IV-4 Massimi e minimi delle funzioni di più variabili

Indice

In questa dispensa vediamo alcuni risultati teorici riguardanti i punti di massimo e di minimo delle funzioni di più variabili. Vediamo anche come, in conseguenza di tali risultati, si procede nella ricerca dei punti di massimo e di minimo. Considereremo due situazioni generali ben distinte: la ricerca dei massimi e minimi cosiddetti *liberi* e la ricerca dei massimi e minimi *vincolati*. Questa parte dell'Analisi matematica è detta *ottimizzazione libera o vincolata*. Anche se i risultati teorici sono enunciati in generale per funzioni definite in $\mathbb{R}^n$ o in sottoinsiemi di $\mathbb{R}^n$, gli esempi riguarderanno, sia per i massimi e minimi liberi sia per quelli vincolati, soltanto funzioni definite in $\mathbb{R}^2$ o suoi sottoinsiemi.

1 Massimi e minimi liberi

Iniziamo con un paio di definizioni fondamentali, già incontrate nella seconda parte del corso per le funzioni di una variabile.

Definizione Sia $f : A \subset \mathbb{R}^n \to \mathbb{R}$ una funzione. Diciamo che il punto $\boldsymbol{x}^0 \in A$ è un **punto di massimo globale di** f se vale

$$f(\boldsymbol{x}^0) \geq f(\boldsymbol{x}), \quad \text{per ogni } \boldsymbol{x} \in A.$$

Analogamente diciamo che il punto $\boldsymbol{x}^0 \in A$ è un **punto di minimo globale di** f se vale

$$f(\boldsymbol{x}^0) \leq f(\boldsymbol{x}), \quad \text{per ogni } \boldsymbol{x} \in A.$$

Osservazione Un punto di massimo (minimo) globale è un punto in cui la funzione assume il suo valore massimo (minimo). L'aggettivo globale sta a significare che prendo in considerazione tutto il dominio in cui è definita la funzione. Come sempre non si deve fare confusione tra il *massimo* della funzione e il *punto di massimo*: il punto di massimo è un punto del dominio di f, mentre il massimo è il massimo dei valori che la funzione assume, quindi nel codominio di f. Pertanto, se ad esempio $f : \mathbb{R}^2 \to \mathbb{R}$, il punto di massimo appartiene ad $\mathbb{R}^2$, mentre il massimo è un numero reale.

Osservazione Non è detto ovviamente che tutte le funzioni abbiano punti di massimo o di minimo globale. Lo studente sa già fornire controesempi con funzioni di una sola variabile.

Definizione Sia $f : A \subset \mathbb{R}^n \to \mathbb{R}$ una funzione definita in un insieme A aperto. Diciamo che il punto $\boldsymbol{x}^0 \in A$ è un **punto di massimo locale di** f se esiste un intorno U di $\boldsymbol{x}^0$ tale che

$$f(\boldsymbol{x}^0) \geq f(\boldsymbol{x}), \quad \text{per ogni } \boldsymbol{x} \in U.$$

Analogamente diciamo che il punto $\boldsymbol{x}^0 \in A$ è un **punto di minimo locale di** f se esiste un intorno U di $\boldsymbol{x}^0$ tale che

$$f(\boldsymbol{x}^0) \leq f(\boldsymbol{x}), \quad \text{per ogni } \boldsymbol{x} \in U.$$

Quello che fa la differenza con la definizione precedente (massimo globale) è che ora ci si accontenta che la disequazione valga in un intorno del punto $\boldsymbol{x}^0$ e non in tutto il dominio di f.

Possiamo dare la definizione di punto di massimo (minimo) locale di una funzione definita in un generico sottoinsieme di $\mathbb{R}^n$ (anche non aperto), con una piccola avvertenza in più. Possiamo cioè dire che $\boldsymbol{x}^0 \in A$ è un punto di massimo locale di f se esiste un intorno U di $\boldsymbol{x}^0$ tale che

$$f(\boldsymbol{x}^0) \geq f(\boldsymbol{x}), \quad \text{per ogni } \boldsymbol{x} \in U \cap A \qquad \text{(e analogamente per il minimo).}$$

Osservazione Lo studente rifletta sull'importanza dell'ipotesi sull'insieme A. Consideri in particolare che, se ad esempio l'insieme A fosse chiuso e se $\boldsymbol{x}^0$ fosse un punto di frontiera per A, la prima definizione data non avrebbe senso: nessun punto di frontiera potrebbe essere punto di massimo o di minimo locale.

Osservazione Un punto di massimo (minimo) globale è anche di massimo (minimo) locale. Non vale il viceversa. Non è comunque una novità: questo si ha anche con le funzioni di una sola variabile. Si pensi ad esempio alla funzione $x \mapsto 2x^3 - 3x^2 + 1$, definita in tutto $\mathbb{R}$. Essa ha in 0 un punto di massimo locale e in 1 un punto di minimo locale, ma non ha punti né di massimo né di minimo globali. Può essere un utile esercizio di ripasso verificare il tutto con un semplice studio della funzione in questione.

Esempi La funzione $f : \mathbb{R}^2 \to \mathbb{R}$, definita da $f(x, y) = x^2 + y^2$, ha in $\boldsymbol{x}^0 = (0, 0)$ un punto di minimo globale.[1]

La funzione $f : \mathbb{R}^2 \to \mathbb{R}$, definita da $f(x, y) = e^{-x^2-y^2}$, ha in $\boldsymbol{x}^0 = (0, 0)$ un punto di massimo globale.[2]

Diamo ora un'altra importante definizione, che non dovrebbe sorprendere lo studente che ricorda le cose viste con le funzioni di una sola variabile.

Definizione Sia $f : A \subset \mathbb{R}^n \to \mathbb{R}$ una funzione di classe $\mathscr{C}^1$ nell'insieme A.[3] Se $\boldsymbol{x}^0$ è un punto *interno* ad A, diciamo che $\boldsymbol{x}^0$ è un **punto stazionario di** f se $\nabla f(\boldsymbol{x}^0) = \mathbf{0}$.

Osservazione Lo studente non confonda i concetti di punto di massimo (minimo) e di punto stazionario. Si noti che nelle definizioni di punto di massimo (minimo) non si fa nessuna ipotesi sulla derivabilità, e che invece nella definizione di punto stazionario la derivabilità è ovviamente necessaria. Si ricordi a tale proposito che ad esempio la funzione (di una variabile) $f(x) = |x|$ ha in 0 un punto di minimo (locale e globale) ma non è in tale punto derivabile.

Per comprendere e ricordare meglio i risultati che ora enuncerò, lo studente potrebbe rivedere rapidamente gli analoghi risultati visti nella seconda parte del corso per le funzioni definite in $\mathbb{R}$.

Proposizione Se f è una funzione di classe $\mathscr{C}^1$ nell'insieme A e $\boldsymbol{x}^0$ è un punto di massimo (o minimo) locale per f *interno* ad A, allora $\boldsymbol{x}^0$ è un punto stazionario di f.

Osservazione Tale risultato viene talvolta indicato come la *condizione necessaria del primo ordine* per l'esistenza di un punto di massimo o di minimo locali (per funzioni derivabili).[4]

Osservazione Si noti che, come avveniva per le funzioni di una sola variabile, la condizione necessaria vale nei punti interni. Lo studente vada a rivedere l'analoga proposizione incontrata nella seconda parte del corso e le osservazioni che la seguono.

Osservazione Ricordo che non vale il viceversa della proposizione appena vista. Non è vero cioè che, se $\boldsymbol{x}^0$ è stazionario, allora esso sia necessariamente o di massimo o di minimo. Si ricordi il classico controesempio (in una variabile) dato da $x \mapsto x^3$, che in $x_0 = 0$ ha un punto stazionario, che però non è né di massimo né di minimo.

Esempio Come esempio in $\mathbb{R}^2$ della stessa situazione possiamo considerare la funzione $f(x, y) = x^2 - y^2$. Si vede facilmente che l'origine $\boldsymbol{x}^0 = (0, 0)$ è un punto stazionario di f. Infatti si ha $\nabla f(x, y) = (2x, -2y)$ e risulta $\nabla f(\boldsymbol{x}^0) = \nabla f(0, 0) = (0, 0)$. Possiamo vedere altrettanto facilmente che l'origine non è punto né di massimo né di minimo di f.[5] A tale proposito ci possono tornare utili le restrizioni di f. Se consideriamo la restrizione di f all'asse x, cioè la retta di equazione $y = 0$, otteniamo la funzione

$$f\big|_{y=0} = f(x, 0) = x^2,$$

per cui $\boldsymbol{x}^0 = (0, 0)$ non può essere punto di massimo. Se invece consideriamo la restrizione di f all'asse y, cioè la retta di equazione $x = 0$, otteniamo la funzione

$$f\big|_{x=0} = f(0, y) = -y^2,$$

per cui $\boldsymbol{x}^0 = (0, 0)$ non può essere nemmeno punto di minimo.

Osservazione Possiamo in generale fare uso di un fatto abbastanza intuitivo: se f ha in $\boldsymbol{x}^0$ un punto di massimo (minimo) locale, allora ogni restrizione "passante per $\boldsymbol{x}^0$" deve avere anch'essa in $\boldsymbol{x}^0$ un punto di massimo (minimo) locale.

Osservazione Per la ricerca dei punti di massimo o di minimo, valendo la condizione del primo ordine, conviene quindi cercare intanto gli eventuali punti stazionari (si parla sempre di funzioni derivabili). La condizione però non

[1] Il risultato si trova immediatamente pensando che la funzione assume sempre valori non negativi e si annulla nell'origine.

[2] Basta osservare che nell'origine la funzione vale 1 e che in tutto il piano si ha $e^{-x^2-y^2} \le 1$, dato che questo equivale a $-x^2 - y^2 \le 0$.

[3] Ricordo che dicendo che f è di classe $\mathscr{C}^1$ in A si intende che f ha derivate continue in A. Più avanti, dicendo che f è di classe $\mathscr{C}^2$ in A si intende che f ha derivate seconde continue in A.

[4] Il risultato esprime infatti che la stazionarietà, cioè l'annullarsi della derivata, è condizione necessaria per avere un punto di massimo o di minimo. Dicendo primo ordine ci si riferisce all'ordine di derivazione: qui si usa infatti la derivata prima.

[5] Basterebbe ricordare il grafico di questa funzione (è una forma quadratica), che abbiamo incontrato parlando appunto di forme quadratiche.

consente di decidere se i punti stazionari trovati siano di massimo o di minimo (o nessuna delle due cose). Si osservi che non è una novità: anche con le funzioni di una variabile si procede nello stesso modo. Anche con quelle, inoltre, sapere che $f'(x_0) = 0$ non ci consente di dire nulla sul punto x_0. Con una variabile occorre studiare o il segno della derivata prima in prossimità di x_0 oppure la derivata seconda in x_0. Per le funzioni di più variabili non ci sono condizioni analoghe basate sul segno delle derivate parziali prime. Le condizioni per poter dire qualcosa sulla natura dei punti stazionari vengono dalle derivate seconde: i risultati generali che dicono come stanno le cose sono le *condizioni del secondo ordine*. Ecco una prima condizione, anche questa necessaria per la presenza di un punto di massimo o di minimo.

Proposizione Sia f una funzione di classe $\mathscr{C}^2$ nell'insieme A e sia $\boldsymbol{x}^0$ un punto stazionario di f interno ad A. Si può dimostrare che

- se $\boldsymbol{x}^0$ è punto di massimo locale, allora $\nabla^2 f(\boldsymbol{x}^0)$ è definito negativo o semidefinito negativo;
- se $\boldsymbol{x}^0$ è punto di minimo locale, allora $\nabla^2 f(\boldsymbol{x}^0)$ è definito positivo o semidefinito positivo.

Osservazione Ovviamente $\nabla^2 f(\boldsymbol{x}^0)$ è il gradiente secondo (matrice hessiana) calcolato nel punto $\boldsymbol{x}^0$. Qui occorre ricordare le definizioni viste un paio di lezioni fa: lo studente, se necessario, vada a rivedere le definizioni di forma quadratica (o matrice simmetrica) semidefinita, definita e indefinita.

Ecco ora le condizioni sufficienti per poter dire che un punto è di massimo o di minimo locale.

Proposizione Sia f una funzione di classe $\mathscr{C}^2$ nell'insieme A e sia $\boldsymbol{x}^0$ un punto stazionario di f interno ad A. Si può dimostrare che

- se $\nabla^2 f(\boldsymbol{x}^0)$ è definita negativa, allora $\boldsymbol{x}^0$ è punto di massimo locale;
- se $\nabla^2 f(\boldsymbol{x}^0)$ è definita positiva, allora $\boldsymbol{x}^0$ è punto di minimo locale;
- se $\nabla^2 f(\boldsymbol{x}^0)$ è indefinita, allora $\boldsymbol{x}^0$ non è né di massimo né di minimo locale.

Osservazione Per lo studio dei massimi e dei minimi locali di una funzione di più variabili (di classe $\mathscr{C}^2$) si inizia quindi con la ricerca dei punti stazionari. Poi occorre studiare la natura degli (eventuali) punti stazionari trovati. Si calcola il gradiente secondo e si studia il segno della forma quadratica associata al gradiente secondo. Se si è fortunati il gradiente secondo risulta o definito o indefinito, e quindi si può concludere. Se non si è fortunati, il gradiente secondo risulta semidefinito e non si può concludere nulla. Si noti quindi che non sempre si riesce a stabilire la natura del punto stazionario con le derivate seconde. Ma nemmeno questa è una novità: anche con funzioni di una variabile, se troviamo un punto x_0 tale che $f'(x_0) = f''(x_0) = 0$, da questa sola informazione non possiamo concludere nulla sul fatto che x_0 sia di massimo o di minimo.

Osservazione Il comportamento locale di una funzione in prossimità di un suo punto stazionario dipende quindi dal segno del suo gradiente secondo. In pratica succede che se il gradiente secondo è indefinito (si ricordi qual è il significato originale di questo termine) ci sono direzioni lungo le quali i valori della funzione aumentano e ci sono altre direzioni lungo le quali i valori di f diminuiscono. A questa situazione, in cui non c'è né massimo né minimo, si dà il nome di *punto di sella*. Se invece il gradiente secondo è definito, ad esempio positivo, (e anche qui si riveda la definizione del termine) in tutte le direzioni i valori della funzione aumentano e pertanto si ha la presenza di un punto di minimo.

Osservazione Un'importante applicazione dell'uso delle derivate parziali per la ricerca dei punti di massimo e di minimo in più variabili è il metodo dei minimi quadrati, che vedrete l'anno prossimo in Statistica. Si veda l'appendice A alla fine di questa dispensa.

È ora il momento di vedere qualche esempio di studio dei massimi e minimi di una funzione di due variabili.

Esempi

- Cerchiamo i punti di massimo e di minimo locale della funzione

$$f(x, y) = 2x^3 - 6xy + 3y^2.$$

I punti stazionari si trovano calcolando anzitutto il gradiente di f:

$$\nabla f(x, y) = (6x^2 - 6y, -6x + 6y),$$

ponendolo poi uguale al vettore nullo e risolvendo cioè il sistema

$$\begin{cases} 6x^2 - 6y = 0 \\ -6x + 6y = 0 \end{cases} \quad \text{, che ha le due soluzioni } (0,0) \text{ e } (1,1).$$ [6]

Questo significa che, in tutto $\mathbb{R}^2$, la funzione f ha soltanto questi due punti stazionari. Ora dobbiamo studiare la natura di questi punti. Ci serve intanto il gradiente secondo, che risulta

$$\nabla^2 f(x,y) = \begin{pmatrix} 12x & -6 \\ -6 & 6 \end{pmatrix}.$$

Studiamo la natura del punto $(0,0)$ calcolando

$$\nabla^2 f(0,0) = \begin{pmatrix} 0 & -6 \\ -6 & 6 \end{pmatrix}. \qquad \text{Si tratta di una matrice indefinita.}$$ [7]

Pertanto l'origine non è né di massimo né di minimo.
Studiamo la natura del punto $(1,1)$ calcolando

$$\nabla^2 f(1,1) = \begin{pmatrix} 12 & -6 \\ -6 & 6 \end{pmatrix}. \qquad \text{Si tratta di una matrice definita positiva.}$$ [8]

Pertanto $(1,1)$ è punto di minimo locale.

- Consideriamo i semplici casi delle funzioni

$$f(x,y) = x^2 + y^2 \qquad \text{e} \qquad g(x,y) = x^2 - y^2.$$

Abbiamo

$$\nabla f(x,y) = (2x, 2y) \qquad \text{e} \qquad \nabla g(x,y) = (2x, -2y).$$

Entrambe hanno, come unico punto stazionario, l'origine $(0,0)$.
I gradienti secondi sono

$$\nabla^2 f(x,y) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \qquad \text{e} \qquad \nabla^2 g(x,y) = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}.$$

Il primo è definito positivo, il secondo è indefinito. Pertanto f ha in $(0,0)$ un punto di minimo locale (in realtà è di minimo globale) e g ha in $(0,0)$ un punto che non è né di massimo né di minimo (classico punto di sella, come avevamo già concluso prima, utilizzando le restrizioni).

- Ora un caso in cui non si riesce a concludere. Consideriamo la funzione $f(x,y) = x^2 + y^4$. Il gradiente è

$$\nabla f(x,y) = (2x, 4y^3), \quad \text{da cui l'unico punto stazionario è l'origine.}$$

Il gradiente secondo è

$$\nabla^2 f(x,y) = \begin{pmatrix} 2 & 0 \\ 0 & 12y^2 \end{pmatrix} \qquad \text{e quindi} \qquad \nabla^2 f(0,0) = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix},$$

che però è semidefinito. Quindi le condizioni del secondo ordine non ci consentono di concludere. In realtà non è un caso particolarmente complicato dato che, osservando che la funzione è non negativa in tutto $\mathbb{R}^2$ e che nell'origine si annulla, possiamo dire facilmente che l'origine è punto di minimo globale.

[6] Attenzione in generale qui. I sistemi che si ottengono annullando un gradiente non sono in genere dei sistemi lineari, e quindi qui non valgono i risultati visti nella parte III, come i teoremi di Cramer o di Rouché-Capelli. Per risolvere un sistema come quello dell'esempio si può però semplicemente ricavare una delle incognite da una delle due equazioni e sostituire quanto trovato nell'altra. Qui ad esempio dalla seconda equazione si ricava $y = x$; sostituendo nella prima equazione si ha $x^2 - x = 0$, che dà per soluzioni $x = 0$ oppure $x = 1$. Da $x = 0$ si trova quindi $y = 0$ (da cui il punto $(0,0)$) e da $x = 1$ si trova $y = 1$ (da cui il punto $(1,1)$).

[7] Abbiamo imparato che lo studio del segno della f.q. può essere condotto attraverso il calcolo dei minori principali. In questo caso $\det \nabla^2 f(0,0) < 0$, e quindi, trattandosi di un minore principale di ordine pari, la f.q. è indefinita.

[8] I minori principali di NO sono entrambi positivi e quindi la f.q. è definita positiva.

Osservazione Generalmente la ricerca dei punti di massimo o minimo globali è più complicata. Una funzione potrebbe avere ad esempio due punti di massimo locale e un punto di minimo locale, ma per sapere chi sono i punti di massimo e di minimo globali (se ci sono) non possiamo concludere con il semplice confronto dei valori. Se la funzione è definita in tutto $\mathbb{R}^2$, non possiamo trascurare quello che succede ad esempio all'infinito.

Si consideri ancora l'esempio (già visto) di $f(x,y) = 2x^3 - 6xy + 3y^2$, che ha un minimo locale in $(1,1)$, oltre ad un punto stazionario né di massimo né di minimo in $(0,0)$. Non possiamo certo dire che $(1,1)$ è punto di minimo globale se la motivazione è che si tratta dell'unico punto di minimo trovato.

Possiamo però dire che f non ha punti di massimo globale, dato che se ne avesse, questi sarebbero anche punti di massimo locale, mentre di questi non ne abbiamo trovati.

In realtà f non ha nemmeno punti di minimo globale, in quanto è illimitata inferiormente. Infatti, se consideriamo la restrizione sull'asse x otteniamo $f\big|_{y=0} = f(x,0) = 2x^3$ che, variando x in tutto $\mathbb{R}$, non è limitata, né inferiormente, né superiormente. Quindi non lo è nemmeno f.

Esempi

- Cerchiamo i punti di massimo e di minimo della funzione $f(x,y) = x(e^y - 1)$, in tutto $\mathbb{R}^2$. Il gradiente è

$$\nabla f(x,y) = (e^y - 1, xe^y).$$

Per trovare i punti stazionari dobbiamo annullare il gradiente e considerare quindi il sistema

$$\begin{cases} e^y - 1 = 0 \\ xe^y = 0 \end{cases} \quad \text{cioè} \quad \begin{cases} e^y = 1 \\ x = 0. \end{cases}$$

Quindi c'è un solo punto stazionario: $(0,0)$. Il gradiente secondo è

$$\nabla^2 f(x,y) = \begin{pmatrix} 0 & e^y \\ e^y & xe^y \end{pmatrix} \qquad \text{e quindi} \qquad \nabla^2 f(0,0) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

La matrice è indefinita e pertanto $(0,0)$ non è né di massimo né di minimo. Quindi con le condizioni del secondo ordine in questo caso possiamo concludere. Proviamo a convincerci del risultato con le restrizioni.

Possiamo osservare che ad esempio le restrizioni lungo gli assi non dicono molto. Infatti

$$f\big|_{y=0} = f(x,0) = 0 \quad \text{e} \quad f\big|_{x=0} = f(0,y) = 0.$$

Proviamo con la restrizione lungo la retta di equazione $y = x$:

$$f\big|_{y=x} = f(x,x) = x(e^x - 1).$$

Lungo questa retta l'origine è un punto di minimo, dato che per ogni $x \neq 0$ risulta $x(e^x - 1) > 0$. Ma lungo la retta di equazione $y = -x$ si ha

$$f\big|_{y=-x} = f(x,-x) = x(e^{-x} - 1),$$

e qui l'origine è invece un punto di massimo, dato che risulta $x(e^{-x} - 1) < 0$ per ogni $x \neq 0$. Abbiamo quindi il tipico comportamento di un punto di sella.

Si vede facilmente che f non è limitata: basta calcolare i limiti all'infinito lungo le due restrizioni sulle bisettrici.[9]

Possiamo anche osservare che, quando una funzione si annulla in un punto stazionario, anche lo studio del segno della funzione può essere utile per stabilire la natura del punto. Nel nostro caso

$$x(e^y - 1) > 0 \quad \Leftrightarrow \quad \begin{cases} x > 0 \\ e^y - 1 > 0 \end{cases} \vee \begin{cases} x < 0 \\ e^y - 1 < 0 \end{cases} \quad \Leftrightarrow \quad \begin{cases} x > 0 \\ y > 0 \end{cases} \vee \begin{cases} x < 0 \\ y < 0. \end{cases}$$

Allora la funzione è positiva nel 1° e 3° quadrante, mentre è negativa nel 2° e 4°. Significa che in ogni intorno dell'origine può assumere valori maggiori e minori di zero e quindi l'origine non può essere né di massimo né di minimo.

[9] Due restrizioni particolarmente "convenienti" per studiare la natura dell'origine e la limitatezza di f sono quelle di equazione $y = \ln(x+1)$ e $y = \ln(1-x)$: infatti lungo la prima la funzione diventa $x \mapsto x^2$ e lungo la seconda $x \mapsto -x^2$.

- Cerchiamo i punti di massimo e minimo della funzione $f(x,y) = e^{-x^2-y^2}$, in tutto $\mathbb{R}^2$. Il gradiente è

$$\nabla f(x,y) = \left(-2xe^{-x^2-y^2}, -2ye^{-x^2-y^2}\right), \quad \text{da cui l'unico punto stazionario è } (0,0).$$

Il gradiente secondo è

$$\nabla^2 f(x,y) = \begin{pmatrix} -2(1+2x^2)e^{-x^2-y^2} & 4xye^{-x^2-y^2} \\ 4xye^{-x^2-y^2} & -2(1+2y^2)e^{-x^2-y^2} \end{pmatrix} \quad \text{e quindi} \quad \nabla^2 f(0,0) = \begin{pmatrix} -2 & 0 \\ 0 & -2 \end{pmatrix}.$$

Si tratta di una matrice definita negativa e quindi l'origine è un punto di massimo locale. È in realtà un punto di massimo globale: come già osservato in precedenza, basta considerare che $f(0,0) = 1$ e

$$e^{-x^2-y^2} \le 1 \quad \text{per ogni } (x,y) \in \mathbb{R}^2.$$

Gli esempi visti riguardano funzioni definite in tutto $\mathbb{R}^2$. Le condizioni enunciate però valgono in generale in insiemi aperti. Possono quindi essere utilizzate ad esempio nell'insieme dei punti *interni* al dominio di una funzione.

Esempi

- Consideriamo la funzione $f(x,y) = x \ln y$.

La funzione f è definita nel semipiano delle $y > 0$, che è un insieme aperto. Cerchiamo i punti stazionari calcolando intanto

$$\nabla f(x,y) = \left(\ln y, \frac{x}{y}\right), \quad \text{che si annulla nel punto } (0,1).$$

Il gradiente secondo è

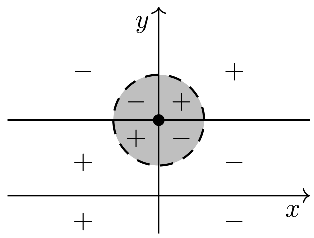
$$\nabla^2 f(x,y) = \begin{pmatrix} 0 & \frac{1}{y} \\ \frac{1}{y} & -\frac{x}{y^2} \end{pmatrix} \quad \text{e quindi} \quad \nabla^2 f(0,1) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

La matrice è indefinita e quindi il punto non è né di massimo né di minimo. Si può confermare questa conclusione utilizzando ad esempio le restrizioni sulle rette di equazione $y = 1 + x$ e $y = 1 - x$.

Anche in questo caso possiamo trovare conferma della natura del punto stazionario considerando il segno della funzione (dopo aver osservato che la funzione si annulla in $(0,1)$).

Si ha

$$x \ln y > 0 \quad \Leftrightarrow \quad \begin{cases} x > 0 \\ \ln y > 0 \end{cases} \vee \begin{cases} x < 0 \\ \ln y < 0 \end{cases} \quad \Leftrightarrow \quad \begin{cases} x > 0 \\ y > 1 \end{cases} \vee \begin{cases} x < 0 \\ y < 1 \end{cases}$$



Il segno della funzione è rappresentato qui a fianco. Come si vede in ogni intorno del punto stazionario ci sono punti in cui la funzione è positiva e punti in cui è negativa. Pertanto il punto non può essere né di massimo né di minimo.

- Consideriamo la funzione $f : C \to \mathbb{R}$, con $f(x,y) = x + y$ e $C = \left\{(x,y) \in \mathbb{R}^2 : x^2 + y^2 \le 1\right\}$.

Qui l'insieme in cui viene considerata la funzione è un insieme chiuso.[10]

Le condizioni studiate valgono nei punti interni a C, cioè nell'insieme $\left\{(x,y) \in \mathbb{R}^2 : x^2 + y^2 < 1\right\}$.[11]

Osservando che $\nabla f = (1,1)$, possiamo dire che non ci sono punti stazionari interni a C, e quindi non ci sono punti di massimo o di minimo interni a C. Attenzione a non pensare a questo punto che non ci siano punti di massimo o di minimo in C: certamente ce ne sono, dato che, essendo C un insieme chiuso e limitato ed essendo f continua, il teorema di Weierstrass dice che f ha almeno un punto di massimo e almeno un punto di minimo in C. Evidentemente tali punti non sono interni a C (se lo fossero sarebbero punti stazionari), e devono stare quindi sulla frontiera di C, cioè sulla circonferenza di equazione $x^2 + y^2 = 1$. Impareremo più avanti come trovarli.

[10] La funzione esiste anche all'esterno del cerchio C, ma la consideriamo solo nel cerchio.

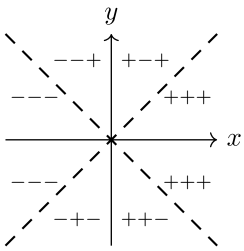
[11] Si rifletta su questo particolare: le condizioni, mettiamo quelle del primo ordine, non sono necessarie sul bordo del cerchio, cioè in altre parole non è detto che in un punto di massimo che sta sul bordo il gradiente debba essere nullo, dato che il fatto che il punto sia di massimo potrebbe dipendere non tanto dalle caratteristiche della funzione quanto da quelle del bordo e la condizione del primo ordine non tiene in nessun conto come è fatto il bordo.

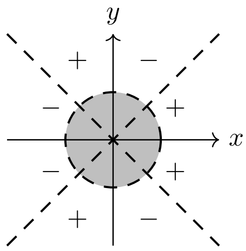
- Consideriamo la funzione $f(x,y) = x^2 - y^2$ nel quadrato $Q = [-1,1] \times [-1,1]$.

  Q è un insieme chiuso e limitato di $\mathbb{R}^2$, e quindi anche qui il teorema di Weierstrass garantisce che esistono un punto di massimo e un punto di minimo di f in Q. Annullando il gradiente di f, come già visto, si trova l'origine come unico punto stazionario, che però non è né di massimo né di minimo. Anche in questo caso allora i punti di massimo e minimo globali stanno sul bordo. Per trovarli si potrebbero utilizzare le restrizioni di f al bordo del quadrato. Lo studente provi a farlo:[12] troverà che ci sono due punti di massimo globale in $(1,0)$ e $(-1,0)$ (dove la funzione vale 1), e due punti di minimo globale in $(0,1)$ e $(0,-1)$ (dove la funzione vale -1). Si noti anche che la funzione si annulla nei punti di Q che stanno lungo le rette di equazioni $y = x$ e $y = -x$.

- Consideriamo la funzione $f(x,y) = x^3 - xy^2$ in tutto il piano.

  Le derivate parziali sono
  $$f'_x = 3x^2 - y^2 \qquad \text{e} \qquad f'_y = -2xy.$$
  Per trovare i punti stazionari dobbiamo considerare quindi il sistema
  $$\begin{cases} 3x^2 - y^2 = 0 \\ -2xy = 0. \end{cases} \quad \Leftrightarrow \quad \begin{cases} x = 0 \\ y^2 = 0 \end{cases} \vee \begin{cases} y = 0 \\ x^2 = 0. \end{cases}$$
  Entrambi forniscono come punto stazionario la stessa unica soluzione: $(0,0)$. Il gradiente secondo è
  $$\nabla^2 f(x,y) = \begin{pmatrix} 6x & -2y \\ -2y & -2x \end{pmatrix} \qquad \text{e quindi} \qquad \nabla^2 f(0,0) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$
  In questo caso le condizioni del secondo ordine non ci permettono di concludere, dato che la matrice è semidefinita. Con lo studio del segno si riesce a risolvere, osservando che la funzione si annulla nell'origine. Si ha
  $$x^3 - xy^2 > 0 \quad \Leftrightarrow \quad x(x-y)(x+y) > 0.$$
  Trattandosi di tre fattori possiamo studiare separatamente il segno di ciascuno, riportare i segni in un grafico e poi operare il prodotto dei segni. Si tenga conto che il primo fattore è positivo nel primo e quarto quadrante, il secondo è positivo al di sotto della retta di equazione $y = x$ ed il terzo è positivo al di sopra della retta di equazione $y = -x$. Si ottiene la rappresentazione qui sotto a sinistra, dove le terne di segni forniscono nell'ordine il segno di primo, secondo e terzo fattore. A destra è indicato il segno del prodotto, cioè il segno della funzione.





  Come si vede in ogni intorno del punto stazionario ci sono punti in cui la funzione è positiva e punti in cui è negativa. Pertanto il punto non può essere né di massimo né di minimo.

Esercizio 1.1 Si determinino i punti stazionari delle seguenti funzioni e si stabilisca la loro natura con le condizioni del secondo ordine. Nei casi in cui tali condizioni non sono sufficienti, si stabilisca la natura dei punti stazionari con lo studio del segno.

(a) $f(x,y) = x^2 - xy + y^2 - x + y$, in tutto $\mathbb{R}^2$

(b) $f(x,y) = x^2 - 3xy + y^2 + x + y + 1$, in tutto $\mathbb{R}^2$

(c) $f(x,y) = x^3 - xy + y^3$, in tutto $\mathbb{R}^2$

(d) $f(x,y) = x^2(e^y - 1)$, in tutto $\mathbb{R}^2$

(e) $f(x,y) = x^2 \ln^2 y$, nel dominio di f, cioè $\mathbb{R} \times (0,+\infty)$

(f) $f(x,y) = x^4 + y^4 - 4xy$, in tutto $\mathbb{R}^2$

(g) $f(x,y) = x + y - \frac{1}{3}(x^3 + y^3)$, in tutto $\mathbb{R}^2$

(h) $f(x,y) = (1 - x^2)y + x^2$, in tutto $\mathbb{R}^2$

[12] Ad esempio, per trovare la restrizione di f sul lato superiore del quadrato, basta usare la restrizione alla retta di equazione $y = 1$, con $x \in [-1,1]$. La restrizione di f è allora la funzione $f\big|_{y=1} = x^2 - 1$, con $x \in [-1,1]$.

2 Massimi e minimi vincolati

Passiamo ora all'ottimizzazione vincolata. Qui mi limito al caso bidimensionale, cioè di una funzione di due variabili. Si parla dei suoi massimi e minimi vincolati quando la funzione non viene considerata su tutto il suo dominio, ma soltanto sui punti di una curva del piano e qui si cercano i punti che risultano di massimo o di minimo rispetto ai valori che la funzione assume sulla curva.[13] Anche se il discorso potrebbe essere più generale, consideriamo solo il caso in cui tale curva è definita attraverso un'equazione. La curva in questione si chiama anche il *vincolo*.

Supponiamo quindi di avere una funzione $f : A \subset \mathbb{R}^2 \to \mathbb{R}$ di classe $\mathscr{C}^1$ sull'aperto A e una curva in A definita dall'equazione $g(x, y) = 0$. Supponiamo che anche la funzione di vincolo g abbia una certa regolarità, in particolare che sia anch'essa di classe $\mathscr{C}^1$. Vogliamo studiare come si determinano i punti di massimo/minimo locale della funzione f sul vincolo.

Consideriamo prima il caso in cui nell'equazione di vincolo sia possibile esplicitare una delle due variabili in funzione dell'altra. Poi esaminiamo un risultato più generale, applicabile anche se non è possibile esplicitare nessuna delle due variabili.

2.1 Vincolo "esplicitabile"

Supponiamo allora che nell'equazione $g(x, y) = 0$ si possa ricavare una delle due variabili in funzione dell'altra. Ci sarà modo quindi di scrivere il vincolo come

$$y = g_1(x) \qquad \text{oppure} \qquad x = g_2(y).$$

A questo punto sappiamo che i valori della funzione f sul vincolo sono dati nei due casi dalle restrizioni

$$f\big|_{y=g_1(x)} = f(x, g_1(x)) \qquad \text{oppure} \qquad f\big|_{x=g_2(y)} = f(g_2(y), y).$$

Definizione Nel primo caso diciamo che P è **punto stazionario vincolato** di f sul vincolo se

$$Df\big|_{y=g_1(x)}(x_0) = Df\big(x, g_1(x)\big)(x_0) = 0.$$ [14]

Osservazione Naturalmente, se x_0 è punto di massimo o di minimo per la funzione $f(x, g_1(x))$, allora $(x_0, y_0) = (x_0, g_1(x_0))$ è punto stazionario vincolato di f.

Osservazione Nel caso sia possibile ricavare la x con $x = g_2(y)$, diciamo che P è punto stazionario vincolato di f sul vincolo se

$$Df\big|_{x=g_2(y)}(y_0) = Df\big(g_2(y), y\big)(y_0) = 0.$$

Esempio Cerchiamo i punti stazionari vincolati della funzione $f(x, y) = x^2 + y^2$ sulla retta definita dall'equazione $x + y - 1 = 0$.

Possiamo esplicitare la y nel vincolo scrivendo $y = 1 - x$, con $x \in \mathbb{R}$. Quindi consideriamo la restrizione

$$f\big|_{y=1-x} = f(x, 1 - x) = 2x^2 - 2x + 1.$$

La derivata di questa è $4x - 2$ e si annulla in $x = \frac{1}{2}$, da cui $y = \frac{1}{2}$. Pertanto c'è il solo punto stazionario vincolato $(\frac{1}{2}, \frac{1}{2})$.

Esempio Cerchiamo i punti stazionari vincolati della funzione $f(x, y) = x^2-y^2$ sulla parabola di equazione $y-x^2 = 0$. Dopo aver esplicitato il vincolo nella forma $y = x^2$, consideriamo la restrizione

$$f\big|_{y=x^2} = f(x, x^2) = x^2 - x^4.$$

La derivata di questa è $2x - 4x^3 = 2x(1 - 2x^2)$ e si annulla in $x = 0$ oppure $x = \pm\frac{1}{\sqrt{2}}$. Pertanto i punti stazionari vincolati sono i punti $(0, 0)$, $(\frac{1}{\sqrt{2}}, \frac{1}{2})$ e $(-\frac{1}{\sqrt{2}}, \frac{1}{2})$.

Esempio Cerchiamo i punti stazionari vincolati della funzione $f(x, y) = x+y$ sulla parabola di equazione $x-y^2 = 0$.

[13] Non si cada in questo equivoco: non stiamo parlando di punti di massimo/minimo della funzione in generale e che appartengono alla curva, ma di punti che sono di massimo/minimo considerando soltanto i punti della curva. Come vedremo la situazione più comune è quella che un punto di massimo/minimo sul vincolo non è nemmeno un punto stazionario per la funzione nel suo intero dominio.

[14] Significa che la derivata della restrizione, che è una funzione della sola x, si annulla per $x = x_0$.

Possiamo esplicitare la x scrivendo $x = y^2$, con $y \in \mathbb{R}$. Quindi consideriamo la restrizione

$$f\big|_{x=y^2} = f(y^2, y) = y^2 + y.$$

La derivata di questa è $2y + 1$ e si annulla in $y = -\frac{1}{2}$, da cui $x = \frac{1}{4}$. Pertanto c'è il solo punto stazionario vincolato $(-\frac{1}{2}, \frac{1}{4})$.

Esempio Cerchiamo i punti stazionari vincolati della funzione $f(x, y) = x^2 - xy + y^2$ sulla parabola di equazione $x - y^2 + 1 = 0$. Dopo aver ricavato la x scrivendo $x = y^2 - 1$, consideriamo la restrizione

$$f\big|_{x=y^2} = f(y^2, y) = (y^2 - 1)^2 - y(y^2 - 1) + y^2 = y^4 - y^3 - y^2 + y + 1.$$

Calcolando la derivata di questa e ponendola uguale a zero otteniamo

$$4y^3 - 3y^2 - 2y + 1 = 0, \quad \text{cioè} \quad (y - 1)(4y^2 + y - 1) = 0.$$

Abbiamo tre possibili soluzioni, che sono $y = 1$, oppure $y = \frac{-1\pm\sqrt{17}}{8}$. Dall'equazione $x = y^2 - 1$ otteniamo i tre punti della parabola che sono i punti stazionari vincolati. Il primo è $(0, 1)$ e gli altri due sono $\big((\frac{-1-\sqrt{17}}{8})^2 - 1, \frac{-1-\sqrt{17}}{8}\big)$ e $\big((\frac{-1+\sqrt{17}}{8})^2 - 1, \frac{-1+\sqrt{17}}{8}\big)$.

Osservazione Può sembrare che il determinare i punti stazionari vincolati non sia più difficile che trovare i punti stazionari di una funzione di una variabile, dato che, negli esempi visti, in effetti abbiamo sempre dovuto annullare la derivata di una funzione di una sola variabile. Ovviamente questo è vero, a patto però di essere in grado di esplicitare il vincolo. Si pensi che già nel semplice caso in cui il vincolo è una circonferenza definita ad esempio dall'equazione $x^2 + y^2 = 1$, ricavare una variabile in funzione dell'altra non è semplicissimo come nei casi visti sopra.

Osservazione Sempre nei casi in cui si può esplicitare la relazione tra x e y che definisce il vincolo non è difficile stabilire se i punti stazionari vincolati trovati sono punti di massimo o di minimo vincolati: basta infatti usare le tecniche studiate in precedenza per le funzioni di una variabile.

Così, ad esempio, non è difficile capire che il punto stazionario $(\frac{1}{4}, -\frac{1}{2})$ del terzo esempio è un punto di minimo locale vincolato per la funzione $f(x, y) = x + y$ sulla parabola di equazione $x - y^2 = 0$. La restrizione di f al vincolo, cioè la funzione "parabolica" $y \mapsto y^2 + y$, ha in $y = -\frac{1}{2}$ un punto di minimo.

Analogamente il punto stazionario $(\frac{1}{2}, \frac{1}{2})$ del primo esempio è un punto di minimo locale vincolato per la funzione $f(x, y) = x^2 + y^2$ sulla retta di equazione $x + y - 1 = 0$. Anche qui la restrizione di f al vincolo, cioè la funzione $x \mapsto 2x^2 - 2x + 1$, ha in $x = \frac{1}{2}$ un punto di minimo.

Esercizio 2.1 Si determinino i punti stazionari vincolati delle seguenti funzioni, sul vincolo indicato, espresso in forma di equazione:

(a) $f(x, y) = x + y^2$, sul vincolo dato da $y - x^2 = 0$

(b) $f(x, y) = x^2 + 2xy + 2y^2$, sul vincolo dato da $2x - y + 1 = 0$

(c) $f(x, y) = x^3 - xy + y^3$, sul vincolo dato da $x + y + 1 = 0$

(d) $f(x, y) = (x^2 + 1) \ln y$, sul vincolo dato da $x^2 - y + 1 = 0$

2.2 Condizioni di Lagrange e funzione Lagrangiana

Per finire, vediamo un fondamentale risultato generale. Supponiamo che il vincolo sia dato in forma implicita come insieme delle soluzioni di un'equazione del tipo $g(x, y) = 0$, con g funzione di classe $\mathscr{C}^1$.[15]

Si può dimostrare che se (x_0, y_0) è un punto di massimo o di minimo vincolato di f e se e $\nabla g(x_0, y_0) \neq (0, 0)$, allora $\nabla f(x_0, y_0)$ è un multiplo di $\nabla g(x_0, y_0)$.[16]

La ricerca dei punti stazionari vincolati in questi casi si può effettuare quindi scrivendo il sistema

$$\begin{cases} \nabla f(x, y) = \lambda \nabla g(x, y) \\ g(x, y) = 0 \end{cases} \quad \text{che equivale a} \quad \begin{cases} \frac{\partial f}{\partial x}(x, y) = \lambda \frac{\partial g}{\partial x}(x, y) \\ \frac{\partial f}{\partial y}(x, y) = \lambda \frac{\partial g}{\partial y}(x, y) \\ g(x, y) = 0 \end{cases} \tag{1}$$

[15] Ribadisco che, se siamo in grado di esplicitare il vincolo, allora abbiamo già un modo per procedere. Quanto segue assume quindi particolare interesse nei casi in cui risulta complicato o non siamo in grado di esplicitare il vincolo.

[16] Significa quindi che i due vettori ∇f e ∇g sono proporzionali nel punto (x_0, y_0), cioè in altre parole che esiste uno scalare λ tale che $\nabla f = \lambda \nabla g$. Questo risultato è dovuto a Lagrange.

in cui le prime due equazioni richiedono che il punto cercato renda proporzionali i gradienti e la terza chiede invece che tale punto appartenga al vincolo. Anche il coefficiente di proporzionalità λ, detto **moltiplicatore di Lagrange** è un'incognita nel sistema. Quelle espresse nel sistema sono dette le **condizioni di Lagrange** per l'ottimalità vincolata. Il metodo si chiama anche **metodo dei moltiplicatori di Lagrange**. Si tratta di un metodo importantissimo, tanto nella teoria quanto nelle applicazioni. Esso ha una validità che va ben oltre l'ambito che stiamo qui considerando. Si tratta di uno dei metodi più generali della matematica e delle sue applicazioni.

Alla scrittura del sistema (1) si può arrivare anche definendo la cosiddetta **funzione Lagrangiana** (o semplicemente **Lagrangiana**) del problema:

$$\mathcal{L}(x, y, \lambda) = f(x, y) - \lambda g(x, y).$$

Al sistema (1), che raccoglie le *condizioni di Lagrange*, si arriva annullando il gradiente della Lagrangiana, e cioè

$$\nabla\mathcal{L}(x, y, \lambda) = \left(\frac{\partial f}{\partial x}(x, y) - \lambda \frac{\partial g}{\partial x}(x, y)\ ,\ \frac{\partial f}{\partial y}(x, y) - \lambda \frac{\partial g}{\partial y}(x, y)\ ,\ -g(x, y) \right).$$

È importante osservare che, dopo aver definito la Lagrangiana, la condizione di ottimalità vincolata si traduce nella stazionarietà (senza vincoli) della funzione Lagrangiana. Quindi possiamo dire che i punti di massimo/minimo vincolati vanno cercati tra i punti stazionari della funzione Lagrangiana.

Vediamo ora qualche esempio. Alcuni casi si potrebbero anche risolvere esplicitando una delle due variabili.

Esempio Consideriamo la funzione $f(x, y) = xy$ sul vincolo dato dall'equazione $x + y - 1 = 0$ e cerchiamo i suoi punti stazionari vincolati.

Osserviamo che il gradiente del vincolo non si annulla mai. La funzione Lagrangiana del problema è

$$\mathcal{L}(x, y, \lambda) = xy - \lambda(x + y - 1).$$

Il gradiente della Lagrangiana è

$$\nabla\mathcal{L}(x, y, \lambda) = \big(y - \lambda, x - \lambda, -x - y + 1\big)$$

e quindi le condizioni di Lagrange sono

$$\begin{cases} y - \lambda = 0 \\ x - \lambda = 0 \\ x + y = 1 \end{cases} \quad \text{cioè} \quad \begin{cases} y = \lambda \\ x = \lambda \\ x + y = 1 \end{cases} \quad \text{cioè} \quad \begin{cases} y = \lambda \\ x = \lambda \\ 2\lambda = 1 \end{cases}$$

dal quale si ottiene $\lambda = x = y = \frac{1}{2}$.

Pertanto si ha l'unico punto stazionario vincolato $(\frac{1}{2}, \frac{1}{2})$. Si noti che il punto trovato sta sul vincolo e in questo punto i gradienti di $f(x, y) = xy$ e di $g(x, y) = x + y - 1$ sono rispettivamente $\nabla f(\frac{1}{2}, \frac{1}{2}) = (\frac{1}{2}, \frac{1}{2})$ e $\nabla g(\frac{1}{2}, \frac{1}{2}) = (1, 1)$, chiaramente proporzionali.

Esempio Consideriamo l'esempio già incontrato della funzione $f(x, y) = x^2 - y^2$ sul vincolo dato dall'equazione $x^2 - y = 0$.

Osserviamo che il gradiente del vincolo si annulla soltanto nell'origine. La funzione Lagrangiana del problema è

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

Il gradiente della Lagrangiana è

$$\nabla\mathcal{L}(x, y, \lambda) = \big(2x - 2\lambda x, -2y + \lambda, -x^2 + y\big)$$

e quindi le condizioni di Lagrange sono

$$\begin{cases} 2x - 2\lambda x = 0 \\ -2y + \lambda = 0 \\ x^2 = y \end{cases} \quad \text{cioè} \quad \begin{cases} 2x(1 - \lambda) = 0 \\ 2y = \lambda \\ x^2 = y. \end{cases}$$

Ragionando sulla prima equazione si ha che deve essere o $x = 0$ oppure $\lambda = 1$. Con $x = 0$ si ottiene $y = 0$ dalla terza e $\lambda = 0$ dalla seconda. Altrimenti con $\lambda = 1$ si ha $y = \frac{1}{2}$ dalla seconda e $x = \pm\frac{1}{\sqrt{2}}$ dalla terza.

Pertanto si trovano tre punti stazionari vincolati: $(0, 0)$, $(\frac{1}{\sqrt{2}}, \frac{1}{2})$ e $(-\frac{1}{\sqrt{2}}, \frac{1}{2})$, già trovati prima per altra via.

Esempio Altro caso già visto: troviamo i punti stazionari vincolati della funzione $f(x, y) = x + y$ sul vincolo dato dall'equazione $x - y^2 = 0$.

Osserviamo che il gradiente del vincolo non si annulla mai. La funzione Lagrangiana del problema è

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

Le condizioni di Lagrange sono

$$\begin{cases} 1 - \lambda = 0 \\ 1 + 2\lambda y = 0 \\ x - y^2 = 0 \text{ [17]} \end{cases} \quad \text{cioè} \quad \begin{cases} \lambda = 1 \\ 1 + 2y = 0 \\ x = y^2 \end{cases} \quad \text{cioè} \quad \begin{cases} \lambda = 1 \\ y = -\frac{1}{2} \\ x = \frac{1}{4}. \end{cases}$$

Unico punto stazionario vincolato è quindi $(\frac{1}{4}, -\frac{1}{2})$.

Esempio Riprendiamo la funzione $f(x, y) = x^2 + y^2$ sul vincolo dato dall'equazione $x + y - 1 = 0$.
Osserviamo che il gradiente del vincolo non si annulla mai. La funzione Lagrangiana del problema è

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

Le condizioni di Lagrange sono

$$\begin{cases} 2x - \lambda = 0 \\ 2y - \lambda = 0 \\ x + y - 1 = 0 \end{cases} \quad \text{cioè} \quad \begin{cases} x = \lambda/2 \\ y = \lambda/2 \\ \lambda/2 + \lambda/2 - 1 = 0 \end{cases} \quad \text{cioè} \quad \begin{cases} \lambda = 1 \\ x = 1/2 \\ y = 1/2. \end{cases}$$

Unico punto stazionario vincolato è quindi $(\frac{1}{2}, \frac{1}{2})$.

Esempio Consideriamo la funzione $f(x, y) = 2x + 2y$ sul vincolo dato dall'equazione $x^2 + y^2 - 2 = 0$, notando che qui esplicitare la x o la y non è così banale.[18]

Osserviamo che il gradiente del vincolo si annulla soltanto nell'origine e questo punto non appartiene al vincolo. La funzione Lagrangiana del problema è

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

Le condizioni di Lagrange sono

$$\begin{cases} 2 - 2\lambda x = 0 \\ 2 - 2\lambda y = 0 \\ x^2 + y^2 - 2 = 0 \end{cases} \quad \text{cioè} \quad \begin{cases} \lambda x = 1 \\ \lambda y = 1 \\ x^2 + y^2 - 2 = 0. \end{cases}$$

Osserviamo che $\lambda = 0$ non porta a soluzioni.[19] Quindi, dato che deve essere $\lambda \neq 0$, possiamo dividere per λ e otteniamo

$$\begin{cases} x = 1/\lambda \\ y = 1/\lambda \\ 1/\lambda^2 + 1/\lambda^2 - 2 = 0 \end{cases} \quad \text{cioè} \quad \begin{cases} x = 1/\lambda \\ y = 1/\lambda \\ \lambda^2 = 1 \end{cases} \quad \text{e quindi} \quad \begin{cases} \lambda = \pm 1 \\ x = 1/\lambda \\ y = 1/\lambda. \end{cases}$$

Le soluzioni sono pertanto i due punti $(1, 1)$ e $(-1, -1)$.

Si noterà che in questi esempi abbiamo parlato solo di punti stazionari (vincolati), e non di punti di massimo o di minimo. Lo studente lungimirante avrà capito che le condizioni di Lagrange che abbiamo visto sono a tutti gli effetti condizioni del primo ordine. Ci si aspetta allora anche condizioni del secondo ordine, che usano cioè le derivate seconde. Ci sono infatti, ma presumibilmente il tempo a mia disposizione per questo corso sta per finire. Non vediamo queste condizioni.

Esempio In questo esempio conclusivo vediamo che la condizione che il gradiente del vincolo non si annulli è importante. Consideriamo la funzione $f(x, y) = x^2 - y$ sul vincolo dato dall'equazione $y^2 = 0$. Il vincolo individua semplicemente l'asse x. La restrizione è

$$f\big|_{y=0} = f(x, 0) = x^2$$

[17] Si ricordi che l'ultima condizione di Lagrange è semplicemente l'appartenenza al vincolo, che quindi può essere riscritta direttamente, senza fare il calcolo della derivata parziale rispetto al moltiplicatore.

[18] Non è difficile ma occorre considerare due casi possibili. Infatti, se vogliamo ad esempio esplicitare y in funzione di x possiamo riscrivere il vincolo nella forma $y^2 = 2 - x^2$ ma ora dobbiamo tenere in conto che non c'è un solo modo di scrivere y in funzione di x, dato che può essere $y = \sqrt{2 - x^2}$ oppure $y = -\sqrt{2 - x^2}$. Quindi per questa strada dovremmo considerare le due possibilità separatamente. Col metodo dei moltiplicatori di Lagrange questo non serve.

[19] Con $\lambda = 0$ le prime due equazioni non sono certamente soddisfatte.

e quindi l'origine è un punto di minimo vincolato. Osserviamo che il gradiente del vincolo è $\nabla g(x, y) = (0, 2y)$ e quindi $\nabla g(0, 0) = (0, 0)$.

Se scriviamo la funzione Lagrangiana abbiamo

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

e quindi le condizioni di Lagrange sono

$$\begin{cases} 2x = 0 \\ -1 - 2\lambda y = 0 \\ y^2 = 0 \end{cases} \quad \text{cioè} \quad \begin{cases} x = 0 \\ -1 - 2\lambda y = 0 \\ y = 0. \end{cases}$$

Il sistema è chiaramente impossibile. Pertanto il punto di minimo vincolato non è stazionario per la funzione Lagrangiana. Venendo a mancare la condizione sul gradiente del vincolo le ipotesi non sono tutte rispettate e la tesi è falsa.

Esercizio 2.2 Si determinino i punti stazionari vincolati delle seguenti funzioni, sul vincolo indicato, espresso in forma di equazione: si utilizzi il metodo dei moltiplicatori di Lagrange.

(a) $f(x, y) = xy$, sul vincolo dato da $2x - y + 1 = 0$

(b) $f(x, y) = x^2 + y^2$, sul vincolo dato da $x - y + 1 = 0$

(c) $f(x, y) = x$, sul vincolo dato da $x^2 + y^2 - 1 = 0$

(d) $f(x, y) = x^2 - y^2$, sul vincolo dato da $x + y - 1 = 0$

3 Soluzioni degli esercizi

Esercizio 1.1

(a) Il gradiente della funzione $f(x, y) = x^2 - xy + y^2 - x + y$ è

$$\nabla f(x, y) = (2x - y - 1, -x + 2y + 1)\,.$$

Quindi scriviamo il sistema

$$\begin{cases} 2x - y - 1 = 0 \\ -x + 2y + 1 = 0 \end{cases} \quad \text{cioè} \quad \begin{cases} 2x - y = 1 \\ x - 2y = 1. \end{cases}$$

Si tratta di un sistema di equazioni lineari. Con le tecniche imparate nella parte III di questo corso si trova facilmente l'unica soluzione $(\frac{1}{3}, -\frac{1}{3})$, che è quindi l'unico punto stazionario della funzione.

Per stabilire la natura di tale punto calcoliamo il gradiente secondo

$$\nabla^2 f(x, y) = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$$

che risulta definita positiva. Il punto $(\frac{1}{3}, -\frac{1}{3})$ è pertanto un punto di minimo locale.

(b) Il gradiente della funzione $f(x, y) = x^2 - 3xy + y^2 + x + y + 1$ è

$$\nabla f(x, y) = (2x - 3y + 1, -3x + 2y + 1)\,.$$

Quindi scriviamo il sistema

$$\begin{cases} 2x - 3y + 1 = 0 \\ -3x + 2y + 1 = 0 \end{cases} \quad \text{cioè} \quad \begin{cases} 2x - 3y = -1 \\ -3x + 2y = -1. \end{cases}$$

Si trova l'unica soluzione $(1, 1)$, che è l'unico punto stazionario della funzione.

Per stabilire la natura del punto calcoliamo il gradiente secondo

$$\nabla^2 f(x, y) = \begin{pmatrix} 2 & -3 \\ -3 & 2 \end{pmatrix}$$

che risulta indefinita. Il punto $(1, 1)$ non è pertanto né di massimo né di minimo locale (è un punto di sella).

(c) Il gradiente della funzione $f(x,y) = x^3 - xy + y^3$ è

$$\nabla f(x,y) = \left(3x^2 - y, -x + 3y^2\right).$$

Quindi scriviamo il sistema

$$\begin{cases} 3x^2 - y = 0 \\ -x + 3y^2 = 0 \end{cases} \quad \text{cioè} \quad \begin{cases} y = 3x^2 \\ -x + 27x^4 = 0 \end{cases} \quad \text{cioè} \quad \begin{cases} x(27x^3 - 1) = 0 \\ y = 3x^2. \end{cases}$$

Con $x = 0$ si trova $y = 0$ e quindi la soluzione $(0,0)$. Con $x = 1/3$ si trova $y = 1/3$ e quindi la soluzione $(\frac{1}{3}, \frac{1}{3})$. Vi sono dunque due punti stazionari: l'origine e $(\frac{1}{3}, \frac{1}{3})$.

Per stabilire la natura dei punti stazionari calcoliamo il gradiente secondo

$$\nabla^2 f(x,y) = \begin{pmatrix} 6x & -1 \\ -1 & 6y \end{pmatrix}.$$

Calcolando questa matrice nei due punti si ha allora

$$\nabla^2 f(0,0) = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}, \quad \text{che è indefinita e}$$

$$\nabla^2 f(\tfrac{1}{3}, \tfrac{1}{3}) = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}, \quad \text{che è definita positiva.}$$

Il punto $(0,0)$ non è né di massimo né di minimo locale, mentre il punto $(\frac{1}{3}, \frac{1}{3})$ è di minimo locale.

(d) Il gradiente della funzione $f(x,y) = x^2(e^y - 1)$ è

$$\nabla f(x,y) = \left(2x(e^y - 1), x^2 e^y\right).$$

Quindi scriviamo il sistema

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

Dalla seconda equazione, dato che non può essere $e^y = 0$, deve necessariamente essere $x = 0$. La prima equazione è allora necessariamente soddisfatta. Pertanto le soluzioni sono tutti i punti con $x = 0$, e cioè tutti i punti dell'asse y. Questi sono i punti stazionari.

Per stabilirne la natura calcoliamo il gradiente secondo

$$\nabla^2 f(x,y) = \begin{pmatrix} 2(e^y - 1) & 2xe^y \\ 2xe^y & x^2 e^y \end{pmatrix}.$$

Nei punti del tipo $(0,y)$ il gradiente secondo vale

$$\nabla^2 f(0,y) = \begin{pmatrix} 2(e^y - 1) & 0 \\ 0 & 0 \end{pmatrix}, \quad \text{che è semidefinita (positiva o negativa a seconda di } y\text{).}$$

Le condizioni del secondo ordine non consentono di concludere. Possiamo allora studiare il segno della funzione f: la funzione è nulla sugli assi, positiva sul primo e sul secondo quadrante e negativa sul terzo e quarto quadrante. Quindi i punti del tipo $(0,y)$ con $y > 0$ sono di minimo locale, quelli del tipo $(0,y)$ con $y < 0$ sono di massimo locale, mentre l'origine non è né di massimo né di minimo.

(e) La funzione $f(x,y) = x^2 \ln^2 y$ è ovviamente definita sulle $y > 0$. Il gradiente della funzione è

$$\nabla f(x,y) = \left(2x \ln^2 y, \frac{2x^2 \ln y}{y}\right).$$

Quindi scriviamo il sistema

$$\begin{cases} 2x \ln^2 y = 0 \\ \frac{2x^2 \ln y}{y} = 0. \end{cases}$$

Dalla prima, se $x = 0$ anche la seconda è soddisfatta, e quindi sono soluzioni tutti i punti del tipo $(0,y)$, con $y > 0$. Poi, sempre dalla prima, se $y = 1$ anche la seconda è soddisfatta e pertanto sono soluzioni anche tutti i punti del tipo $(x, 1)$, con x qualunque. Allora i punti stazionari sono quelli del semiasse positivo delle y e quelli della retta di equazione $y = 1$.

Per stabilirne la natura calcoliamo il gradiente secondo

$$\nabla^2 f(x,y) = \begin{pmatrix} 2\ln^2 y & 4x\ln y \cdot \frac{1}{y} \\ 4x\ln y \cdot \frac{1}{y} & 2x^2 \cdot \frac{1-\ln y}{y^2} \end{pmatrix}.$$

Nei punti del tipo $(0, y)$ il gradiente secondo vale

$$\nabla^2 f(0,y) = \begin{pmatrix} 2\ln^2 y & 0 \\ 0 & 0 \end{pmatrix}, \quad \text{che è semidefinita positiva.}$$

Nei punti del tipo $(x, 1)$ il gradiente secondo vale

$$\nabla^2 f(x,1) = \begin{pmatrix} 0 & 0 \\ 0 & 2x^2 \end{pmatrix}, \quad \text{che è ancora semidefinita positiva.}$$

Le condizioni del secondo ordine non consentono di concludere in nessun caso. Lo studio del segno della funzione, che è nulla sui punti stazionari e per il resto positiva, porta a dire che tutti i punti stazionari sono di minimo (globale).

(f) Il gradiente della funzione $f(x, y) = x^4 + y^4 - 4xy$ è

$$\nabla f(x,y) = \big(4x^3 - 4y, 4y^3 - 4x\big).$$

Quindi scriviamo il sistema

$$\begin{cases} 4x^3 - 4y = 0 \\ 4y^3 - 4x = 0 \end{cases} \quad \text{cioè} \quad \begin{cases} 4x^3 = 4y \\ 4y^3 = 4x \end{cases} \quad \text{cioè} \quad \begin{cases} x^3 = y \\ x^9 = x \end{cases}$$

e quest'ultimo fornisce le soluzioni $(0, 0)$, $(1, 1)$ e $(-1, -1)$, che sono i punti stazionari.

Il gradiente secondo è

$$\nabla^2 f(x,y) = \begin{pmatrix} 12x^2 & -4 \\ -4 & 12y^2 \end{pmatrix}.$$

Pertanto si ha

$$\nabla^2 f(0,0) = \begin{pmatrix} 0 & -4 \\ -4 & 0 \end{pmatrix}, \quad \text{che è indefinita.}$$

Si ha poi

$$\nabla^2 f(1,1) = \nabla^2 f(-1,-1) = \begin{pmatrix} 12 & -4 \\ -4 & 12 \end{pmatrix}, \quad \text{che è definita positiva.}$$

Pertanto possiamo concludere che $(1, 1)$ e $(-1, -1)$ sono punti di minimo locale, mentre $(0, 0)$ non è né di massimo né di minimo.

(g) Il gradiente della funzione $f(x, y) = x + y - \frac{1}{3}(x^3 + y^3)$ è

$$\nabla f(x,y) = \big(1 - x^2, 1 - y^2\big).$$

Quindi scriviamo il sistema

$$\begin{cases} 1 - x^2 = 0 \\ 1 - y^2 = 0 \end{cases} \quad \text{cioè} \quad \begin{cases} x^2 = 1 \\ y^2 = 1 \end{cases}$$

e quest'ultimo fornisce le soluzioni $(1, 1)$, $(1, -1)$, $(-1, 1)$ e $(-1, -1)$.

Il gradiente secondo è

$$\nabla^2 f(x,y) = \begin{pmatrix} -2x & 0 \\ 0 & -2y \end{pmatrix}.$$

Pertanto si ha

$$\nabla^2 f(1,1) = \begin{pmatrix} -2 & 0 \\ 0 & -2 \end{pmatrix}, \quad \text{che è definita negativa;}$$

poi
$$\nabla^2 f(1,-1) = \begin{pmatrix} -2 & 0 \\ 0 & 2 \end{pmatrix}, \quad \text{che è indefinita;}$$
poi
$$\nabla^2 f(-1,1) = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}, \quad \text{che è indefinita;}$$
e infine
$$\nabla^2 f(-1,-1) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}, \quad \text{che è definita positiva.}$$
Pertanto possiamo concludere che $(1,1)$ è un punto di massimo locale, $(-1,-1)$ è punto di minimo locale, $(1,-1)$ e $(-1,1)$ non sono né di massimo né di minimo.

(h) Il gradiente della funzione $f(x,y) = (1-x^2)y + x^2$ è
$$\nabla f(x,y) = \big(-2xy + 2x, 1 - x^2\big)\,.$$
Quindi scriviamo il sistema
$$\begin{cases} -2xy + 2x = 0 \\ 1 - x^2 = 0 \end{cases} \quad \text{cioè} \quad \begin{cases} 2x(1-y) = 0 \\ x^2 = 1 \end{cases}$$
e quest'ultimo fornisce le soluzioni $(1,1)$ e $(-1,1)$. Il gradiente secondo è
$$\nabla^2 f(x,y) = \begin{pmatrix} -2y+2 & -2x \\ -2x & 0 \end{pmatrix}.$$
Pertanto si ha
$$\nabla^2 f(1,1) = \begin{pmatrix} 0 & -2 \\ -2 & 0 \end{pmatrix}, \quad \text{che è indefinita}$$
e
$$\nabla^2 f(-1,1) = \begin{pmatrix} 0 & 2 \\ 2 & 0 \end{pmatrix}, \quad \text{che è indefinita.}$$
Pertanto possiamo concludere che i due punti stazionari non sono né di massimo né di minimo.

Esercizio 2.1

(a) Possiamo esplicitare il vincolo nella forma $y = x^2$. Allora consideriamo la restrizione
$$f\big|_{y=x^2} = f(x,x^2) = x + x^4.$$
La derivata si annulla in $x = -\frac{1}{\sqrt[3]{4}} = -4^{-1/3}$. Il punto stazionario vincolato è quindi $(-4^{-1/3}, 4^{-2/3})$.

(b) Possiamo esplicitare il vincolo nella forma $y = 2x+1$. Allora consideriamo la funzione
$$f\big|_{y=2x+1} = f(x,2x+1) = x^2 + 2x(2x+1) + 2(2x+1)^2 = 13x^2 + 10x + 2.$$
La derivata si annulla in $x = -\frac{5}{13}$. Il punto stazionario vincolato è quindi $(-\frac{5}{13}, \frac{3}{13})$.

(c) Possiamo esplicitare il vincolo nella forma $y = -x-1$. Allora consideriamo la funzione
$$f\big|_{y=-x-1} = f(x,-x-1) = x^3 - x(-x-1) + (-x-1)^3 = -2x^2 - 2x - 1.$$
La derivata si annulla in $x = -\frac{1}{2}$. Il punto stazionario vincolato è quindi $(-\frac{1}{2}, -\frac{1}{2})$.

(d) Possiamo esplicitare il vincolo nella forma $y = x^2+1$. Allora consideriamo la funzione
$$f\big|_{y=x^2+1} = f(x,x^2+1) = (x^2+1)\ln(x^2+1).$$
La derivata è $x \mapsto 2x\ln(x^2+1) + (x^2+1)\frac{2x}{x^2+1} = 2x\big(\ln(x^2+1)+1\big)$ e si annulla soltanto per $x=0$. Quindi c'è il solo punto stazionario vincolato $(0,1)$.

Esercizio 2.2

(a) Abbiamo $f(x,y) = xy$ e poniamo $g(x,y) = 2x - y + 1$.

La funzione Lagrangiana del problema è

$$\mathcal{L}(x,y,\lambda) = xy - \lambda(2x - y + 1).$$

Le condizioni di Lagrange sono

$$\begin{cases} y - 2\lambda = 0 \\ x + \lambda = 0 \\ 2x - y + 1 = 0 \end{cases} \quad \text{cioè} \quad \begin{cases} y = 2\lambda \\ x = -\lambda \\ -2\lambda - 2\lambda + 1 = 0 \end{cases} \quad \text{cioè} \quad \begin{cases} \lambda = \frac{1}{4} \\ x = -\frac{1}{4} \\ y = \frac{1}{2}. \end{cases}$$

Pertanto si ha l'unico punto stazionario vincolato $(-\frac{1}{4}, \frac{1}{2})$.

(b) Abbiamo $f(x,y) = x^2 + y^2$ e poniamo $g(x,y) = x - y + 1$.

La funzione Lagrangiana del problema è

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

Le condizioni di Lagrange sono

$$\begin{cases} 2x - \lambda = 0 \\ 2y + \lambda = 0 \\ x - y + 1 = 0 \end{cases} \quad \text{cioè} \quad \begin{cases} x = \lambda/2 \\ y = -\lambda/2 \\ \lambda/2 + \lambda/2 + 1 = 0 \end{cases} \quad \text{cioè} \quad \begin{cases} \lambda = -1 \\ x = -\frac{1}{2} \\ y = \frac{1}{2}. \end{cases}$$

Pertanto si ha l'unico punto stazionario vincolato $(-\frac{1}{2}, \frac{1}{2})$.

(c) Abbiamo $f(x,y) = x$ e poniamo $g(x,y) = x^2 + y^2 - 1$.

La funzione Lagrangiana del problema è

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

Le condizioni di Lagrange sono

$$\begin{cases} 1 - 2\lambda x = 0 \\ -2\lambda y = 0 \\ x^2 + y^2 - 1 = 0. \end{cases}$$

Fissiamo l'attenzione sulla seconda equazione, che è soddisfatta o con $\lambda = 0$ o con $y = 0$. Il valore $\lambda = 0$ non è accettabile, in quanto per tale valore la prima equazione non è soddisfatta. Allora proviamo con $y = 0$. Sostituendo nella terza equazione, otteniamo $x^2 = 1$, che è verificata con $x = \pm 1$. Pertanto, con $x = 1$ otteniamo la soluzione $(1,0)$ (e $\lambda = \frac{1}{2}$) e con $x = -1$ otteniamo la soluzione $(-1,0)$ (e $\lambda = -\frac{1}{2}$).

Abbiamo dunque due punti stazionari vincolati: $(1,0)$ e $(-1,0)$.

(d) Abbiamo $f(x,y) = x^2 - y^2$ e poniamo $g(x,y) = x + y - 1$.

La funzione Lagrangiana del problema è

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

Le condizioni di Lagrange sono

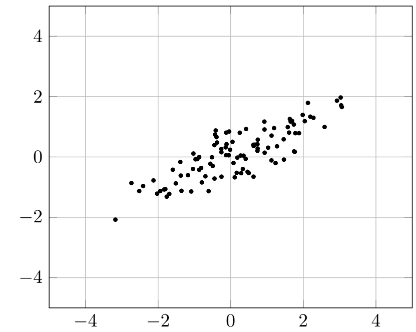
$$\begin{cases} 2x - \lambda = 0 \\ -2y - \lambda = 0 \\ x + y - 1 = 0 \end{cases} \quad \text{cioè} \quad \begin{cases} x = \lambda/2 \\ y = -\lambda/2 \\ \lambda/2 - \lambda/2 = 1 \end{cases}$$

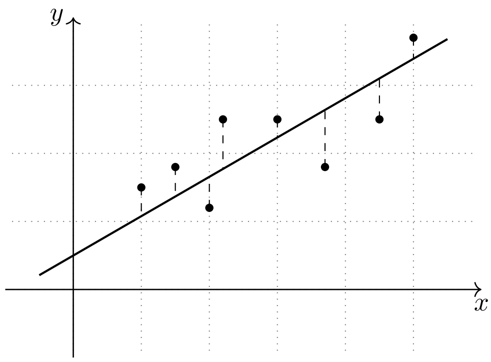
che è evidentemente impossibile. Pertanto non ci sono punti stazionari vincolati.

4 Appendice – Interpolazione con il metodo dei Minimi quadrati

Concludo questa dispensa con un'applicazione di quanto fin qui visto, relativamente alla ricerca dei massimi e minimi delle funzioni di più variabili. Si tratta di un argomento molto importante nella Statistica, che quindi ritroverete l'anno prossimo. Presento la questione da un punto di vista prettamente matematico. Nel corso di Statistica vedrete tutte le relative considerazioni, legate a quella disciplina.

Supponiamo di avere un certo numero di punti nel piano, diciamo N punti, quindi N coppie del tipo (x_i, y_i), con $i = 1, 2, \dots, N$. Una raffigurazione di tali punti potrebbe essere qualcosa che assomiglia alla figura qui sotto a sinistra.





Consideriamo ora una retta nel piano, di equazione $y = b_0 + b_1 x$. Per misurare "quanto la retta si adatta ai punti" possiamo considerare la quantità

$$\sum_{i=1}^{N} (b_0 + b_1 x_i - y_i)^2. \tag{2}$$

Serve qualche commento. Quello che abbiamo fatto è considerare, per ogni punto (x_i, y_i), la differenza tra la sua ordinata y_i e quella del corrispondente punto sulla retta, cioè del punto sulla retta con la stessa ascissa. Tale punto è ovviamente $(x_i, b_0 + b_1 x_i)$. Questa differenza può essere positiva o negativa, a seconda che il punto si trovi rispettivamente al di sotto o al di sopra della retta. Il motivo dell'elevamento al quadrato è che vogliamo valutare gli scostamenti dalla retta indipendentemente dal segno.[20] Nella figura sopra a destra ho evidenziato, con pochi punti, le differenze di cui parlavo.

Si intuisce che la quantità indicata in (2) sarà grande se la retta si adatta male ai punti e piccola se invece si adatta bene. Si intuisce forse anche che solo in casi estremamente particolari questa quantità può essere zero.[21]

Ora veniamo al dunque. Se cerchiamo la retta che meglio si adatta ai punti siamo interessati a minimizzare la quantità (2). Formalizzando quindi il problema possiamo definire la funzione

$$f(b_0, b_1) = \sum_{i=1}^{N} (b_0 + b_1 x_i - y_i)^2, \tag{3}$$

funzione ovviamente delle due variabili b_0, b_1. Dobbiamo trovare quali sono i valori di b_0, b_1 che rendono minima la funzione f, cioè dobbiamo risolvere un problema di minimo (non vincolato) per una funzione di due variabili. Abbiamo visto che la tecnica analitica è anzitutto annullare il gradiente di f, cioè le sue derivate parziali, che ora calcoliamo. Faccio notare che per calcolare le derivate in presenza di una sommatoria basta ricordare che la derivata di una somma è la somma delle derivate e quindi basterà fare la sommatoria delle derivate dei singoli termini. Quindi abbiamo (per semplicità evito d'ora in poi di scrivere l'indice delle sommatorie, ricordando però quando servirà che si tratta di N termini)

$$f'_{b_0}(b_0, b_1) = \sum 2(b_0 + b_1 x_i - y_i) \qquad \text{e} \qquad f'_{b_1}(b_0, b_1) = \sum 2(b_0 + b_1 x_i - y_i) \cdot x_i.$$

Nel punto cercato queste due derivate devono annullarsi. In particolare dovrà essere nulla la derivata parziale di f rispetto a b_0, cioè dovrà essere

$$\sum (b_0 + b_1 x_i - y_i) = 0 \qquad \text{e cioè} \qquad N b_0 + b_1 \sum x_i - \sum y_i = 0.$$

[20] Qualcuno potrebbe osservare che si poteva anche considerare il valore assoluto di quella differenza, cioè considerare la quantità $\sum_{i=1}^{N} |b_0 + b_1 x_i - y_i|$, e sarebbe un'osservazione giustissima. Però il valore assoluto ha lo spiacevole difetto di rendere non derivabili le funzioni, come dovreste ricordare. E la non derivabilità, per quello che stiamo per fare nel seguito, è assolutamente deleteria.

[21] L'unico caso è quando tutti i punti sono allineati, cioè in altre parole quando c'è una retta che passa per tutti i punti. Questo perché l'unico modo per annullare l'intera sommatoria è annullare tutti i suoi termini, essendo questi non negativi. Quindi si deve avere $b_0 + b_1 x_i = y_i$ per ogni i.

Dividendo questa equazione per N otteniamo

$$b_0 + b_1 \frac{\sum x_i}{N} - \frac{\sum y_i}{N} = 0.$$

Se indichiamo con

$$\overline{x} = \frac{1}{N} \sum x_i \qquad \text{e} \qquad \overline{y} = \frac{1}{N} \sum y_i$$

($\overline{x}$ e $\overline{y}$ non sono altro che le medie aritmetiche delle x_i e delle y_i e il punto $(\overline{x}, \overline{y})$ si dice il *baricentro* dell'insieme di punti), l'ultima equazione può essere scritta come $b_0 + b_1\overline{x} - \overline{y}$, e cioè $\overline{y} = b_0 + b_1\overline{x}$. Questo dice una cosa importante e cioè che la retta cercata deve passare per il baricentro dell'insieme dei punti.

Tenendo conto allora che deve essere $b_0 = \overline{y} - b_1\overline{x}$, la nostra funzione diventa

$$f(b_0, b_1) = \sum \big(\overline{y} - b_1\overline{x} + b_1 x_i - y_i\big)^2 = \sum \big[b_1(x_i - \overline{x}) - (y_i - \overline{y})\big]^2.$$

Ponendo per comodità $\Delta x_i = x_i - \overline{x}$ e $\Delta y_i = y_i - \overline{y}$ si può scrivere

$$f(b_0, b_1) = \sum \big[b_1 \Delta x_i - \Delta y_i\big]^2.$$

La derivata parziale rispetto a b_1 di f è allora

$$f'_{b_1}(b_0, b_1) = \sum 2(b_1 \Delta x_i - \Delta y_i) \cdot \Delta x_i$$

per cui, annullandola, avremo

$$\sum (b_1 \Delta x_i - \Delta y_i) \cdot \Delta x_i = 0 \qquad \text{cioè} \qquad b_1 \sum (\Delta x_i)^2 - \sum \Delta x_i \Delta y_i = 0,$$

da cui infine si ricava

$$b_1 = \frac{\sum \Delta x_i \Delta y_i}{\sum (\Delta x_i)^2}.$$

Quindi la retta che meglio si adatta ai punti dati (retta che minimizza gli scarti quadratici, quindi metodo dei *minimi quadrati*) è la retta di equazione

$$y = b_0 + b_1 x \qquad \text{con} \quad b_0 = \overline{y} - m\overline{x} \quad \text{e} \quad b_1 = \frac{\sum \Delta x_i \Delta y_i}{\sum (\Delta x_i)^2},\text{[22]} \tag{4}$$

detta dalla Statistica *retta dei minimi quadrati* o *retta di regressione*.

È doverosa una precisazione finale. Abbiamo trovato un punto stazionario della funzione f: occorrerebbe accertarsi che si tratta in effetti di un punto di minimo per la funzione stessa. Questo si può fare, o considerando le derivate parziali seconde oppure scrivendo la matrice di rappresentazione della parte quadratica della funzione stessa.[23] Devo dire onestamente però che, a volerlo fare questo controllo, non è proprio facilissimo. Nell'esempio numerico che presento tra breve cercherò di convincervi che si tratta effettivamente di un minimo disegnando il grafico della funzione f.

Può essere forse il caso di accennare ad una generalizzazione di tutto questo. Il modello scelto per operare l'interpolazione è chiaramente un modello lineare, cioè abbiamo determinato quale sia la retta che meglio si accosta ai dati. A priori o nel caso in cui la retta non ci soddisfi, possiamo decidere di utilizzare un altro tipo di modello. Potremmo pensare ad una relazione di tipo quadratico tra x e y (quindi una parabola di interpolazione) o in generale ad una funzione $y = \varphi(x)$. Tenendosi nella massima generalità il nostro problema diventa la minimizzazione della quantità

$$\sum_{i=1}^{N} (\varphi(x_i) - y_i)^2.$$

È chiaro che per poter operare la minimizzazione, scrivendo le condizioni di ottimalità come abbiamo fatto nel caso lineare, occorre avere a disposizione un modello preciso, cioè una certa forma funzionale con dei parametri incogniti.

[22] La Statistica scriverà più semplicemente $b_1 = \frac{\sigma_{xy}}{\sigma_x^2}$, indicando con σ_{xy} la *covarianza* di x e y e con σ_x^2 la *varianza* di x.

[23] Attenzione che f non è una forma quadratica, in quanto contiene anche una parte lineare e una parte costante. Possiamo dire infatti che f è un polinomio di secondo grado in due variabili, con termini di grado uno e di grado zero (basta sviluppare i conti relativi al quadrato). Per quanto riguarda la verifica delle condizioni del secondo ordine quello che conta però è la sola parte quadratica, dato che le altre, derivando due volte, vengono eliminate.

Nel caso del modello quadratico, tale forma potrebbe essere $y = b_0+b_1x+b_2x^2$, per la quale la quantità da minimizzare sarebbe

$$f(b_0, b_1, b_2) = \sum_{i=1}^{N}(b_0 + b_1x_i + b_2x_i^2 - y_i)^2.$$

Non facciamo i calcoli, ma evidentemente quello che ci si aspetta è di dover scrivere tre derivate parziali e di dover poi risolvere un sistema di tre equazioni in tre incognite. Più sono i parametri del modello scelto, più è grande il sistema da risolvere.

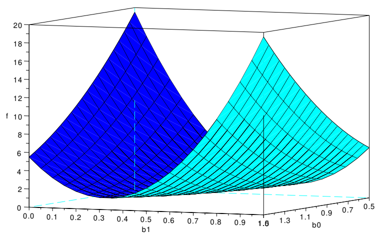
Penso sia opportuno un semplice esempio numerico in cui mettiamo in pratica quanto ottenuto in generale.

Supponiamo di considerare lo studio della dipendenza di una variabile y da una variabile x, cioè di voler trovare quale sia il modello lineare che permette di rappresentare meglio la dipendenza di y da x. Abbiamo a disposizione delle osservazioni, cioè delle coppie (x_i, y_i). Tali coppie sono elencate nella tabella a fianco.

i	x_i	y_i
1	1	1
2	1	2
3	2	2
4	3	2
5	3	3
6	4	3

Scegliendo un modello lineare, le formule (4) ci forniscono direttamente i valori ottimali (nel senso dei minimi quadrati) per i parametri b_0 e b_1 della retta di regressione. Voglio prima però cercare di convincervi del senso di quanto stiamo facendo.

Posso ad esempio proporvi il grafico della funzione



$$f(b_0, b_1) = \sum_{i=1}^{N}(b_0 + b_1x_i - y_i)^2.$$

Nel nostro caso $N = 6$ e le x_i, y_i sono fornite nella tabella qui sopra. Chiaramente, essendo queste componenti dei numeri fissati, la funzione è delle due variabili b_0, b_1. Utilizzando un software in grado di fare il “plot 3d” del grafico di funzioni di due variabili si ottiene quanto compare qui a sinistra. Si vede chiaramente che si tratta di una sorta di paraboloide con la concavità rivolta verso l'alto.[24]

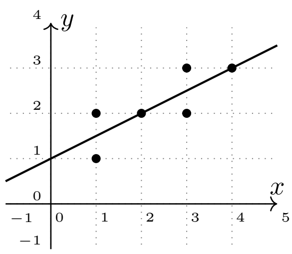
Aiutandosi (eventualmente) con un foglio di calcolo si ottiene una tabella come quella qui sotto in cui ci procuriamo le varie quantità utili per il calcolo dei parametri.

i	x_i	y_i	Δx_i	Δy_i	$\Delta x_i \Delta y_i$	$(\Delta x_i)^2$
1	1	1	-1.333	-1.1667	1.556	1.778
2	1	2	-1.333	-0.167	0.222	1.778
3	2	2	-0.333	-0.167	0.056	0.111
4	3	2	0.667	-0.167	-0.111	0.444
5	3	3	0.667	0.833	0.556	0.444
6	4	3	1.667	0.833	1.389	2.778
	$2.333(\overline{x})$	$2.167(\overline{y})$			$\Sigma = 3.667$	$\Sigma = 7.333$

Applicando le formule (4) si ottengono i valori ottimali

$$b_0 = \overline{y} - m\overline{x} = 1 \qquad \text{e} \qquad b_1 = \frac{\sum \Delta x_i \Delta y_i}{\sum (\Delta x_i)^2} = 0.5,$$

con i quali possiamo disegnare la retta di regressione, rappresentata qui sotto.



Un aspetto interessante è valutare quanto la retta sia un buon modello per i dati osservati, cioè in che misura la dipendenza di y da x sia lineare. Ma questo è un terreno tipicamente statistico e avrete modo di studiare la questione l'anno prossimo.

[24] Per il grafico ho utilizzato SCILAB, ovviamente un free software. Per la verità non è facile identificare nel grafico il punto di minimo, perché il paraboloide è caratterizzato da una “valle” di pendenza quasi nulla. Dopo aver trovato con le formule il valore dei parametri, si può osservare che i valori ottimi sono ragionevolmente quelli che rendono minima la funzione.