & 0 & 12 \end{Vmatrix} = 8 > 0 .$$

Torniamo al problema della definizione delle condizioni di concavità/convessità e al connesso problema della ricerca dei massimi/minimi. In R (ovvero con $z = f(x)$, ad esempio) per un massimo si richiedeva $d^2z < 0$ e per un minimo $d^2z > 0$. Se avevamo già fatto il test di concavità/convessità della funzione già conoscevamo il segno di d^2z. Allora in presenza di una funzione già riconosciuta come concava (come da segno della derivata seconda ovvero del differenziale di secondo ordine) potevamo dire che nel punto che azzera la derivata prima la funzione è massima e che la derivata prima azzerata è condizione **sia necessaria sia sufficiente** per il massimo. Stesso ragionamento vale (adattando) per un minimo di funzione convessa. Quindi: se una funzione è concava (convessa) la derivata prima azzerata è condizione necessaria e sufficiente per un massimo (minimo). Vi è quindi una stretta corrispondenza tra concavità/convessità della funzione e condizioni necessarie e sufficienti per il suo massimo/minimo.
Ma come ce la caviamo in R^2? In R^2, dx e dy devono ovviamente essere considerate delle variabili mentre le derivate seconde, che assumeranno specifici valori nei punti di massimo/minimo da studiare, altrettanto ovviamente saranno considerate delle costanti (vedi sopra). Noi siamo quindi interessati al segno di $q = d^2z$ in modo da poter tornare ad applicare anche in R^2 lo stesso ragionamento fatto per R (in R^2: sufficienza dell'azzeramento di $\nabla_f(x, z)$) e ci poniamo la domanda: che restrizioni occorre imporre ad a, b e h (ovvero alle derivate seconde e alla seconda mista o incrociata) per avere un segno definito di q visto che u e v (ovvero dx e dy) possono variare? Diciamo che

- La forma quadratica q è detta positiva definita se $\boldsymbol{H}$ (A nell'esempio) è invariabilmente positiva **> 0**
- La forma quadratica q è detta positiva semi-definita se $\boldsymbol{H}$ (A nell'esempio) è invariabilmente non-negativa **≥ 0**
- La forma quadratica q è detta negativa semi-definita se $\boldsymbol{H}$ (A nell'esempio) è invariabilmente non-positiva **≤ 0**
- La forma quadratica q è detta negativa definita se $\boldsymbol{H}$ (A nell'esempio) è invariabilmente negativa **< 0**

Se q cambia segno quando le variabili assumono valori diversi si dice che q è non definita.
Chiaramente il caso della forma definita o semi-definita **positiva** è associato alle condizioni per la **convessità** della funzione e quindi, di conseguenza, a quelle richieste per un **minimo** nei valori che azzerano $\nabla_f(x, z)$ e quello della forma definita o semi-definita **negativa** alle condizioni per la **concavità** o quasi concavità della funzione e quindi, di conseguenza, a quelle per un **massimo** nei valori che azzerano $\nabla_f(x, z)$. Quindi, come anticipato nel paragrafo precedente, il segno di q è associato alla convessità o concavità della *f(x, y)* e alle condizioni per i massimi e i minimi Quando dovremo studiare funzioni utilità o di produzione tornerà utile disporre di questo strumento.

Iniziamo con la definizione della condizione per la definizione del segno di q.

Caso concavità.

La *f(x, y)* è concava se $q < 0$ (ovvero è definita negativa). Quando q è definita negativa (strettamente)? **Quando *H* è a sua volta definita negativa**. Ovvero, **seguendo il criterio di Silvester,** quando il primo minore di Nord-Ovest o di testa (lo chiamo H_1) è negativo e il secondo di Nord-Ovest o di testa (che in R^2 corrisponde al determinante di H) è positivo. Quindi, in R^2, quando

$$\|H_1\| = \|f_{xx}\| < 0$$

$$\|H_2\| = \left\| \begin{matrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{matrix} \right\| > 0$$

In R^N il criterio richiede che:

a) Tutti i minori di Nord-Ovest o di testa di indice dispari (1, 3, 5, …) siano negativi
b) Tutti i minori di Nord-Ovest o di testa di indice pari (2, 4, 6, …) siano positivi

Occorre in pratica un'alternanza dei segni dei determinanti dei minori di testa ma sempre partendo da $\|H_1\| < 0$ e proseguendo sino a $(-1)^N \|H_N\|$.

ESEMPIO

Sia, per X e Y maggiori di zero, $f(X, Y) = \sqrt[3]{XY}$ di classe C^2 su un insieme aperto e convesso che appartiene a R^2. Il grafico della funzione è

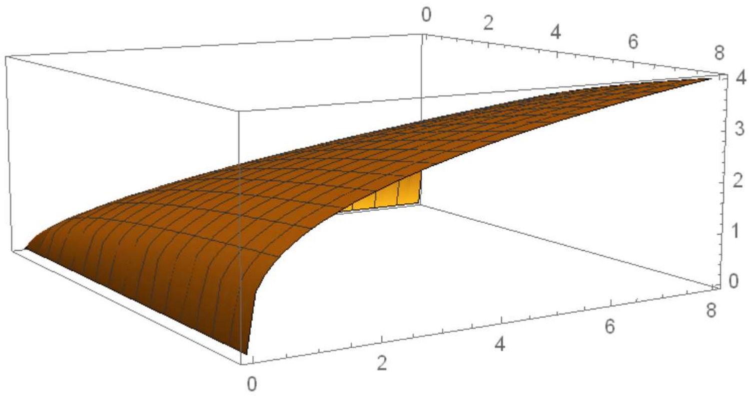


Fig. 7 Grafico di $f(X, Y) = \sqrt[3]{XY}$

Valutare se essa è concava o convessa. Abbiamo

$$H = \begin{bmatrix} -\dfrac{2Y^2}{9(XY)^{5/3}} & \dfrac{1}{9(XY)^{2/3}} \\ \dfrac{1}{9(XY)^{2/3}} & -\dfrac{2X^2}{9(XY)^{5/3}} \end{bmatrix}$$

E quindi

$\|H_1\| = -\frac{2Y^2}{9(XY)^{5/3}} < 0$ (Δ del primo minore di Nord-Ovest dispari)

$\|H_2\| = \frac{(XY)^{2/3}}{27\ {}^2Y^2} > 0$ (Δ del rimo minore di Nord-Ovest pari)

H è negativa definita e quindi *f(x, y)* è strettamente concava. Ripetere l'esercizio con $\ln(x_1x_2)^2$.

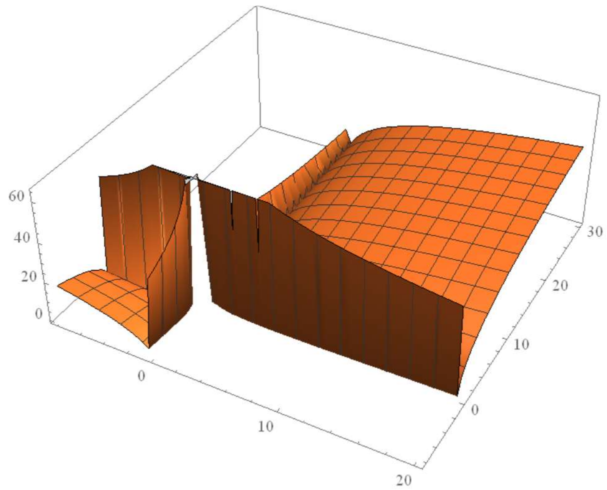


Fig. 9 Grafico di $f(x_1x_2) = \ln(x_1x_2)^2$

Caso Convessità.

La *f(x, y)* è convessa se *q* > 0 (definita positiva). Quando *q* è positiva definita (strettamente)? **Quando *H* è a sua volta definita positiva**. Ovvero, **seguendo il criterio di Silvester,** quando tutti i minori di Nord-Ovest o di testa (li chiamo H_i) sono strettamente positivi. Quindi in R^2 quando

$$\|H_1\| = \|f_{xx}\| > 0$$

$$\|H_2\| = \left\|\begin{matrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{matrix}\right\| > 0$$

In R^N il criterio richiede che:

c) Tutti i minori di Nord-Ovest o di testa di indice dispari (1, 3, 5, …) siano positivi
d) Tutti i minori di Nord-Ovest o di testa di indice pari (2, 4, 6, …) siano positivi

Occorre in pratica che non ci sia alternanza dei segni dei determinanti dei minori di testa ma partendo in questo caso da $\|H_1\| > 0$ e proseguendo sino a $\|H_N\| > 0$.

Esempio

Sia $f(x_1, x_2) = 2x_1^3 + x_2^4$ di classe C^2 su un set aperto e convesso S. Valutare se essa è concava o convessa nell'intervallo ($0, \infty$). In questo caso:

$$H = \begin{bmatrix} 12x_1 & 0 \\ 0 & 12x_2 \end{bmatrix}$$

E quindi

$\|H_1\| = \|f_{x_1x_1}\| = 12x_1 > 0$ (primo minore di Nord-Ovest dispari)
$\|H_2\| = \left\| \begin{matrix} 12x_1 & 0 \\ 0 & 12x_2 \end{matrix} \right\| = 144x_1x_2^2 > 0$ (primo minore di Nord-Ovest pari)

Tutti minori sono positivi. La funzione è convessa nell'intervallo considerato.

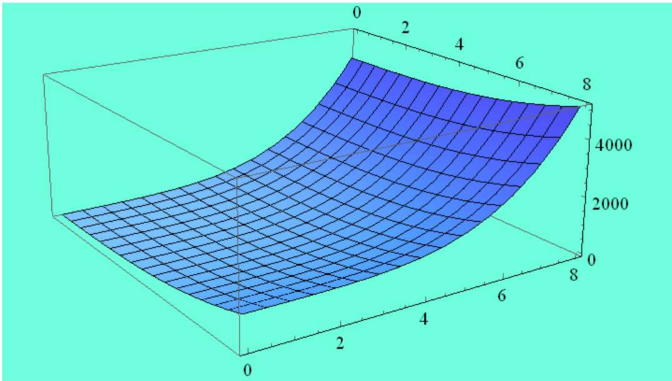


Fig. 9 Grafico di $f(x_1, x_2) = 2x_1^3 + x_2^4$

Caso quasi-concavità.

In R una funzione *f(x)* è quasi-concava (*quasi* vuol dire "come se fosse") se

$$f(\lambda x + (1-\lambda)y) \geq \min\{f(x), f(y)\} \; \forall \lambda \in (0,1).$$

In altre parole *f(x)* sta sempre sopra il valore minimo che essa può assumere nel dominio. La concavità e la convessità o la linearità sono "sotto casi". In R l'esempio più noto è la densità della normale (mono variata) che è quasi concava ma non concava. In R^2 un esempio è la densità della distribuzione congiunta normale bivariata. Se applicata alla teoria dell'utilità, quella di quasi-concavità è un'ipotesi piuttosto pilatesca sulla struttura ordinata delle preferenze; se applicata alla teoria delle funzioni di produzione, quella della quasi concavità permette di ipotizzare che la funzione mostri qualsiasi regime di scala. Per questi motivi la quasi concavità è **molto usata in microeconomia**.
La *f(x, y)* è quasi concava se $q \leq 0$ (semi definita negativa). Quando q è semi definita negativa? **Quando *H* è semi definita negativa**. Ovvero, **seguendo il criterio di Silvester,** quando tutti i minori di Nord-Ovest o di Testa di ordine pari sono non negativi e tutti i minori di Nord-Ovest o di Testa di ordine dispari sono non positivi. Quindi in R^2

$$\|H_1\| = \|f_{xx}\| \leq 0$$

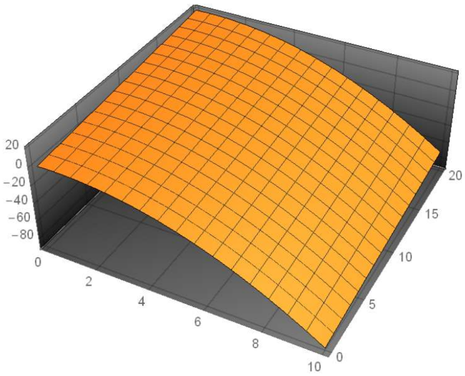
$$\|H_2\| = \left\| \begin{matrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{matrix} \right\| \geq 0$$

In R^N il criterio richiede che:

e) Tutti i minori di Nord-Ovest o di testa di indice dispari (1, 3, 5, …) siano non positivi
f) Tutti i minori di Nord-Ovest o di testa di indice pari (2, 4, 6, …) siano non negativi.

ESEMPIO

Sia $f(x_1, x_2) = x_1 - x_2 - x_1^2$ di classe C^2 su un set aperto e convesso S .



Valutare se essa è concava o convessa in S. Abbiamo

$\|H_1\| = \|f_{x_1x_1}\| = -2$ (primo minore dispari)
$\|H_2\| = \begin{Vmatrix} -2 & 0 \\ 0 & 0 \end{Vmatrix} = 0$ (primo minore pari)

H è negativa semi-definita e quindi $f(x_1, x_2)$ è quasi-concava.

Caso quasi-convessità.

In R una funzione $f(x)$ è quasi-convessa (*quasi* vuol dire "come se fosse") se comunque scelti x e y nel dominio (convesso)

$$f(\lambda x + (1-\lambda)y) \leq \max\{f(x), f(y)\} \ \forall \lambda \in (0,1) .$$

Ovvero la funzione è quasi convessa se le immagini di qualsiasi punto $\lambda x + (1-\lambda)y$ mediante f non supera alcun valore tra $f(x)$ e $f(y)$. In altre parole la funzione sta sempre al di sotto del valore massimo che essa può assumere nel dominio.
La $f(x, y)$ è quasi convessa se $q \geq 0$ (semi definita positiva). Quando q è semi definita positiva? **Quando H è semi definita positiva**. Ovvero, **seguendo il criterio di Silvester,** quando tutti i minori di Nord-Ovest o di testa di qualsiasi ordine sono non negativi. Quindi in R^2

$$\|H_1\| = \|f_{xx}\| \geq 0$$

$$\|H_2\| = \begin{Vmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{Vmatrix} \geq 0$$

In R^N il criterio richiede che:

g) Tutti i minori di Nord-Ovest o di testa di qualsiasi indice siano non positivi.

Esempio da microeconomia

Sia, su un dominio tutto strettamente positivo, la seguente funzione di utilità $U = 100 + 2\sqrt[3]{XY}$ dove X sono pere e Y camicie. La matrice hessiana è

$$H = \begin{bmatrix} -\frac{4Y^2}{9(XY)^{5/3}} & \frac{2}{9(XY)^{2/3}} \\ \frac{2}{9(XY)^{2/3}} & -\frac{4X^2}{9(XY)^{5/3}} \end{bmatrix}$$

Il primo minore di N-O dispari è certamente negativo.
Il primo minore di N-O pari (determinante) invece è

$$\Delta = \frac{4(XY)^{2/3}}{27X^2Y^2} > 0$$

Confrontare con la precedente funzione di utilità $U = \sqrt[3]{XY}$.
Che tipo di trasformazione è la $U = 100 + 2\sqrt[3]{XY}$?

Una (e una sola) indicazione bibliografica in regalo

Le/gli studentesse/ti economiste/i che ritengono che stiamo parlando di cose serie (mi riferisco all'applicabilità in economia non allo studio della concavità/convessità in matematica) non dovrebbero fare a meno del seguente testo (vecchio, e perciò affidabile…)

M. Avriel, W. E. Diewert, S. Schaible e I. Zang, *Generalized Concavity*, siam (society for industrial and applied mathematics) Philadelphia 2010 (prima edizione 1988).
In particolare le persone di cui sopra dovrebbero studiare il capitolo di Diewert.

2.3 La ricerca di un massimo/minimo vincolato in R^2

Supponiamo adesso che *x* e *z* siano "legate" tra loro in qualche modo e che la funzione *f* debba sempre (quindi anche nel suo massimo e nel suo minimo, se ci sono) rispettare questo legame che esprimiamo come una nuova funzione delle due variabili. Chiamiamo

$g(x, z)$

tale nuova funzione che potrebbe essere posta pari ad una costante *C*, eventualmente pari a zero. A questo punto la ricerca di un massimo/minimo di *f* non è più libera ma deve rispettare il legame imposto alla *x* e alla *z* attraverso la *g*, che da adesso in poi chiamiamo vincolo. Poniamo che la *g* sia

$g(x, z) \equiv 4x - 3z - 10 = 0$

e la *f* sia

$f(x, z) = 2x^2 + 3zx - 2z$

Come si vede il vincolo è espresso (in questo caso!) come uguaglianza. In pratica è come se la *g* ci dicesse: la *x* e la *z* devono stare <u>sempre</u> tra loro in un rapporto determinato in modo esatto dalla *g*. Non trattiamo qui il caso in cui il vincolo viene espresso in termini di disuguaglianza (del tipo $x \geq (3/4)z + 10/4$).
L'esistenza del vincolo trasforma il nostro problema iniziale dalla ricerca di un massimo/minimo libero (non vincolato) nella ricerca di un massimo/minimo vincolato.
Come procedere? Trattiamo solo i due modi principali: Sostituzione e funzione Lagrangiana.

<u>Sostituzione</u>

Sostituiamo *g* in *f*, dopo aver risolto la *g* per una delle due variabili (ad esempio esprimendo nella *g* la *x* in termini di *z*). Dopo la sostituzione la *f* diventa

$$f(x(z), z) = 2\left[\underbrace{\frac{3}{4}z + \frac{5}{2}}_{\substack{\text{Valore di } x \text{ in termini di } z \\ \text{come da funzione g}}}\right]^2 + 3z\left[\underbrace{\frac{3}{4}z + \frac{5}{2}}_{\substack{\text{Valore di } x \text{ in termini di } z \\ \text{come da funzione g}}}\right] - 2z$$

La *f* è adesso funzione della sola *z* e derivandola rispetto a *z* otteniamo

$$\frac{d}{dz}f(x(z), z) = 2\left[2\left(\frac{3}{4}z + \frac{5}{2}\right)\frac{3}{4}\right] + 3\left(\frac{3}{4}z + \frac{5}{2}\right) + 3z\left(\frac{3}{4}\right) - 2\,.$$

Uguagliando la derivata a zero otteniamo

$z\hat{} = -\,52/27$.

Sostituendo $z\hat{} = -\,52/27$ nel vincolo *g* otteniamo $x\hat{} = 19/18$. Come si vede $z\hat{}$ è diverso da $z^* = -22/9$ (da dove sbuca?) e $x\hat{}$ è a sua volta diverso da $x^* = 11/6$ (da dove sbuca?). Nel punto (– 52/27, 19/18) abbiamo trovato un massimo o un minimo? Aspettiamo a rispondere. Per il momento ripetete il test della derivata seconda per valutare "intuitivamente", se P_V = (– 52/27, 19/18) genera un massimo o un minimo e poi (se avete fretta) andate a consultare le condizioni di cui alla tabella 2. Naturalmente, anche il valore di *f* in P_V sarà diverso (maggiore o minore?) di *f* in P_L = (– 22/9, 11/6). (Farlo per esercizio).
Il metodo della sostituzione sembra semplice ma non lo è in generale. Ad esempio quando la *g* non è lineare o quando il legame tra le variabili (che, per inciso, non è sempre detto che siano solo due…) è una diseguaglianza. Occorre conoscere almeno un altro metodo.

<u>Funzione Lagrangiana</u>

La *f* e la *g* devono essere "messe assieme" per formare una nuova funzione da massimizzare/minimizzare rispetto a *x* e *z*. In che modo? Sommandole attraverso l'imposizione di un particolare legame tra di loro in modo da formare **una nuova funzione** (composta dalla somma di *f* e di *g*) che sarà derivata rispetto a *x* e *z* per trovare il massimo/minimo. La somma algebrica delle due funzioni avviene usando una **costante positiva** (chiamiamola λ) che moltiplica *g* (opportunamente riscritta) e la "aggiunge" alla *f* (che non viene modificata). In sintesi, si forma la nuova funzione, chiamiamola Λ, in questi due possibili modi (equivalenti nel risultato).

Versione a) riscrivo *g* portando tutti i termini a destra del segno di eguaglianza

$g(x, z) \equiv 0 = -4x + 3z + 10$

e formiamo la funzione Lagrangiana in questo modo

$$\Lambda(x, z, \lambda) = 2x^2 + 3xz - 2z + \lambda\left[-4x + 3z + 10\right]$$

Versione b) riscriviamo *g* portando tutti i termini a Sx del segno di eguaglianza

e formo la funzione Lagrangiana in questo modo

$$\Lambda(x, z, \lambda) = 2x^2 + 3xz - 2z - \lambda\left[4x - 3z - 10\right]$$

Ovviamente a) e b) sono del tutto equivalenti. ***In generale la funzione Lagrangiana*** sarà

$$\begin{aligned} \Lambda(x, z, \lambda) &= f(x, z) - \lambda\left[g(x, z)\right] \qquad a) \\ \Lambda(x, z, \lambda) &= f(x, z) + \lambda\left[-g(x, z)\right] \qquad b) \end{aligned}$$

I valori di *x* e z che danno un massimo/minimo della Λ si trovano risolvendo il sistema di 3 equazioni in 3 incognite generato derivando Λ rispetto a *x*, *z* e λ. Quindi per risolvere il problema della ricerca del MAX o del MIN vincolati si procede così:

MAX o MIN di *f(x, z)* con il vincolo *g(x, z)* = 0 :

1) si forma la funzione Lagrangiana,
2) la si deriva parzialmente rispetto a *x*, *z*, e λ
3) si risolve il sistema in 3 equazioni e 3 incognite definito azzerando le derivate parziali di Λ
4) si trovano così i valori di *x* e *z* che massimizzano (minimizzano) la f
5) **ma non è finita** ….

Nel nostro caso (usando la versione *a*)

$$\Lambda(x,z,\lambda) = 2x^2 + 3xz - 2z + \lambda\left[-4x + 3z + 10\right] \quad (1)$$

$$\frac{\partial \Lambda}{\partial x} = 4x + 3z - 4\lambda = 0 \quad (2)$$

$$\frac{\partial \Lambda}{\partial z} = 3x - 2 + 3\lambda = 0 \quad (3)$$

$$\frac{\partial \Lambda}{\partial \lambda} = -4x + 3z + 10 = 0 \quad (4)$$

Ovvero

$$\begin{bmatrix} 4 & 3 & -4 \\ 3 & 0 & 3 \\ -4 & 3 & 0 \end{bmatrix}\begin{bmatrix} x \\ z \\ \lambda \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \\ -10 \end{bmatrix}$$

Il determinante della matrice dei coefficienti è – 108 ≠ 0 e i valori cercati di x, z e λ sono $x^\wedge$ = 19/18, $z^\wedge$ = – 52/27 e λ^ = – 7/18. Sostituendo questi valori in *Λ(x, z, λ)* otteniamo che

$$\Lambda(x^\wedge, z^\wedge, \lambda) = \underbrace{2\left(\frac{19}{18}\right)^2 - 3\left(\frac{19}{18}\right)\left(\frac{52}{27}\right) + 2\left(\frac{52}{27}\right)}_{f(x^\wedge, z^\wedge)} + \underbrace{\left(\frac{2}{3} - \frac{19}{18}\right)}_{\lambda(x^\wedge)}\left[\underbrace{-4\left(\frac{19}{18}\right) + 3\left(-\frac{52}{27}\right) + 10}_{g(x^\wedge, z^\wedge)}\right]$$
$$\underbrace{-1/54}_{f(x^\wedge, z^\wedge)} + \underbrace{0}_{\lambda(x^\wedge) g(x^\wedge, z^\wedge)}$$
$$= -1/54$$

La funzione ottenuta impiegando i valori "ottimi" $f(x^\wedge, z^\wedge, \lambda^\wedge)$ prende il nome di funzione di Massimo (o Minimo) valore e tornerà utile in vari semplici casi. Possiamo semplicemente chiamarla Funzione di Valore (*value function*) di un problema di ottimo vincolato poiché essa indica il valore raggiunto dalla massimizzanda o minimizzanda in funzione del parametro del problema (10 nel nostro caso; ma al numero 10 potevamo sostituire una qualsiasi costante "che varia" quale la R del vincolo di bilancio del consumatore degli esercizi di microeconomia). In questo caso, avremmo avuto

$$x^\wedge = \frac{8 + 3R}{36} \qquad z^\wedge = \frac{8 - 6R}{27} \qquad \lambda^\wedge = \frac{12 + 4R}{108}$$

che corrispondono ai precedenti per R = 10. Quindi usando R al posto di 10:

$$\Lambda(\hat{x},\hat{z},\lambda) = \underbrace{2\left(\frac{19}{18}\right)^2 - 3\left(\frac{19}{18}\right)\left(\frac{52}{27}\right) + 2\left(\frac{52}{27}\right)}_{f(\hat{x},\hat{z})} + \underbrace{\left(-\frac{7}{18}\right)}_{\lambda(\hat{x})}\left[\underbrace{-4\left(\frac{19}{18}\right) + 3\left(-\frac{52}{27}\right) + R}_{g(\hat{x},\hat{z},R)}\right]$$
$$= \frac{209 - 21R}{54}$$

che coincide con il precedente valore se poniamo $R = 10$. Da notare che

$$\frac{d}{dR}\Lambda(\hat{x},\hat{z},\lambda) = -\frac{21}{54} = -\frac{7}{18} = \hat{\lambda}$$

2.1 Massimi o minimi? (Ovvero: perché non è finita al punto 4)

Torniamo al problema iniziale. Nei due casi (assenza/presenza di vincolo) abbiamo ottenuto un minimo o un massimo? Ci si ricorderà che nel primo caso trattato dopo il calcolo delle derivate seconde avevamo lasciato il giudizio un po' in sospeso anche se dal grafico delle *f* era abbastanza chiaro che eravamo in presenza di un minimo. Adesso però valutiamo la cosa con un pizzico di rigore in più. Le condizioni che cerchiamo per verificare se il punto ottenuto costituisce un massimo o un minimo cambiano a seconda che il problema sia formulato con o senza vincolo. Separiamo i due casi.

3 Caso del massimo/minimo vincolato

Le condizioni di secondo ordine nel problema con il vincolo (espresso come uguaglianza) sono solo apparentemente più incasinate, ma per l'appunto è solo apparenza. Ciò nonostante, le condizioni di secondo ordine in un massimo/minimo con vincolo non sono la riproduzione esatta delle condizioni per gli estremi non vincolati.

3.1 Caso R^2

Siano continue, C^2 e definite sullo stesso campo di esistenza le funzioni $f(x,y)$ e $g(x,y)$. Cerchiamo i punti di singolarità della prima con il vincolo della seconda.

Formiamo la lagrangiana

$$\Lambda = f(x,y) + \lambda g(x,y)$$

Nei punti del piano *xy* appartenenti al vincolo la funzione Λ coincide con la f e ciò trasforma un problema di massimo/minimo vincolato in un problema non vincolato in 3 variabili.

Ricaviamo il/i punto/i di singolarità dalla Λ

$$\begin{cases} \Lambda'_x = 0 \\ \Lambda'_y = 0 \\ \Lambda'_\lambda = 0 \end{cases} \Rightarrow \begin{cases} f'_x + \lambda g'_x = 0 \\ f'_y + \lambda g'_y = 0 \\ g = 0 \end{cases} \Rightarrow P_0(x_0, y_0, \lambda_0)$$

Si costruisce una nuova matrice Hessiana

$$\tilde{H} = \begin{bmatrix} \Lambda''_{xx} & \Lambda''_{xy} & \Lambda''_{x\lambda} \\ \Lambda''_{yx} & \Lambda''_{yy} & \Lambda''_{y\lambda} \\ \Lambda''_{\lambda x} & \Lambda''_{\lambda y} & \Lambda''_{\lambda\lambda} \end{bmatrix} \text{ scritta anche nella forma } \begin{bmatrix} \Lambda''_{\lambda\lambda} & \Lambda''_{\lambda x} & \Lambda''_{\lambda z} \\ \Lambda''_{x\lambda} & \Lambda''_{xx} & \Lambda''_{xz} \\ \Lambda''_{z\lambda} & \Lambda''_{zx} & \Lambda''_{zz} \end{bmatrix}$$

Si può dimostrare che:

SE det[$\tilde{H}$] valutato in $P_0 < 0$ allora P_0 è un minimo vincolato

SE det[$\tilde{H}$] valutato in $P_0 > 0$ allora P_0 è un massimo vincolato

Nel nostro precedente esempio

$$\Lambda(x,z,\lambda) = 2x^2 + 3xz - 2z - \lambda\left[4x - 3z - 10\right]$$

i valori di x e z dati da x^ = 19/18 e z^ = -52/27 definiscono un minimo. Rifacciamo la valutazione della natura di tali punti di singolarità con il determinate Hessiano che verrà chiamato "orlato".

$$\tilde{H} = \begin{bmatrix} \Lambda''_{\lambda\lambda} & \Lambda''_{\lambda x} & \Lambda''_{\lambda z} \\ \Lambda''_{x\lambda} & \Lambda''_{xx} & \Lambda''_{xz} \\ \Lambda''_{z\lambda} & \Lambda''_{zx} & \Lambda''_{zz} \end{bmatrix} = \begin{bmatrix} 0 & -4 & 3 \\ -4 & 4 & 3 \\ 3 & 3 & 0 \end{bmatrix} \Rightarrow \left\| \tilde{H} \right\| = -108 < 0$$

Emerge subito che le condizioni per un massimo/minimo vincolato sono diverse da quelle per il massimo/minimo libero). Occorre allora ricavare più in generale le condizioni che stiamo cercando. Procediamo passo passo alla costruzione di $\tilde{H}$.

La prima cosa da ricordare nella costruzione di una tabella per il caso vincolato è che le derivate parziali che calcoliamo non sono derivate parziale di f ma di Λ. Detto questo, la prima riga della tabella delle condizioni esposte per il problema non vincolato resta inalterato. Per costruire gli elementi della seconda riga procediamo come segue

a) Calcoliamo le derivate seconde delle condizioni di primo ordine del problema. Ovvero, deriviamo ancora la (2) la (3) e la (4) come segue

$$\frac{\partial^2\Lambda}{\partial x^2} = 4, \frac{\partial^2\Lambda}{\partial z^2} = 0, \underbrace{\frac{\partial^2\Lambda}{\partial x\partial z} = \frac{\partial^2\Lambda}{\partial z\partial x}}_{\text{Come mai?}} = 3$$

e le usiamo per costruire una matrice Hessiana delle derivate seconde:

$$\underbrace{\begin{bmatrix} \frac{\partial^2\Lambda}{\partial x^2} & \frac{\partial^2\Lambda}{\partial x\partial z} \\ \frac{\partial^2\Lambda}{\partial z\partial x} & \frac{\partial^2\Lambda}{\partial z^2} \end{bmatrix}}_{2\times 2} = \underbrace{\begin{bmatrix} 4 & 3 \\ 3 & 0 \end{bmatrix}}_{2\times 2}$$

b) Completiamo la matrice Hessiana con l'aggiunta di una colonna e di una riga passanti per la posizione (1,1) inserendo 0 al vertice sinistro in alto e otteniamo la $\tilde{H}$:

$$\underbrace{\begin{bmatrix} 0 & ? & ? \\ ? & \frac{\partial^2 \Lambda}{\partial x^2} & \frac{\partial^2 \Lambda}{\partial x \partial z} \\ ? & \frac{\partial^2 \Lambda}{\partial z \partial x} & \frac{\partial^2 \Lambda}{\partial z^2} \end{bmatrix}}_{3\times 3} = \underbrace{\begin{bmatrix} 0 & ? & ? \\ ? & 4 & 3 \\ ? & 3 & 0 \end{bmatrix}}_{3\times 3}$$

Che ci mettiamo al posto dei ?, e perché?

c) Usiamo il vincolo *g(x, z)* e calcoliamo le derivate prime parziali rispetto a x e z come segue (nel nostro esempio): $\frac{\partial g}{\partial x} = -4$, $\frac{\partial g}{\partial z} = 3$, e riscriviamo la matrice Hessiana (che si chiamerà da adesso in poi *Hessiana orlata*) sostituendo nella prima riga e nella prima colonna i valori ottenuti:

$$\tilde{H} = \underbrace{\begin{bmatrix} 0 & \frac{\partial g}{\partial x} & \frac{\partial g}{\partial z} \\ \frac{\partial g}{\partial x} & \frac{\partial^2 \Lambda}{\partial x^2} & \frac{\partial^2 \Lambda}{\partial x \partial z} \\ \frac{\partial g}{\partial z} & \frac{\partial^2 \Lambda}{\partial z \partial x} & \frac{\partial^2 \Lambda}{\partial z^2} \end{bmatrix}}_{3\times 3} = \underbrace{\begin{bmatrix} 0 & -4 & 3 \\ -4 & 4 & 3 \\ 3 & 3 & 0 \end{bmatrix}}_{3\times 3}$$

d) Calcoliamo il determinante della Hessiana orlata che è

$$\left\| \tilde{H}_2 \right\| = \begin{Vmatrix} 0 & \frac{\partial g}{\partial x} & \frac{\partial g}{\partial z} \\ \frac{\partial g}{\partial x} & \frac{\partial^2 \Lambda}{\partial x^2} & \frac{\partial^2 \Lambda}{\partial x \partial z} \\ \frac{\partial g}{\partial z} & \frac{\partial^2 \Lambda}{\partial z \partial x} & \frac{\partial^2 \Lambda}{\partial z^2} \end{Vmatrix} = \begin{Vmatrix} 0 & -4 & 3 \\ -4 & 4 & 3 \\ 3 & 3 & 0 \end{Vmatrix} = 0 \begin{Vmatrix} 4 & 3 \\ 3 & 0 \end{Vmatrix} + 4 \begin{Vmatrix} -4 & 3 \\ 3 & 0 \end{Vmatrix} + 3 \begin{Vmatrix} -4 & 4 \\ 3 & 3 \end{Vmatrix} = -108 < 0$$

Se il problema fosse stato definito su un numero di variabili indipendenti maggiore di 2 ovviamente avremmo avuto più derivate parziali e una dimensione dell'hessiana orlata maggiore. E quindi avremmo avuto anche $\left\| \bar{H}_3 \right\|, \left\| \bar{H}_4 \right\|$, ecc.

e) A questo punto studiamo il segno dei determinanti (se ne abbiamo più di uno). Per un **Massimo** dobbiamo avere $\left\| \bar{H}_2 \right\| > 0, \left\| \bar{H}_3 \right\| < 0, \left\| \bar{H}_4 \right\| > 0$ ecc. (alternanza dei segni partendo dal positivo). Per un **Minimo** dobbiamo avere $\left\| \bar{H}_2 \right\|, \left\| \bar{H}_3 \right\|, \left\| \bar{H}_4 \right\|, \ldots < 0$ (tutti negativi). Nel nostro esempio abbiamo solo $\left\| \bar{H}_2 \right\| < 0$.

I risultati possono essere sintetizzati nella seguente tabella.

Tab.3 *Massimi e minimi vincolati del problema* $Max\,/\,Min\,f(x,z)$ soggetta a vincolo $g(x,z)=C$ *da cui* $\Lambda(x,z,\lambda)=f(x,z)+\lambda\left[C-g(x,z)\right]$ dove C è una costante (eventualmente pari a 0)

Condizione	Massimo vincolato di *f*	Minimo vincolato di *f*
Punto di singolarità	$\frac{\partial\Lambda}{\partial x}=\frac{\partial\Lambda}{\partial z}=\frac{\partial\Lambda}{\partial\lambda}=0$ + le altre derivate parziali se ci sono N variabili indipendenti	$\frac{\partial\Lambda}{\partial x}=\frac{\partial\Lambda}{\partial z}=\frac{\partial\Lambda}{\partial\lambda}=0$ + altre derivate parziali se ci sono N variabili indipendenti
Natura del punto	$\lVert\bar{H}_2\rVert>0,\lVert\bar{H}_3\rVert<0,\lVert\bar{H}_4\rVert>0,....,(-1)^N\lVert\bar{H}_N\rVert>0$ se ci sono *N* variabili indipendenti	$\lVert\bar{H}_2\rVert,\lVert\bar{H}_3\rVert,\lVert\bar{H}_4\rVert,...,\lVert\bar{H}_N\rVert<0$ se ci sono *N* variabili indipendenti)

N.B: Tutte le derivate seconde sono valutate nel punto di singolarità (cioè dove si annullano le derivate prime).

Esempio con l'utilità da consumo di beni

Concludiamo con un esempio di microeconomia relativo al consumatore. Supponiamo che Tizia/o trae utilità dal consumo di X e Y secondo una funzione di utilità $U(X,Y)$ avente le usuali comode proprietà. Supponiamo anche che la spesa totale possibile R sia data, e che Tizia/o non possa spendere né di più né di meno. Il vincolo di bilancio sarà

$$R=p_X X+p_Y Y$$

Tizia/o prima di uscire di casa per andare al mercato imposta la funzione lagrangiana

$$\Lambda(x,z,\lambda)=U(X,Y)+\lambda[R-p_X X-p_Y Y]$$

e calcola in primo luogo le derivate parziali (condizioni del primo ordine)

$$\frac{\partial\Lambda}{\partial\lambda}=R-p_X X-p_Y Y=0$$
$$\frac{\partial\Lambda}{\partial X}=\frac{\partial U}{\partial X}-\lambda p_X \qquad =0$$
$$\frac{\partial\Lambda}{\partial Y}=\frac{\partial U}{\partial Y}-\lambda p_Y \qquad =0$$

Dalle condizioni del primo ordine Tizia/o apprende che la sua utilità sarà massima se consumerà X e Y in quella quantità tale che

$$\frac{\frac{\partial U}{\partial X}}{\frac{\partial U}{\partial Y}}=\frac{p_X}{p_Y}$$

Tale condizione (per specifiche funzioni U, vedi le note E-Learning sulla dualità nel consumo) gli daranno due valori di X^ e di Y^ che azzerano le derivate. Ma Tizia/o, avendo letto i paragrafi precedenti, sa che non è finita qui e, spingendo il carrello tra i corridoi del supermercato, procede a calcolare l'Hessiano orlato (con le 2 variabili di scelta valutate ai loro valori di stazionarietà):

$$\bar{H}_2 = \underbrace{\begin{bmatrix} 0 & \frac{\partial g}{\partial X} & \frac{\partial g}{\partial Y} \\ \frac{\partial g}{\partial X} & \frac{\partial^2 \Lambda}{\partial X^2} & \frac{\partial^2 \Lambda}{\partial X \partial Y} \\ \frac{\partial g}{\partial Y} & \frac{\partial^2 \Lambda}{\partial Y \partial X} & \frac{\partial^2 \Lambda}{\partial Y^2} \end{bmatrix}}_{3\times 3} = \underbrace{\begin{bmatrix} 0 & p_X & p_Y \\ p_X & \frac{\partial^2 U}{\partial X^2} & \frac{\partial^2 U}{\partial X \partial Y} \\ p_Y & \frac{\partial^2 U}{\partial X \partial Y} & \frac{\partial^2 U}{\partial Y^2} \end{bmatrix}}_{3\times 3}$$

ma

$$\begin{Vmatrix} 0 & p_X & p_Y \\ p_X & \frac{\partial^2 U}{\partial X^2} & \frac{\partial^2 U}{\partial X \partial Y} \\ p_Y & \frac{\partial^2 U}{\partial X \partial Y} & \frac{\partial^2 U}{\partial Y^2} \end{Vmatrix} = 2 p_X p_Y \frac{\partial^2 U}{\partial X \partial Y} - p_Y^2 \frac{\partial^2 U}{\partial X^2} - p_X^2 \frac{\partial^2 U}{\partial Y^2} > 0$$

Se le restrizioni che vengono usualmente poste alla funzione $U(X, Y)$ sono valide (ovvero: utilità marginale decrescente e quindi derivate seconde dirette negative e derivate miste non negative[9]) segue che

$$\frac{\partial^2 U}{\partial X \partial Y} > \frac{1}{2}\left[\underbrace{\frac{p_Y}{p_X}\frac{\partial^2 U}{\partial X^2} + \frac{p_X}{p_Y}\frac{\partial^2 U}{\partial Y^2}}_{<0}\right]$$

e quindi il determinante è positivo. In questo caso Tizia/o avrà individuato una coppia di valori X^ e Y^ che effettivamente generano un massimo di utilità vincolata dal suo bilancio. E si avvia felice verso la cassa.

Adesso affrontiamo il problema che abbiamo lasciato in sospeso: **perché le derivate parziali del vincolo sono inserite nell'Hessiana per formare l'Hessiana "orlata"?** A questo fine allarghiamo la "dimensione" del problema supponendo N variabili indipendenti e $p < N$ vincoli. Sia allora:

$$f(x_1, x_2, \ldots, x_N)$$

una funzione di classe C^2 in N variabili non indipendenti tra loro ma legate dai seguenti vincoli riuniti in sistema

$$\begin{aligned} \varphi_1(x_1, x_2, \ldots, x_N) &= 0 \\ \varphi_2(x_1, x_2, \ldots, x_N) &= 0 \\ \vdots \qquad & \quad \vdots \\ \varphi_p(x_1, x_2, \ldots, x_N) &= 0 \end{aligned}$$

[9] Quest'ultima cosa vuol dire che l'utilità marginale del consumo di X aumenta se aumenta Y. In tale caso, che relazione si suppone che esista tra X e Y? Sono beni complementi o sostituti?

Si supponga che $f, \varphi_1, ..., \varphi_p$ siano definite in un campo A dello spazio R^N e che in detto campo esistano dei punti $P(x_1, x_2, ..., x_N)$ che sono soluzioni del sistema dei vincoli. Questi punti costituiscono un insieme $B \subseteq A$. Considerando solo la restrizione della $f(x_1, x_2, ..., x_N)$ all'insieme $B \subseteq A$ si dice che il punto $Q_0(x_1^0, x_2^0, ..., x_N^0) \in B$ è un punto di massimo (minimo) relativo vincolato per la $f(x_1, x_2, ..., x_N)$ se esiste un intorno I circolare di $Q_0(x_1^0, x_2^0, ..., x_N^0) \in B$ nel quale risulta

$$f(x_1, x_2, ..., x_N) \le f(x_1^0, x_2^0, ..., x_N^0) \text{ oppure } f(x_1, x_2, ..., x_N) \ge f(x_1^0, x_2^0, ..., x_N^0)$$

per tutti i punti di I appartenenti a *B*. Se le diseguaglianze valgono per tutti i punti di *B* il punto $Q_0(x_1^0, x_2^0, ..., x_N^0) \in B$ si dirà di massimo (minimo) assoluto vincolato. Possiamo allora enunciare il seguente

TEOREMA 3

Siano $f, \varphi_1, ..., \varphi_p$ continue insieme alle loro derivate prime in ogni punto di un campo *A* e la matrice jacobiana

$$J = \underbrace{\begin{bmatrix} \frac{\partial \varphi_1}{\partial x_1} & \frac{\partial \varphi_1}{\partial x_2} & \cdots & \frac{\partial \varphi_1}{\partial x_N} \\ \frac{\partial \varphi_2}{\partial x_1} & \frac{\partial \varphi_2}{\partial x_2} & \cdots & \frac{\partial \varphi_2}{\partial x_N} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial \varphi_p}{\partial x_1} & \frac{\partial \varphi_p}{\partial x_2} & \cdots & \frac{\partial \varphi_p}{\partial x_N} \end{bmatrix}}_{p \times N}$$

abbia caratteristica[10] *p* in ogni punto di A, ovvero che il massimo ordine dei minori aventi determinante diverso da zero sia p. In queste ipotesi se il massimo o il minimo relativo della $f(x_1, x_2, ..., x_N)$ in un punto nell'insieme B costituito dai punti soluzione del sistema riportato sopra, è assunto in un punto Q_0 interno ad A**, esisterà un numero p di costanti** $\lambda_1, \lambda_2, ,,, \lambda_p$ tali che la nuova funzione

$$f(x_1, x_2, ..., x_N) + \sum_{i=1}^{p} \lambda_i \varphi_i$$

abbia nel punto Q_0 derivate parziali prime tutte nulle. Allora

$$\begin{aligned} \frac{\partial f}{\partial x_1} + \lambda_1 \frac{\partial \varphi_1}{\partial x_1} + \lambda_2 \frac{\partial \varphi_2}{\partial x_1} + \cdots + \lambda_p \frac{\partial \varphi_p}{\partial x_1} &= 0 \\ \frac{\partial f}{\partial x_2} + \lambda_1 \frac{\partial \varphi_1}{\partial x_2} + \lambda_2 \frac{\partial \varphi_2}{\partial x_2} + \cdots + \lambda_p \frac{\partial \varphi_p}{\partial x_2} &= 0 \\ \vdots \qquad\qquad\qquad \vdots \quad \vdots & \\ \frac{\partial f}{\partial x_N} + \lambda_1 \frac{\partial \varphi_1}{\partial x_N} + \lambda_2 \frac{\partial \varphi_2}{\partial x_N} + \cdots + \lambda_p \frac{\partial \varphi_p}{\partial x_N} &= 0 \end{aligned}$$

[10] Ricordarsi che essendo la *J* una $p \times N$, vale $0 \le r(J) \le \min\{p, N\}$. Ciò vuol dire che avranno rango (caratteristica) nullo/a i minori di ordine $p + 1$ (ed eventualmente superiori).

Questo sistema in *p* incognite unito con il sistema dei vincoli (in N incognite) formano un sistema in N + p incognite che risolve il problema della ricerca del massimo (minimo). Se B è un compatto (vedi sopra), Weirstrass assicura l'esistenza di un massimo (minimo).

A questo punto il **TEOREMA 3 ha trasformato il problema vincolato in N variabili in un problema non vincolato in N + p variabili** ed è questa la chiave logica di quello che stiamo facendo (meglio: di quello che ha fatto Lagrange). *Casualmente* chiamiamo Λ la nuova funzione e "moltiplicatori di Lagrange" i singoli λ. Adesso valutiamo i punti di singolarità azzerando le derivate parziali prime e poniamo

$$
\begin{cases}
\dfrac{\partial \Lambda}{\partial \lambda_1} = 0 \\
\dfrac{\partial \Lambda}{\partial \lambda_2} = 0 \\
\vdots \\
\dfrac{\partial \Lambda}{\partial \lambda_p} = 0 \\
\dfrac{\partial \Lambda}{\partial x_1} = 0 \\
\dfrac{\partial \Lambda}{\partial x_2} = 0 \\
\vdots \\
\dfrac{\partial \Lambda}{\partial x_N} = 0
\end{cases}
\Rightarrow
\begin{cases}
\varphi_1(x_1, x_2, ..., x_N) = 0 \\
\varphi_2(x_1, x_2, ..., x_N) = 0 \\
\vdots \\
\varphi_p(x_1, x_2, ..., x_N) = 0 \\
\dfrac{\partial f}{\partial x_1} + \lambda_1 \dfrac{\partial \varphi_1}{\partial x_1} + \lambda_2 \dfrac{\partial \varphi_2}{\partial x_1} + \cdots + \lambda_p \dfrac{\partial \varphi_p}{\partial x_1} = 0 \\
\dfrac{\partial f}{\partial x_2} + \lambda_1 \dfrac{\partial \varphi_1}{\partial x_2} + \lambda_2 \dfrac{\partial \varphi_2}{\partial x_2} + \cdots + \lambda_p \dfrac{\partial \varphi_p}{\partial x_2} = 0 \\
\vdots \\
\dfrac{\partial f}{\partial x_N} + \lambda_1 \dfrac{\partial \varphi_1}{\partial x_N} + \lambda_2 \dfrac{\partial \varphi_2}{\partial x_N} + \cdots + \lambda_p \dfrac{\partial \varphi_p}{\partial x_N} = 0
\end{cases}
$$

A questo punto per determinare la natura dei punti di singolarità guardiamo al determinante della matrice Hessiana (senza orli…) delle derivate seconde e seconde miste. Per semplificare l'esposizione supponiamo N = 3 e p = 1.

$$
H(\lambda; x_1, x_2, x_3) = \begin{bmatrix}
\dfrac{\partial^2 \Lambda}{\partial \lambda^2} & \dfrac{\partial^2 \Lambda}{\partial x_1 \partial \lambda} & \dfrac{\partial^2 \Lambda}{\partial x_2 \partial \lambda} & \dfrac{\partial^2 \Lambda}{\partial x_3 \partial \lambda} \\
\dfrac{\partial^2 \Lambda}{\partial \lambda \partial x_1} & \dfrac{\partial^2 \Lambda}{\partial x_1^2} & \dfrac{\partial^2 \Lambda}{\partial x_1 \partial x_2} & \dfrac{\partial^2 \Lambda}{\partial x_1 \partial x_3} \\
\dfrac{\partial^2 \Lambda}{\partial \lambda \partial x_2} & \dfrac{\partial^2 \Lambda}{\partial x_2 \partial x_1} & \dfrac{\partial^2 \Lambda}{\partial x_2^2} & \dfrac{\partial^2 \Lambda}{\partial x_2 \partial x_3} \\
\dfrac{\partial^2 \Lambda}{\partial \lambda \partial x_3} & \dfrac{\partial^2 \Lambda}{\partial x_3 \partial x_1} & \dfrac{\partial^2 \Lambda}{\partial x_3 \partial x_2} & \dfrac{\partial^2 \Lambda}{\partial x_3^2}
\end{bmatrix}
$$

Ma dalle condizioni per il punto di singolarità ricaviamo che la derivata seconda rispetto al moltiplicatore è nulla e che la derivata seconda mista della lagrangiana corrisponde alla derivata del vincolo, ovvero

$\dfrac{\partial^2 \Lambda}{\partial x_i \partial \lambda} = \dfrac{\partial \varphi_1}{\partial x_i} = \dfrac{\partial^2 \Lambda}{\partial \lambda \partial x_i}$ con i = 1,2,3 e quindi l'Hessiana diventa (adesso spunta l'orlo …)

$$H(\lambda;x_1,x_2,x_3)=\begin{bmatrix} 0 & \frac{\partial\varphi_1}{\partial x_1} & \frac{\partial\varphi_1}{\partial x_2} & \frac{\partial\varphi_1}{\partial x_3} \\ \frac{\partial\varphi_1}{\partial x_1} & \frac{\partial^2\Lambda}{\partial x_1^2} & \frac{\partial^2\Lambda}{\partial x_1\partial x_2} & \frac{\partial^2\Lambda}{\partial x_1\partial x_3} \\ \frac{\partial\varphi_1}{\partial x_2} & \frac{\partial^2\Lambda}{\partial x_2\partial x_1} & \frac{\partial^2\Lambda}{\partial x_2^2} & \frac{\partial^2\Lambda}{\partial x_2\partial x_3} \\ \frac{\partial\varphi_1}{\partial x_3} & \frac{\partial^2\Lambda}{\partial x_3\partial x_1} & \frac{\partial^2\Lambda}{\partial x_3\partial x_2} & \frac{\partial^2\Lambda}{\partial x_3^2} \end{bmatrix}$$

Per valutare la natura del punto di singolarità trovato azzerando le derivate prime procediamo (ovviamente!) come nel caso non vincolato. Chiamiamo il punto $P_0(\lambda_0;x_{1,0},x_{2,0},x_{3,0})$ e diciamo che se $P_0(\lambda_0;x_{1,0},x_{2,0},x_{3,0})$ è un punto di stazionarietà per la lagrangiana (**problema diventato non vincolato**) allora la funzione $f(x_1,x_2,x_3)$ nel punto $P_0(x_{1,0},x_{2,0},x_{3,0})$ ha un massimo o un minimo a seconda del segno dei determinanti di cui alla Tab. 2.

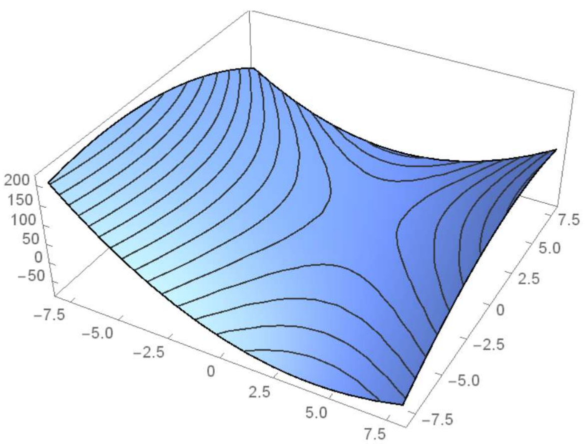
Esercizi svolti

ES.1 Sia assegnata la funzione $f(x,y)=2x^2-y^2+xy-8x+10$ la cui massimizzazione è soggetta a vincolo lineare

$$2x+y-2=0$$

Determinare gli eventuali punti di massimo e di minimo relativo. Svolgiamo l'esercizio utilizzando sia la sostituzione sia la funzione di Lagrange.

Per prima cosa presentiamo il grafico della funzione (lo studente cerchi di interpretare il significato delle curve tracciate sulla superficie. Diciamo che il grafico è solo la "foto" (tra -8 < x < 8) e -8 < y < 8) della funzione che sostituisce lo studio del campo di esistenza, asintoti, curve di livello, ecc.



Già un'idea dovremmo farcela…

Metodo della Sostituzione. Se esistono, i punti cercati sono soluzione di

$$\begin{cases} z = 2x^2-y^2+xy-8x+10 \\ 2x+y-2 = 0 \end{cases} \Rightarrow \begin{cases} z = 2x^2-y^2+xy-8x+10 \\ y = 2-2x \end{cases} \Rightarrow \begin{cases} z = -4x^2+2x+6 \\ y = 2-2x \end{cases}$$

Deriviamo la funzione *z* ricavata dopo la sostituzione e otteniamo

$$\frac{dz}{dx} = -8x + 2 = 0 \Rightarrow x = \frac{1}{4}$$

$$\frac{d^2 z}{dx^2} = -8 < 0$$

Il punto cercato è quindi (1/4, 3/2) ed è di massimo relativo vincolato.

Metodo di Lagrange. Sia la funzione

$\Lambda = 2x^2 - y^2 + xy - 8x + 10 - \lambda\left[2x + y - 2\right]$ da cui

$$\begin{cases} \dfrac{\partial \Lambda}{\partial x} = 4x + y - 8 - 2\lambda = 0 \\ \dfrac{\partial \Lambda}{\partial y} = -2y + x - \lambda = 0 \\ \dfrac{\partial \Lambda}{\partial \lambda} = 2x + y - 2 = 0 \end{cases} \Rightarrow \begin{bmatrix} 4 & 1 & -2 \\ 1 & -2 & -1 \\ 2 & 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ \lambda \end{bmatrix} = \begin{bmatrix} 8 \\ 0 \\ 2 \end{bmatrix}$$

da cui il determinante della matrice dei coefficienti è pari a – 8 (diverso da zero) e quindi (con Cramer) i valori sono $x = 1/4$, $y = 3/2$, $\lambda = -11/4$. E' davvero un massimo relativo? Formiamo l'Hessiana "orlata"

$\bar{H}_2 = \begin{bmatrix} 0 & 2 & 1 \\ 2 & 4 & 1 \\ 1 & 1 & -2 \end{bmatrix}$ da cui $\left\|\bar{H}_2\right\| = 8 > 0$ e quindi il punto P (1/4, 3/2, -11/4) è un massimo relativo e la funzione lagrangiana massimizzata vale

$$\Lambda^* = 2\left(\frac{1}{4}\right)^2 - \left(\frac{3}{2}\right)^2 + \frac{1}{4} \times \frac{3}{2} - 8\frac{1}{4} + 10 - (-\frac{11}{4})\left[2\frac{1}{4} + \frac{3}{2} - 2\right] = \frac{25}{4} = f^*(x, y)$$

Osservazione. Supponiamo per un momento che il termine noto del vincolo non sia –2 ma un parametro variabile *K*. La lagrangiana massimizzata (funzione di massimo valore) sarebbe

$$\Lambda^* = 2\left(\frac{1}{4}\right)^2 - \left(\frac{3}{2}\right)^2 + \frac{1}{4} \times \frac{3}{2} - 8\frac{1}{4} + 10 - (-\frac{11}{4})\left[2\frac{1}{4} + \frac{3}{2} + K\right]$$

Allora

$\dfrac{d\Lambda^*}{dK} = \dfrac{11}{4} = \lambda^*$. **Il moltiplicatore sembra avere una parte in commedia non del tutto posta in evidenza dalla mera procedura della ricerca del massimo/minimo**. Ciò era già emerso in precedenza ma bisogna tornarci su più avanti.

ES.2 (2 vincoli) Sia assegnata la funzione $f(x,y) = x$, da massimizzare sotto le condizioni $g_1 : x^2 + y^2 + z^2 = 5/2$ e $g_2 : y + z = 1$. La funzione esiste ovunque e quindi A = R^3, quindi i vincoli hanno senso. Formiamo la lagrangiana associata al problema

$$\Lambda = x + \lambda_1 \left[x^2 + y^2 + z^2 - \frac{5}{2} \right] + \lambda_2 \left[y + z - 1 \right]$$

Allora derivando

$$\begin{cases} \dfrac{\partial\Lambda}{\partial\lambda_2} = y + z - 1 = 0 \\ \dfrac{\partial\Lambda}{\partial\lambda_1} = x^2 + y^2 + z^2 - \dfrac{5}{2} = 0 \\ \dfrac{\partial\Lambda}{\partial z} = 2\lambda_1 z + \lambda_2 = 0 \\ \dfrac{\partial\Lambda}{\partial y} = 2\lambda_1 y + \lambda_2 = 0 \\ \dfrac{\partial\Lambda}{\partial x} = 1 + 2\lambda_1 x = 0 \end{cases} \Rightarrow \begin{cases} -\dfrac{\lambda_2}{\lambda_1} - \dfrac{\lambda_2}{\lambda_1} - 1 = 0 \\ \dfrac{1}{4\lambda_1^2} + \dfrac{\lambda_2^2}{4\lambda_1^2} + \dfrac{\lambda_2^2}{4\lambda_1^2} - \dfrac{5}{2} = 0 \\ z = -\dfrac{\lambda_2}{2\lambda_1} \\ y = -\dfrac{\lambda_2}{2\lambda_1} \\ x = -\dfrac{1}{2\lambda_1} \end{cases} \Rightarrow \begin{cases} \lambda_2 = -\lambda_1 \\ \lambda_1 = \pm\dfrac{1}{2\sqrt{2}} \\ z = -\dfrac{\lambda_2}{2\lambda_1} \\ y = -\dfrac{\lambda_2}{2\lambda_1} \\ x = -\dfrac{1}{2\lambda_1} \end{cases} \Rightarrow \begin{cases} \lambda_2 = \mp\dfrac{1}{2\sqrt{2}} \\ \lambda_1 = \pm\dfrac{1}{2\sqrt{2}} \\ z = \dfrac{1}{2} \\ y = \dfrac{1}{2} \\ x = \pm\sqrt{2} \end{cases}$$

Emergono 2 punti di singolarità

$$P_1\left(-\sqrt{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2\sqrt{2}}, -\frac{1}{2\sqrt{2}}\right) \quad P_2\left(\sqrt{2}, \frac{1}{2}, \frac{1}{2}, -\frac{1}{2\sqrt{2}}, \frac{1}{2\sqrt{2}}\right)$$

Data la funzione $f(x,y) = x$ avremo che ovviamente P_2 si riferisce al massimo e P_1 al minimo. L'analisi della matrice Hessiana orlata (che orlata non è …) dovrebbe confermarci il risultato. Formiamo la matrice Hessiana di dimensione (nel nostro caso) 5×5:

$$\overline{H}_{(5\times5)} = \begin{bmatrix} \Lambda''_{xx} & \Lambda''_{xy} & \Lambda''_{xz} & \Lambda''_{x\lambda_1} & \Lambda''_{x\lambda_1} \\ \Lambda''_{yx} & \Lambda''_{yy} & \Lambda''_{yz} & \Lambda''_{y\lambda_1} & \Lambda''_{y\lambda_2} \\ \Lambda''_{zx} & \Lambda''_{zy} & \Lambda''_{zz} & \Lambda''_{z\lambda_1} & \Lambda''_{z\lambda_2} \\ \Lambda''_{\lambda_1 x} & \Lambda''_{\lambda_1 y} & \Lambda''_{\lambda_1 z} & \Lambda''_{\lambda_1\lambda_1} & \Lambda''_{\lambda_1\lambda_2} \\ \Lambda''_{\lambda_2 x} & \Lambda''_{\lambda_2 y} & \Lambda''_{\lambda_2 z} & \Lambda''_{\lambda_2\lambda_1} & \Lambda''_{\lambda_2\lambda_2} \end{bmatrix} = \begin{bmatrix} \Lambda''_{xx} & \Lambda''_{xy} & \Lambda''_{xz} & g'_{1x} & g'_{2x} \\ \Lambda''_{yx} & \Lambda''_{yy} & \Lambda''_{yz} & g'_{1y} & g'_{2y} \\ \Lambda''_{zx} & \Lambda''_{zy} & \Lambda''_{zz} & g'_{1z} & g'_{2z} \\ g'_{1x} & g'_{1y} & g'_{1z} & 0 & 0 \\ g'_{2x} & g'_{2y} & g'_{2z} & 0 & 0 \end{bmatrix}$$

E oplà, come si vede adesso la matrice ha ben 2 orli (come mai?). Per valutarla nel punto P_2 ci vogliono le derivate seconde e incrociate di Λ^* (o quelle dei vincoli, direbbe sbrigativamente qualcuno).

$$\begin{cases} \dfrac{\partial \Lambda}{\partial \lambda_2} = y + z - 1 = 0 \\ \dfrac{\partial \Lambda}{\partial \lambda_1} = x^2 + z^2 - \dfrac{5}{2} = 0 \\ \dfrac{\partial \Lambda}{\partial z} = 2\lambda_1 z + \lambda_2 = 0 \\ \dfrac{\partial \Lambda}{\partial y} = 2\lambda_1 y + \lambda_2 = 0 \\ \dfrac{\partial \Lambda}{\partial x} = 1 + 2\lambda_1 x = 0 \\ \dfrac{\partial^2 \Lambda}{\partial x \partial y} = \dfrac{\partial^2 \Lambda}{\partial x \partial z} = \dfrac{\partial^2 \Lambda}{\partial y \partial x} = \dfrac{\partial^2 \Lambda}{\partial y \partial z} = \dfrac{\partial^2 \Lambda}{\partial z \partial x} = \dfrac{\partial^2 \Lambda}{\partial z \partial y} = 0 \end{cases} \Rightarrow \begin{cases} \dfrac{\partial^2 \Lambda}{\partial \lambda_2^2} = 0 \\ \dfrac{\partial^2 \Lambda}{\partial \lambda_1^2} = 0 \\ \dfrac{\partial^2 \Lambda}{\partial z^2} = 2\lambda_1 \\ \dfrac{\partial^2 \Lambda}{\partial y^2} = 2\lambda_1 \\ \dfrac{\partial^2 \Lambda}{\partial x^2} = 2\lambda_1 \end{cases}$$

E quindi, nel punto P_2:

$$\overline{H}_{(5\times 5)} = \begin{bmatrix} 2\lambda_1 & 0 & 0 & 2x & 0 \\ 0 & 2\lambda_1 & 0 & 2y & 1 \\ 0 & 0 & 2\lambda_1 & 2z & 1 \\ 2x & 2y & 2z & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 2(-\frac{1}{2\sqrt{2}}) & 0 & 0 & 2\sqrt{2} & 0 \\ 0 & 2(-\frac{1}{2\sqrt{2}}) & 0 & 2\frac{1}{2} & 1 \\ 0 & 0 & 2(-\frac{1}{2\sqrt{2}}) & 2\frac{1}{2} & 1 \\ 2\sqrt{2} & 2\frac{1}{2} & 2\frac{1}{2} & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \end{bmatrix}$$

$$= \begin{bmatrix} -\frac{1}{\sqrt{2}} & 0 & 0 & 2\sqrt{2} & 0 \\ 0 & -\frac{1}{\sqrt{2}} & 0 & 1 & 1 \\ 0 & 0 & -\frac{1}{\sqrt{2}} & 1 & 1 \\ 2\sqrt{2} & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \end{bmatrix}$$

Calcolando il determinate abbiamo $\left\| \overline{H} \right\| = -8\sqrt{2} < 0$. L'ordine del determinante è dispari e quindi P^2 è un massimo vincolato. Ripetere per esercizio con P_1 e verificare che è un minimo.

APPENDICI

Appendice di Pro-memoria 1

Forme quadratiche e matrici

Definiamo Forma Quadratica un polinomio in n variabili con tutti i termini di grado 2.

Es. 1 in R^2: $q(x,y) = ax^2 + 2bxy + cy^2$ dove il 2 che moltiplica il coefficiente b mostra chiaramente che la forma quadratica deriva dal termine di grado 2 del polinomio di Taylor della funzione.

Che relazione esiste tra forme quadratiche e matrici? Usiamo l'esempio 1 e riscriviamo la funzione usando la matrice che formiamo con i coefficienti (okkio al 2 che moltiplica b):

$$\begin{aligned} q(x,y) &= \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} a & b \\ b & c \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \\ &= \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} ax+by \\ bx+cy \end{bmatrix} \\ &= x(ax+by) + y(bx+cy) \\ &= ax^2 + bxy + byx + cy^2 = ax^2 + 2bxy + cy^2 \end{aligned}$$

Più in generale, ad ogni forma quadratica in n variabili corrisponde **una matrice simmetrica** $n \times n$ e viceversa. Es. 2 in R^3: $q(x,y,z) = 3x^2 + 6xy + z^2 - 3yz$. Scriviamo la 3×3 corrispondente secondo il seguente criterio

$$\begin{bmatrix} 3 & 3 & 0 \\ 3 & 0 & -\frac{3}{2} \\ 0 & -\frac{3}{2} & 1 \end{bmatrix} = \begin{bmatrix} \boxed{\textit{coefficiente di } x^2} & \frac{\textit{coefficiente di } xy}{2} & \frac{\textit{coefficiente di } xz}{2} \\ \frac{\textit{coefficiente di } yx}{2} & \boxed{\textit{coefficiente di } y^2} & \frac{\textit{coefficiente di } yz}{2} \\ \frac{\textit{coefficiente di } zx}{2} & \frac{\textit{coefficiente di } zy}{2} & \boxed{\textit{coefficiente di } z^2} \end{bmatrix}$$

I termini sulla diagonale principale sono i coefficienti dei quadrati e la simmetria è garantita dal teorema di Young che già avevamo richiamato implicitamente con il riferimento al polinomio di Taylor. Quindi

$$\begin{aligned} q(x,y,z) &= \underbrace{\begin{bmatrix} x & y & z \end{bmatrix}}_{1\times 3} \underbrace{\begin{bmatrix} 3 & 3 & 0 \\ 3 & 0 & -\frac{3}{2} \\ 0 & -\frac{3}{2} & 1 \end{bmatrix}}_{3\times 3} \underbrace{\begin{bmatrix} x \\ y \\ z \end{bmatrix}}_{3\times 1} \\ &= \begin{bmatrix} x & y & z \end{bmatrix} \begin{bmatrix} 3x+3y \\ 3x-\frac{3}{2}z \\ -\frac{3}{2}y+z \end{bmatrix} \\ &= 3x^2 + 3xy + 3yx - \frac{3}{2}yz - \frac{3}{2}zy + z^2 = 3x^2 + 6xy - 3yz + z^2 \end{aligned}$$

Naturalmente la strada si può percorrere al contrario. Ovvero data una matrice simmetrica si risale alla q che le corrisponde. Sia

$$\begin{bmatrix} 0 & 3 & 5 \\ 3 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix} = \begin{bmatrix} \text{coefficiente di } x^2 & \dfrac{\text{coefficiente di } xy}{2} & \dfrac{\text{coefficiente di } xz}{2} \\ \dfrac{\text{coefficiente di } yx}{2} & \text{coefficiente di } y^2 & \dfrac{\text{coefficiente di } yz}{2} \\ \dfrac{\text{coefficiente di } zx}{2} & \dfrac{\text{coefficiente di } zy}{2} & \text{coefficiente di } z^2 \end{bmatrix}$$

Otteniamo la funzione

$q(x, y, z) = y^2 + 2z^2 + 6xy + 10xz$.

Diciamo adesso che una forma quadratica (o una matrice simmetrica; è lo stesso) si dice

Categorie	Condizione	Valori di x e y
DEFINITA POSITIVA	Se $q(x, y) > 0$	Per ogni $(x, y) \neq (0,0)$
DEFINITA NEGATIVA	Se $q(x, y) < 0$	Per ogni $(x, y) \neq (0,0)$
SEMI DEFINITA POSITIVA	Se $q(x, y) \geq 0$	Per ogni (x, y)
SEMIDEFINTA NEGATIVA	Se $q(x, y) \leq 0$	Per ogni (x, y)
NON DEFINITA	➢ > 0 per certi valori di x e y; ➢ < 0 per altri valori di x e y	Esempio: $x^2 - y^2$ è positiva in (1, 0) ed è negativa in (0, 1)

Le prima 4 categorie non sono mutuamente escludentesi, date le disuguaglianze.
Esistono 3 metodi per valutare in quale caso ricadiamo.

$$\text{Metodi} \begin{cases} \textbf{\textit{Completamento dei quadrati}} & 1 \\ \textbf{\textit{Autovalori}} & 2 \\ \textbf{\textit{Minori}} & 3 \end{cases}$$

Metodo 1. Completamento dei quadrati. Per i problemi mediamente trattati dagli *economisti* il metodo (1) è poco più che un *curiosum*. Basta fare un paio di esempi per mostrare come funziona. Sia

Es. 1
$q(x, y, z) = x^2 + 2xy + 3y^2$

Bisogna cercare il modo per riscrivere la q come somma o differenza di quadrati. Il perché sarà chiarito meglio dopo. Se riscrivo l'ultimo termine come $y^2 + 2y^2$ ottengo

$$q(x, y, z) = \underbrace{x^2 + 2xy + y^2}_{(x+y)^2} + 2y^2 = (x + y)^2 + 2y^2$$

La q si è trasformata in una somma di quadrati che non sarà mai negativa e quindi sarà o positiva definita o semi positiva definita. Essa si annullerà (semi definita positiva) se e solo se $x + y = 0$ e $y = 0$ ovvero con $(x, y) = (0,0)$.

Es. 2
Come trasformiamo $q(x, y, z) = x^2 + xy + y^2$ in una somma di prodotti? Moltiplico e divido per 2 il termine xy:

$q(x,y,z) = x^2 + 2x\frac{y}{2} + y^2$ e poi scindo il termine y^2 in due parti che sommate diano y^2:

$q(x,y,z) = x^2 + 2x\frac{y}{2} + y^2 = x^2 + 2x\frac{y}{2} + \frac{y^2}{4} + \frac{3y^2}{4} = \left(x + \frac{y}{2}\right)^2 + \frac{3y^2}{4}$. Quindi la q è positiva definita.

Fino a che otteniamo una somma di quadrati è tutto chiaro: il risultato della soma è positivo. Che succede se otteniamo una differenza di quadrati. In generale la forma non è definita. Per vedere come mai facciamo l'ultimo esempio.

Es. 3
Come trasformiamo $q(x,y,z) = x^2 + 3xy + y^2$ in una somma di prodotti? Moltiplico e divido per 2 il termine xy:

$q(x,y,z) = x^2 + 2x\frac{3y}{2} + y^2$ e poi scindo il termine y^2 in due parti che sommate diano y^2:

$q(x,y,z) = x^2 + 2x\frac{3y}{2} + y^2 = x^2 + 2x\frac{3y}{2} + 9\frac{y^2}{4} - 5\frac{y^2}{4} = \left(x + \frac{3y}{2}\right)^2 - 5\frac{y^2}{4}$. Essendoci il meno questa forma non è definita. Se voglio che la differenza dei quadrati sia positiva devo prendere y = 0 con x qualunque. Se voglio che la differenza sia negativa basta prendere $\left(x + \frac{3y}{2}\right) = 0$ con y qualunque.

Metodo 2. Calcolo degli Autovalori. Per apprezzare questo metodo occorre ricordare che per una matrice simmetrica gli autovalori sono sempre reali. La tabella riassume

Categorie	Condizione
DEFINITA POSITIVA	Se *tutti gli autovalori sono* > 0
DEFINITA NEGATIVA	Se *tutti gli autovalori sono* < 0
SEMI DEFINITA POSITIVA	Se *tutti gli autovalori sono* ≥ 0
SEMIDEFINTA NEGATIVA	Se *tutti gli autovalori sono* ≤ 0
NON DEFINITA	Segni alternati

Esempio che non richiede calcolo delle radici del polinomio caratteristico per ottenere gli autovalori (dire perché)

$q(x,y,z) = x^2 - y^2 = [x \quad y]\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}$. Gli autovalori della matrice sono 1 e -1 e quindi la matrice non è definita.

Esempi ulteriori nel Promemoria 2 e nelle dispense di microeconomia.

Metodo 3. Calcolo dei Minori. Già trattate nel testo.

Appendice di Pro-memoria 2

Caso di matrici simmetriche a coefficienti reali e matrici hermitiane (non trattate)

In quanto segue

x è un vettore riga $1 \times n$

M è una matrice quadrata $n \times n$

$\boldsymbol{x^T}$ è il vettore colonna n × 1 ottenuto trasponendo è un vettore riga $\boldsymbol{x}$

Risulta quindi intuitivo che $\boldsymbol{xMx^T}$ è un'operazione consistente e che corrisponde ad un numero reale di cui ci si chiederà se è positivo, negativo o nullo.

Allora:

- una matrice **M** si dice **definita positiva** se la forma quadratica $\boldsymbol{xMx^T > 0}$ per ogni vettore $\boldsymbol{x}$
- una matrice **M** si dice **semidefinita positiva** se la forma quadratica $\boldsymbol{xMx \geq 0}$ per ogni vettore $\boldsymbol{x}$

- una matrice **M** si dice **definita negativa** se la forma quadratica $\boldsymbol{xMx^T < 0}$ per ogni vettore $\boldsymbol{x}$
- una matrice **M** si dice **semidefinita negativa** se la forma quadratica $\boldsymbol{xMx^T \leq 0}$ per ogni vettore $\boldsymbol{x}$

Una funzione continua a derivate continue f è **convessa** se e solo se la matrice hessiana $\mathbf{H_{f(x)}}$ è **semidefinita positiva**

Una funzione continua a derivate continue f è **concava** se e solo se la matrice hessiana $\mathbf{H_{f(x)}}$ è **semidefinita negativa.**

"Scorciatoie" in $\boldsymbol{R^2}$ (con possibili alternative di Cartesio a Silvester) per chi ha fretta

1 **Usando gli autovalori**

Di solito basta guardare gli autovalori della matrice. Sia infatti M con $n = 2$ di autovalori λ_1, λ_2 allora M è

- **definita > 0** se $\lambda_1 > 0$ e $\lambda_2 > 0$
- **definita < 0** se $\lambda_1 < 0$ e $\lambda_2 < 0$
- **semi definita > 0** se $\lambda_1 \geq 0$ e $\lambda_2 \geq 0$ (entrambi non negativi)
- **semi definita < 0** se $\lambda_1 \leq 0$ e $\lambda_2 \leq 0$ (entrambi non positivi)
- **indefinita** se ha sia autovalore strettamente positivo sia autovalore strettamente negativo.

In generale si guardano determinante e traccia della matrice poiché il determinante è il prodotto degli autovalori e la traccia è la loro somma. Infatti (sempre per semplicità in R^2)

1) **det > 0** implica che gli **autovalori hanno segno concorde**; se **tr[M] > 0** vuol dire che sono entrambi positivi (**quindi la matrice def > 0**)
2) **det > 0** implica che gli **autovalori hanno segno concorde**; se **tr[M] < 0** vuol dire che sono entrambi negativi (**quindi la matrice def > 0**)

3) **det < 0** implica che gli autovalori hanno segno discorde **(matrice non def > 0)**

4) **det = 0** implica che un autovalore è nullo; se **tr[M] > 0** vuol dire che l'autovalore non nullo è positivo (**matrice semidef > 0**)
5) **det = 0** implica che un autovalore è nullo; se **tr[M] < 0** indica che allora l'autovalore non nullo è negativo (**matrice semidef < 0**)

Quindi, usando gli autovalori:

Definizione. Una matrice A qualsiasi di ordine n è:

1. *definita positiva* ⇔ i suoi autovalori sono tutti strettamente positivi;
2. *semi definita positiva* ⇔ i suoi autovalori sono tutti non negativi;
3. *definita negativa* ⇔ i suoi autovalori sono tutti strettamente negativi;
4. *semi definita negativa* ⇔ i suoi autovalori sono tutti non positivi;
5. *indefinita* ⇔ presenta sia autovalori strettamente positivi che strettamente negativi (almeno 2 autovalori discordi per $n > 2$).

Nota: la *matrice nulla* è sia semidefinita positiva che semidefinita negativa.

Teorema. Sia $H_{f(P0)}$ una la *matrice hessiana* di ordine n, segue che se nel punto P_0 essa è:

1. definita positiva, allora P_0 è un punto di minimo relativo per la f;
2. definita negativa, allora P_0 è un punto di massimo relativo per la f;
3. indefinita, allora P_0 è un punto di sella per la f.

Nota: per i casi semi definiti nulla si può dire a priori sulla natura di P_0.

Corollario 1. Sia $Hf(P_0)$ una *matrice hessiana* di ordine n, segue che se nel punto P_0 vale:

1. $det(H_k) > 0 \ \forall k \in [1, n]$, P_0 è un punto di minimo relativo per f;
2. $(-1)^k det(H_k) > 0 \ \forall k \in [1, n]$, P_0 è un punto di massimo relativo per f;
3. se non si verificano nemmeno 1. e 2. con $\geq$, P_0 è un punto di sella per f;
4. se si verificano 1. o 2. con $\geq$ essendo nei casi semidefiniti nulla può dirsi a priori.

Corollario 2. Sia $H_{f(P0})$ una *matrice hessiana* di ordine 2, segue che se:

1. $det(H_2) > 0 \wedge det(H_1) > 0$, P_0 è un punto di minimo relativo per f;
2. $det(H_2) > 0 \wedge det(H_1) < 0$, P_0 è un punto di massimo relativo per f;
3. $det(H_2) < 0$, P_0 è un punto di sella per f;

Più precisamente, se:

3.1. $H_{(1,1)} \cdot H_{(2,2)} < 0$, P_0 è un punto di sella (nel senso proprio della parola);
3.2. $H_{(1,1)} \cdot H_{(2,2)} = 0$, P_0 è un punto di sella o un punto di flesso per f: caso ambiguo;
3.3. $H_{(1,1)} \cdot H_{(2,2)} > 0$, P_0 è un punto di flesso per f.

4. $det(H_2) = 0$, nulla si può dire a priori sulla natura di P_0.

2 Usando minori e determinante

a) Sia H una matrice simmetrica 2 per 2

$$H = \begin{bmatrix} a & \mathrm{b} \\ \mathrm{b} & c \end{bmatrix}$$

H è detta **Definita Positiva** se e solo se $a > 0$ e $\|H\| > 0$. Ad esempio con $H = \begin{bmatrix} 2 & 3 \\ 3 & 6 \end{bmatrix}$, $a = 2 > 0$ e $\|H\| = 3 > 0$.

b) Sia H una matrice simmetrica 2 per 2

$$H = \begin{bmatrix} a & \mathrm{b} \\ \mathrm{b} & c \end{bmatrix}$$

H è detta **Definita Negativa** se e solo se $a < 0$ e $\|H\| > 0$. Ad esempio con $H = \begin{bmatrix} -1 & 2 \\ 2 & -9 \end{bmatrix}$, $a = -1 < 0$ e $\|H\| = 5 > 0$.

c) Sia H una matrice simmetrica 2 per 2

$$H = \begin{bmatrix} a & \mathrm{b} \\ \mathrm{b} & c \end{bmatrix}$$

H è detta **Semi Definita Positiva** se $a \geq 0$, c ≥ 0 e $\|H\| \geq 0$. Ad esempio con $H = \begin{bmatrix} 6 & 6 \\ 6 & 6 \end{bmatrix}$, $a = 6 \geq 0$, c = $6 \geq 0$ e $\|H\| = 0 \geq 0$.

d) Sia H una matrice simmetrica 2 per 2

$$H = \begin{bmatrix} a & \mathrm{b} \\ \mathrm{b} & c \end{bmatrix}$$

H è detta **Semi Definita Negativa** se $a \leq 0$, c ≤ 0 e $\|H\| \geq 0$. Ad esempio con $H = \begin{bmatrix} 0 & 0 \\ 0 & -9 \end{bmatrix}$, $a = 0 \leq 0$, c = $-9 \leq 0$ e $\|H\| = 0 \geq 0$.

NB Se H è Definita Positiva è anche Semi Definita Positiva
Se H è Definita Negativa è anche Semi Definita Negativa
Se H non è Semi Definita Positiva né Semi Definita Negativa essa è Indefinita
Se a e c sono discordi H è Indefinita
Se $\|H\| < 0$ allora H è Indefinita, a prescindere dal segno concorde o discorde di a e c.

Esempio

$$H = \begin{bmatrix} 3 & 10 \\ 10 & 8 \end{bmatrix}$$

a e c sono concordi ma $\|H\| = -76 < 0$.

Esercizio 1 (familiare a chi viene da triennali di economia)

Sia $f(X, Y) = 10X^{\frac{1}{3}}Y^{\frac{1}{3}}$ definita su R^2_{++}. La sua matrice Hessiana è

$$H = \begin{bmatrix} -\frac{20Y^{1/3}}{9X^{5/3}} & \frac{10}{9X^{2/3}Y^{2/3}} \\ \frac{10}{9X^{2/3}Y^{2/3}} & -\frac{20X^{1/3}}{9Y^{5/3}} \end{bmatrix}$$

Quindi $a = -\frac{20Y^{1/3}}{9X^{5/3}} < 0, c = -\frac{20Y^{1/3}}{9X^{5/3}} < 0$ e il determinante è

$$\frac{100}{27X^{4/3}Y^{4/3}} > 0$$

Allora H è Semi Definita Negativa. Ci attendiamo autovalori di segno concorde (determinante positivo):

$$\lambda_1 = \frac{10\left(-X^2 - Y^2 - \sqrt{X^4 - X^2Y^2 + Y^4}\right)}{9X^{5/3}Y^{5/3}}$$
$$\lambda_2 = \frac{10(-X^2 - Y^2 + \sqrt{X^4 - X^2Y^2 + Y^4})}{9X^{5/3}Y^{5/3}}$$

La traccia è $-\frac{20(X^2+Y^2)}{9X^{5/3}Y^{5/3}} < 0$.

Sulla base di quanto ottenuto sopra, valutare se $f(X,Y)$ è concava o convessa in P(2, 3).
In *P* il determinante *a* < 0, *c* < 0 e il determinante è (circa) 0.34 > 0. In *P* l'Hessiana è Semi Definita Negativa. La funzione è concava.

Esercizio 2

Sia $f(x,y) = x + \operatorname{sen} x + y^2$definita su R_+^2 . La sua matrice Hessiana è

$$H_f = \begin{bmatrix} -senx & 0 \\ 0 & 2 \end{bmatrix}$$

il cui determinante è

$$-2senx$$

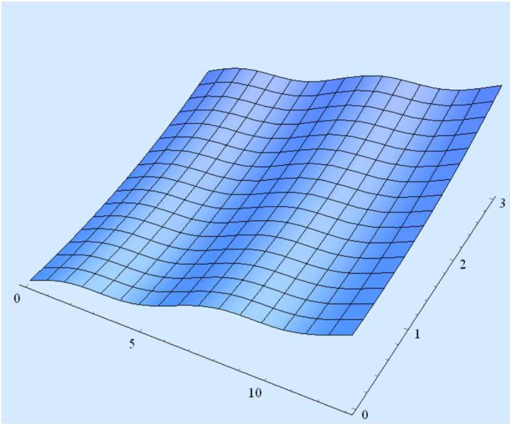
con autovalori ricavati da (per esercizio; altrimenti è sufficiente guardare determinante)

$$\|H_f - \lambda I\| = 0$$

pari a

$$\lambda_1 = 2$$
$$\lambda_2 = -\text{senx}$$

Definiamo il punto $P_0 = (2\pi, 0)$. Il **determinante diventa = 0** e $\lambda_2 = 0$. Gli **autovalori sono uno pari a zero e un altro sempre positivo (2) e la traccia è quindi positiva**. La H è quindi semi definita positiva e la *f* è convessa nel punto Po. Attenzione ai sono punti di sella, come lascia pensare il grafico. Si possono trovare?



Esercizio 3

Sia $f(x,y) = 5\sqrt{xy}$ definita su R^2_{++}. La sua matrice Hessiana è

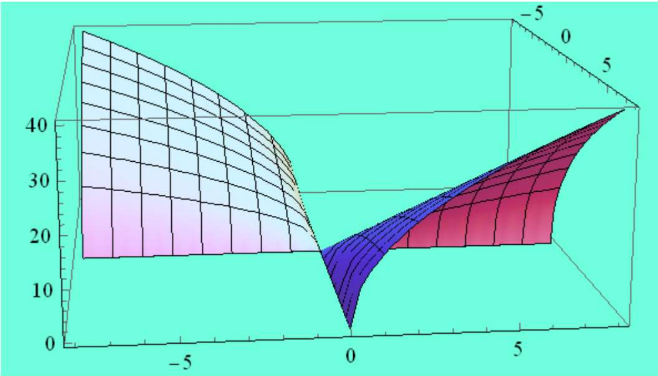
$$H_f = \begin{bmatrix} -\frac{5y^2}{4(xy)^{3/2}} & -\frac{5xy}{4(xy)^{3/2}} + \frac{5}{2\sqrt{xy}} \\ -\frac{5xy}{4(xy)^{3/2}} + \frac{5}{2\sqrt{xy}} & -\frac{5x^2}{4(xy)^{3/2}} \end{bmatrix}$$

Il cui determinante è = 0. Gli autovalori sono

$$\lambda_1 = 0$$
$$\lambda_2 = -\frac{5(x^2+y^2)}{4(xy)^{3/2}}$$

Quindi la traccia è negativa. Siamo nel caso in cui **det[H] = 0**, un autovalore è nullo e **tr[H] < 0**. La **matrice H è semi definita < 0** e la funzione è concava.

Il grafico che segue si riferisce alla funzione in R^2.



Come valutiamo la funzione nell'origine?

Esempio microeconomia. Produzione.

Sia la funzione di produzione C-D su R_+^2 : $Q = 5K^{0.4}L^{0.6}$. La matrice hessiana è

$$H = \begin{bmatrix} -\dfrac{1.2L^{0.6}}{K^{1.6}} & \dfrac{1.2}{K^{0.6}L^{0.4}} \\ \dfrac{1.2}{K^{0.6}L^{0.4}} & -\dfrac{1.2K^{0.4}}{L^{1.4}} \end{bmatrix}$$

Quindi $a \leq 0$, $c \leq 0$ e $\|H\| \geq 0$. H è Semi Definita Negativa. La funzione è concava.
Criterio dei autovalori:

$$\lambda_1 = 0$$
$$\lambda_2 = -\frac{K^2 + L^2}{2L^{1.4}K^{1.6}} < 0$$

Allora H è *semi definita negativa* perché i suoi autovalori sono tutti non positivi. La funzione è quasi concava.

Esempio con "confronto" dei 2 metodi (Silvester e Cartesio)

Sia

$\boldsymbol{A} = \begin{bmatrix} 2 & 1 & -5 \\ 1 & 7 & 4 \\ -5 & 4 & 19 \end{bmatrix}$ Che segno ha? Proviamo con Cartesio. Costruiamo il polinomio associato ad $\boldsymbol{A}$:

$$\overline{\boldsymbol{A}} = (\boldsymbol{A} - \lambda \boldsymbol{I}) = \begin{bmatrix} 2-\lambda & 1 & -5 \\ 1 & 7-\lambda & 4 \\ -5 & 4 & 19-\lambda \end{bmatrix} \text{ da cui } \det(\overline{\boldsymbol{A}}) = \begin{Vmatrix} 2-\lambda & 1 & -5 \\ 1 & 7-\lambda & 4 \\ -5 & 4 & 19-\lambda \end{Vmatrix} = 0 \text{ che implica}$$

$$\begin{aligned} p_{\bar{A}}(\lambda) &= -\lambda^3 + 28\lambda^2 - 143\lambda \\ &= \lambda(-\lambda^2 + 28\lambda - 143) \end{aligned}$$

Una radice è nulla. Possiamo quindi dire che **la matrice non è definita**. Potrà essere **semi definita**, positiva o negativa. Guardiamo il termine tra parentesi

$$-\lambda^2 + 28\lambda - 143$$

Ci sono due variazioni di segno nel polinomio e ciò permette di concludere che esso ammette due radici positive[11]. Quindi ***A*** è **semi definita** positiva.

Vediamo se Silvester produce lo stesso risultato.

$$\|H_1\| = 2 > 0 \quad \|H_2\| = 13 > 0 \quad \|H_3\| = 0$$

Non vi è alternanza di segno (positivo) ma il terzo minore di testa ha determinante nullo. D'altra parte quest'ultimo determinante è il determinante della matrice ***A*** che essendo pari al prodotto dei tre autovalori non può che essere nullo. La matrice è semi definita positiva.
Ognuna/o usi il metodo che le/gli appare più comodo.

Esercizio

Sia $\boldsymbol{V} = \begin{bmatrix} 2 & 0 & -5 \\ 1 & 7 & 4 \\ -1 & 4 & 19 \end{bmatrix}$

a) determinare con Cartesio e Silvester se $\boldsymbol{V}$ è definita positiva (lo è)
b) Mostrare che $\boldsymbol{V}^{1}$ è anch'essa positiva definita
c) Lo è anche $\boldsymbol{V}^{T}$?
d) Mostrare che $\boldsymbol{V}$ ha rango 3 e cercare di argomentare che ogni matrice positiva definita ha rango massimo (è utile Silvester)
e) Mostrare che dato un qualsiasi scalare $z > 0$, $z\boldsymbol{V}$ è definita positiva (è utile Silvester)
f) Mostrare che essendo $\boldsymbol{V}$ definita positiva esiste una matrice $\boldsymbol{C}$ tale che $\boldsymbol{V} = \boldsymbol{C}^T\boldsymbol{C}$

[11] Per la cronaca esse sono $\lambda_2 = 14 - \sqrt{53}$ e $\lambda_3 = 14 + \sqrt{53}$.

Appendice di Pro-memoria 3

Tabella riassuntiva della struttura matematica dei problemi di ottimizzazione *più popolari* in microeconomia (no problemi dinamici)

FUNZIONE DI “PERFORMANCE” *P*	CASI TIPICI di *fo*	VINCOLO	NOTE
$P = f_0(\boldsymbol{x})$ dove $\boldsymbol{x} = [x_1, x_2, \ldots, x_n]'$ deve essere scelto per l’ottimo e $f_0(\boldsymbol{x})$ è una funzione assegnata	Forma generale con $\boldsymbol{x}$ che non dipende da t	Nessuno	Max-Min liberi. Punti di stazionarietà
		$f_1(\boldsymbol{x}) = c_i$ dove c_i è una costante	Max-Min vincolati da eguaglianza (moltiplicatori di Lagrange)
		$f_1(\boldsymbol{x}) \leq c_i$ dove c_I è una costante Oppure $f_1(\boldsymbol{x}) \geq c_i$ dove c_I è una costante	Richiedono l’introduzione di una *slack variable* in modo da convertire il vincolo di diseguaglianza in un vincolo di eguaglianza. Condizioni di K-T.
	$P = \int_{t_\alpha(\boldsymbol{x})}^{t_\beta(\boldsymbol{x})} f_0(\boldsymbol{x}, t)dt$ dove $t_\alpha(\boldsymbol{x})$ e $t_\beta(\boldsymbol{x})$ sono funzioni date di $\boldsymbol{x}$	Nessuno	Come sopra primo rigo
		$\boldsymbol{x}$ indipendente da t e $c_i = \int_{t_\alpha(\boldsymbol{x})}^{t_\beta(\boldsymbol{x})} f_1(\boldsymbol{x}, t)dt$	*Max-Min* vincolati da eguaglianza (condizioni di Lagrange valide anche qui)
	$\boldsymbol{c}'\boldsymbol{x} = \sum_{\boldsymbol{j=1}}^{\boldsymbol{n}} c_j x_j$ Dove $\boldsymbol{c}' = [c_1, c_2, \ldots, c_n]$ è un vettore riga e $\boldsymbol{x}' = [x_1, x_2, \ldots, x_n]$	$0 \leq x_j$ $\sum_{\boldsymbol{j=1}}^{\boldsymbol{n}} a_{ij} x_j = b_i$ dove a_{ij} e b_i sono costanti	Genera, con modifiche, varie versioni di LP