IL TEOREMA DI FILIPPOV - CESARI

Si tratta di una coppia di teoremi che forniscono l'esistenza di un controllo ottimale per una classe molto vasta di problemi che adesso descriviamo.

<u>Problema</u> Minimizzare il funzionale

$$J(y_0,u) = \int_{t_0}^{t_1} L(t, y(t), u(t))\, dt + \phi(e)$$

nella classe delle funzioni $u: [t_0,t_1] \to U$ sommabili in un certo intervallo $[t_0,t_1]$, dipendente da u, sotto le condizioni

$$\begin{cases} y'(t) = f(t, y(t), u(t)), & t \in [t_0,t_1], \\ y(t_0) = y_0, & e = (t_0, t_1, y(t_0), y(t_1)) \in S. \end{cases}$$

L'intervallo $[t_0,t_1]$ si considera <u>parte</u> della definizione di u (è il suo dominio).
Dunque: sono dati gli insiemi $U \subseteq \mathbb{R}^m$ (valori dei controlli) e $S \subseteq \mathbb{R}^{2n+2}$ (stati e tempi iniziali e finali), nonché le funzioni f, L, ϕ; il controllo è (y_0, u); lo stato è definito dall'equazione differenziale ma è anche vincolato dalla condizione $e \in S$, e il controllo è vincolato dalle condizioni $u(t) \in U$ $\forall t \in [t_0,t_1]$ ed $e \in S$.

Le ipotesi di base sono le seguenti:

(a) $f: \mathbb{R}^{n+1} \times U \to \mathbb{R}^n$, $L: \mathbb{R}^{n+1} \times U \to \mathbb{R}$, $\phi: S \to \mathbb{R}$ sono funzioni continue;

(b) $|f(t,y,u)| \le c_1 [1 + |y| + |u|]$ $\forall (t,y,u) \in \mathbb{R}^{n+1} \times U$;

(c) $|f(t,y,u) - f(t,\bar{y},u)| \le c_2 [1 + |u|] \, |y - \bar{y}|$ $\forall (t,y,u), (t,\bar{y},u) \in \mathbb{R}^{n+1} \times U$.

Osserviamo subito che se U è un insieme limitato, in (b) e (c) si può omettere a secondo membro la dipendenza da $|u|$.

Lo stato $y(\cdot)$ è univocamente determinato, in virtù del teorema 25 di McShane e e dell'ipotesi (c). Esso si scrive implicitamente in forma integrale:

$$(*) \quad y(t) = y_0 + \int_{t_0}^{t} f(s, y(s), u(s))\, ds, \quad t \in [t_0, t_1].$$

Di conseguenza, è valida la stima

$$(d) \quad |y(t) - y_0| \leq C \int_{t_0}^{t} [1 + |y_0| + |u(s)|]\, ds \quad \forall t \in [t_0, t_1].$$

Infatti, per (b),

$$0 \leq |y(t) - y_0| = \left| \int_{t_0}^{t} f(s, y(s), u(s))\, ds \right| \leq$$

$$\leq C_1 \int_{t_0}^{t} [1 + |y(s) - y_0| + |y_0| + |u(s)|]\, ds =$$

$$= C_1 \int_{t_0}^{t} |y(s) - y_0|\, ds + C_1 \int_{t_0}^{t} [1 + |y_0| + |u(s)|]\, ds;$$

utilizzando il classico lemma di Grönwall, si ricava la tesi con

$$C = e^{C_1(t_1 - t_0)} - 1 + C_1. \quad \square$$

Introduciamo i seguenti insiemi:

$$\mathcal{J} = \{ (y_0, u) : y_0 \in \mathbb{R}^n,\ u \in L^1(t_0, t_1, U) \text{ per un opportuno } [t_0, t_1] \subset \mathbb{R},$$
$$y \text{ verifica } (*),\ e = (t_0, t_1, y_0, y(t_1)) \in S \}$$

($\mathcal{J}$ è la classe dei controlli ammissibili),

$$F(t, y) = \{ f(t, y, u) : u \in U \}, \quad (t, y) \in \mathbb{R}^{n+1},$$

$$\tilde{F}(t, y) = \{ (z, w) \in \mathbb{R}^{n+1} : z = f(t, y, u),\ w \geq L(t, y, u),\ u \in U \}, \quad (t, y) \in \mathbb{R}^{n+1}.$$

Valgono allora i seguenti Teoremi:

<u>Teorema 27</u> (di Filippov). Nelle ipotesi (a), (b), (c), supponiamo $L \equiv 0$ e inoltre:

(i) $\mathcal{J} \neq \phi$,

(ii) U compatto,

(iii) S compatto,

(iv) $F(t,y)$ convesso (e, ovviamente, chiuso) per ogni $(t,y) \in \mathbb{R}^{n+1}$.

Allora esiste $(y_0^*, u^*) \in \mathcal{J}$ tale che $J(y_0^*, u^*) = \min_{\mathcal{J}} J$.

<u>Teorema 28</u> (di Cesàri). Nelle ipotesi (a), (b), (c), supponiamo inoltre:

(i) $\mathcal{J} \neq \phi$,

(ii) U chiuso,

(iii) S compatto,

(iv) $\tilde{F}(t,y)$ convesso per ogni $(t,y) \in \mathbb{R}^{n+1}$,

(v) esiste $g : U \to \mathbb{R}$ continua, tale che

$$L(t,y,u) \geq g(u) \quad \forall (t,y,u) \in \mathbb{R}^{n+1} \times U, \qquad \lim_{\substack{|u| \to +\infty \\ u \in U}} \frac{g(u)}{|u|} = +\infty .$$

Allora esiste $(y_0^*, u^*) \in \mathcal{J}$ tale che $J(y_0^*, u^*) = \min_{\mathcal{J}} J$.

<u>Osservazioni</u> (1) Nei due teoremi l'ipotesi (iv) può essere sostituita da

(iv)' U convesso, $f(t,y,u) = \alpha(t,y) + \beta(t,y)u$, $u \mapsto L(t,y,u)$ convessa;

perché da (iv)' segue che $F(t,y)$ e $\tilde{F}(t,y)$ sono convessi.

(2) Nei due teoremi l'ipotesi (iii) può essere sostituita da

(iii)' $\exists\, \mu_1 > \inf_{\mathcal{J}} J$, $\exists\, S' \subseteq S$, S' compatto, tale che

$$e \in S,\ \exists\, u : J(y_0 u) \leq \mu_1 \implies e \in S'.$$

Infatti, ai fini della minimizzazione di J si può riformulare il problema prendendo S' in luogo di S.

(3) Nel teorema 28, se U è compatto, la condizione (V) si riduce ad una limitazione inferiore per L della forma

$$L(t,y,u) \geq K \quad \forall (t,y,u) \in \mathbb{R}^{n+1} \times U;$$

infatti g non ha vincoli all'infinito ma è continua sul compatto U, dunque è inferiormente limitata.

In particolare, se nel teorema 28 assumiamo U compatto e $L \equiv 0$, sono automaticamente verificate le ipotesi del teorema 27. Perciò, se dimostriamo il teorema 28 sarà provato anche il teorema 27.

(4) Nel teorema 28, in conseguenza di (V), esiste $b \geq 0$ tale che

$$g(u) \geq -b \quad \forall u \in U.$$

Infatti $g(u) \geq M|u|$ se $|u| > R_0$, mentre sul compatto $U \cap B(0,R_0)$ la g è continua e dunque inferiormente limitata.

(5) Poiché S è compatto, esiste $[T_0,T_1]$ tale che $[t_0,t_1] \subseteq [T_0,T_1]$ per ogni $(t_0,t_1,y_0,y_1) \in S$. Fissato un elemento $u_0 \in U$, possiamo estendere ogni controllo $u : [t_0,t_1] \to U$ a tutto l'intervallo $[T_0,T_1]$

ponendo $u(t) = u_0$ $\forall t \in [T_0, T_1] \setminus [t_0, t_1]$. Di conseguenza, anche lo stato y è definito da ⊛ in tutto $[T_0, T_1]$, e possiamo estendere la relazione (d) a $[T_0, T_1]$:

$$(d) \quad |y(t) - y_0| \leq c \left| \int_{t_0}^{t} [1 + |y_0| + |u(s)|] \, ds \right| \quad \forall t \in [T_0, T_1].$$

(6) Se U è compatto, esistono $M_1, M_2 \geq 0$ tali che

(α) $|y(t)| \leq M_1$ $\forall t \in [T_0, T_1]$, $\forall (y_0, u) \in \mathcal{F}$

(β) $|y(t) - y(s)| \leq M_2 |t-s|$ $\forall t, s \in [T_0, T_1]$, $\forall (y_0, u) \in \mathcal{F}$.

Infatti, essendo S compatto, $|y_0| \leq M_0$ $\forall (y_0, u) \in \mathcal{F}$; per la compattezza di U, si ha $|u(t)| \leq B$ $\forall (y_0, u) \in \mathcal{F}$. Da (d) e (b),

$$|y(t)| \leq |y(t) - y_0| + |y_0| \leq c_1 (T_1 - T_0) [1 + M_0 + B] + M_0 =: M_1,$$

$$|y(t) - y(s)| = \left| \int_s^t f(\tau, y(\tau), u(\tau)) \, d\tau \right| \leq \left| \int_s^t c_1 [1 + M_1 + B] \, d\tau \right| \leq$$

$$\leq c_1 [1 + M_1 + B] \, |t-s| =: M_2 |t-s|. \quad \square$$

Prima di dimostrare il teorema di Cesari vediamo qualche esempio e controesempio.

<u>Esempio</u> Come caso particolare del teorema 27, fissiamo t_0 e y_0, imponendo solamente la condizione finale $(t_1, y(t_1)) \in S_1$, con $S_1 \subseteq \mathbb{R}^{n+1}$. Dunque $S = \{ e = (t_0, t_1, y_0, y_1) : (t_1, y_1) \in S_1 \}$. Si ha allora $(y_0, u) \in \mathcal{F} \iff \exists\, t_1 \geq t_0$ tale che $(t_1, y(t_1)) \in S_1$. Poniamo $J(u) = t_1 - t_0$; minimizzare J equivale a portare $(t, y(t))$ in S_1 in tempo minimo.

Corollario 30 (problema del tempo minimo) Se valgono le ipotesi (a), (b), (c) e (i), (ii), (iv) del teorema 28, e se inoltre S_1 è chiuso, allora esiste $u^*:[t_0,t^*]\to U$ tale che $J(u^*)=\min J$.

dim. Si ha $\phi(e)=t_1-t_0$. Sia $u:[t_0,\tau]$ un controllo ammissibile, che esiste per (i); sia $\mu_1=\tau-t_0=J(u)$; se $\mu_1=\inf J$, la tesi è provata. Supponiamo dunque $\mu_1>\inf J$. Dall'osservazione (3) si ha

$$|y(t)|\le M_1 \quad \forall t\in[t_0,\tau].$$

Poniamo

$$S'=\{(t_0,t_1,y_0,y_1): t_1\in[t_0,\tau],\ |y_1|\le M_1,\ (t_1,y_1)\in S_1\};$$

allora S' è compatto, è incluso in S, e verifica la condizione dell'osservazione (2): infatti, se $e\in S$ ed esiste $\bar u$ tale che $J(y_0,\bar u)\le\mu_1$, ciò significa che $\bar u:[t_0,\theta]\to U$ è tale che $(\theta,\bar y(\theta))\in S_1$, con

$$J(y_0,\bar u)=\theta-t_0\le\mu_1=\tau-t_0.$$

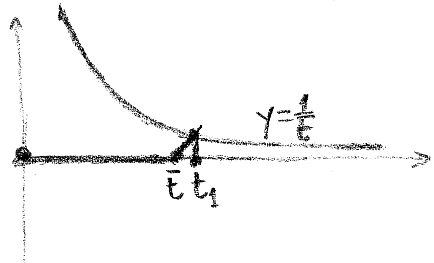
Dunque $\theta\in[t_0,\tau]$, $|y(\theta)|\le M_1$ e $(\theta,y(\theta))\in S_1$, ossia $e\in S'$. Pertanto il teorema 27 implica l'esistenza di un controllo ottimale $u^*:[t_0,t^*]\to U$, tale che $(t^*,y^*(t^*))\in S'$ e $J(t^*)=\min J$. □

Controesempio 31 Il teorema 27 è in generale falso quando S non è compatto. Siano $n=m=1$, $U=[-1,1]$, $t_0=0$, $y_0=0$; consideriamo

$$f(t,y,u)=u,\qquad S=\{(0,0,t,y): y=\tfrac{1}{t},\ t>0\},$$

$$J(u)=y(t_1),\quad t_1=\inf\{t>0: y(t)=\tfrac{1}{t}\}.$$

L'insieme S non è compatto. Proviamo che $\inf J=0$: definendo



$\bar u(t) = \begin{cases} 0 & \text{in } [0, \bar t] \\ 1 & \text{in } [\bar t, \infty[\end{cases}$, si ha $y(t) = \begin{cases} 0 & \text{in } [0, \bar t] \\ t - \bar t & \text{in } [\bar t, \infty[\end{cases}$.

e quindi si ha $\frac{1}{t} = t - \bar t \iff t^2 - \bar t t - 1 = 0 \iff t = \frac{\bar t + \sqrt{\bar t^2 + 4}}{2}$.

Perciò

$$\bar J(\bar u) = y(t) = \frac{1}{t} = \frac{2}{\bar t + \sqrt{\bar t^2 + 4}} < \varepsilon \quad \text{se } \bar t \geq T_\varepsilon \text{ opportuno.}$$

Poi è ovvio che $J(u) = y(t_1) > 0$ visto che $y(t_1) = \frac{1}{t_1} > 0$.

Si noti che S non verifica la condizione dell'osservazione (2): infatti per ogni $\varepsilon > 0$ la successione di controlli $\{u_n\}$, ove $u_n = I_{[n, \infty[}$, è tale che $J(y_n) < \varepsilon$ definitivamente, ma $e_n = (0, t_n, 0, y_n(t_n))$ non appartiene ad alcun compatto $\subseteq S$, visto che $t_n > n \to \infty$.

<u>Controesempio 32</u> Il teorema 27 è in generale falso quando $F(t,y)$ non è convesso per qualche (t,y). Siano $n = 2$, $m = 1$, $U = [-1,1]$, $t_0 = 0$, $\underline{y}_0 = (-1, 0)$, $C = \{(y_1, y_2) : y_1^2 + y_2^2 = a^2\}$, con $0 < a < 1$, e

$$f(t, y, u) = \begin{pmatrix} -y_2^2 + u^2 \\ u \end{pmatrix}.$$

Siano $S = \{(0, t_1, \underline{y}_0, \underline{y}_1) : \underline{y}_1 \in C\}$, e $J(u) = t_1 = \inf\{t > 0 : \underline{y}(t) \in C\}$.

È ancora un problema di tempo minimo; S è chiuso ma si può verificare che esiste S' compatto, incluso in S, dotato della proprietà dell'osservazione (2).

Tuttavia,

$$F(t,\underline{y}) = \{ (-y_2^2+u^2, u) : u \in [-1,1] \}$$

non è convesso, essendo un arco di parabola. Il sistema, naturalmente, è

$$\begin{cases} y_1'(t) = -y_2(t)^2 + u(t)^2 \\ y_2'(t) = u(t), \end{cases} \qquad t \in [t_0, t_1],$$

Proveremo che:

(I) $t_1 \geq 1-a$ per ogni controllo u,

(II) $\inf J = 1-a$,

(III) $t_1 > 1-a$ per ogni controllo u,

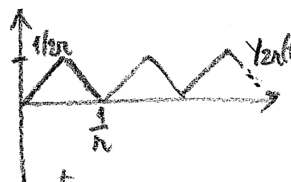
e dunque J non ha minimo.

Proviamo (I): si ha $y_1(t) \leq -1 + \int_0^t u(s)^2 ds \leq -1+t \leq -a \quad \forall t \leq 1-a$; pertanto deve essere $t_1 \geq 1-a$.

Proviamo (II): per $n \in \mathbb{N}^+$ sia $u_n(t) = \begin{cases} 1 & \text{se } t \in \left[\frac{2i-2}{2n}, \frac{2i-1}{2n}\right[\\ -1 & \text{se } t \in \left[\frac{2i-1}{2n}, \frac{2i}{2n}\right[, \end{cases} \quad i \in \mathbb{N}^+.$

Detto y_n lo stato corrispondente, si ha

$y_{2n}(t) = \int_0^t u(s) ds$, il grafico è:



mentre

$$y_{1n}(t) = -1 + \int_0^t [-y_{2n}(s)^2 + u(s)^2] ds = -1+t-\int_0^t y_{2n}(s)^2 ds.$$

Poniamo $t_{1n} = J(u_n)$ e proviamo che $\lim_{n\to\infty} t_{1n} = 1-a$.

Sappiamo da (I) che $1-a \leq t_{1n}$. Si ha

$$0 \leq y_{2n}(t) \leq \frac{1}{2n} \qquad \forall t \geq 0,$$

da cui

$$Y_{1n}(t) = -1+t - \int_0^t Y_{2n}(s)^2\,ds \Longrightarrow -1+t \quad \text{per } n\to\infty;$$

e inoltre, scelto $t=1$, e fissato $\varepsilon \in]0,a[$, si ha

$$\begin{cases} Y_{1n}(1) > -\varepsilon > -a \\ |Y_{2n}(1)| \le \dfrac{1}{2n}, \end{cases} \qquad \text{per } n \text{ grande},$$

il che implica $(Y_{1n}(1), Y_{2n}(1))$ interno a C per n grande. Dunque

$$1-a \le t_{1n} < 1 \quad \text{per } n \text{ grande}.$$

D'altronde

$$Y_{1n}(t_{1n}) = -1 + t_{1n} - \int_0^{t_{1n}} Y_{2n}(s)^2\,ds = -1 + t_{1n}\left(1 + O\left(\tfrac{1}{n^2}\right)\right);$$

ed essendo $Y_{1n}(t_{1n})^2 + Y_{2n}(t_{1n})^2 = a^2$, si ricava

$$\lim_{n\to\infty} Y_{1n}(t_{1n})^2 = a^2$$

ma $Y_{1n}(t_{1n}) \le -1 + t_{1n} \le 0$, e dunque $\lim\limits_{n\to\infty} Y_{1n}(t_{1n}) = -a$.

Perciò $$\lim_{n\to\infty} t_{1n} = \lim_{n\to\infty} \frac{1+Y_{1n}(t_{1n})}{1+O\left(\frac{1}{n^2}\right)} = 1-a.$$

Ciò prova che $\inf J = 1-a$.

Proviamo (III): sia u un controllo ammissibile

(A) Se $u \equiv 1$ q.o. in $[0,1-a]$, allora per lo stato y corrispondente si ha

$$\begin{cases} Y_1(t) = -1 + \int_0^t (-s^2+1)\,ds = -1 + t - \dfrac{t^3}{3} \\ Y_2(t) = t, \end{cases} \qquad t\in[0,1-a]$$

e se fosse $a^2 = Y_1(1-a)^2 + Y_2(1-a)^2$, ricaveremmo

$$a^2 = \left(-1 + (1-a) - \frac{(1-a)^3}{3}\right)^2 + (1-a)^2,$$

e dopo qualche calcolo

$$0 = \frac{(1-a)^6}{9} + (1-a)^2 - \frac{2}{3}a(1-a)^3,$$

il che è impossibile se $0<a<1$ (la funzione a secondo membro è nulla per $a=1$ e ha derivata negativa in $[a,1]$).

(B) Se esiste $E \subset [0,1-a]$, di misura positiva, tale che $u<1$ q.o. in E, allora

$$y_1(t) \le -1 + \int_0^t u(s)^2\,ds < -1+t \le -a \qquad \forall t \in [0,1-a]$$

dunque è impossibile che $(y_1(t), y_2(t)) \in C$ per alcun $t \le 1-a$. □

<u>Controesempio 33</u> Il teorema 28 di Cesari è in generale falso se manca l'ipotesi (V). Siano $n=m=1$, $t_0=0$, $y_0=1$, $t_1=1$, $y_1=0$ (dunque $S=\{(0,1,1,0)\}$). Poniamo $f(t,x,u)=u$, $L(t,x,u)=tu^2$. Stiamo cercando, fra i controlli che portano lo stato da $y_0=1$ a $y_1=0$ in tempo unitario, quello che minimizza

$$J(u) = \int_0^1 s\,u(s)^2\,ds.$$

Proviamo che $\inf J = 0$ ma che $J(u)>0$ per ogni $u \in \mathcal{J}$.
Notiamo che $u\equiv 0$ q.o. non è un controllo ammissibile: lo stato y corrispondente sarebbe $y(t)\equiv 1$). Quindi $J(u)>0$ per ogni u ammissibile.
Per $n\in\mathbb{N}^+$ sia $u_n(t) = -\frac{1}{n}t^{\frac{1}{n}-1}$; allora $y_n(t) = 1-t^{1/n}$, e

$$J(u_n) = \int_0^1 \frac{s}{n^2} s^{\frac{2}{n}-2}\,ds = \frac{1}{n^2}\left[\frac{s^{\frac{2}{n}}}{\frac{2}{n}}\right]_0^1 = \frac{1}{2n} \longrightarrow 0;$$

perciò $\inf J = 0$, e il minimo non esiste. □