Funzioni Lipschitziane, Teorema del punto fisso, e Teorema della funzione inversa

Funzioni Lipschitziane

Definizione 1. *Sia K un sottoinsieme di $\mathbb{R}^n$. Data una costante $L > 0$, diciamo che una funzione $F : K \to \mathbb{R}^m$ è L-Lipschitz, se*

$$|F(X) - F(Y)| \le L|X - Y| \quad \textit{per ogni} \quad X, Y \in K.$$

Proposizione 2. *Sia Ω un aperto di $\mathbb{R}^n$ e $F : \Omega \to \mathbb{R}^m$ una funzione di classe C^1 su Ω. Sia $\overline{B}_R(X_0)$ una palla chiusa inlcusa in Ω. Allora,*

$$|F(X) - F(Y)| \le L|X - Y| \quad \textit{per ogni} \quad X, Y \in \overline{B}_R(X_0)\,,$$

dove

$$L = \sup_{X\in\overline{B}_R(X_0)} \|DF(X)\| \qquad e \qquad \|DF(X)\| := \Big(\sum_{i=1}^n\sum_{j=1}^m \big(\partial_i F_j(X)\big)^2\Big)^{1/2}.$$

Dimostrazione. Siano X, Y due punti in $\overline{B}_R(X_0)$. Allora, per ogni $t \in [0, 1]$, abbiamo che

$$tX + (1-t)Y \in \overline{B}_R(X_0).$$

Consideriamo la funzione

$$f : [0,1] \to \mathbb{R}^m\,, \quad f(t) := F\big(tX + (1-t)Y\big).$$

Allora $f(0) = F(X)$, $f(1) = F(Y)$ e

$$f'(t) = DF\big(tX + (1-t)Y\big)[X - Y].$$

Di conseguenza,

$$F(Y) - F(X) = \int_0^1 DF\big(tX + (1-t)Y\big)[X - Y]\,dt.$$

Quindi, possiamo stimare

$$\begin{aligned}
|F(Y) - F(X)| &= \left|\int_0^1 DF\big(tX + (1-t)Y\big)[X - Y]\,dt\right| \\
&\le \int_0^1 \big|DF\big(tX + (1-t)Y\big)[X - Y]\big|\,dt \\
&\le \int_0^1 |X - Y|\,\big\|\nabla F\big(tX + (1-t)Y\big)\big\|\,dt \le L|X - Y|,
\end{aligned}$$

dove L è il massimo di $\|\nabla F\|$ in $\overline{B}_R(X_0)$ e dove abbiamo usato che per ogni matrice A, $n \times m$ a coefficienti reali, e per ogni vettore $X \in \mathbb{R}^n$ vale la disuguaglianza

$$|AX| \le |X|\|A\|.$$

□

Teorema 3. *Siano Ω un insieme chiuso in $\mathbb{R}^d$ ed $F : \Omega \to \mathbb{R}^d$ una funzione tale che:*

- *per ogni $X \in \Omega$, $F(X) \in \Omega$;*
- *esiste una costante*
$$0 < L < 1,$$
tale per cui F sia L-Lipschitz su Ω, ovvero
$$|F(X) - F(Y)| \leq L|X - Y| \quad \textit{per ogni} \quad X, Y \in \Omega\,.$$

Allora, esiste un unico punto $Y \in \Omega$ tale che
$$F(Y) = Y.$$

Dimostrazione.

Unicità. Dimostriamo prima l'unicita del punto fisso. Supponiamo che Y_1 e Y_2 siano due punti di Ω tali che
$$Y_1 = F(Y_1) \qquad \text{e} \qquad Y_2 = F(Y_2).$$
Siccome L è L-Lipschitz, abbiamo che
$$|Y_2 - Y_1| = |F(Y_2) - F(Y_1)| \leq L|Y_2 - Y_1|,$$
e quindi
$$(1 - L)|Y_2 - Y_1| \leq 0.$$
Siccome $L < 1$, otteniamo che $Y_2 = Y_1$, ovvero se un punto fisso esiste, allora è unico.

Esistenza. Scegliamo un qualsiasi punto $X_0 \in \Omega$ e definiamo per ricorrenza la successione
$$X_1 = F(X_0)\ ; \qquad X_{n+1} := F(X_n) \quad \text{per ogni} \quad n \geq 1.$$
Osserviamo che per ogni $k \geq 1$, abbiamo
$$|X_k - X_{k-1}| \leq L^{k-1}|X_1 - X_0|$$
Dati $N \geq n \geq 0$, osserviamo che
$$\begin{aligned} |X_N - X_n| = \left|\sum_{j=n+1}^{N} (X_j - X_{j-1})\right| &\leq \sum_{j=n+1}^{N} |X_j - X_{j-1}| \\ &\leq |X_1 - X_0| \sum_{j=n+1}^{N} L^{j-1} \\ &\leq |X_1 - X_0| \sum_{j=n+1}^{+\infty} L^{j-1} \\ &= L^n|X_1 - X_0| \sum_{i=0}^{+\infty} L^i = \frac{|X_1 - X_0|}{1 - L} L^n. \end{aligned}$$
Di conseguenza, $(X_n)_{n\geq 1}$ è una successione di Cauchy in $\mathbb{R}^n$ e quindi ammette un limite in $\mathbb{R}^n$
$$X_\infty = \lim_{n\to\infty} X_n.$$
Inoltre, siccome Ω è chiuso, abbiamo che
$$X_\infty \in \Omega.$$
Usando la contuinuità di F, abbiamo che
$$X_\infty = \lim_{n\to\infty} X_n = \lim_{n\to\infty} X_{n+1} = \lim_{n\to\infty} F(X_n) = F(X_\infty),$$
e quindi X_∞ è un punto fisso per la mappa $F : \Omega \to \Omega$. $\square$

Teorema della funzione inversa

Teorema 4 (Teorema della funzione inversa)**.** *Siano* Ω *un aperto in* $\mathbb{R}^d$ *e* $F:\Omega\to\mathbb{R}^d$ *una funzione di classe* C^1 *su* Ω*. Sia* $X_0\in\Omega$ *un punto tale che*

$$\det DF(X_0)\neq 0.$$

Allora, esistono una palla $B_R(X_0)$ *contenuta in* Ω *ed un raggio* $r>0$*, tali che la funzione*

$$F: B_R(X_0)\cap F^{-1}\big(B_r(Y_0)\big)\to B_r(Y_0)\quad \textit{dove}\quad Y_0:=F(X_0),$$

è un diffeomorfismo di classe C_1*.*

Dimostrazione. Sceglieremo i raggi $R>0$ e $r>0$ nel corso della dimostrazione.
Dato un $Y\in B_r(Y_0)$ mostreremo che esiste un unico punto $X\in B_R(X_0)$ tale che

$$F(X)=Y.$$

Osserviamo che, siccome la matrice $DF(X_0)$ è invertibile,

$$F(X)=Y\quad\Leftrightarrow\quad DF(X_0)^{-1}\Big[F(X)-Y\Big]=0,$$

e quindi

$$F(X)=Y\quad\Leftrightarrow\quad X=X-DF(X_0)^{-1}\Big[F(X)-Y\Big].$$

Ora, fissato $Y\in B_r(Y_0)$, definiamo ora la mappa

$$\Phi(X):=X-DF(X_0)^{-1}\Big[F(X)-Y\Big].$$

Mostreremo che possiamo scegliere $r>0$ e $R>0$ (indipendentemente da Y) in modo d'avere

(1) $\Phi(X)\in\overline{B}_R(X_0)$ per ogni $X\in\overline{B}_R(X_0)$;

(2) Φ è $\frac{1}{2}$-Lipschitz in $\overline{B}_R(X_0)$.

(e quindi in modo da poter applicare il teorema del punto fisso alla mappa $\Phi:\overline{B}_R(X_0)\to\overline{B}_R(X_0)$.)

Scelta di R e dimostrazione di (2). Scegliamo $R>0$ abbastanza piccolo tale che $\overline{B}_R(X_0)\subset\Omega$ e

$$\|D\Phi(X)\|=\|\mathrm{Id}-DF(X_0)^{-1}DF(X)\|\le\frac{1}{2}\qquad\text{per ogni}\qquad X\in\overline{B}_R(X_0). \tag{1}$$

Dalla Proposizione 2 abbiamo che Φ è $1/2$-Lipschitz su $\overline{B}_R(X_0)$.

Scelta di r e dimostrazione di (1). Osserviamo che per ogni $X\in\overline{B}_R(X_0)$ abbiamo

$$\begin{aligned}|\Phi(X)-X_0|&=\Big|X-DF(X_0)^{-1}\big[F(X)-Y\big]-X_0\Big|\\&=\Big|X-DF(X_0)^{-1}\big[F(X)-F(X_0)+F(X_0)-Y\big]-X_0\Big|\\&\le\Big|\Big(X-DF(X_0)^{-1}[F(X)]\Big)-\Big(X_0-DF(X_0)^{-1}[F(X_0)]\Big)\Big|+\Big|DF(X_0)^{-1}\big[Y_0-Y\big]\Big|.\end{aligned}$$

Usando di nuovo (1), abbiamo che

$$\Big|\Big(X-DF(X_0)^{-1}[F(X)]\Big)-\Big(X_0-DF(X_0)^{-1}[F(X_0)]\Big)\Big|\le\frac{1}{2}|X-X_0|\le\frac{1}{2}R.$$

D'altra parte, abbiamo

$$\Big|DF(X_0)^{-1}\big[Y_0-Y\big]\Big|\le\|DF(X_0)^{-1}\||Y-Y_0|\le r\|DF(X_0)^{-1}\|.$$

Scegliendo quindi,

$$r = \frac{R}{2\|DF(X_0)^{-1}\|},$$

otteniamo che

$$\begin{aligned}|\Phi(X) - X_0| &\le \Big|\Big(X - DF(X_0)^{-1}[F(X)]\Big) - \Big(X_0 - DF(X_0)^{-1}[F(X_0)]\Big)\Big| + \Big|DF(X_0)^{-1}\big[Y_0 - Y\big]\Big| \\ &< \frac{R}{2} + \frac{R}{2} = R,\end{aligned}$$

il che dimostra (1). Ora, per il teorema del punto fisso, per ogni $Y \in B_r(Y_0)$, esiste un unico punto fisso $X_Y \in \overline{B}_R(X_0)$ of Φ. Osserviamo inoltre, che siccome

$$|X_Y - X_0| = |\Phi(X_Y) - X_0| < R,$$

abbiamo che necessariamente $X_Y \in B_R(X_0)$. Definiamo quindi la funzione

$$G : B_r(Y_0) \to \mathbb{R}^d \qquad \text{come} \qquad G(Y) = X_Y.$$

e osserviamo che per costruzione la mappa

$$G : B_r(Y_0) \to F^{-1}\big(B_r(Y_0)\big) \cap B_R(X_0)$$

è bigettiva.

Continuità di G. Dimostreremo che G è differenziabile in ogni punto $Y \in B_r(Y_0)$. Senza perdita di generalità, possiamo supporre che $Y = Y_0$. Dato $\varepsilon > 0$ consideriamo

$$\delta = \frac{\varepsilon}{2\|DF(X_0)^{-1}\|}.$$

Abbiamo quindi che per ogni $Y \in B_\delta(Y_0)$, la mappa

$$\Phi(X) := X - DF(X_0)^{-1}\Big[F(X) - Y\Big]$$

è tale che:

(1) $\Phi(X) \in \overline{B}_\varepsilon(X_0)$ per ogni $X \in \overline{B}_\varepsilon(X_0)$;

(2) Φ è $\frac{1}{2}$-Lipschitz in $\overline{B}_\varepsilon(X_0)$.

Per il teorema del punto fisso abbiamo quindi che esiste un unico punto $X \in B_\varepsilon(X_0)$ tale che

$$F(X) = Y.$$

Ma allora, siccome X è unico anche in $B_R(X_0)$, abbiamo che

$$X = G(Y).$$

Abbiamo quindi dimostrato che

$$Y \in B_\delta(Y_0) \quad \Rightarrow \quad G(Y) \in B_\varepsilon(X_0).$$

Differenziabilità di G. Dimostreremo che G è differenziabile in ogni punto $Y \in B_r(Y_0)$ e che

$$DG(Y) = DF(X)^{-1} \qquad \text{dove} \qquad X = G(Y).$$

Senza perdita di generalità, possiamo supporre che $Y = Y_0$ e $X = X_0$. Per ogni

$$Y \in B_r(Y_0)$$

abbiamo l'identità

$$F(G(Y)) = Y.$$

Scrivendo

$$G(Y) = G(Y_0) + G(Y) - G(Y_0) = X_0 + G(Y) - G(Y_0),$$

e usando la differenziabilità di F:

$$F(X_0 + H) = F(X_0) + DF(X_0)[H] + \varepsilon(H) \qquad \text{dove} \qquad \lim_{H\to 0} \frac{\varepsilon(H)}{|H|} = 0,$$

abbiamo che

$$Y = F(G(Y)) = F\Big(X_0 + G(Y) - G(Y_0)\Big) = F(X_0) + DF(X_0)[G(Y) - G(Y_0)] + \varepsilon\big(G(Y) - G(Y_0)\big),$$

che possiamo scrivere anche come

$$Y - Y_0 = DF(X_0)[G(Y) - G(Y_0)] + \varepsilon\big(G(Y) - G(Y_0)\big), \tag{2}$$

$$DF(X_0)^{-1}[Y - Y_0] = G(Y) - G(Y_0) + DF(X_0)^{-1}\big[\varepsilon\big(G(Y) - G(Y_0)\big)\big],$$

Siccome G è continua e $\varepsilon(H) = o(|H|)$, questo implica che per Y abbastanza piccolo,

$$\begin{aligned}
|G(Y) - G(Y_0)| &\le \Big|DF(X_0)^{-1}[Y - Y_0]\Big| + \Big|DF(X_0)^{-1}\big[\varepsilon\big(G(Y) - G(Y_0)\big)\big]\Big| \\
&\le \|DF(X_0)^{-1}\|\,|Y - Y_0| + \|DF(X_0)^{-1}\|\,\Big|\varepsilon\big(G(Y) - G(Y_0)\big)\Big| \\
&\le \|DF(X_0)^{-1}\|\,|Y - Y_0| + \frac{1}{2}\big|G(Y) - G(Y_0)\big|,
\end{aligned}$$

e di conseguenza,

$$\frac{1}{2}|G(Y) - G(Y_0)| \le \|DF(X_0)^{-1}\|\,|Y - Y_0|,$$

o in altre parole

$$G(Y) - G(Y_0) = O(|Y - Y_0|).$$

Questo implica che

$$\varepsilon\big(G(Y) - G(Y_0)\big) = o(|Y - Y_0|).$$

Usando di nuovo (2), ottaniamo che

$$G(Y) - G(Y_0) = DF(X_0)^{-1}[Y - Y_0] + o(|Y - Y_0|),$$

il che implica che G è differenziabile in Y_0 e che $DG(Y_0) = DF(X_0)^{-1}$. □

Teorema della funzione implicita

Il seguente teorema è una versione più generale del teorema di Dini, nel quale useremo la notazione seguente. Dati $N > n$, possiamo scrivere ogni vettore

$$X = (x_1, \dots, x_n, x_{n+1}, \dots, x_N) \in \mathbb{R}^n$$

come

$$X = (X_1, X_2),$$

dove

$$X_1 = (x_1, \dots, x_n) \in \mathbb{R}^n \qquad \text{e} \qquad X_2 = (x_{n+1}, \dots, x_N) \in \mathbb{R}^{N-n}.$$

Teorema 5 (Teorema della funzione implicita)**.** *Siano $N > n$ due numeri naturali, Ω un aperto in $\mathbb{R}^N$,*

$$F = \begin{pmatrix} F_1 \\ \vdots \\ F_n \end{pmatrix} : \Omega \to \mathbb{R}^n$$

una funzione di classe C^1 su Ω. Supponiamo che la matrice

$$\begin{pmatrix} \partial_1 F_1 & \dots & \partial_n F_1 \\ \vdots & \ddots & \vdots \\ \partial_1 F_n & \dots & \partial_n F_n \end{pmatrix}$$

calcolata nel punto $\underline{X} = (\underline{X}_1, \underline{X}_2) \in \Omega$, sia invertibile. Allora esistono $r > 0$, $R > 0$ ed una funzione di $N - n$ variabili, definita sulla palla $B_r(\underline{X}_2) \subset \mathbb{R}^{N-n}$,

$$\eta : B_r(\underline{X}_2) \to \mathbb{R}^n,$$

tale che

(i) *per ogni $X_2 \in B_r(\underline{X}_2)$, $\eta(X_2) \in B_R(\underline{X}_1)$;*

(ii) *per ogni punto $X = (X_1, X_2) \in B_R(\underline{X}_1) \times B_r(\underline{X}_2)$ si ha che*

$$F(X_1, X_2) = F(\underline{X}_1, \underline{X}_2) \qquad \Leftrightarrow \qquad X_1 = \eta(X_2);$$

(iii) *$\eta : B_r(X_2) \to \mathbb{R}^n$ è una funzione differenziabile di classe C^1.*

Dimostrazione. Consideriamo la funzione

$$\Phi : \Omega \to \mathbb{R}^N \ , \qquad \Phi(X_1, X_2) = \big(F(X_1, X_2), X_2\big).$$

Allora, la matrice $D\Phi(\underline{X}_1, \underline{X}_2)$ è invertibile e possiamo applicare il teorema della funzione inversa, ottenendo una funzione di classe C^1 (definita su una palla $B_\rho(\underline{Y}_1, \underline{Y}_2) \subset \mathbb{R}^N$)

$$\Psi : B_\rho(\underline{Y}_1, \underline{Y}_2) \to \mathbb{R}^N$$

con $\Psi(\underline{Y}_1, \underline{Y}_2) = (\underline{X}_1, \underline{X}_2)$ e tale che la mappa

$$\Psi : B_\rho(\underline{Y}_1, \underline{Y}_2) \to \Psi\big(B_\rho(\underline{Y}_1, \underline{Y}_2)\big)$$

è l'inversa di

$$\Phi : \Psi\big(B_\rho(\underline{Y}_1, \underline{Y}_2)\big) \to B_\rho(\underline{Y}_1, \underline{Y}_2).$$

Ora, siccome Φ lascia invariata la seconda componente X_2, abbiamo che Ψ è della forma

$$\Psi(Y_1, Y_2) = \big(G(Y_1, Y_2), Y_2\big).$$

Si ha quindi che G è di classe C^1 su $B_\rho(\underline{Y}_1, \underline{Y}_2)$ e

$$F\big(G(Y_1, Y_2), Y_2\big) = Y_1 \quad \text{per ogni} \quad (Y_1, Y_2) \in B_\rho(\underline{Y}_1, \underline{Y}_2).$$

Quindi, basta definire η come

$$\eta(X_2) = G(\underline{Y}_2, X_2).$$

□