3 Il teorema di minimax

3.1 Il teorema di passo montano

Sia H uno spazio di Hilbert reale e sia f una funzione reale differenziabile su H. Supponiamo che un certo sottolivello $\{f < a\}$ non sia connesso, diciamo $\{f < a\} = A \cup B$ con A e B aperti disgiunti non vuoti. Possiamo pensare ad A e B come due valli e considerare l'insieme dei cammini che vanno da una valle all'altra, ossia l'insieme

$$\Gamma := \{\text{curve in } H \text{ con un estremo in } A \text{ e l'altro in } B\}.$$

Qua per curva intendiamo l'immagine di un'applicazione continua da un intervallo $[a,b]$ in H e per estremi l'immagine di a e b. Definiamo il valore di minimax di f relativo alla classe Γ come

$$c := \inf_{\gamma \in \Gamma} \max_{x \in \gamma} f(x),$$

e notiamo che $a \leq c < +\infty$, poiché Γ è non vuoto ed ogni suo elemento è un compatto che interseca $H \setminus (A \cup B) = \{f \geq a\}$. Ci si potrebbe aspettare che questo livello di passo montano c sia un punto critico di f. Il seguente semplice esempio mostra che questo potrebbe non essere vero:

Esempio 3.1. *Si consideri la funzione $f : \mathbb{R}^2 \to \mathbb{R}$ definita da*

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

Allora $\{f < 0\}$ ha due componenti connesse, $c = 0$, ma f non ha punti critici. Il problema qui è che il punto critico si trova "all'infinito": più precisamente, vi sono successioni $(z_n) \subset \mathbb{R}^2$, per esempio $z_n = (-n, 0)$, tali che $f(z_n)$ tende al valore di passo montano $c = 0$ e $df(z_n)$ tende a zero.

Questo esempio suggerisce la seguente definizione, che diamo in generale per funzioni differenziabili con continuità su spazi di Banach.

Definizione 3.2. *Siano X uno spazio di Banach reale e $f \in C^1(X)$. Una successione $(x_n) \subset X$ si dice successione di Palais-Smale al livello c per f (in breve, successione $(PS)_c$) se*

$$\lim_{n\to\infty} f(x_n) = c \quad e \quad \lim_{n\to\infty} df(x_n) = 0.$$

Diciamo che la funzione f soddisfa la condizione di Palais-Smale al livello c (in breve, soddisfa $(PS)_c$) se tutte le successioni $(PS)_c$ possiedono una sottosuccessione convergente. Diciamo che f soddisfa (PS) se soddisfa $(PS)_c$ per ogni $c \in \mathbb{R}$.

La convergenza di $df(x_n)$ è da intendersi nella topologia di X^* indotta dalla norma duale. Dal fatto che f è C^1, deduciamo che i punti limite di successioni $(\mathrm{PS})_c$ sono punti critici di f al livello c. Possiamo enunciare adesso il teorema del passo montano di Ambrosetti-Rabinowitz:

Teorema 3.3 (Passo montano). *Sia H uno spazio di Hilbert reale, sia $f \in C^{1,1}(H)$ tale che $\{f < a\}$ non sia connesso e sia c definito come sopra. Allora f possiede una successione $(PS)_c$. In particolare, se f soddisfa $(PS)_c$, allora c è un valore critico di f.*

Qui $C^{1,1}$ indica l'insieme delle funzioni il cui differenziale è localmente Lipschitziano.

Dimostrazione. Per assurdo, supponiamo che esista un $\epsilon > 0$ tale che $\|df\| \geq \epsilon$ sull'insieme $\{|f - c| \leq \epsilon\}$. Indichiamo con ∇f il gradiente di f e supponiamo, per semplicità, che il campo localmente Lipschitziano $-\nabla f$ sia positivamente completo, ossia che il suo flusso locale ϕ sia definito per tutti i tempi positivi. Si veda l'Osservazione 3.4 per informazioni su come rimuovere questa ipotesi supplementare.

Si noti che

$$\frac{d}{dt} f\big(\phi_t(u)\big) = df\big(\phi_t(u)\big)\big[-\nabla f\big(\phi_t(u)\big)\big] = -\big\|df\big(\phi_t(u)\big)\big\|^2 \leq 0,$$

quindi la funzione $t \mapsto f(\phi_t(u))$ è decrescente per ogni $u \in H$ (strettamente se $u \notin \operatorname{crit} f$). Se $|f(\phi_t(u)) - c| \leq \epsilon$ per ogni $t \in [0,T]$, si ha

$$2\epsilon \geq f(u) - f(\phi_T(u)) = -\int_0^T \frac{d}{dt} f(\phi_t(u))dt = \int_0^T \big\|df\big(\phi_t(u)\big)\big\|^2 dt \geq \epsilon^2 T,$$

da cui concludiamo che $T \leq 2/\epsilon$. Ne segue che se $T > 2/\epsilon$, allora

$$\phi_T\big(\{f \leq c+\epsilon\}\big) \subset \{f < c-\epsilon\}. \tag{1}$$

Fissiamo un numero $T > 2/\epsilon$, scegliamo $\gamma \in \Gamma$ tale che $\max_\gamma f \leq c+\epsilon$ e poniamo

$$\tilde{\gamma} := \phi_T(\gamma).$$

Il fatto che f decresce lungo le orbite di ϕ implica che per ogni $t \geq 0$ la curva $\phi_t(\gamma)$ appartiene a Γ. In particolare, $\tilde{\gamma}$ appartiene a Γ. Dal fatto che $\gamma \subset \{f \leq c+\epsilon\}$ e da (1) segue che $\tilde{\gamma} \subset \{f < c-\epsilon\}$, il che contraddice la definizione di c. $\square$

Osservazione 3.4. *Se il campo $-\nabla f$ non è positivamente completo, possiamo rimpiazzarlo con il campo completo $-\nabla f/\sqrt{\|\nabla f\|^2+1}$ (è completo poiché limitato). Lavorando con questo campo, la dimostrazione richiede soltanto piccoli aggiustamenti.*

Osservazione 3.5. *Il teorema di passo montano vale anche per $f \in C^{1,1}(M)$, dove (M,g) è una varietà Hilbertiana connessa munita di una metrica Riemanniana completa. In questo caso, $(x_n) \subset M$ si dice successione $(PS)_c$ se $f(x_n) \to c$ e $\|df(x_n)\| \to 0$, dove $\|\cdot\|$ indica la norma duale indotta da g. Si noti che la condizione di Palais-Smale e la completezza di g sono richieste antagoniste. Si può sempre realizzare la completezza di una metrica Riemanniana arbitraria g moltiplicandola per una funzione positiva che diverga all'infinito così rapidamente da eliminare tutte le successioni di Cauchy non convergenti. Invece la condizione di Palais-Smale può essere realizzata moltiplicando g per una funzione positiva sufficientemente infinitesima all'infinito: dato che la norma duale è moltiplicata per l'inversa di tale funzione, questa operazione riduce l'insieme delle successioni di Palais-Smale.*

Osservazione 3.6. *Il teorema di passo montano vale anche se f è solamente di classe C^1 ed è definita su uno spazio di Banach (o più in generale su una varietà di Banach). In questo caso il campo gradiente non è definito (caso Banach), o comunque non definisce un flusso continuo (caso Hilbert ma f solo C^1), però si può ovviare a questo problema costruendo un campo pseudo-gradiente per f che sia localmente Lipschitziano sul complementare dei punti critici. La costruzione si basa sul fatto che su spazi di Banach esistono sempre partizioni dell'unità localmente Lipschitziane (sugli spazi di Hilbert esistono anche partizioni dell'unità C^∞ ed è possibile costruire campi pseudo-gradiente di classe C^∞ sul complementare dei punti critici). Si veda [Str00, Lemma 3.2] per i dettagli di questa costruzione.*

Osservazione 3.7. *Trattando con funzioni definite su varietà talvolta è utile poter disporre di una formulazione del teorema di passo montano che non coinvolga la scelta di una metrica. Ecco una tale formulazione. Sia f una funzione di classe C^1 su una varietà di Banach connessa e sia V un campo vettoriale localmente Lipschitziano e positivamente completo su M tale che $df[V] < 0$ su $M \setminus \operatorname{crit} f$. Allora il teorema di passo montano vale, a patto che per successione $(PS)_c$ si intenda una successione $(x_n) \subset M$ tale che $f(x_n) \to c$ e $df(x_n)[V(x_n)] \to 0$. Adesso l'antagonismo contrappone questa forma della condizione di Palais-Smale alla positiva completezza di V.*

3.2 Il teorema di minimax generale

Nella dimostrazione del teorema di passo montano non abbiamo usato il fatto che Γ sia un insieme di curve, ma piuttosto il fatto che Γ sia positivamente invariante per il flusso gradiente ϕ di f, ossia che $\phi_t(\gamma) \in \Gamma$ per ogni $\gamma \in \Gamma$ e per ogni $t \geq 0$. Qui ϕ è il flusso di $-\nabla f$, qualora questo campo vettoriale sia positivamente completo, oppure il flusso di un campo positivamente completo conformalmente equivalente, come $-\nabla f/\sqrt{\|\nabla f\|^2+1}$, nel caso generale. Questa semplice osservazione porta alla seguente potente generalizzazione del teorema di passo montano:

Teorema 3.8 (Principio di minimax generale). *Siano H uno spazio di Hilbert, $f \in C^{1,1}(H)$ e Γ una classe di sottoinsiemi di H positivamente invariante per il flusso gradiente di f. Se il numero*

$$c := \inf_{\gamma\in\Gamma} \sup_{x\in\gamma} f(x)$$

è finito, allora f possiede una successione $(PS)_c$. In particolare, se f soddisfa $(PS)_c$, allora c è un valore critico di f.

La dimostrazione è pressoché identica quella del Teorema 3.3. Anche l'utilizzo di classi Γ piuttosto banali fornisce risultati interessanti. Se per esempio Γ consiste del solo insieme H, c coincide con l'estremo superiore di f su H. Quando questo estremo superiore è finito, il Teorema 3.8 assicura che esistano successioni massimizzanti lungo le quali il differenziale è infinitesimo. Simmetricamente, se Γ è l'insieme di tutti i punti di H, c coincide con l'estremo inferiore e quando questo è finito esistono successioni minimizzanti lungo le quali il differenziale è infinitesimo. Il lettore è invitato a provare a dare una dimostrazione diretta di questi fatti.

Il teorema di passo montano può essere visto come il caso $k = 0$ del seguente enunciato: supponiamo che $f \in C^{1,1}(H)$ possieda un sottolivello $\{f < a\}$ che non sia k-connesso, ossia tale che $\pi_k(\{f < a\}) \neq 0$, per un certo $k \in \mathbb{N}$; allora f possiede successioni $(\mathrm{PS})_c$. Infatti si può prendere come classe Γ l'insieme delle immagini delle applicazioni continue

$$u : \big(D^{k+1}, \partial D^{k+1}\big) \to \big(H, \{f < a\}\big),$$

la cui restrizione al bordo determini una classe non banale in $\pi_k(\{f < a\})$. Qua D^{k+1} indica la palla chiusa di dimensione $k+1$. Questa classe è ovviamente positivamente invariante per il flusso gradiente di f, poiché per tempi positivi questo manda $\{f < a\}$ in sé stesso. Dato che Γ è non vuota ed è composta da compatti, il valore di minimax c non è $+\infty$. Inoltre il fatto che $[u|_{\partial D^{k+1}}] \neq 0$ implica che $u(D^{k+1})$ interseca in complementare di $\{f < a\}$ e quindi $c \geq a > -\infty$. Il Teorema 3.8 implica dunque l'esistenza di successioni $(PS)_c$.

Il teorema di minimax generale può essere esteso a varietà, di Hilbert e di Banach, e a funzioni solamente C^1. Si vedano le osservazioni che seguono il teorema di passo montano, alle quali aggiungiamo la seguente:

Osservazione 3.9. *Talvolta è utile rimpiazzare il flusso gradiente con un flusso che fissa un certo sottolivello di f. Sia $\rho : \mathbb{R} \to [0, +\infty[$ una funzione di classe C^∞ liminitata tale che $\rho = 0$ su $[-\infty, b]$ e $\rho > 0$ su $]b, +\infty[$. Consieriamo il campo vettoriale $V = -\rho(f) \cdot \nabla f$ (oppure $V = -\rho(f) \cdot \nabla f / \sqrt{\|\nabla f\|^2 + 1}$ nel caso non positivamente completo) ed indichiamo il suo flusso con ϕ. Si tratta di un flusso gradiente troncato sotto il livello b: la funzione $t \mapsto f(\phi_t(u))$ è costante se $u \in \operatorname{crit} f \cup \{f \leq b\}$ ed è strettamente decrescente altrimenti. Se Γ è positivamente invariante rispetto a questo flusso e risulta $c > b$, allora f possiede successioni $(\mathrm{PS})_c$. La dimostrazione di questo fatto è lasciata al lettore.*

3.3 Esistenza di una geodetica chiusa sulla sfera

Sia g una matrica Riemanniana di classe C^∞ sulla sfera S^2. Le geodetiche sono le soluzioni $x : \mathbb{R} \to S^2$ dell'equazione

$$\nabla_t \dot{x}(t) = 0.$$

dove ∇_t indica la derivata covariante lungo x. Equivalentemente, se consideriamo un'embedding di S^2 in $\mathbb{R}^N$ e muniamo S^2 della metrica indotta dalla struttura Euclidea di $\mathbb{R}^N$, le geodetiche sono quelle curve che prendono valori in S^2 e la cui derivata seconda sia normale alla superficie.

Siamo interessati alle geodetiche chiuse, ossia a quelle geodetiche non costanti $x : \mathbb{R} \to S^2$ tali che $x(t + T) = x(t)$ per ogni $t \in \mathbb{R}$, per un certo $T > 0$. A meno di riparametrizzare la geodetica, possiamo sempre assumere che T sia 1. È ben noto che le geodetiche chiuse sono gli estremali del funzionale energia

$$E(x) = \frac{1}{2} \int_{\mathbb{T}} g_{x(t)}\big(\dot{x}(t), \dot{x}(t)\big)\, dt,$$

definito su un opportuno spazio di curve chiuse $x : \mathbb{T} \to S^2$. Qui $\mathbb{T} := \mathbb{R}/\mathbb{Z}$.

È possibile definire una varietà Hilbertiana M di curve $x : \mathbb{T} \to S^2$ tale che:

1. M si immerge con continuità in $C^0(\mathbb{T}, S^2)$, munito della topologia compatto aperta, e questa immersione è un'equivalenza omotopica (ossia esiste una mappa continua r da $C^0(\mathbb{T}, S^2)$ in M tale che, detta i l'immersione, le mappe $r \circ i$ e $i \circ r$ sono omotope all'identità);

2. il funzionale E è di classe C^∞ su M;

3. i punti critici di E sono tutte e sole le geodetiche chiuse su S^2;

4. M possiede una struttura Riemanniana completa e E soddisfa la condizione di Palais-Smale rispetto a tale struttura.

La varietà M consiste delle curve chiuse di classe di Sobolev H^1, ossia è definita da

$$M := \left\{ x : \mathbb{T} \to S^2 \mid x \text{ è assolutamente continua e } \int_{\mathbb{T}} g(\dot{x}, \dot{x})\, dt < +\infty \right\}.$$

Equivalentemente, può essere definita come il sottoinsieme dello spazio di Sobolev $H^1(\mathbb{T}, \mathbb{R}^3)$ che consiste delle curve x tali che $|x(t)| = 1$ per ogni $t \in \mathbb{T}$. Qua non intendiamo approfondire i dettagli analitici che stanno dietro alla dimostrazione dei punti 1-4 (per i quali si può consultare [Str00]), ma piuttosto mostrare l'idea geometrica dell'argomento di minimax, dovuta a G. Birkhoff.

Il funzionale E è non negativo, il suo minimo è 0 ed è raggiunto sul sottoinsieme M_0 di M composto dalle curve costanti. Il nostro scopo è dimostrare l'esistenza di un livello critico positivo per E.

Sia $I = [0, 1]$. Una mappa continua

$$u : (I, \partial I) \to (M, M_0)$$

determina una mappa continua $\tilde{u} : S^2 \to S^2$ che in coordinate sferiche

$$x = \sin\theta \cos\varphi, \quad y = \sin\theta \sin\varphi, \quad z = \cos\theta, \qquad 0 \le \theta \le \pi, \quad 0 \le \varphi \le 2\pi,$$

assume la forma

$$\tilde{u}(\theta, \varphi) := u(\theta/\pi)\big(\varphi/(2\pi)\big).$$

Indichiamo con Γ l'insieme delle immagini delle mappe continue $u : (I, \partial I) \to (M, M_0)$ tali che $\tilde{u}$ è omotopa all'identità su S^2. Per esempio, l'immagine della mappa

$$u(s)(t) := \big(\sin(\pi s)\cos(2\pi t), \sin(\pi s)\sin(2\pi t), \cos(\pi s)\big)$$

appartiene a Γ, poiché $\tilde{u}$ è proprio l'indentità su S^2. Il fatto che M_0 è composto da punti fissi per il flusso gradiente di E implica che Γ è invariante per tale flusso. Consideriamo il livello di minimax

$$c := \inf_{\gamma \in \Gamma} \max_{x \in \gamma} E(x).$$

Ovviamente $0 \le c < +\infty$. Il Teorema 3.8 implica che c è un valore critico di E. È quindi sufficiente dimostrare che $c > 0$.

Sia $r > 0$ un numero inferiore al raggio di iniettività della varietà Riemanniana (S^2, g): per ogni $x_0 \in S^2$ la palla $B_r(x_0) \subset S^2$ composta dai punti che distano da x_0 meno di r nella distanza indotta da g è diffeomorfa ad un disco aperto di raggio r nello spazio tangente $T_{x_0}S^2$ tramite l'inversa della mappa esponenziale

$$\exp_{x_0} : T_{x_0}S^2 \to S^2.$$

Mostriamo che se $u : (I, \partial I) \to (M, M_0)$ soddisfa

$$\max_{s \in I} E(u(s)) < \frac{r^2}{2}, \tag{2}$$

allora $\tilde{u}$ è omotopa ad una costante. Dato che l'identità su S^2 non è omotopa ad una costante, questo implica che $c \ge r^2/2 > 0$.

Se $s \in I$, la lunghezza di $u_s = u(s\cdot)$ può essere maggiorata grazie alla disuguaglianza di Cauchy-Schwarz da

$$\ell(u_s) = \int_{\mathbb{T}} \sqrt{g\Big(\frac{d}{dt}u_s(t), \frac{d}{dt}u_s(t)\Big)}\, dt \le \left(\int_{\mathbb{T}} g\Big(\frac{d}{dt}u_s(t), \frac{d}{dt}u_s(t)\Big)\, dt \right)^{1/2} = \big(2E(u_s)\big)^{1/2} < r.$$

Allora per ogni $s \in I$ risulta

$$u_s(t) \in B_r\big(u_s(0)\big).$$

Usando la mappa esponenziale di centro $u_s(0)$ e la convessità della palla di raggio r in $T_{u_s(0)}S^2$, possiamo costruire un'omotopia

$$h_\lambda : (I, \partial I) \to (M, M_0), \quad \lambda \in [0,1],$$

tra u e la mappa

$$v : (I, \partial I) \to (M, M_0), \quad v(s)(t) := u_s(0),$$

la quale può essere facilmente omotopizzata alla costante

$$w : (I, \partial I) \to (M, M_0), \quad w(s)(t) := u_0(0).$$

Corrispondentemente, la mappa $\tilde{u}$ è omotopa alla costante $\tilde{w}$, come affermato. Abbiamo così dimostrato il seguente:

Teorema 3.10 (Birkhoff). *La sfera S^2 munita di un'arbitraria metrica Riemanniana possiede una geodetica chiusa non costante.*

Osservazione 3.11. *In effetti, qualunque varietà Riemanniana compatta (W, g) di dimensione $n \geq 1$ possiede una geodetica chiusa non costante. Se W non è semplicemente connessa, allora è possibile trovare una geodetica chiusa non costante minimizzando l'energia E su una componente connessa di lacci non contrattili su W (per esempio usando lo spazio di Sobolev $H^1(\mathbb{T}, W)$ e sfruttando il fatto che E soddisfa (PS) su tale spazio, oppure usando il metodo diretto del calcolo delle variazioni). Se W è semplicemente connessa, si usa un argomento di minimax analogo a quello visto sopra, partendo dal fatto che deve esistere un $k \geq 2$ tale che $\pi_k(W) \neq 0$ (se fosse $\pi_k(W) = 0$ per ogni $k \geq 1$, allora, per un teorema di Whitehead, W sarebbe omotopicamente equivalente ad un punto, ma una varietà compatta non lo è mai, ad esempio poiché il suo n-esimo gruppo di omologia a coefficienti in $\mathbb{Z}_2$ è non banale).*

Riferimenti bibliografici

[Str00] M. Struwe, *Variational methods*, third ed., Springer-Verlag, Berlin, 2000.