Teorema di Dini ed estremi vincolati

Teorema di Dini

Diamo l'enunciato e la dimostrazione del teorema di Dini per funzioni di due variabili reali.

Teorema 1 *Siano $X \subseteq \mathbb{R}^2$ aperto, $f : X \to \mathbb{R}$ funzione continua e $(x_0, y_0) \in X$. Se $\partial_y f$ è funzione continua e $\partial_y f(x_0, y_0) \neq 0$, allora esistono un intorno $\mathcal{U} =]x_0 - \varepsilon_1, x_0 + \varepsilon_1[$ di x_0, un intorno $\mathcal{V} =]y_0 - \varepsilon_2, y_0 + \varepsilon_2[$ di y_0 ed un'unica funzione $g : \mathcal{U} \to \mathcal{V}$ continua e tale che*

$$\big\{(x, y) \in \mathcal{U} \times \mathcal{V} : f(x, y) = f(x_0, y_0)\big\} = \big\{(x, y) \in \mathcal{U} \times \mathcal{V} : y = g(x)\big\} .$$

Inoltre, se $f \in \mathcal{C}^1(X)$, allora $g \in \mathcal{C}^1(\mathcal{U})$ e vale

$$g'(x) = -\frac{\partial_x f(x, g(x))}{\partial_y f(x, g(x))} . \tag{1}$$

Dim. Per fissare le idee supponiamo che $f(x_0, y_0) = 0$ e $\partial_y f(x_0, y_0) > 0$, gli altri casi si trattano in modo del tutto analogo.

Esistenza e unicità di g. Poiché $\partial_y f$ è continua in X per il teorema della permanenza del segno esistono $\varepsilon_0, \varepsilon_2 > 0$ tali che

$$\partial_y f(x, y) > 0 \quad \forall\, (x, y) \in [x_0 - \varepsilon_0, x_0 + \varepsilon_0] \times [y_0 - \varepsilon_2, y - 0 + \varepsilon_2] \subset X . \tag{2}$$

In particolare la funzione $y \mapsto f(x_0, y)$ è strettamente crescente in $[y_0 - \varepsilon_2, y_0 + \varepsilon_2]$ e si annulla per $y = y_0$, che quindi risulta il suo unico zero nell'intervallo $[y_0 - \varepsilon_2, y_0 + \varepsilon_2]$. Se ne deduce che

$$f(x_0, y_0 - \varepsilon_2) < 0 < f(x_0, y_0 + \varepsilon_2) .$$

Ancora per il teorema della permanenza del segno (ricordiamo che $f \in \mathcal{C}^0(X)$) esiste $0 < \varepsilon_1 \leq \varepsilon_0$ tale che

$$f(x, y_0 - \varepsilon_2) < 0 < f(x, y_0 + \varepsilon_2) \quad \forall\, x_0 - \varepsilon_1 \leq x \leq x_0 + \varepsilon_1 . \tag{3}$$

Siano $\mathcal{U} \doteq]x_0 - \varepsilon_1, x_0 + \varepsilon_1[$ e $\mathcal{V} \doteq]y_0 - \varepsilon_2, y_0 + \varepsilon_2[$, con $\varepsilon_1, \varepsilon_2 > 0$ scelti come sopra. Per ogni $x \in \mathcal{U}$ grazie a (2) la funzione $y \mapsto f(x, y)$ risulta strettamente crescente nell'intervallo $\mathcal{V}$ ed assume valori di segno opposto agli estremi per (3). Per il teorema di Bolzano e grazie alla stretta monotonia essa dunque ammette un unico zero, e quindi è individuata un'unica funzione $g : \mathcal{U} \to \mathcal{V}$ tale che $f(x, g(x)) = 0$ per ogni $x \in \mathcal{U}$.

Continuità di g. Dobbiamo dimostrare che, fissato $\overline{x} \in \mathcal{U}$, per ogni $\varepsilon > 0$ esiste $\delta > 0$ tale che se $|x - \overline{x}| < \delta$, $x \in \mathcal{U}$, allora $|g(x) - g(\overline{x})| < \varepsilon$. Sia allora $\varepsilon > 0$ fissato, e non è restrittivo assumere che $[g(\overline{x}) - \varepsilon, g(\overline{x}) + \varepsilon] \subset \mathcal{V}$. Poiché $y \mapsto f(\overline{x}, y)$ è strettamente crescente e $f(\overline{x}, g(\overline{x})) = 0$ si ha

$$f(\overline{x}, g(\overline{x}) - \varepsilon) < 0 < f(\overline{x}, g(\overline{x}) + \varepsilon) .$$

Grazie alla continuità di f, esiste $\delta > 0$ tale che

$$f(x, g(\overline{x}) - \varepsilon) < 0 < f(x, g(\overline{x}) + \varepsilon) \quad \forall\, |x - \overline{x}| < \delta\, ,\ x \in \mathcal{U} .$$

Poiché $y = g(x)$ è l'unica soluzione dell'equazione $f(x, y) = 0$ nell'intervallo $\mathcal{V} \supseteq [g(\overline{x}) - \varepsilon, g(\overline{x}) + \varepsilon]$ si ottiene che vale

$$g(\overline{x}) - \varepsilon < g(x) < g(\overline{x}) + \varepsilon \quad \forall\, |x - \overline{x}| < \delta\,,\ x \in \mathcal{U}\,,$$

come si voleva.

Derivabilità di g. Sia ora $f \in \mathcal{C}^1(X)$ e, fissato $x \in \mathcal{U}$, sia $h \in \mathbb{R}$ tale che $x + h \in \mathcal{U}$. Si osservi che il segmento di estremi $(x, g(x))$ e $(x + h, g(x + h))$ è interamente contenuto in $\mathcal{U} \times \mathcal{V}$. Per il teorema del valor medio esiste un punto (ξ, η) su questo segmento tale che

$$0 = f(x + h, g(x + h)) - f(x, g(x)) = \partial_x f(\xi, \eta)h + \partial_y f(\xi, \eta)\big(g(x + h) - g(x)\big)\,,$$

da cui, grazie al fatto che vale (2), si ottiene

$$\frac{g(x + h) - g(x)}{h} = -\frac{\partial_x f(\xi, \eta)}{\partial_y f(\xi, \eta)}\,. \tag{4}$$

Poiché per $h \to 0$ si ha $(\xi, \eta) \to (x, g(x))$ e data la continuità delle derivate parziali di f, passando al limite per $h \to 0$ in (4) si ottiene (1), come si voleva. Il fatto che poi g' sia continua e quindi g di classe $\mathcal{C}^1$ segue direttamente da (1) sempre per la continuità delle derivate parziali di f. □

Derivate della funzione implicita

Siano $1 \leq m < n$ numeri naturali, $X \subseteq \mathbb{R}^n$ aperto, $\mathbf{f} : X \to \mathbb{R}^m$, $\mathbf{f} = (f_1, \ldots, f_m)$, funzione di classe $\mathcal{C}^1$. Sia $\overline{\mathbf{x}} = (\overline{x}_1, \ldots, \overline{x}_n) \in X$ tale che

$$\det \begin{pmatrix} \partial_{x_1} f_1(\overline{\mathbf{x}}) & \cdots & \partial_{x_m} f_1(\overline{\mathbf{x}}) \\ \vdots & \ddots & \vdots \\ \partial_{x_1} f_m(\overline{\mathbf{x}}) & \cdots & \partial_{x_m} f_m(\overline{\mathbf{x}}) \end{pmatrix} \neq 0\,. \tag{5}$$

Grazie al teorema di Dini, esistono un intorno $\mathcal{U}$ di $(\overline{x}_{m+1}, \ldots, \overline{x}_n)$ in $\mathbb{R}^{n-m}$, un intorno $\mathcal{V}$ di $(\overline{x}_1, \ldots, \overline{x}_m)$ in $\mathbb{R}^m$ e una sola funzione $\mathbf{g} : \mathcal{U} \to \mathcal{V}$, $\mathbf{g} = (g_1, \ldots, g_m)$, tale che

$$\big\{\mathbf{x} \in \mathcal{U} \times \mathcal{V} : \mathbf{f}(\mathbf{x}) = \mathbf{f}(\overline{\mathbf{x}})\big\} = \big\{\mathbf{x} = (x_1, \ldots, x_n) \in \mathcal{U} \times \mathcal{V} : x_i = g_i(x_{m+1}, \ldots, x_n)\,, i = 1, \ldots, m\big\}\,. \tag{6}$$

Poiché $\mathbf{g}$ risulta essere di classe $\mathcal{C}^1$, ci poniamo il problema di calcolare le sue derivate parziali rispetto a $x_{m+1}, \ldots, x_n$. Sia allora $i \in \{m + 1, \ldots, n\}$ fissato e calcoliamo $\partial_{x_i} g_k$, $k = 1, \ldots, m$. Sappiamo che vale

$$f_j(\mathbf{g}(x_{m+1}, \ldots, x_n), x_{m+1}, \ldots, x_n) = f_j(\overline{\mathbf{x}}) \quad \forall\, (x_{m+1}, \ldots, x_n) \in \mathcal{U}\,, \forall\, j = 1, \ldots, m\,,$$

e quindi le funzioni

$$\varphi_j(x_{m+1}, \ldots, x_n) \doteq f_j(\mathbf{g}(x_{m+1}, \ldots, x_n), x_{m+1}, \ldots, x_n) \qquad (x_{m+1}, \ldots, x_n) \in \mathcal{U}$$

sono tutte costanti. Allora

$$\frac{\partial \varphi_j}{\partial x_i}(x_{m+1}, \ldots, x_n) = 0 \quad \forall\, (x_{m+1}, \ldots, x_n) \in \mathcal{U}\,, \forall\, j = 1, \ldots, m\,, \forall\, i = m + 1, \ldots, n\,.$$

Poiché usando la regola della catena si ricava che

$$\frac{\partial \varphi_j}{\partial x_i}(x_{m+1},\dots,x_n) = \sum_{k=1}^{m} \frac{\partial f_j}{\partial x_k}(\mathbf{g}(x_{m+1},\dots,x_n),x_{m+1},\dots,x_n)\frac{\partial g_k}{\partial x_i}(x_{m+1},\dots,x_n)+$$
$$+\frac{\partial f_j}{\partial x_i}(\mathbf{g}(x_{m+1},\dots,x_n),x_{m+1},\dots,x_n)$$

si ricava che $\partial g_k/\partial x_i$, $k=1,\dots,m$, sono soluzioni del sistema lineare in m equazioni e nelle m incognite $t_k = \partial_{x_i} g_k$

$$\begin{cases} \displaystyle\sum_{k=1}^{m} \frac{\partial f_1}{\partial x_k}(\mathbf{g}(x_{m+1},\dots,x_n),x_{m+1},\dots,x_n)t_k = -\frac{\partial f_1}{\partial x_i}(\mathbf{g}(x_{m+1},\dots,x_n),x_{m+1},\dots,x_n) \\ \qquad\qquad\vdots \qquad\qquad\qquad\qquad \vdots \qquad\qquad\qquad\qquad \vdots \\ \displaystyle\sum_{k=1}^{m} \frac{\partial f_m}{\partial x_k}(\mathbf{g}(x_{m+1},\dots,x_n),x_{m+1},\dots,x_n)t_k = -\frac{\partial f_m}{\partial x_i}(\mathbf{g}(x_{m+1},\dots,x_n),x_{m+1},\dots,x_n)\,. \end{cases}$$

Quindi si ottiene

$$\begin{pmatrix} \partial_{x_i} g_1 \\ \vdots \\ \partial_{x_i} g_m \end{pmatrix} = -\begin{pmatrix} \partial_{x_1} f_1 & \cdots & \partial_{x_m} f_1 \\ \vdots & \ddots & \vdots \\ \partial_{x_1} f_m & \cdots & \partial_{x_m} f_m \end{pmatrix}^{-1} \begin{pmatrix} \partial_{x_i} f_1 \\ \vdots \\ \partial_{x_i} f_m \end{pmatrix},$$

dove le derivate di g_k sono calcolate in $(x_{m+1},\dots,x_n)$, mentre quelle delle componenti f_j di $\mathbf{f}$ sono calcolate in $(\mathbf{g}(x_{m+1},\dots,x_n),x_{m+1},\dots,x_n)$.

Esempio 1 Verifichiamo che l'equazione

$$\arctan z + xy^2 + xz - y^3 = 1$$

definisce implicitamente una funzione $z = g(x,y)$ in un intorno di $(0,-1,0)$. Scriviamo poi l'equazione del piano tangente al grafico della funzione implicita $z = g(x,y)$ in $(0,-1,0)$.
Sia $f(x,y,z) = \arctan z + xy^2 + xz - y^3$, cosicché l'equazione si scrive $f(x,y,z) = 1$, che ha $(0,-1,0)$ come soluzione. Poiché $\partial_z f(x,y,z) = 1/(1+z^2) + x$ e quindi $\partial_z f(0,-1,0) = 1 \neq 0$, con il teorema di Dini si deduce l'esistenza della funzione implicita $z = g(x,y)$. Essendo f di classe $\mathcal{C}^\infty$, tale è anche g. L'equazione del piano tangente al grafico di g nel punto $(0,-1,0)$ si scrive

$$z = \partial_x g(0,-1)x + \partial_y g(0,-1)(y+1)\,,$$

e quindi dobbiamo calcolare le derivate parziali di g in $(0,-1)$. Si ottiene

$$\partial_x g(0,-1) = -\frac{\partial_x f(0,-1,0)}{\partial_z f(0,-1,0)} = -1\,, \qquad \partial_y g(0,-1) = -\frac{\partial_y f(0,-1,0)}{\partial_z f(0,-1,0)} = 3\,,$$

e l'equazione del piano tangente è $z = -x + 3(y+1)$.

Esempio 2 Verifichiamo che il sistema

$$\begin{cases} x + \ln y + z = 2 \\ 2x - y^2 + z = 1 \end{cases}$$

definisce implicitamente due funzioni $y = g(x)$ e $z = h(x)$ in un intorno di $x = 0$ tali che $g(0) = 1$ e $h(0) = 2$. Successivamente scriviamo la formula di MacLaurin di ordine 2 con il resto di Peano per le funzioni g e h.

Se

$$\mathbf{f}(x, y, z) = (x + \ln y + z, 2x - y^2 + z) = \big(f_1(x, y, z), f_2(x, y, z)\big)$$

il sistema si riscrive $\mathbf{f}(x, y, z) = (2, 1)$. Dopo aver osservato che $(0, 1, 2)$ è soluzione, scriviamo la matrice jacobiana di f:

$$J_f(x, y, z) = \begin{pmatrix} 1 & 1/y & 1 \\ 2 & -2y & 1 \end{pmatrix}$$

e quindi $J_f(0, 1, 2)$ ha rango 2. In particolare

$$\det \begin{pmatrix} \partial_y f_1(0, 1, 2) & \partial_z f_1(0, 1, 2) \\ \partial_y f_2(0, 1, 2) & \partial_z f_2(0, 1, 2) \end{pmatrix} = \det \begin{pmatrix} 1 & 1 \\ -2 & 1 \end{pmatrix} \neq 0\,,$$

da cui deduciamo l'esistenza delle funzioni g e h, che risultano essere di classe $\mathcal{C}^\infty$ essendo tale f. Inoltre

$$\begin{aligned} \begin{pmatrix} g'(x) \\ h'(x) \end{pmatrix} &= -\begin{pmatrix} \partial_y f_1(x, g(x), h(x)) & \partial_z f_1(x, g(x), h(x)) \\ \partial_y f_2(x, g(x), h(x)) & \partial_z f_2(x, g(x), h(x)) \end{pmatrix}^{-1} \begin{pmatrix} \partial_x f_1(x, g(x), h(x)) \\ \partial_x f_2(x, g(x), h(x)) \end{pmatrix} \\ &= -\begin{pmatrix} 1/g(x) & 1 \\ -2g(x) & 1 \end{pmatrix}^{-1} \begin{pmatrix} 1 \\ 2 \end{pmatrix} \\ &= -\frac{1}{1/g(x) + 2g(x)} \begin{pmatrix} 1 & -1 \\ 2g(x) & 1/g(x) \end{pmatrix} \begin{pmatrix} 1 \\ 2 \end{pmatrix} \end{aligned}$$

e quindi

$$g'(x) = \frac{g(x)}{2g^2(x) + 1}\,, \qquad h'(x) = -1 - \frac{1}{2g^2(x) + 1}\,, \tag{7}$$

da cui si ricava che $g'(0) = 1/3$ e $h'(0) = -4/3$. Derivando ulteriormente le espressioni delle derivate prime in (7) ricaviamo le derivate seconde:

$$g''(x) = \frac{g'(x)(2g^2(x) + 1) - 4g^2(x)g'(x)}{(2g^2(x) + 1)^2}\,, \qquad h''(x) = \frac{4g(x)g'(x)}{(2g^2(x) + 1)^2}\,,$$

e quindi $g''(0) = -1/27$ e $h''(0) = 4/27$. Allora le formule di MacLaurin del secondo ordine con il resto di Peano per g e h si scrivono

$$g(x) = 1 + \frac{1}{3}x - \frac{1}{54}x^2 + o(x^2)\,, \qquad h(x) = 2 - \frac{4}{3}x + \frac{2}{27}x^2 + o(x^2)\,.$$

Spazio tangente

Siano al solito $1 \leq m < n$ numeri naturali, $X \subseteq \mathbb{R}^n$ aperto e $\mathbf{f} : X \to \mathbb{R}^m$, $\mathbf{f} = (f_1, \ldots, f_m)$, funzione di classe $\mathcal{C}^1$. Consideriamo il vincolo

$$\Gamma = \big\{\mathbf{x} \in X : \mathbf{f}(\mathbf{x}) = \mathbf{b}\big\}\,,$$

dove $\mathbf{b}$ è un fissato elemento di $\mathbb{R}^m$.

Definizione 1 *Un punto $\overline{\mathbf{x}} \in \Gamma$ si dice regolare se la matrice jacobiana di $\mathbf{f}$ in $\overline{\mathbf{x}}$, $J_{\mathbf{f}}(\overline{\mathbf{x}})$, ha rango massimo (e quindi pari a m). In altre parole, $\overline{\mathbf{x}} \in \Gamma$ è punto regolare se gli m gradienti $\nabla f_1(\overline{\mathbf{x}}), \ldots, \nabla f_m(\overline{\mathbf{x}})$ di $f_1, \ldots, f_m$ in $\overline{\mathbf{x}}$ sono linearmente indipendenti.*

In un punto regolare del vincolo Γ è possibile dare una nozione di spazio tangente.

Definizione 2 *Sia $\overline{\mathbf{x}} \in \Gamma$ punto regolare. Si definisce lo spazio tangente a Γ in $\overline{\mathbf{x}}$, denotato con $T_\Gamma(\overline{\mathbf{x}})$, in questo modo: il vettore $\mathbf{v} \in \mathbb{R}^n$ appartiene a $T_\Gamma(\overline{\mathbf{x}})$ se esistono un $\delta > 0$ ed una curva $\gamma : [-\delta, \delta] \to \mathbb{R}^n$ di classe $\mathcal{C}^1$ con sostegno contenuto in Γ tali che $\gamma(0) = \overline{\mathbf{x}}$ e $\gamma'(0) = \mathbf{v}$.*

In altre parole, $T_\Gamma(\overline{\mathbf{x}})$ è l'insieme che contiene tutti i vettori del tipo $\gamma'(0)$ con γ curva $\mathcal{C}^1$ con sostegno contenuto in Γ tale che $\gamma(0) = \overline{\mathbf{x}}$. Lo spazio tangente al vincolo Γ si può completamente caratterizzare tramite i gradienti delle componenti di $\mathbf{f}$.

Teorema 2 *Siano Γ come sopra e $\overline{\mathbf{x}} \in \Gamma$ punto regolare. Allora $T_\Gamma(\overline{\mathbf{x}})$ è uno spazio vettoriale di dimensione $n - m$ e si ha*

$$T_\Gamma(\overline{\mathbf{x}}) = \big[\mathrm{span}\big\{\nabla f_1(\overline{\mathbf{x}}), \ldots, \nabla f_m(\overline{\mathbf{x}})\big\}\big]^\perp\,, \tag{8}$$

dove span$\{\nabla f_1(\overline{\mathbf{x}}), \ldots, \nabla f_m(\overline{\mathbf{x}})\}$ indica lo spazio vettoriale generato dai vettori $\nabla f_1(\overline{\mathbf{x}}), \ldots, \nabla f_m(\overline{\mathbf{x}})$.

Dim. Dividiamo la dimostrazione in due parti.

$T_\Gamma(\overline{\mathbf{x}})$ è spazio vettoriale. Siano $\mathbf{v} \in T_\Gamma(\overline{\mathbf{x}})$ e $\alpha \in \mathbb{R}$ e proviamo che $\alpha\mathbf{v} \in T_\Gamma(\overline{\mathbf{x}})$. Possiamo assumere $\alpha \neq 0$, altrimenti $\alpha\mathbf{v} \in T_\Gamma(\overline{\mathbf{x}})$ banalmente. Per ipotesi $\mathbf{v} = \gamma'(0)$ con γ curva di classe $\mathcal{C}^1$ definita in un intervallo $[-\delta, \delta]$, $\delta > 0$, con $\gamma(0) = \overline{\mathbf{x}}$. Sia $\widetilde{\gamma}(t) \doteq \gamma(\alpha t)$, $t \in [-\delta/|\alpha|, \delta/|\alpha|]$. Allora anche $\widetilde{\gamma}$ è curva di classe $\mathcal{C}^1$ con sostegno contenuto in Γ e soddisfa $\widetilde{\gamma}(0) = \gamma(0) = \overline{\mathbf{x}}$. Allora $\widetilde{\gamma}'(0) = \alpha\gamma'(0) = \alpha\mathbf{v} \in T_\Gamma(\overline{\mathbf{x}})$. Siano ora $\mathbf{v}, \widetilde{\mathbf{v}} \in T_\Gamma(\overline{\mathbf{x}})$ e proviamo che $\mathbf{v} + \widetilde{\mathbf{v}} \in T_\Gamma(\overline{\mathbf{x}})$. Siano γ, $\widetilde{\gamma}$ curve di classe $\mathcal{C}^1$ definite in un intervallo $[-\delta, \delta]$, $\delta > 0$, tali che $\gamma(0) = \widetilde{\gamma}(0) = \overline{\mathbf{x}}$, $\gamma'(0) = \mathbf{v}$ e $\widetilde{\gamma}'(0) = \widetilde{\mathbf{v}}$. Poiché $\overline{\mathbf{x}} = (\overline{x}_1, \ldots, \overline{x}_n)$ è punto regolare, la matrice $J_f(\overline{\mathbf{x}})$ ha rango m, e possiamo assumere che valga (5), cosicché per il teorema di Dini troviamo una funzione implicita $\mathbf{g}$ per cui vale (6). In particolare, a patto al più di prendere $\delta > 0$ sufficientemente piccolo, le componenti $(\gamma_1, \ldots, \gamma_n)$, $(\widetilde{\gamma}_1, \ldots, \widetilde{\gamma}_n)$ di γ e $\widetilde{\gamma}$ soddisfano

$$\gamma_j(t) = g_j(\gamma_{m+1}(t), \ldots, \gamma_n(t))\,, \quad \widetilde{\gamma}_j(t) = g_j(\widetilde{\gamma}_{m+1}(t), \ldots, \widetilde{\gamma}_n(t)) \qquad \forall\, j = 1, \ldots, m\,,$$

e quindi le componenti $(v_1, \ldots, v_n)$, $(\widetilde{v}_1, \ldots, \widetilde{v}_n)$ dei vettori $\mathbf{v} = \gamma'(0)$ e $\widetilde{\mathbf{v}} = \widetilde{\gamma}'(0)$ soddisfano

$$\begin{aligned} v_j &= \big\langle \nabla g_j(\overline{x}_{m+1}, \ldots, \overline{x}_n), (v_{m+1}, \ldots, v_n)\big\rangle \\ \widetilde{v}_j &= \big\langle \nabla g_j(\overline{x}_{m+1}, \ldots, \overline{x}_n), (\widetilde{v}_{m+1}, \ldots, \widetilde{v}_n)\big\rangle \end{aligned} \qquad j = 1, \ldots, m\,. \tag{9}$$

Definiamo ora la curva $\overline{\gamma}(t) = (\overline{\gamma}_1(t), \dots, \overline{\gamma}_n(t))$, $t \in [-\delta/2, \delta/2]$,

$$\begin{aligned} \overline{\gamma}_i(t) &\doteq \frac{\gamma_i(2t) + \widetilde{\gamma}_i(2t)}{2} \qquad && i = m+1, \dots, n\,, \\ \overline{\gamma}_i(t) &\doteq g_i(\overline{\gamma}_{m+1}(t), \dots, \overline{\gamma}_n(t)) && i = 1, \dots, m\,. \end{aligned} \tag{10}$$

che risulta essere di classe $\mathcal{C}^1$ essendo tali γ, $\widetilde{\gamma}$ e $\mathbf{g}$. Per costruzione, grazie a (6), il sostegno di $\overline{\gamma}$ è contenuto in Γ. Inoltre

$$\begin{aligned} \overline{\gamma}_i(0) &= \frac{\gamma_i(0) + \widetilde{\gamma}_i(0)}{2} = \overline{x}_i && i = m+1, \dots, n\,, \\ \overline{\gamma}_i(0) &= g_i(\overline{\gamma}_{m+1}(0), \dots, \overline{\gamma}_n(0)) = g_i(\overline{x}_{m+1}, \dots, \overline{x}_n) = \overline{x}_i \qquad && i = 1, \dots, m\,, \end{aligned}$$

Allora $\overline{\gamma}'(0) \in T_\Gamma(\overline{\mathbf{x}})$: in particolare, se $j = m+1, \dots, n$, vale

$$\overline{\gamma}_j'(0) = 2 \cdot \frac{\gamma_j'(0) + \widetilde{\gamma}_j'(0)}{2} = v_j + \widetilde{v}_j\,,$$

mentre, se $j = 1, \dots, m$, da (10) si ottiene

$$\begin{aligned} \overline{\gamma}_j(0) &= \langle \nabla g_j(\overline{\gamma}_{m+1}(0), \dots, \overline{\gamma}_n(0)), (\overline{\gamma}_{m+1}'(0), \dots, \overline{\gamma}_n'(0)) \rangle \\ &= \big\langle \nabla g_i(\overline{x}_{m+1}, \dots, \overline{x}_n), (v_{m+1}, \dots, v_n) \big\rangle + \big\langle \nabla g_i(\overline{x}_{m+1}, \dots, \overline{x}_n), (\widetilde{v}_{m+1}, \dots, \widetilde{v}_n) \big\rangle \\ &= v_j + \widetilde{v}_j\,, \end{aligned}$$

dove si è usata (9). Allora $\overline{\gamma}'(0) = \mathbf{v} + \widetilde{\mathbf{v}}$ e quindi $\mathbf{v} + \widetilde{\mathbf{v}} \in T_\Gamma(\overline{\mathbf{x}})$, come si voleva.

$T_\Gamma(\overline{\mathbf{x}})$ ha dimensione $n-m$ e vale (8). Innanzitutto cerchiamo $n-m$ vettori di $T_\Gamma(\overline{\mathbf{x}})$ linearmente indipendenti, cosicché possiamo dedurre che $\dim T_\Gamma(\overline{\mathbf{x}}) \geq n-m$. Per farlo dobbiamo trovare $n-m$ curve γ^k, $k = 1, \dots, n-m$, di classe $\mathcal{C}^1$ definite in un intervallo $[-\delta, \delta]$ con sostegno contenuto in Γ tali che $\gamma^k(0) = \overline{\mathbf{x}}$ per ogni k e ${\gamma^1}'(0), \dots, {\gamma^{n-m}}'(0)$ siano linearmente indipendenti. Come già osservato in precedenza, supponendo che valga (5) e detta $\mathbf{g}$ la funzione implicita definita dal vincolo $\mathbf{f}(\mathbf{x}) = \mathbf{b}$ in un intorno di $\overline{\mathbf{x}}$, le componenti $(\gamma_1^k, \dots, \gamma_n^k)$ di ciascuna delle curve γ^k devono soddisfare

$$\gamma_j^k(t) = g_j(\gamma_{m+1}^k(t), \dots, \gamma_n^k(t)) \quad \forall\, j = 1, \dots, m\,, \tag{11}$$

e quindi basta assegnare le componenti $\gamma_{m+1}^k(t), \dots, \gamma_n^k(t)$ di $\gamma^k(t)$ per assegnare la curva. Siano allora

$$\gamma_i^k(t) = \begin{cases} \overline{x}_i & \text{se} \quad i \neq m+k\,, \\ \overline{x}_i + t & \text{se} \quad i = m+k\,, \end{cases} \qquad i = m+1, \dots, n\,,$$

e quindi ricaviamo $\gamma_j^k(t)$, $j = 1, \dots, m$, da (11). Si osservi che poiché

$$\overline{x}_j = g_j(\overline{x}_{m+1}, \dots, \overline{x}_n) \quad \forall\, j = 1, \dots, m\,,$$

e

$$\gamma_i^k(0) = \overline{x}_i \quad \forall\, i = m+1, \dots, n\,,$$

si ha $\gamma^k(0) = \overline{\mathbf{x}}$ per ogni $k = 1, \dots, n-m$. Allora ${\gamma^k}'(0) \in T_\Gamma(\mathbf{x})$ per ogni k. Calcoliamo quindi ${\gamma^k}'(0)$, $k = 1, \dots, n-m$: si ottiene

$$\begin{aligned}
{\gamma^1}'(0) &= \big({\gamma_1^1}'(0), \dots, {\gamma_m^1}'(0), 1, 0, \dots, 0\big) \\
{\gamma^2}'(0) &= \big({\gamma_1^2}'(0), \dots, {\gamma_m^2}'(0), 0, 1, 0, \dots, 0\big) \\
\vdots \quad & \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \\
{\gamma^{n-m}}'(0) &= \big({\gamma_1^{n-m}}'(0), \dots, {\gamma_m^{n-m}}'(0), 0, \dots, 0, 1\big)\,,
\end{aligned}$$

che sono $n-m$ vettori linearmente indipendenti. Dunque $\dim T_\Gamma(\overline{\mathbf{x}}) \geq n-m$: se dimostriamo che

$$T_\Gamma(\overline{\mathbf{x}}) \subseteq \big[\mathrm{span}\big\{\nabla f_1(\overline{\mathbf{x}}), \dots, \nabla f_m(\overline{\mathbf{x}})\big\}\big]^\perp\,, \tag{12}$$

allora possiamo dedurre (8). Infatti, $T_\Gamma(\overline{\mathbf{x}})$ risulta essere un sottospazio vettoriale di dimensione almeno $n-m$ di uno spazio vettoriale di dimensione esattamente $n-m$, perché, essendo i vettori $\nabla f_1(\mathbf{x}), \dots, \nabla f_m(\mathbf{x})$ linearmente indipendenti, si ha

$$\dim \mathrm{span}\big\{\nabla f_1(\overline{\mathbf{x}}), \dots, \nabla f_m(\overline{\mathbf{x}})\big\} = m \qquad \Longrightarrow \qquad \dim \big[\mathrm{span}\big\{\nabla f_1(\overline{\mathbf{x}}), \dots, \nabla f_m(\overline{\mathbf{x}})\big\}\big]^\perp = n-m\,.$$

Dimostrare (12) significa provare che se $\mathbf{v} \in T_\Gamma(\overline{\mathbf{x}})$, allora

$$\langle \nabla f_i(\overline{\mathbf{x}}), \mathbf{v}\rangle = 0\,.$$

Sia $\gamma : [-\delta, \delta] \to \mathbb{R}^n$ curva di classe $\mathcal{C}^1$ con sostegno contenuto in Γ tale che $\gamma(0) = \overline{\mathbf{x}}$ e $\gamma'(0) = \mathbf{v}$. Allora $\mathbf{f}(\gamma(t)) = \mathbf{b}$ per ogni $t \in [-\delta, \delta]$, e quindi $f_i(\gamma(t)) = b_i$ per ogni $t \in [-\delta, \delta]$ e per ogni $i = 1, \dots, m$, dove $\mathbf{b} = (b_1, \dots, b_m)$. Allora

$$0 = \frac{d}{dt} f_i(\gamma(t)) = \langle \nabla f_i(\gamma(t)), \gamma'(t)\rangle\,,$$

ed in particolare per $t = 0$ si ottiene

$$0 = \langle \nabla f_i(\gamma(0)), \gamma'(0)\rangle = \langle \nabla f_i(\overline{\mathbf{x}}), \mathbf{v}\rangle\,,$$

come si voleva. □

I moltiplicatori di Lagrange

Nel seguito $X \subseteq \mathbb{R}^n$ è un aperto e $\Gamma \subset X$ un vincolo definito da

$$\Gamma = \big\{\mathbf{x} \in X : \mathbf{g}(\mathbf{x}) = \mathbf{b}\big\}\,, \tag{13}$$

dove $\mathbf{b} = (b_1, \dots, b_m) \in \mathbb{R}^m$ è fissato e $\mathbf{g} : X \to \mathbb{R}^m$ è funzione di classe $\mathcal{C}^1$ di componenti $\mathbf{g} = (g_1, \dots, g_m)$.

Definizione 3 *Un punto $\overline{\mathbf{x}} \in \Gamma$ si dice di massimo (risp. minimo) relativo o locale vincolato a Γ per una funzione $f : X \to \mathbb{R}$ se esiste un intorno $\mathcal{U}$ di $\overline{\mathbf{x}}$ tale che $f(\overline{\mathbf{x}}) \geq f(\mathbf{x})$ (risp. $f(\overline{\mathbf{x}}) \leq f(\mathbf{x})$) per ogni $\mathbf{x} \in \mathcal{U} \cap \Gamma$. Un punto di massimo o minimo relativo vincolato si dice anche di estremo vincolato.*

Usando la definizione 2 di spazio tangente ad un vincolo, si dimostra il seguente

Teorema 3 *Sia* $\overline{\mathbf{x}} \in \Gamma$ *punto regolare di* Γ *di estremo locale vincolato per una funzione* $f : X \to \mathbb{R}$ *differenziabile in* $\overline{\mathbf{x}}$. *Allora*

$$\langle \nabla f(\overline{\mathbf{x}}), \mathbf{v} \rangle = 0 \quad \forall\, \mathbf{v} \in T_\Gamma(\overline{\mathbf{x}})\,. \tag{14}$$

Dim. Per fissare le idee assumiamo che $\overline{\mathbf{x}}$ sia di massimo locale vincolato per f, cosicché troviamo un intorno $\mathcal{U}$ di $\overline{\mathbf{x}}$ tale che $f(\overline{\mathbf{x}}) \geq f(\mathbf{x})$ per ogni $\mathbf{x} \in \mathcal{U} \cap \Gamma$. Sia $\mathbf{v} \in T_\Gamma(\overline{\mathbf{x}})$ e sia $\gamma : [-\delta, \delta] \to \mathbb{R}^n$, $\delta > 0$, una curva di classe $\mathcal{C}^1$ con sostegno contenuto in Γ tale che $\gamma(0) = \overline{\mathbf{x}}$ e $\gamma'(0) = \mathbf{v}$. Senza perdere generalità, possiamo assumere che $\gamma(t) \in \mathcal{U}$ per ogni $t \in [-\delta, \delta]$. Allora la funzione

$$\varphi(t) \doteq f(\gamma(t)) \quad t \in [-\delta, \delta]\,,$$

è funzione derivabile in $t = 0$ che in tale punto ha un massimo, perché $\gamma(t) \in \mathcal{U}$ e quindi

$$\varphi(0) = f(\gamma(0)) = f(\overline{\mathbf{x}}) \geq f(\gamma(t))\,.$$

Poiché f è differenziabile in $\overline{\mathbf{x}}$ e γ derivabile in $t = 0$, φ è derivabile in $t = 0$ e grazie al teorema di Fermat $\varphi'(0) = 0$. Ma

$$\varphi'(0) = \langle \nabla f(\gamma(0)), \gamma'(0) \rangle = \langle \nabla f(\overline{\mathbf{x}}), \mathbf{v} \rangle\,,$$

da cui si deduce (14) per l'arbitrarietà di $\mathbf{v}$. □

Il teorema 3 è l'analogo del teorema di Fermat per gli estremi liberi di una funzione. Seguendo l'analogia, diamo la seguente

Definizione 4 *Un punto* $\overline{\mathbf{x}} \in \Gamma$ *si dice punto critico vincolato per una funzione* $f : X \to \mathbb{R}$ *differenziabile in* $\overline{\mathbf{x}}$ *se vale* (14)

Con questa definizione, il teorema 3 si può riparafrasare dicendo che un punto di estremo locale vincolato per una funzione differenziabile è un punto critico vincolato per tale funzione.
Dalla caratterizzazione dello spazio tangente ad un vincolo in un suo punto regolare fornita dal teorema 2 e usando il teorema 3, si deduce il seguente

Teorema 4 (sui moltiplicatori di Lagrange) *Sia* $\overline{\mathbf{x}} \in \Gamma$ *punto regolare di* Γ *di estremo locale vincolato per una funzione* $f : X \to \mathbb{R}$ *differenziabile in* $\overline{\mathbf{x}}$. *Allora esistono* $\overline{\lambda}_1, \ldots, \overline{\lambda}_m \in \mathbb{R}$ *tali che*

$$\nabla f(\overline{\mathbf{x}}) = \sum_{i=1}^{m} \overline{\lambda}_i \nabla g_i(\overline{\mathbf{x}})\,. \tag{15}$$

In particolare, $\overline{\lambda}_1, \ldots, \overline{\lambda}_m$ *si dicono moltiplicatori di Lagrange per il problema di estremo vincolato.*

Dim. Per (14), $\nabla f(\overline{\mathbf{x}})$ è ortogonale a tutti gli elementi di $T_\Gamma(\overline{\mathbf{x}})$, cioè $\nabla f(\overline{\mathbf{x}}) \in [T_\Gamma(\overline{\mathbf{x}})]^\perp$. Grazie a (8), si ha

$$[T_\Gamma(\overline{\mathbf{x}})]^\perp = \mathrm{span}\big\{\nabla g_1(\overline{\mathbf{x}}), \ldots, \nabla g_m(\overline{\mathbf{x}})\big\}\,,$$

e quindi $\nabla f(\overline{\mathbf{x}})$ è una combinazione lineare di $\nabla g_1(\overline{\mathbf{x}}), \ldots, \nabla g_m(\overline{\mathbf{x}})$, da cui la tesi del teorema. □

Il teorema 4 implica che i candidati punti di estremo vincolato per una funzione $f : X \to \mathbb{R}$ di classe $\mathcal{C}^1$ cadono necessariamente tra

1. i punti irregolari del vincolo;
2. i punti regolari del vincolo per i quali esistono $\overline{\lambda}_1, \ldots, \overline{\lambda}_m$ per cui (15) è soddisfatta.

In particolare, se il vincolo Γ è fatto di punti regolari e $\overline{\mathbf{x}} \in \Gamma$ è un punto di estremo relativo vincolato di f allora esistono $\overline{\lambda}_1, \ldots, \overline{\lambda}_m \in \mathbb{R}$ tali che $\overline{\mathbf{x}} = (\overline{x}_1, \ldots, \overline{x}_n)$ e $\overline{\lambda} = (\overline{\lambda}_1, \ldots, \overline{\lambda}_m)$ sono soluzioni del sistema di $n+m$ equazioni nelle $n+m$ incognite $x_1, \ldots, x_n, \lambda_1, \ldots, \lambda_m$

$$\begin{cases} \partial_{x_j} f(x_1, \ldots, x_n) = \displaystyle\sum_{i=1}^{m} \lambda_i \partial_{x_j} g_i(x_1, \ldots, x_n) & j = 1, \ldots, n\,, \\ g_i(x_1, \ldots, x_n) = b_i & i = 1, \ldots, m\,, \end{cases} \tag{16}$$

dove le prime n equazioni sono una riscrittura di (15), mentre le ultime m sono una riscrittura dell'equazione $\mathbf{g}(\mathbf{x}) = \mathbf{b}$ che definisce il vincolo Γ in (13). Allora, introdotta la funzione $\mathcal{L} : X \times \mathbb{R}^m \to \mathbb{R}$ definita da

$$\mathcal{L}(\mathbf{x}, \lambda) = f(\mathbf{x}) - \langle \lambda, \mathbf{g}(\mathbf{x}) - \mathbf{b} \rangle = f(x_1, \ldots, x_n) - \sum_{i=1}^{m} \lambda_i \big[g_i(x_1, \ldots, x_n) - b_i \big]\,,$$

e detta *lagrangiana* del problema di estremo vincolato, si può enunciare il seguente

Corollario 1 *Se f è di classe $\mathcal{C}^1$, ha un estremo relativo vincolato a Γ in $\overline{\mathbf{x}}$ e $\overline{\mathbf{x}}$ è punto regolare di Γ, allora esiste $\overline{\lambda} \in \mathbb{R}^m$ tale che $(\overline{\mathbf{x}}, \overline{\lambda})$ è un punto critico (libero) per $\mathcal{L}$.*

Dim. Se $\overline{\mathbf{x}} \in \Gamma$ è punto regolare di stremo relativo vincolato per f, allora esiste $\overline{\lambda} = (\overline{\lambda}_1, \ldots, \overline{\lambda}_m) \in \mathbb{R}^m$ tale che $\overline{x}_1, \ldots, \overline{x}_n, \overline{\lambda}_1, \ldots, \overline{\lambda}_m$ sono soluzioni del sistema (16). Poiché

$$\begin{aligned} \partial_{x_j} \mathcal{L}(\mathbf{x}, \lambda) &= \partial_{x_j} f(x_1, \ldots, x_n) - \sum_{i=1}^{m} \lambda_i \partial_{x_j} g_i(x_1, \ldots, x_n)\,, \\ \partial_{\lambda_i} \mathcal{L}(\mathbf{x}, \lambda) &= g_i(x_1, \ldots, x_n) - b_i\,, \end{aligned}$$

questo equivale a dire che $(\overline{\mathbf{x}}, \overline{\lambda}) \in X \times \mathbb{R}^m$ è punto critico (libero) per $\mathcal{L}$. $\square$

Esercizi

1. Provare che la funzione $f(x,y,z) = x^2 + y^2 + z^2$ ha massimo e minimo nell'insieme Γ delle soluzioni del sistema

$$\begin{cases} x^2 + y^2 = 1\,, \\ z + xy = 1\,, \end{cases}$$

e calcolarne il valore.

2. Data la funzione $f(x,y,z) = x^2 - y^2 - z^2$, calcolarne i punti critici vincolati a $\Gamma = \{(x,y,z) \in \mathbb{R}^3 : \operatorname{sen} x = y\}$ e studiarne la natura.

3. Determinare, se esistono, massimo e minimo della funzione $f(x,y) = \cos^2 y - \operatorname{sen}^2 x$ nell'insieme

$$\Gamma \doteq \big\{ (x,y) \in [0,\pi] \times [0,\pi] : \cos x = \operatorname{sen} y \big\}\,.$$

4. Si verifichi che l'equazione

$$\operatorname{sen}(x-1) - e^z + e^y + xz - y = 0$$

definisce in un intorno del punto $(1, 0, 0)$ una funzione $x = g(y, z)$. Successivamente si provi che in $(0, 0)$ tale funzione ha un punto critico di cui si chiede di determinare la natura.

5. Si provi che il sistema

$$\begin{cases} xe^x + y^2 - z^2 = 0 \\ x^2 - z + u^2 = 0 \end{cases}$$

definisce implicitamente in un intorno di $(0, 0, 0, 0)$ due funzioni $x = g(y, u)$ e $z = h(y, u)$. Si provi che entrambe le funzioni hanno in $(0, 0)$ un punto critico

6. Sia $f : \mathbb{R} \to \mathbb{R}$ una funzione continua tale che $f(0) = 0$. Si verifichi che il sistema

$$\begin{cases} y^2 - z^2 + \displaystyle\int_0^x f(t)\, dt = 0 \\ \ln y - 2 \ln z + x = 0 \end{cases}$$

definisce in un intorno di $x = 0$ due funzioni $y = g(x)$ e $z = h(x)$ tali che $g(0) = 1$ e $h(0) = 1$. Si discuta la derivabilità di tali funzioni e si scriva l'equazione della retta tangente e del piano normale in $(0, 1, 1)$ alla curva definita implicitamente dal sistema.

7. Provare che esistono e calcolare massimo e minimo assoluti della funzione

$$f(x, y) = \ln(y - x + 1)$$

nell'insieme

$$\Omega = \big\{(x, y) \in \mathbb{R}^2 : y \geq \arctan x\,,\ x \geq -1\,,\ y \leq \pi/4\big\}\,.$$

8. Provare che esistono e calcolare massimo e minimo assoluti della funzione

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

nell'insieme

$$\Omega = \big\{(x, y) \in \mathbb{R}^2 : x + 2 \leq y \leq \sqrt{4 - x^2}\big\}\,.$$

9. Sia $n \geq 2$ e

$$f(x_1, \ldots, x_n) = x_1 + \cdots + x_n \qquad x_i > 0\,.$$

1. Determinare il minimo di f sotto la condizione $x_1 \cdots x_n = 1$, dopo aver dimostrato che tale minimo esiste (si osservi che $f(\mathbf{x}) \to +\infty$ per $\mathbf{x} \to \infty$).
2. Dedurre la disuguaglianza
$$\sqrt[n]{x_1 \cdots x_n} \leq \frac{x_1 + \cdots + x_n}{n}$$
valida per $x_i > 0$ per ogni i (calcola f nel vettore di componenti $x_i / \sqrt[n]{x_1 \cdots x_n}$).

10. Sia
$$\Gamma = \left\{(x,y,z) \in R^3 : x^2 + y^2 - z^2 = 1\,,\ \ x + y + 2z^3 = 0\right\}.$$

1. Provare che esistono due funzioni $y = g(x)$ e $z = h(x)$ definite in un intorno $\mathcal{U}$ di $x = 1$, tali che $g(1) = 1$, $h(1) = -1$ e per cui $(x, g(x), h(x)) \in \Gamma$ per ogni $x \in \mathcal{U}$.
2. Provare che g è strettamente decrescente in un intorno di $x = 1$.
3. Provare che h ha in $x = 1$ un punto di estremo.
4. Provare che esiste un punto di Γ di quota (coordinata z) massima.
5. Verificare che tutti i punti di Γ sono regolari.
6. Trovare il punto di Γ di quota (coordinata z) massima.