Massimi e minimi vincolati

Vedremo tra breve un metodo per studiare il problema di trovare il minimo e il massimo di una funzione su di un sottoinsieme dello spazio ambiente che non sia un aperto. Abbiamo già visto che in un tale insieme non è possibile fare le derivate per cui il problema va modificato e si è già visto che un modo è parametrizzare l'insieme. Vediamo ora alcuni casi particolari senza fare ragionamenti rigorosi, ma utili alla comprensione del metodo che illustreremo.
Prima di tutto facciamo un'osservazione importante. Si consideri una funzione $F : A \subset \mathbf{R}^{m+n} \to \mathbf{R}$ di classe C^1, A aperto, e un insieme $V \subset A$ (che potrebbe essere chiuso). Si supponga che $\bar{x}$ sia un punto di minimo (oppure di massimo) per F nell'insieme V. Allora per ogni curva γ di classe C^1, $\gamma : (-\varepsilon, \varepsilon) \to V \subset \mathbf{R}^{m+n}$ soddisfacente $\gamma(0) = \bar{x}$ si ha che

$$t = 0 \quad \text{è di minimo (rispettivamente di massimo) per} \quad F \circ \gamma$$

e si ha

$$\frac{d}{dt}\Big|_{t=0} \big(F(\gamma(t))\big) = \langle \nabla F(\bar{x}), \gamma'(0) \rangle = 0. \tag{1}$$

Definizione 0.1. *Diremo che $\bar{x}$ è critico o stazionario per F nell'insieme V se per ogni curva $\gamma : (-\varepsilon, \varepsilon) \to \mathbf{R}^{m+n}$ di classe C^1, il cui sostegno sia contenuto in V e soddisfacente $\gamma(0) = \bar{x}$ si ha che $t = 0$ è critico o stazionario per la funzione $F \circ \gamma$. Inoltre, se anche F è C^1, vale* (1)*.*

Si osservi che da (1) si ha che il gradiente di F è ortogonale, nel punto $\bar{x}$, alla retta tangente alla curva γ, e questo *per ogni* curva γ passante per il punto $\bar{x}$ e il cui supporto sia contenuto in V.
Si può dimostrare che ciò implica il fatto che il gradiente di F in $\bar{x}$ è ortogonale allo spazio generato dalle rette tangenti a tutte le possibili curve regolari passanti per $\bar{x}$. In particolare, nel caso in cui V è una curva in $\mathbf{R}^2$, $\nabla F(\bar{x})$ è ortogonale alla retta tangente alla curva V, nel caso in cui V è una superficie in $\mathbf{R}^3$ $\nabla F(\bar{x})$ è ortogonale al piano tangente alla superficie V, nel caso in cui V è una curva in $\mathbf{R}^3$ $\nabla F(\bar{x})$ è ortogonale alla retta tangente alla curva V.

Nella sezione dedicata al teorema delle funzioni implicite abbiamo dato la definizione di dimensione per un luogo di zeri che è grafico di una funzione. Se $F : A \subset \mathbf{R}^{m+n} \to \mathbf{R}$ e l'insieme $V = \{x \in A \,|\, G_1(x) =$

$0\} \cap \cdots \cap \{x \in A \,|\, G_n(x) = 0\}$ con $G = (G_1, \dots G_n)$ che soddisfa il teorema delle funzioni implicite diciamo che V ha dimensione m. Con il termine *codimensione* denotiamo il numero n riferito al fatto che V ha dimensione m ed è un sottoinsieme di $\mathbf{R}^{m+n}$.

Caso 1 - Il vincolo è una curva in $\mathbf{R}^2$ (la codimensione è 1) - Si supponga di voler cercare i punti stazionari per una certa funzione $F : A \subset \mathbf{R}^2 \to \mathbf{R}$ di classe C^1, A aperto, in un insieme $V \subset A$ e che V sia il sostegno di una curva. Se tale curva è data come luogo di zeri di una funzione $G : A \to \mathbf{R}$

$$V := \Big\{(x, y) \in A \,\Big|\, G(x, y) = 0\Big\}$$

nell'ipotesi in cui $G_y \neq 0$ in un certo punto $(\bar{x}, \bar{y}) \in V$ si ha l'esistenza di una funzione g tale che

$$G(x, g(x)) = 0.$$

Nel caso in cui il punto $(\bar{x}, \bar{y})$ sia di minimo o di massimo per F sul vincolo V, scegliendo come γ in (1) la curva $x \mapsto (x, g(x))$ si ottiene

$$\frac{d}{dx}\Big|_{x=\bar{x}} \big(F(x, g(x)\big) = \frac{\partial F}{\partial x}(\bar{x}, \bar{y}) + \frac{\partial F}{\partial y}(\bar{x}, \bar{y}) g'(\bar{x}) = 0$$

cioè il gradiente di F in $(\bar{x}, \bar{y})$ è ortogonale alla curva $x \mapsto (x, g(x))$ in $\bar{x}$, cioè è ortogonale al vincolo V. Poiché $G(x, g(x)) = 0$ per ogni x si ha pure che il gradiente di G è ortogonale al vincolo V in ogni punto x. Sintetizziamo queste informazioni dicendo che

$$\nabla F(\bar{x}, \bar{y}) \perp V \qquad \text{e} \qquad \nabla G(x, y) \perp V \;\; \text{per ogni } (x, y) \in V.$$

In particolare si ha $\nabla G(\bar{x}, \bar{y}) \perp V$ per cui, poiché lo spazio ambiente è $\mathbf{R}^2$, i due vettori devono essere linearmente dipendenti ed esiste un numero $\lambda \in \mathbf{R}$ tale che

$$\nabla F(\bar{x}, \bar{y}) + \lambda \nabla G(\bar{x}, \bar{y}) = (0, 0)\,. \tag{2}$$

Caso 2 - Il vincolo è una superficie in $\mathbf{R}^3$ (la codimensione è 1) - Sia ora il caso in cui $A \subset \mathbf{R}^3$ e

$$V := \Big\{(x, y, z) \in A \,\Big|\, G(x, y, z) = 0\Big\}.$$

Nell'ipotesi che $G_z(\bar{x}, \bar{y}, \bar{z}) \neq 0$ dal teorema del Dini si ha l'esistenza di una funzione g il cui grafico è una superficie e

$$G(x, y, g(x, y)) = 0$$

Poiché V è l'insieme di livello 0 di G si ha

$$\Big\langle \nabla G(x,y,f(x,y)),(1,0,g_x(x,y))\Big\rangle = 0,$$
$$\Big\langle \nabla G(x,y,f(x,y)),(0,1,g_y(x,y))\Big\rangle = 0,$$

e il gradiente di G è ortogonale al piano tangente al grafico di g (si veda il capitolo precedente) in ogni punto $(x,y,z) \in V$.
Si considerino ora le due curve, con $t \in \mathbf{R}$ $(\bar{z} = g(\bar{x},\bar{y}))$

$$\begin{aligned}\gamma(t) &= \big(t,\bar{y},\bar{z}+g_x(\bar{x},\bar{y})(t-\bar{x})\big) = \\ &= t\big(1,0,g_x(\bar{x},\bar{y})\big) + \big(0,\bar{y},-\bar{x}\,g_x(\bar{x},\bar{y})\big) \\ \sigma(t) &= \big(\bar{x},t,\bar{z}+g_y(\bar{x},\bar{y})(t-\bar{y})\big) = \\ &= t\big(0,1,g_y(\bar{x},\bar{y})\big) + \big(\bar{x},0,-\bar{y}\,g_y(\bar{x},\bar{y})\big)\end{aligned}$$

le cui derivate sono due vettori paralleli al piano tangente al grafico di g nel punto $(\bar{x},\bar{y},\bar{z})$. Quindi, se $(\bar{x},\bar{y},\bar{z})$ è un punto di minimo o di massimo sul vincolo V e $\bar{z} = g(\bar{x},\bar{y})$, si ha

$$\Big\langle \nabla F(\bar{x},\bar{y},g(\bar{x},\bar{y})),(1,0,g_x(\bar{x},\bar{y}))\Big\rangle = 0,$$
$$\Big\langle \nabla F(\bar{x},\bar{y},g(\bar{x},\bar{y})),(0,1,g_y(\bar{x},\bar{y}))\Big\rangle = 0,$$

per cui il gradiente di F è ortogonale al piano tangente al grafico di g nel punto $(\bar{x},\bar{y},g(\bar{x},\bar{y})) \in V$.
Anche in questo caso valutando ∇G nel punto $(\bar{x},\bar{y},g(\bar{x},\bar{y}))$ si ha che tale vettore è ortogonale allo stesso piano al quale è ortogonale $\nabla F(\bar{x},\bar{y},g(\bar{x},\bar{y}))$. Poiché due vettori in $\mathbf{R}^3$ ortogonali ad un stesso piano giacciono sulla stessa retta e sono quindi linearmente dipendenti, esiste $\lambda \in \mathbf{R}$ tale che

$$\nabla F(\bar{x},\bar{y},g(\bar{x},\bar{y})) + \lambda \nabla G(\bar{x},\bar{y},g(\bar{x},\bar{y})) = (0,0,0). \tag{3}$$

Caso 3 - Il vincolo è una curva in $\mathbf{R}^3$ (la codimensione è 2) - Se il vincolo V è dato da

$$V := \Big\{(x,y,z)\in A \,\Big|\, G_1(x,y,z) = 0\Big\} \cap \Big\{(x,y,z)\in A \,\Big|\, G_2(x,y,z) = 0\Big\}$$

abbiamo visto che, sotto opportune ipotesi, V è l'immagine di una curva (cartesiana). La funzione F, come visto in (1), avrà gradiente ortogonale alla curva in un generico punto stazionario $(\bar{x},\bar{y},\bar{z})$. Si faccia attenzione che in questo caso, poiché siamo in $\mathbf{R}^3$, vi sono infinite direzioni che può avere $\nabla F(\bar{x},\bar{y},\bar{z})$, tutte quelle del piano ortogonale alla curva.

Dal teorema del Dini sappiamo che, se

$$\det\begin{pmatrix} \dfrac{\partial G_1}{\partial y}(\bar{x},\bar{y},\bar{z}) & \dfrac{\partial G_1}{\partial z}(\bar{x},\bar{y},\bar{z}) \\ \\ \dfrac{\partial G_2}{\partial y}(\bar{x},\bar{y},\bar{z}) & \dfrac{\partial G_2}{\partial z}(\bar{x},\bar{y},\bar{z}) \end{pmatrix} \neq 0, \tag{4}$$

esistono y e z funzioni di una variabile, tali che

$$G_1(x, y(x), z(x)) = 0 \qquad \text{e} \qquad G_2(x, y(x), z(x)) = 0$$

da cui, derivando rispetto a x e y,

$$\frac{\partial G_1}{\partial x}(x,y(x),z(x)) + \frac{\partial G_1}{\partial y}(x,y(x),z(x))y'(x) + \frac{\partial G_1}{\partial z}(x,y(x),z(x))z'(x) = 0$$
$$\frac{\partial G_2}{\partial x}(x,y(x),z(x)) + \frac{\partial G_2}{\partial y}(x,y(x),z(x))y'(x) + \frac{\partial G_2}{\partial z}(x,y(x),z(x))z'(x) = 0$$

cioè sia ∇G_1 che ∇G_2 sono ortogonali alla curva $x \mapsto (x, y(x), z(x))$ in ogni punto. In particolare lo saranno nel punto $(\bar{x},\bar{y},\bar{z}) = (\bar{x}, y(\bar{x}), z(\bar{x}))$. Dalla condizione (4) si ha che i due vettori $\nabla G_1(\bar{x}, y(\bar{x}), z(\bar{x}))$ e $\nabla G_2(\bar{x}, y(\bar{x}), z(\bar{x}))$ sono linearmente indipendenti, per cui generano il piano nel quale giacciono. Poiché anche $\nabla F(\bar{x}, y(\bar{x}), z(\bar{x}))$ giace in tale piano esso sarà combinazione lineare dei primi due, cioè esisteranno $\lambda, \mu \in \mathbf{R}$ tali che

$$\begin{aligned} &\nabla F(\bar{x}, y(\bar{x}), z(\bar{x})) + \lambda \nabla G_1(\bar{x}, y(\bar{x}), z(\bar{x})) + \\ &\qquad\qquad + \mu \nabla G_2(\bar{x}, y(\bar{x}), z(\bar{x})) = (0,0,0). \end{aligned} \tag{5}$$

1. Metodo dei moltiplicatori di Lagrange

Si è visto che per cercare i punti di minimo e i punti di massimo per una funzione f su un insieme chiuso V si può utilizzare una parametrizzazione, chiamiamola ϕ.
La difficoltà di questo metodo sta nel fatto che è necessario conoscere l'espressione di ϕ per poter fare i calcoli.
Vediamo ora un metodo alternativo, che prende il nome di *metodo dei moltiplicatori di Lagrange*, tramite il quale si può risolvere il problema della ricerca dei punti stazionari di F in V anche senza conoscere ϕ.

Si consideri il primo caso affrontato nella prima parte, l'insieme V è una curva in $\mathbf{R}^2$. Introduciamo la funzione ausiliaria, definita in $\mathbf{R}^3$,

$$H(x, y, \lambda) := F(x, y) + \lambda G(x, y).$$

Volendo cercare i punti stazionari di H, nell'ipotesi che F e G siano regolari, si studia il sistema

$$\begin{cases} \dfrac{\partial H}{\partial x}(x,y,\lambda)=0 \\ \dfrac{\partial H}{\partial y}(x,y,\lambda)=0 \\ \dfrac{\partial H}{\partial \lambda}(x,y,\lambda)=0. \end{cases}$$

Questo sistema altro non è che, accoppiando le prime due equazioni,

$$\begin{cases} \nabla F(x,y)+\lambda G(x,y)=0 \\ G(x,y)=0. \end{cases}$$

Questo significa che cercando i punti stazionari liberi (cioè su tutto $\mathbf{R}^3$) di H si stanno in realtà chiedendo due condizioni bene precise: la prima è che i punti cercati debbano soddisfare l'equazione (2), la seconda è che tali punti appartenagano all'insieme

$$V := \big\{(x,y)\in A \,\big|\, G(x,y)=0\big\}.$$

Cosa analoga si può fare nel seconda caso, quando invece l'insieme V è una superficie in $\mathbf{R}^3$: si considera la funzione ausiliaria di quattro variabili

$$H(x,y,z,\lambda) := F(x,y,z)+\lambda G(x,y,z)$$

e si procede in maniera anloga a prima.

Nel terzo caso invece il vincolo è dato dall'intersezione di due superfici: si introduce la funzione ausiliaria

$$H(x,y,z,\lambda,\mu) := F(x,y,z)+\lambda G_1(x,y,z)+\mu G_2(x,y,z)$$

e si noti che eguagliando a zero tutte e cinque le derivate parziali di H si ottiene il sistema

$$\begin{cases} \nabla F(x,y,z)+\lambda G_1(x,y,z)+\mu\, G_2(x,y,z)=0 \\ G_1(x,y,z)=0 \\ G_2(x,y,z)=0 \end{cases}$$

cioè che sia soddisfatta l'equazione (5) e che i punti trovati stiano in entrambi i luoghi di zeri che definiscono V.

Enunciato del teorema - Si consideri ora una funzione F definita in un aperto $A \subset \mathbf{R}^{m+n} \to \mathbf{R}$, una funzione $G : A \to \mathbf{R}^n$, l'insieme

$$V := \{x \in A \,|\, G_1(x) = 0\} \cap \cdots \cap \{x \in A \,|\, G_n(x) = 0\}$$

ed un punto $\bar{x} \in V$. Diremo che tale punto è *regolare* per il vincolo V se la matrice jacobiana di G valutata in $\bar{x}$ ha rango n, cioè esiste un minore $n \times n$ con determinante non nullo.

Teorema 1.1. *Nel caso in cui $\bar{x} \in V$ sia un punto regolare, F, $G_1, \ldots G_n$ siano di classe $C^1(A)$ si ha che $\bar{x}$ è un punto critico per F nell'insieme V se e solo se esistono $\lambda_1, \ldots \lambda_n \in \mathbf{R}$ (detti moltiplicatori di Lagrange) tali che*

$$\nabla F(\bar{x}) + \lambda_1 \nabla G_1(\bar{x}) + \cdots + \lambda_n \nabla G_n(\bar{x}) = (0, \ldots 0).$$

Concludiamo che per cercare i punti critici, fra i quali ci saranno il punto di massimo e il punto di minimo, di una funzione C^1 in un vincolo definito da n funzioni anch'esse C^1 è sufficiente trovare i punti critici liberi della funzione ausiliaria

$$H(x, \lambda_1, \ldots \lambda_n) = F(x) + \lambda_1 G_1(x) + \cdots + \lambda_n G_n(x) \qquad (x \in \mathbf{R}^n).$$

Se poi in alcuni punti F oppure le G_i non saranno differenziabili si andranno a considerare come candidati, oltre ai punti critici di H, anche questi punti.

Attenzione!! Se si ha una funzione F definita in un aperto di $\mathbf{R}^3$ e si vuole minimizzare F su di una superficie espressa da $G(x, y, z) = 0$ se si riesce a parametrizzare tale superficie ci si riduce ad avere due variabili (e quindi a risolvere un sistema di due equazioni in due incognite), usando il metodo dei moltiplicatori di Lagrange si hanno quattro variabili.
Se anziché una superficie si ha una curva in $\mathbf{R}^3$, parametrizzandola si ha una sola variabile, con il metodo dei moltiplicatori di Lagrange si arriva a cinque variabili. È sempre conveniente?

Esempio 1.2. - Si trovi la minima distanza dall'origine dei punti appartenenti all'insieme

$$V = \{(x, y) \in \mathbf{R}^2 \,|\, xy = 1\}.$$

Abbiamo già visto che un modo per risolvere il problema è parametrizzare l'insieme. In questo caso è possibile farlo nel modo seguente: l'insieme V è l'unione di due rami di un'iperbole. Limitandoci ad uno dei due possiamo parametrizzare uno dei due come segue (se $xy = 1$

allora $y = 1/x$)

$$\gamma : (0, +\infty) \to \mathbf{R}^2, \qquad \gamma(t) = \left(t, \frac{1}{t}\right).$$

L'altro ramo può essere parametrizzato da $\sigma : (0, +\infty) \to \mathbf{R}^2$, $\sigma(t) = (-t, -1/t)$. Ma limitiamo al primo, tanto per simmetria l'altro sarà simile. La distanza di un punto (x, y) dall'origine è espressa dalla quantità $\sqrt{x^2 + y^2}$, ma come al solito considereremo $x^2 + y^2$. Detta f tale funzione e componendola con γ si ha $f(\gamma(t)) = t^2 + 1/t^2$. Derivando e ponendo uguale a zero si ha

$$\frac{d}{dt}\big(f(\gamma(t))\big) = 2t - 2\frac{1}{t^3} = 2\,\frac{t^4 - 1}{t^3} = 0$$

che si annulla solo per $t = 1$ (!! $t > 0$). Tale punto è l'unico stazionario e corrisponde al punto $(1, 1)$ sul vincolo. Come capire se è di minimo, di massimo o nessuno dei due?
Sappiamo che

$$a^2 + b^2 \geqslant 2ab \qquad \text{per ogni } a, b \in \mathbf{R}$$

per cui considerando $x = t$ e $y = 1/t$ si ha

$$f(\gamma(t)) = t^2 + \frac{1}{t^2} \geqslant 2.$$

Siccome per $t = 1$ si ha $f(\gamma(1)) = 2$ si conclude che il punto $t = 1$, e quindi $(1, 1)$ sul vincolo, è di minimo. Si osservi che la funzione è definita in un insieme illimitato e che

$$\lim_{t \to 0^+} f(\gamma(t)) = \lim_{t \to +\infty} f(\gamma(t)) = +\infty$$

per cui la funzione ammette solo un punto di minimo ed è quello trovato. Analogamente nell'altro ramo si troverà $(-1, -1)$ e la minima distanza sarà $\sqrt{2}$.

Ma vediamo come fare l'esercizio con i moltiplicatori di Lagrange. Si consideri la funzione

$$H(x, y, \lambda) = x^2 + y^2 + \lambda(xy - 1)$$

e si consideri il sistema

$$\begin{cases} \dfrac{\partial H}{\partial x}(x, y, \lambda) = 2x + \lambda y = 0 \\[2mm] \dfrac{\partial H}{\partial y}(x, y, \lambda) = 2y + \lambda x = 0 \\[2mm] \dfrac{\partial H}{\partial \lambda}(x, y, \lambda) = xy - 1 = 0. \end{cases}$$

Poiché x e y non si annullano, dalle prime due equazioni si ricava

$$\lambda = -\frac{2x}{y} = -\frac{2y}{x}$$

cioè

$$\frac{x}{y} = \frac{y}{x}$$

il che implica $x = y = 1$ oppure $x = y = -1$. Si conclude che sono punti di minimo analogamente a prima usando la stima

$$x^2 + y^2 \geqslant 2xy$$

e quindi per $x = y = 1$ e $x = y = -1$ si ha $x^2 + y^2 \geqslant 2$, oppure osservando che

$$\lim_{\substack{|(x,y)| \to +\infty \\ x \in V}} (x^2 + y^2) = +\infty$$

per cui la funzione nel punto trovato ha un punto di minimo.

Figura 1.d