18. Qualche applicazione

Il teorema di invertibilità locale e il teorema delle funzioni implicite sono strumenti utili in molte e svariate questioni. In questa sezione, illustriamo qualche applicazione.

18.a Piccole perturbazioni

Si tratta di una struttura tipica che si presenta in molti contesti.

18.1 Sistemi lineari. Si voglia risolvere il sistema di n equazioni in n incognite

$$\sum_{ij=1}^{n} a^h_{ij} x_i x_j + \sum_{i=1}^{n} b^h_i x_i + c^h = 0, \qquad h = 1, n \tag{18.1}$$

o in forma più compatta

$$X^T \mathbf{A} X + \mathbf{B} X + C = 0$$

con $X \in \mathbb{R}^n$, $\mathbf{A} \in M_{n,n}$, $\mathbf{B} \in M_{n,n}$, $C \in \mathbb{R}^n$. L'applicazione $f : \mathbb{R}^n \to \mathbb{R}^n$ data da $f(X) := X^T \mathbf{A} X + \mathbf{B} X$ è di classe C^∞, $f(0) = 0$, e ha differenziale in 0 dato da $df_0(v) = \mathbf{B}v$. Se $\mathbf{B}$ è invertibile, segue dal teorema di invertibilità locale che esistono due intorni U, V di zero in $\mathbb{R}^n$ tali che la (18.1) ha un'unica soluzione $X \in U$ per ogni scelta di $C \in V$.

18.b Lemma di Morse

18.2 Proposizione (Lemma di Morse). *Sia f una funzione regolare in un aperto di $\mathbb{R}^n$ e sia $0 \in \Omega$ un punto critico non degenere, $\mathbf{D}f(0) = 0$, $\det \mathbf{H}f(0) \neq 0$. Esistono allora un diffeomorfismo $\varphi : U \to \mathbb{R}^n$ da un intorno aperto U di 0 e un intero k tali che*

$$f(\varphi(\xi)) = -(\xi^1)^2 - \cdots - (\xi^k)^2 + (\xi^{k+1})^2 + \cdots + (\xi^n)^2.$$

Dimostrazione. Cominciamo con l'osservare che k non dipende dalla scelta delle coordinate, essendo in particolare la dimensione del più grande autospazio su cui la forma quadratica della matrice hessiana $\mathbf{H}\xi \bullet \xi$ è definita negativa.

Osserviamo ancora che per la formula di Taylor con resto in forma integrale o per il lemma di Hadamard, possiamo scrivere

$$f(x) = \sum_{i,j=1}^{n} h_{ij}(x) x^i x^j$$

in una palla centrata in 0.

Dimostriamo ora la tesi per induzione sulla dimensione. Possiamo supporre, modulo una trasformazione lineare che $h_{11}(0) = 1$. Consideriamo la trasformazione $y := \gamma(x)$ definita da

$$\begin{cases} y^i = x^i & \text{se } i \neq 1, \\ y^1 = x^1 + \sum_{i \geq 1} h_{1i}(x) x^i. \end{cases}$$

Poiché $\mathbf{D}\gamma(0) \neq 0$, γ è invertibile in un intorno di zero e non è difficile mostrare che

$$f(\gamma^{-1}(y)) = \pm (y^1)^2 + \sum_{i,j>1} h_{ij}^{(1)}(y) y^i y^j.$$

Supponendo ora che in un intorno di zero si possa scrivere in opportune coordinate

$$f(y) = \pm (y^1)^2 \pm (y^2)^2 \pm \cdots \pm (y^{r-1})^2 + \sum_{i,j=r}^{n} H_{ij}(x) y^i y^j,$$

con un cambiamento di coordinate lineare facciamo in modo che $H_{rr}(0) \neq 0$; la trasformazione di coordinate

$$\begin{cases} y^i = x^i & \text{se } i \neq r, \\ y^r = x^r + \sum_{i>r} H_{ir}(x) x^i \end{cases}$$

è invertibile in un intorno di zero e si verifica ancora che

$$f(y) = \sum_{i \leq r} \pm (y^i)^2 + \sum_{i,j>r} H_{ij}(y) y^i y^j.$$

□

18.3 Corollario. *I punti critici non degeneri di una funzione di classe C^2 sono isolati.*

18.c Il flusso gradiente

Siano A aperto in $\mathbb{R}^n$ e $f : A \to \mathbb{R}$ una funzione di classe $C^2(A)$. L'assegnazione di un dato iniziale $\overline{x}$ non critico per f produce un'unica soluzione $x(t)$ del sistema

$$\begin{cases} x'(t) = \nabla f(x(t)), \\ x(0) = \overline{x}, \end{cases}$$

detta *traiettoria del flusso gradiente per $\overline{x}$*. Si osservi che

(i) le traiettorie del flusso gradiente sono ortogonali agli insiemi di livello di f,
(ii) f cresce lungo le traiettorie del flusso gradiente in quanto

$$\frac{d}{dt} f(x(t)) = \nabla f(x(t)) \bullet x'(t) = |\nabla f(x(t))|^2.$$

18.4 Proposizione. *Se una traiettoria del flusso gradiente esiste per tutti i tempi $t > 0$ ed ha limite x_0 per $t \to \infty$, allora x_0 è un punto critico per f.*

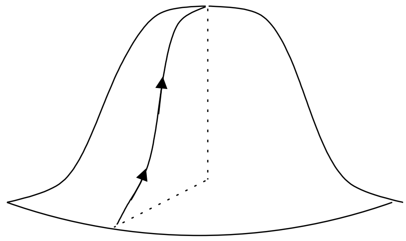

Figura 18.1. Il flusso gradiente.

Dimostrazione. Sia $x(t)$ una traiettoria che supponiamo esistere per ogni t con $x(t) \to x_0 \in \mathbb{R}^n$. Se per assurdo fosse $\nabla f(x_0) \neq 0$, esisterebbe un intorno U limitato di x_0 tale che $|\nabla f(x)| \geq m > 0$. Non è restrittivo supporre che $|f(x)| \leq M$ per qualche $M > 0$ in U e che per $t > \overline{t}$ $x(t) \in U$; per ogni $t > \overline{t}$ si avrebbe allora

$$f(x(t)) - f(x(\overline{t})) = \int_{\overline{t}}^{t} \frac{d}{ds} f(x(s))\, ds \geq m^2(t - \overline{t}),$$

i.e., $f(x(t)) \to +\infty$, in contradizione con il fatto che $\varphi(t) = f(x(t)) \to f(x_0)$. □

18.5 Proposizione. *Sia x_0 un punto critico per f con $\mathbf{H}f(x_0) < 0$. Allora esiste un intorno U di x_0 tale che ogni traiettoria del flusso gradiente che inizia al tempo $t = 0$ in U termina in x_0.*

Dimostrazione. Si ha infatti

$$\nabla f(x) \bullet (x - x_0) = \sum_{i,j=1}^{n} f_{x_i x_j}(x_0 + s(x - x_0))(x - x_0)_i (x - x_0)_j \leq -\frac{m}{2} |x - x_0|^2$$

per ogni x in un intorno di x_0 in cui $\mathbf{H}f(x)v \bullet v < -(m/2)|v|^2$. Ponendo $\psi(t) := |x(t) - x_0|^2$, troviamo allora

$$\psi'(t) = 2\,(x(t) - x_0) \bullet x'(t) \leq -m|x(t) - x_0|^2 = -m\,\psi(t),$$

e integrando $\log(\psi(t)) - \log(\psi(0)) \leq -mt$. □

La Proposizione 18.5 motiva la terminologia di *punto critico stabile* per un x_0 tale che $\nabla f(x_0) = 0$ e $\mathbf{H}f(x_0) < 0$.

18.d Punti critici vincolati

Discutiamo ora alcune condizioni necessarie per i massimi e minimi di funzioni differenziabili in presenza di vincoli sulle variabili indipendenti.

Sia Ω un aperto di $\mathbb{R}^n$ e $S \subset \mathbb{R}^n$. $x_0 \in S$ si dice di *massimo* (risp. *minimo*) *relativo per f vincolato a S* se esiste un intorno aperto W di x_0 in $\mathbb{R}^n$ tale che $f(x_0) \geq f(x)$ (risp. $f(x_0) \leq f(x)$) per ogni $x \in S \cap W$.

Discuteremo il caso dei punti di massimo e minimi *vincolati su vincoli "bilaterali lisci"*, i.e. vincoli che siano *sottovarietà*.

18.6 Definizione. *Siano Ω un aperto in $\mathbb{R}^n$, $f : \Omega \to \mathbb{R}$ una funzione di classe C^1, e Γ una sottovarietà di dimensione r. Un punto $x_0 \in \Gamma$ si dice un* punto critico per f vincolato a Γ *se*

$$\nabla f(x_0) \perp \operatorname{Tan}_{x_0}\Gamma. \tag{18.2}$$

In altre parole x_0 è critico per f su Γ se le derivate nelle direzioni tangenziali a Γ sono nulle,

$$\frac{\partial f}{\partial v}(x_0) = \mathbf{D}f(x_0)(v) = \nabla f(x_0) \bullet v = 0 \qquad \forall v \in \operatorname{Tan}_{x_0}\Gamma.$$

18.7 Proposizione. *Sia $\varphi : \Omega \subset \mathbb{R}^m \to \Gamma \subset \mathbb{R}^n$ una parametrizzazione iniettiva di rango massimo di Γ in un intorno di $x_0 = \varphi(u_0) \in \Gamma$. Allora $x_0 \in \Gamma$ è un punto critico per f vincolato a Γ se e solo se*

$$\mathbf{D}(f \circ \varphi)(u_0) = 0.$$

In particolare, se x_0 è di massimo o minimo relativo per f vincolato a Γ, allora x_0 è un punto critico per f vincolato a Γ.

Dimostrazione. Infatti
$$\mathbf{D}(f \circ \varphi)(u_0) = \mathbf{D}f(\varphi(u_0))\mathbf{D}\varphi(u_0)$$
e, formando le colonne di $\mathbf{D}\varphi(u_0)$ una base di $\operatorname{Tan}_{x_0}\Gamma$, si conclude che $\mathbf{D}(f \circ \varphi)(u_0) = 0$ se e solo se $0 = \mathbf{D}(f \circ \varphi)(u_0) = \nabla f(x_0) \bullet v = 0$ per ogni $v \in \operatorname{Tan}_{x_0}\Gamma$.

Se poi x_0 è un massimo (minimo) relativo per f su Γ, $f \circ \varphi : \Omega \to \mathbb{R}$ ha un massimo (minimo) relativo libero. Segue dal teorema di Fermat che $\mathbf{D}(f \circ \varphi)(u_0) = 0$. □

18.8 Teorema (dei moltiplicatori di Lagrange). *Siano Ω un aperto di $\mathbb{R}^n$, $\phi : \Omega \to \mathbb{R}^m$ di classe C^1, $m < n$. Siano $\Gamma := \{x \in \Omega \,|\, \phi(x) = 0\}$ e $x_0 \in \Gamma$ tale che* $\operatorname{Rank}\mathbf{D}\phi(x_0) = m$. *Sia poi $f : \Omega \to \mathbb{R}$ una funzione di classe C^1. Sono fatti equivalenti:*

(i) *x_0 è un punto critico per f vincolato a Γ.*
(ii) *Esistono costanti $\lambda_1^0, \dots, \lambda_m^0$ tali che*

$$\mathbf{D}f(x_0) = \sum_{i=1}^{m} \lambda_i^0 \mathbf{D}\phi^i(x_0). \tag{18.3}$$

(iii) *Esiste $\lambda^0 \in \mathbb{R}^m$ tale che (x_0, λ^0) è un punto critico libero per la funzione $F : \Omega \times \mathbb{R}^r \to \mathbb{R}$*

$$F(x,\lambda) := f(x) - \sum_{i=1}^{m} \lambda_i \phi^i(x).$$

I numeri $\lambda_1^0, \lambda_2^0, \dots, \lambda_m^0$ si chiamano *moltiplicatori di Lagrange*.

Dimostrazione. (i) $\Leftrightarrow$ (ii). Si ha
$$\operatorname{Tan}_{x_0}\Gamma = \ker \mathbf{D}\phi(x_0)$$
e per definizione x_0 è critico per f su Γ se e solo se
$$\mathbf{D}f(x_0)(v) \in \ker \mathbf{D}\phi(x_0)^{\perp} = \operatorname{Im}(\mathbf{D}\phi(x_0)^T)$$
i.e., se e solo se $\mathbf{D}f(x_0)$ è combinazione lineare delle righe di $\mathbf{D}\varphi(x_0)$.

(ii) $\Leftrightarrow$ (iii). La condizione di gradiente nullo per $F(x,\lambda)$ in (x_0, λ^0) è proprio la (ii)

$$\begin{cases} \mathbf{D}f(x_0) = \sum_{i=1}^{m} \lambda_i^0 \mathbf{D}\phi^i(x_0), \\ \phi(x_0) = 0. \end{cases}$$

□

18.9 Esercizio. Osserviamo esplicitamente che nel Teorema 18.8 l'ipotesi che Γ sia una sottovarietà è essenziale. Infatti scegliendo ad esempio $f(x,y) = y$ in $\mathbb{R}^2$ e scegliendo come vincolo $\Gamma := \{(x,y) \,|\, y^3 - x^2 = 0\}$ si vede subito che f ha un minimo assoluto in $(0,0)$ su Γ, mentre la funzione $F(x,y,\lambda) := y + \lambda(y^3 - x^2)$ non ha punti critici.

Varie diseguaglianze omogenee ottenibili con l'uso essenzialmente della diseguaglianza di Jensen o con l'algebra lineare sono riottenibili usando il teorema dei moltiplicatori di Lagrange. Concludiamo questa sezione con alcuni esempi.

18.10 Proiezione ortogonale su un piano. Siano S un sottospazio vettoriale di $\mathbb{R}^n$, $(\ |\ )$ un prodotto scalare su $\mathbb{R}^n$, $||x||^2 := (x|x)$ la norma relativa e $b \in \mathbb{R}^n$. La funzione $f(x) := ||x - b||^2$, $x \in S$ è continua non negativa e $f(||x||) \to +\infty$ per $||x|| \to +\infty$, $x \in S$. f ha dunque (almeno) un minimo vincolato a S, x_0. Deve essere perciò

$$\nabla(||x_0 - b||^2) = 2(x - b) \perp \operatorname{Tan}_{x_0} S,$$

i.e.,

$$2(x_0 - b|v) = 0 \qquad \forall v \in S$$

quindi x_0 è unico ed è il piede della perpendicolare a S passante per b.

18.11 Minima distanza tra due superfici. Date due sottovarietà S, T compatte senza bordo di $\mathbb{R}^n$, caratterizzare i punti $x \in S$, $y \in T$ a distanza minima.

Si tratta di minimizzare $|x - y|$ con il vincolo $x \in S$, $y \in T$. Poiché la funzione $|x - y|$, $(x,y) \in \mathbb{R}^n \times \mathbb{R}^n$ è continua, e $S \times T \subset \mathbb{R}^n \times \mathbb{R}^n$ è compatto essendo S e T compatti, l'esistenza di almeno una coppia di punti $x_0 \in S$ e $y_0 \in T$ a distanza minima segue dal teorema di Weierstrass. È poi ovvio che $S \cap T \neq \emptyset$ se e solo se la distanza minima di S da T è nulla e in questo caso tutte le coppie di punti in $S \cap T$ sono minimi. Se invece $x_0 \in S$, $y_0 \in T$, $x_0 \neq y_0$ realizzano la distanza minima, allora $y \to |x_0 - y|$, $y \in T$, ha un minimo in y_0, e $x \to |x - y_0|$, $x \in S$ ha un minimo in x_0. Pertanto, cfr. 18.10, il vettore $x_0 - y_0$ è perpendicolare sia a $\operatorname{Tan}_{x_0} S$ che a $\operatorname{Tan}_{y_0} T$.

18.12 Autovalori di una matrice autoaggiunta. Sia $\mathbf{A} \in M_{n,n}(\mathbb{R}^n)$ una matrice autoaggiunta. È noto che per il più grande autovalore positivo L di $\mathbf{A}$ si ha

$$L := \max_{|x|=1} \mathbf{A}x \bullet x$$

e che il massimo è ottenuto sugli autovettori relativi a L di norma 1. Si può ritrovare questa formula guardando al problema vincolato

$$\begin{cases} \text{Massimizzare } \mathbf{A}x \bullet x \\ \text{sul vincolo } |x|^2 = 1. \end{cases}$$

Essendo $S := \{x \,|\, |x| = 1\}$ compatto, $\mathbf{A}x \bullet x$ ha certamente almeno un punto di massimo x_0 in S. Essendo S una sottovarietà di $\mathbb{R}^n$, dovrà essere

$$\nabla(\,\mathbf{A}x \bullet x\,)(x_0) = 2\mathbf{A}x_0 \in (\operatorname{Tan}_{x_0} S)^\perp, \qquad \text{e} \qquad |x|^2 - 1 = 0$$

i.e.,

$$\begin{cases} 2\mathbf{A}x_0 = 2\lambda\, x_0, \\ |x_0| = 1 \end{cases}$$

per qualche $\lambda \in \mathbb{R}$. In altre parole x_0 è un autovettore, e il moltiplicatore λ è l'autovalore corrispondente. D'altra parte $\mathbf{A}x_0 \bullet x_0 = \lambda\,|x_0|^2 = \lambda$. Si conclude quindi che i massimi di $(\mathbf{A}x|x)$ con il vincolo $|x| = 1$ sono tutti gli autovettori relativi all'autovalore massimo L e che

$$\max_{|x|=1} \mathbf{A}x \bullet x = L.$$