Ottimalità

Ci occuperemo adesso dei controlli di un sistema dinamico che sono ottimali rispetto a un determinato criterio. Sia dunque

$$\begin{cases} y'(t) = f(y(t), u(t)), & t \in [0,T] \\ y(0) = x; \end{cases}$$

cerchiamo un controllo $u(\cdot)$ che minimizzi il funzionale costo

$$J(x,u) = \int_0^T g(y(t), u(t))\,dt + G(y(T)),$$

ove f, g, G son funzioni continue sul loro dominio. Questo problema di controllo ottimale è in orizzonte finito, ma si può anche analizzare il problema in orizzonte infinito, vale a dire il caso di un sistema

$$\begin{cases} y'(t) = f(y(t), u(t)), & t \geq 0, \\ y(0) = x \end{cases}$$

e di un costo

$$J(x,u) = \int_0^\infty g(y(t), u(t))\,dt.$$

In questo caso occorrerà fare qualche ulteriore ipotesi su f e g affinché esista la soluzione $y(t)$ in $[0,\infty[$ ed esista almeno un controllo che renda finito il funzionale.

Per questo tipo di problemi forniremo condizioni <u>necessarie</u> per l'esistenza di controlli ottimali, e anche condizioni <u>sufficienti</u>, almeno in dimensione finita.

Il teorema di Pontryagin

Questo risultato vale in dimensione finita o infinita, e fornisce una condizione necessaria che generalizza il teorema dei moltiplicatori di Lagrange.

Occorre però qualche considerazione preliminare.

Sia X uno spazio di Banach. Sia $F \subseteq X$, sia $\bar{x} \in F$.

Definizione Il <u>cono normale</u> a F in $\bar{x}$ è il sottoinsieme del duale X^* definito da

$$N(F,\bar{x}) = \{\varphi \in X^* : \forall \varepsilon > 0 \ \exists \delta > 0 : \ \varphi(x-\bar{x}) \le \varepsilon \|x-\bar{x}\| \quad \forall x \in F \cap B(\bar{x},\delta)\}$$

Se $F = \{\bar{x}\}$, si ottiene $N(F,\{\bar{x}\}) = X^*$; se $\bar{x}$ è punto d'accumulazione per F, possiamo scrivere

$$N(F,\bar{x}) = \left\{\varphi \in X^* : \limsup_{x \to \bar{x},\, x \in F \setminus \{\bar{x}\}} \frac{\varphi(x-\bar{x})}{\|x-\bar{x}\|} \le 0 \right\}.$$

Definizione Il <u>cono tangente</u> a F in $\bar{x} \in F$ è il sottoinsieme di X definito da

$$T(F,\bar{x}) = \{u \in X : \exists \{u_n\} \subseteq X, \exists \{t_n\} \subseteq \mathbb{R} : \ u_n \to u, \ t_n \to 0^+, \ \bar{x} + t_n u_n \in F\} =$$

$$= \{u \in X : \forall \delta > 0 \ \exists t \in]0,\delta[, \ \exists v \in B(u,\delta) : \ \bar{x} + t v \in F\}.$$

Proposizione 15 Risulta per $\bar{x}\in\bar{F}$, $F\subseteq X$:

$$N(F,\bar{x})\subseteq\{\varphi\in X^*:\ \varphi(u)\le 0\ \forall u\in T(F,\bar{x})\}.$$

dim. Sia $\varphi\in N(F,\bar{x})$ e sia $u\in T(F,\bar{x})$. Scelte $\{u_n\}$ e $\{t_n\}$ secondo la definizione, si ha

$$\limsup_{n\to\infty}\frac{\varphi(t_n u_n)}{t_n\|u_n\|}=\limsup_{n\to\infty}\frac{\varphi(u_n)}{\|u_n\|}\le 0.$$

Dunque, per ogni $\varepsilon>0$ esiste $\bar{n}\in\mathbb{N}$ tale che

$$\varphi(u_n)\le\varepsilon\|u_n\|\quad\forall n\ge\bar{n},$$

e per $n\to\infty$

$$\varphi(u)\le\varepsilon\|u\|\quad\forall u\in T(F,\bar{x}).$$

Siccome ε è arbitrario, si deduce

$$\varphi(u)\le 0\quad\forall u\in T(F,\bar{x}).\ \square$$

Proposizione 16 (Fermat) Sia $f:X\to\mathbb{R}$ F-differenziabile in $\bar{x}\in F$. Se $f(\bar{x})=\min_F f$, allora $-f'(\bar{x})\in N(F,\bar{x})$.

dim. Dalla formula di Taylor centrata in $\bar{x}$ segue per $x\in F$

$$f(x)=f(\bar{x})+[f'(\bar{x})](x-\bar{x})+o(\|x-\bar{x}\|)\quad\text{per } x\to\bar{x},$$

quindi

$$-[f'(\bar{x})](x-\bar{x})=f(\bar{x})-f(x)+o(\|x-\bar{x}\|)\le o(\|x-\bar{x}\|),$$

da cui

$$\limsup_{x\to\bar{x},\, x\in F\setminus\{\bar{x}\}} \frac{-[f'(\bar{x})](x-\bar{x})}{\|x-\bar{x}\|} \le 0. \quad \square$$

<u>Osservazione</u> La nozione di cono tangente e di cono normale è <u>locale</u>; in altre parole, se $\bar{x}\in \bar{F}\cap\bar{G}$ e V è un intorno di $\bar{x}$ tale che $F\cap V = G\cap V$, allora $N(F,\bar{x}) = N(G,\bar{x})$. La verifica è immediata per definizione.

<u>Proposizione 17</u> Siano X,Y spazi di Banach. Se $\bar{x}\in\bar{F}$, $F\subseteq X$, e $\bar{y}\in\bar{G}$, $G\subseteq Y$, allora

$$N(F\times G,(\bar{x},\bar{y})) = N(F,\bar{x})\times N(G,\bar{y}).$$

<u>dim.</u> Sia $\psi\in N(F\times G,(\bar{x},\bar{y}))$. Allora sarà $\psi\in(X\times Y)^* = X^*\times Y^*$, cioè $\psi=(\varphi,\eta)$, e

$$\limsup_{\substack{(x,y)\to(\bar{x},\bar{y})\\ (x,y)\in F\times G\setminus\{(\bar{x},\bar{y})\}}} \frac{\varphi(x-\bar{x})+\eta(y-\bar{y})}{\sqrt{\|x-\bar{x}\|^2+\|y-\bar{y}\|^2}} \le 0.$$

Quindi

$$\limsup_{\substack{x\to\bar{x}\\ x\in F\setminus\{\bar{x}\}}} \frac{\varphi(x-\bar{x})}{\|x-\bar{x}\|} \le 0, \quad \limsup_{\substack{y\to\bar{y}\\ y\in G\setminus\{\bar{y}\}}} \frac{\eta(y-\bar{y})}{\|y-\bar{y}\|} \le 0$$

ossia $\psi\in N(F,\bar{x})\times N(G,\bar{y})$. Viceversa, siano $\varphi\in N(F,\bar{x})$ e $\eta\in N(G,\bar{y})$: posto $\psi=(\varphi,\eta)$ si ha

$$\limsup_{\substack{(x,y)\to(\bar{x},\bar{y})\\ (x,y)\in F\times G\setminus\{(\bar{x},\bar{y})\}}} \frac{\varphi(x-\bar{x})+\eta(y-\bar{y})}{\sqrt{\|x-\bar{x}\|^2+\|y-\bar{y}\|^2}} \le \limsup_{\substack{x\to\bar{x}\\ x\in F\setminus\{\bar{x}\}}} \frac{\varphi(x-\bar{x})}{\|x-\bar{x}\|} + \limsup_{\substack{y\to\bar{y}\\ y\in G\setminus\{\bar{y}\}}} \frac{\eta(y-\bar{y})}{\|y-\bar{y}\|} \le 0,$$

dunque $\psi = (\varphi, r) \in N(F \times G, (\bar{x}, \bar{y}))$. □

<u>Proposizione 18</u> Siano $F, G \subseteq X$ con $F \subseteq G$, e sia $\bar{x} \in F$. Allora $N(F, \bar{x}) \supseteq N(G, \bar{x})$. Se inoltre $F = \bigcup_{i=1}^{k} F_i$, e $\bar{x} \in \bigcap_{i=1}^{k} F_i$, allora $N(F, \bar{x}) = \bigcap_{i=1}^{k} N(F_i, \bar{x})$.

<u>dim</u> Il primo enunciato è evidente per definizione. Da ciò segue l'inclusione $\subseteq$ del secondo enunciato; l'altra segue per definizione. □

Dopo queste proprietà elementari, vogliamo vedere come cambia il cono normale per trasformazioni differenziabili. Qui interviene pesantemente la teoria delle mappe invertibili e quella, collegata, della risoluzione di equazioni non lineari.

<u>Teorema 19 (di Kantorovich)</u> Siano X, Y spazi di Banach, sia $W \subseteq X$ un aperto, sia $f : W \to Y$ un'applicazione continua. Sia poi $x_0 \in W$, sia $B(x_0, r) \subseteq W$ e sia infine $A : B(x_0, r) \longrightarrow \mathcal{L}(X, Y)$ tale che

(i) $\forall x \in B(x_0, r)$ $\exists\, B(x) : Y \to X$, inversa destra di $A(x)$, tale che $\|B(x)(u)\|_X \le \beta \|u\|_Y$ $\forall u \in Y$, con $\beta > 0$;

(ii) $\forall x, w \in B(x_0, r)$, $\|f(w) - f(x) - A(x)(w - x)\| \le \alpha \|w - x\|$, con $0 < \alpha < \frac{1}{\beta}$.

Allora se $\|f(x_0)\| < \frac{1}{\beta}(1 - \alpha\beta) r$, esiste $\bar{x} \in B(x_0, r)$ tale che $f(\bar{x}) = 0$; tale $\bar{x}$ viene selezionato col metodo di Newton: cioè, la successione

$$x_{n+1} = x_n - B(x_n)(f(x_n)), \quad n \in \mathbb{N}$$

è ben definita, converge a $\bar{x}$ e verifica

$$\|x_n - \bar{x}\| \le r\,(\alpha\beta)^n \quad \forall n \in \mathbb{N},$$

$$\|x_0 - \bar{x}\| \le \frac{\beta}{1-\alpha\beta}\,\|f(x_0)\| < r.$$

<u>dim</u> Proviamo per induzione che

$$\{x_n\} \subseteq B(x_0, r), \quad \|x_{n+1} - x_n\| \le \beta\,(\alpha\beta)^n\,\|f(x_0)\|, \quad \|f(x_n)\| \le (\alpha\beta)^n\,\|f(x_0)\|.$$

Per $n=0$ è tutto ovvio. Se le tre condizioni valgono per $0 \le n < N$, allora

$$\|x_N - x_0\| \le \sum_{n=0}^{N-1} \|x_{n+1} - x_n\| \le \beta \sum_{n=0}^{N-1} \|f(x_n)\| \le \beta \sum_{n=0}^{N-1} (\alpha\beta)^n\,\|f(x_0)\| \le$$

$$\le \beta\,\frac{1}{1-\alpha\beta}\,\|f(x_0)\| < r\,;$$

inoltre, essendo

$$f(x_{N-1}) + A(x_{N-1})\,(x_N - x_{N-1}) = 0,$$

$$\|f(x_N)\| = \|f(x_N) - f(x_{N-1}) - A(x_{N-1})\,(x_N - x_{N-1})\| \le \alpha\,\|x_N - x_{N-1}\| \le$$

$$\le \alpha\beta\,\|f(x_{N-1})\| \le (\alpha\beta)^N\,\|f(x_0)\|,$$

e infine

$$\|x_{N+1} - x_N\| \le \beta\,\|f(x_N)\| \le \beta\,(\alpha\beta)^N\,\|f(x_0)\|.$$

Le tre condizioni sono dimostrate. Essendo $\alpha\beta < 1$, la successione $\{x_n\}$ è di Cauchy in X e quindi ha limite $\bar{x}$, si ha

$$\|\bar{x} - x_0\| = \lim_{n\to\infty} \|x_n - x_0\| \le \beta\,\frac{1}{1-\alpha\beta}\,\|f(x_0)\| < r,$$

e per continuità

$$f(x_n) \to 0 = f(\bar{x}).$$

Infine

$$\|x_n - \bar{x}\| = \lim_{p\to\infty} \|x_n - x_p\| \le \lim_{p\to\infty} \sum_{k=n}^{p-1} \|x_{k+1} - x_k\| \le$$

$$\le \sum_{k=n}^{\infty} \|f(x_0)\| \, \beta (\alpha\beta)^k = \beta(\alpha\beta)^n \frac{1}{1-\alpha\beta} \|f(x_0)\| < r(\alpha\beta)^n . \quad \square$$

<u>Teorema 20 (di Lyusternik-Graves)</u> Siano X, Y spazi di Banach, sia $W \subseteq X$ un aperto, sia $g : W \to Y$ un'applicazione F-differenziabile in $B(\bar{x}, r_0) \subseteq W$, con $g'(x)$ continua in $\bar{x}$ e $g'(\bar{x})$ surgettiva. Allora esistono $\rho \in]0, \frac{r_0}{2}]$, $\sigma > 0$ e $c > 0$ tali che

$\forall w \in B(\bar{x}, \rho)$, $\forall y \in B(g(\bar{x}), \sigma)$ $\exists x \in B(\bar{x}, \frac{r_0}{2})$: $g(x) = y$, $\|x - w\| \le c \|y - g(w)\|$

<u>dim</u> Proviamo anzitutto che esiste $\beta > 0$ tale che, posto $A = g'(\bar{x})$,

$$\forall y \in Y \; \exists w \in X : \; Aw = y, \quad \|w\|_X \le \beta \|y\|_Y .$$

Se $B = \{x \in X : \|x\| < 1\}$, $S_\delta = \{y \in Y : \|y\| < \delta\}$, essendo A un'applicazione aperta, esiste $\delta > 0$ tale che $S_\delta \subseteq A(B)$. Se $y = 0$, allora $w = 0$ verifica la relazione sopra scritta; se $y \ne 0$, sia $t = \frac{\delta}{2\|y\|}$; allora $ty \in S_\delta$ e quindi esiste $z \in B$ tale che $Az = ty$. Dunque se $w = \frac{z}{t}$ si ha $w \in X$, $Aw = y$ e $\|w\| \le \frac{1}{t} = \frac{2}{\delta}\|y\|$. Ne segue che la relazione vale con $\beta = \frac{2}{\delta}$.

Si noti che la mappa $y \longmapsto w$ definisce un'inversa destra $h : Y \to X$ di $A = g'(\bar{x})$, tale che $\|h(y)\| \le \beta \|y\|$ per ogni $y \in Y$.

Ciò premesso, sia $\alpha \in]0, \frac{1}{\beta}[$ e sia $r \in]0, \frac{r_0}{2}[$ tale che

$$\|g(w) - g(x) - g'(\bar{x})(w - x)\| \le \alpha \|w - x\| \quad \forall w, x \in B(\bar{x}, 2r)$$

(tale r esiste per la continuità di $g'(x)$ in $\bar{x}$).

Scegliamo $\sigma, \tau > 0$ con $\sigma + \tau < \frac{1}{\beta}(1-\alpha\beta)r$; per continuità di g, esiste $\rho \in]0, r]$ tale che $g(w) \in B(g(\bar{x}), \tau)$ per $w \in B(\bar{x}, \rho)$. Siano adesso $w \in B(\bar{x}, \rho)$, $y \in B(g(\bar{x}), \sigma)$. Poniamo

$$f : B(\bar{x}, r) \to Y, \quad f(x) = g(x) - y,$$

e mostriamo che f verifica le ipotesi del teorema 19 di Kantorovich. Ovviamente f è continua, e

$$\|f(w)\| = \|g(w) - g(\bar{x}) + g(\bar{x}) - y\| < \tau + \sigma < \frac{1}{\beta}(1-\alpha\beta)r.$$

Inoltre, per ogni $x, x' \in B(w, r)$

$$\|f(x') - f(x) - g'(\bar{x})(x'-x)\| = \|g(x') - g(x) - g'(\bar{x})(x'-x)\| < \alpha \|x'-x\|$$

in quanto $\|x' - \bar{x}\| \le \|x' - w\| + \|w - \bar{x}\| < r + \rho \le 2r$, e similmente $\|x - \bar{x}\| < 2r$.

Dunque, per il teorema di Kantorovich, esiste $x \in B(w, r)$ tale che $f(x) = 0$, ossia $g(x) = y$, e inoltre $\|x - w\| \le \beta \frac{1}{1-\alpha\beta} \|f(w)\| < r$, vale a dire

$$\|x - w\| \le \frac{\beta}{1-\alpha\beta} \|y - g(w)\|.$$

Per l'arbitrarietà di $w \in B(\bar{x}, \rho)$ e $y \in B(g(\bar{x}), \sigma)$, si ha la tesi con $c = \frac{\beta}{1-\alpha\beta}$. $\square$

<u>Proposizione 21</u> Siano X, Y spazi di Banach, sia $W \subseteq X$ aperto, e sia $g : W \to Y$ un'applicazione F-differenziabile in $\bar{x} \in W$. Sia $B \subseteq W$ tale che $\bar{x} \in B$. Allora:

(i) Se $g^{-1}(C) \supseteq B$, si ha $[g'(\bar{x})]^{-1}(N(B, \bar{x})) \supseteq N(C, g(\bar{x}))$.

(ii) Se $g^{-1}(C) = B$ e g è F-differenziabile in $B(\bar{x}, r_0) \subseteq W$, con g' continua in $\bar{x}$

[e $g'(\bar{x})$ surgettiva, si ha $[g'(\bar{x})^*]^{-1}(N(B,\bar{x})) = N(C, g(\bar{x}))$.

dim (i) Sia $\varphi \in N(C, g(\bar{x}))$: dunque

$$\varphi(y - g(\bar{x})) \leq \omega(\|y - g(\bar{x})\|) \quad \forall y \in C,$$

con $\frac{\omega(\|z\|)}{\|z\|} \to 0$ per $z \to 0$ in Y. Poiché g è differenziabile in $\bar{x}$,

$$g(x) - g(\bar{x}) = [g'(\bar{x})](x - \bar{x}) + \theta(\|x - \bar{x}\|) \quad \forall x \in B,$$

ove $\frac{\theta(\|z\|)}{\|z\|} \to 0$ per $z \to 0$ in X. Sia $x \in B$: allora $g(x) \in C$, e dunque

$$[g'(\bar{x})^*\varphi](x - \bar{x}) = \varphi(g'(\bar{x})(x - \bar{x})) = \varphi(g(x) - g(\bar{x}) - \theta(\|x - \bar{x}\|)) =$$

$$= \varphi(g(x) - g(\bar{x})) + \theta_1(\|x - \bar{x}\|) \leq$$

$$\leq \omega(\|g(x) - g(\bar{x})\|) + \theta_1(\|x - \bar{x}\|).$$

con $\frac{\theta_1(\|z\|)}{\|z\|} \to 0$ per $z \to 0$ in X. Dato che per x vicino a $\bar{x}$ risulta

$$\|g(x) - g(\bar{x})\| \leq [\|g'(\bar{x})\| + 1] \|x - \bar{x}\|,$$

si vede immediatamente che esiste $K > 0$ per cui

$$\frac{[g'(\bar{x})^*\varphi](x - \bar{x})}{\|x - \bar{x}\|} \leq \frac{\omega(K\|x - \bar{x}\|)}{\|x - \bar{x}\|} + \frac{\theta_1(\|x - \bar{x}\|)}{\|x - \bar{x}\|}$$

da cui

$$\limsup_{x \to \bar{x},\, x \in B \setminus \{\bar{x}\}} \frac{[g'(\bar{x})^*\varphi](x - \bar{x})}{\|x - \bar{x}\|} \leq 0.$$

Perciò $\varphi \in [g'(\bar{x})^*]^{-1}(N(B,\bar{x}))$. Ciò implica (i).

(ii) Sia $g^{-1}(C) = B$. Per il teorema 20 di Lyusternik-Graves, esistono $c > 0$, $\rho \in]0, \frac{r_0}{2}]$, $r > 0$ tali che

$\forall w \in B(\bar{x}, \rho)$, $\forall y \in B(g(\bar{x}), r)$ $\exists x \in B(\bar{x}, \frac{r_0}{2})$: $g(x) = y$, $\|x - w\| \leq c\|y - g(w)\|$.

Scelto $w = \bar{x}$, per ogni $y \in B(g(\bar{x}), r)$ denotiamo con x_y il punto x per cui vale quanto sopra: dunque

$$x_y \in B \cap B(\bar{x}, \tfrac{r_0}{2}), \quad g(x_y) = y, \quad \|x_y - \bar{x}\| \leq c\|y - g(\bar{x})\|.$$

Sia $\varphi \in [g'(\bar{x})^*]^{-1}(N(B, \bar{x}))$, vale a dire $g'(\bar{x})^* \varphi \in N(B, \bar{x})$: ciò significa

$$\varphi(g'(\bar{x})(x - \bar{x})) = g'(\bar{x})^* \varphi\,(x - \bar{x}) \leq \omega(\|x - \bar{x}\|) \quad \forall x \in B,$$

ove $\frac{\omega(\|z\|)}{\|z\|} \to 0$ per $z \to 0$ in X. Usando la differenziabilità di g in $\bar{x}$, si ha per $y \in C \cap B(g(\bar{x}), r)$

$$\varphi(y - g(\bar{x})) = \varphi(g(x_y) - g(\bar{x})) = \varphi\big(g'(\bar{x})(x_y - \bar{x}) + \theta(\|x_y - \bar{x}\|)\big) \leq \theta_1(\|x_y - \bar{x}\|) \leq$$
$$\leq \theta_1(2c\|y - g(\bar{x})\|) = \theta_2(\|y - g(\bar{x})\|),$$

ove $\frac{\theta_2(\|z\|)}{\|z\|} \to 0$ per $z \to 0$ in Y. Ciò prova che $\varphi \in N(C \cap B(g(\bar{x}), r), g(\bar{x}))$.

Dunque $[g'(\bar{x})^*]^{-1}(N(B, \bar{x})) \subseteq N(C \cap B(g(\bar{x}), r), g(\bar{x})) = N(C, g(\bar{x}))$ per l'osservazione di pag. 40. $\square$

<u>Teorema 22 (di Lyusternik)</u> Siano X, Y spazi di Banach, sia $W \subseteq X$ un aperto, sia $g: W \to Y$ una applicazione F-differenziabile in $\bar{x} \in W$, con $g'(\bar{x})$ surgettiva. Posto $\bar{y} = g(\bar{x})$ e $S = \{x \in W : g(x) = \bar{y}\}$, risulta

$$N(S, \bar{x}) = [g'(\bar{x})^*](Y^*).$$

<u>dim.</u> Scegliamo $C=\{\bar{y}\}=g(S)$. Dalla proposizione 21(i) si ha
$$g'(\bar{x})^*(Y^*)=g'(\bar{x})^*\left(N(\{\bar{y}\},\bar{y})\right)\subseteq N(S,\bar{x}).$$
Viceversa, sia $\varphi\in N(S,\bar{x})$. Osserviamo che $T(S,x)=\ker g'(\bar{x})$: infatti
$$u\in T(S,\bar{x})\iff \exists\, t_n\to 0^+,\ \exists\, u_n\to u \text{ in } X, \text{ con } g(\bar{x}+t_nu_n)=\bar{y}.$$
Ma
$$g(\bar{x}+t_nu_n)=g(\bar{x})+[g'(\bar{x})]t_nu_n+o(t_nu_n) \quad \text{per } n\to\infty;$$
pertanto
$$u\in T(S,\bar{x}) \iff g'(\bar{x})t_nu_n=o(t_nu_n) \quad \text{per } n\to\infty$$
$$\iff g'(\bar{x})=0.$$
Ricordando la proposizione 15, si ha
$$\varphi(u)\le 0 \quad \forall u\in\ker g'(\bar{x})=T(S,x),$$
e scambiando u con $-u$, si trova $\varphi(u)=0$ $\forall u\in\ker g'(\bar{x})$.
Utilizziamo adesso un risultato di analisi funzionale.

<u>Lemma 23 (di fattorizzazione)</u> Siano X,Y spazi di Banach, e sia $A\in\mathcal{L}(X,Y)$ surgettivo. Se $\varphi\in X^*$ e se $\varphi x=0$ per ogni $x\in\ker A$, allora esiste $T\in Y^*$ tale che
$$\varphi x=T(Ax)\quad\forall x\in X,$$
ossia $\varphi=A^*T$.

<u>dim.</u> Sia $y\in Y$ e siano $x,x'\in X$ tali che $Ax=Ax'=y$ (x e x' esistono per la surgettività). Allora per ipotesi $\varphi x=\varphi x'$, cioè φ è costante su $A^{-1}(y)$.
Poniamo allora
$$Ty=\varphi x\quad\forall x\in A^{-1}(y).$$

La definizione è buona e T è lineare. Proviamo che $T \in Y^*$. Se $G \subseteq \mathbb{R}$ è aperto,

$$T^{-1}(G) = \{y \in Y : Ty \in G\} = \{Ax : x \in X,\ \varphi x \in G\} =$$
$$= A\{x \in X : \varphi x \in G\} = A\varphi^{-1}(G).$$

Dato che φ è continuo, $\varphi^{-1}(G)$ è un aperto; per il teorema dell'applicazione aperta, $A\varphi^{-1}(G)$ è aperto. Dunque $T \in Y^*$ e risulta

$$\varphi x = T(Ax) \quad \forall x \in X. \ \square$$

Applichiamo il lemma 22 al nostro $\varphi \in N(S, \bar{x})$: poiché $g'(\bar{x})$ è surgettiva, possiamo dedurre che esiste $T \in Y^*$ tale che

$$\varphi = g'(\bar{x})^* \circ T.$$

Ciò prova che $\varphi \in g'(\bar{x})^*(Y^*)$. $\square$

<u>Osservazione</u> La completezza degli spazi X, Y è stata utilizzata soltanto per dimostrare i teoremi da 19 in poi. La parte precedente, cioè le proprietà del cono normale e del cono tangente, si può ripetere tale e quale nel caso di spazi normati.

Veniamo finalmente al teorema di Pontrjagin, che riguarda l'ottimizzazione di un funzionale sotto un vincolo di uguaglianza e uno di disuguaglianza.

Teorema 24 (di Pontryagin) Siano X, Y spazi di Banach, sia $J : X \to \mathbb{R}$ un funzionale, siano $\phi : X \to Y$ e $g : X \to \mathbb{R}$ i vincoli. Supponiamo J, ϕ, g F-differenziabili. Posto

$$K = \{x \in X : \phi(x) = 0,\ g(x) \in [a,b]\},$$

sia $\bar{x} \in K$ e supponiamo che l'applicazione $\psi : X \to Y \times \mathbb{R}$,

$$\psi(x) = \begin{pmatrix} \phi(x) \\ g(x) \end{pmatrix}$$

abbia differenziale continuo in $\bar{x}$, e soddisfi

$$R(\psi'(\bar{x})) = Y \times \mathbb{R} \quad \text{(condizione di qualificazione)}.$$

Se x_0 è punto di massimo o minimo relativo per J su K, allora esistono $\varphi \in Y^*$, $\lambda \in \mathbb{R}$ (i "moltiplicatori") tali che

$$J'(\bar{x}) + \phi'(\bar{x})^* \circ \varphi + g'(\bar{x})^* \lambda = 0 \in X^*.$$

dim Cominciamo con l'osservare che

$$K = \{x \in X : \psi(x) \in \{0\} \times [a,b]\} = \psi^{-1}(\{0\} \times [a,b]).$$

Per la Proposizione 16 si ha, supponendo $\bar{x}$ punto di minimo su K,

$$-J'(\bar{x}) \in N(K, \bar{x}).$$

In virtù delle ipotesi fatte su ψ', utilizzando la proposizione 21(ii) possiamo scrivere

$$N(K,\bar{x}) = \psi'(\bar{x})^* \left(N(\{0\}\times[a,b], \psi(\bar{x})) \right) =$$
$$= \psi'(\bar{x})^* \left(N(\{0\}\times[a,b], (0, g(\bar{x}))) \right) =$$
$$= \psi'(\bar{x})^* \left(N(\{0\}, 0) \times N([a,b], g(\bar{x})) \right) =$$
$$= \psi'(\bar{x})^* \left(X^* \times N([a,b], g(\bar{x})) \right).$$

D'altra parte, per facile conseguenza della definizione,

$$N([a,b], g(\bar{x})) = \begin{cases}]-\infty, 0] & \text{se } g(\bar{x}) = a, \\ \{0\} & \text{se } g(\bar{x}) \in]a,b[, \\ [0,\infty[& \text{se } g(\bar{x}) = b. \end{cases}$$

Quindi la condizione $-J'(\bar{x}) \in N(K,\bar{x})$ significa che esiste $(\varphi, \lambda) \in X \times \mathbb{R}$ tale che

$$-J'(\bar{x}) = \psi'(\bar{x})^*(\varphi, \lambda), \qquad \lambda \begin{cases} \leq 0 & \text{se } g(\bar{x}) = a, \\ = 0 & \text{se } g(\bar{x}) \in]a,b[, \\ \geq 0 & \text{se } g(\bar{x}) = b. \end{cases}$$

ossia

$$0 = J'(\bar{x}) + \phi'(\bar{x})^*\varphi + g'(\bar{x})^*\lambda. \quad \square$$

Nei problemi concreti si tratterà di esplicitare questa condizione, cercando di dedurre proprietà specifiche del controllo e dello stato ottimale, che sono descritti dal punto $\bar{x}$.

Applicazione al problema di Bolza

Torniamo al problema di minimizzare

$$J(x,u) = \int_0^T g(y(t),u(t))\,dt + G(y(T))$$

sotto il vincolo

$$\begin{cases} y'(t) = f(y(t),u(t)), & t \in [0,T], \\ y(0) = x, \end{cases}$$

supponendo $g: \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$, $G: \mathbb{R}^n \to \mathbb{R}$, $f: \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ funzioni differenziabili con continuità.

Per la presenza del termine $G(y(T))$, scegliamo

$$X = C([0,T],\mathbb{R}^n) \times C([0,T],\mathbb{R}^m), \quad Y = C([0,T],\mathbb{R}^n),$$

$$J(y,u) = \int_0^T g(y(t),u(t))\,dt + G(y(T)),$$

$$[\Phi(y,u)](t) = y(t) - x - \int_0^t f(y(s),u(s))\,ds, \quad t \in [0,T].$$

Rispetto al problema astratto, sono assenti i vincoli di disuguaglianza: $g \equiv 0$.

Calcoliamo le derivate di J e Φ:

$$[J'(y_0,u_0)](h,k) = \int_0^T \left[g_y(y_0(t),u_0(t)) \cdot h(t) + g_u(y_0(t),u_0(t)) \cdot k(t) \right] dt + DG(y_0(T)) \cdot h(T),$$

$$[[\Phi'(y_0,u_0)](h,k)](t) = h(t) - \int_0^t \left[f_y(y_0(s),u_0(s)) \cdot h(s) + f_u(y_0(s),u_0(s)) \cdot k(s) \right] ds, \quad t \in [0,T]$$

Vediamo se l'applicazione $\Phi'(y_0,u_0): X \to Y$ è surgettiva: fissata

$\ell \in C([0,T],\mathbb{R}^n)$, occorre trovare $(h,k) \in C([0,T],\mathbb{R}^n) \times C([0,T],\mathbb{R}^m)$ tale che, per ogni $t \in [0,T]$, risulti

$$h(t) - \int_0^t \left[f_y(y_0(s),v_0(s)) \cdot h(s) + f_u(y_0(s),v_0(s)) \cdot k(s) \right] ds = \ell(t) .$$

Questa è una equazione integrale lineare di Volterra. Si può addirittura scegliere $k=0$ e risolverla rispetto a h: è una equazione del tipo

$$h(t) - [Kh](t) = \ell(t)$$

ove

$$[Kh](t) = \int_0^t A(s) \cdot h(s)\, ds , \qquad A \in C([0,T],\mathbb{R}^{n^2})$$

Si verifica facilmente che

$$|[Kh](t)| \le \|A\|_\infty\, t\, \|h\|_\infty ,$$

e induttivamente

$$|[K^m h](t)| \le \frac{\|A\|_\infty^m t^m}{m!} \|h\|_\infty$$

per cui

$$\|K^m\|_{\mathcal{L}(C([0,T],\mathbb{R}^n))} \le \frac{\|A\|_\infty^m T^m}{m!} .$$

Dunque

$$h(t) = [(1-K)^{-1}\ell](t) = \sum_{m=0}^\infty [K^m \ell](t)$$

è ben definita, e

$$\|h\|_\infty \le \sum_{m=0}^\infty \frac{\|A\|_\infty^m T^m}{m!} \|\ell\|_\infty .$$

Dunque $\Phi'(y_0,u_0)$ è surgettiva.

Se (y_0,u_0) è una coppia ottimale, o anche soltanto un punto di massimo o minimo relativo per J, esiste un elemento $\mu \in [C([0,T], \mathbb{R}^n)]^*$ tale che

$$J'(y_0,u_0) + \Phi'(y_0,u_0)^* \mu = 0 \in [C([0,T],\mathbb{R}^n) \times C([0,T],\mathbb{R}^m)]^*.$$

Ci occorre a questo punto una digressione sul duale di $C[0,T]$. Lo spazio $(C[0,T])^*$ è costituito dalle misure (con segno) μ, associate a funzioni φ che sono differenze di funzioni monotone crescenti e continue a sinistra, salvo al più nel punto T. Per tali misure μ si ha

$$\mu([c,d[) = \varphi(d) - \varphi(c) \quad \text{se } d<T, \quad \mu([c,T]) = \varphi(T^-) - \varphi(c).$$

Si ha allora, per ogni $h \in C[0,T]$,

$$\mu(h) := \int_{[0,T]} h(t)\, d\mu(t),$$

e se $h \in C^1[0,T]$ vale la formula di integrazione per parti

$$\int_{[s,T]} h(t)\, d\mu(t) = \varphi(T)h(T) - \varphi(s)h(s) - \int_s^T h'(\tau)\varphi(\tau)\, d\tau.$$

Infatti, per il teorema di Fubini,

$$\int_{[s,T]} h(t)\, d\mu(t) = \int_{[s,T]} [h(t) - h(s)]\, d\mu(t) + h(s)\,[\varphi(T) - \varphi(s)] =$$

$$= \int_{[s,T]} \int_s^t h'(\tau)\,d\tau \, d\mu(t) + h(s)\,[\varphi(T) - \varphi(s)] =$$

$$= \int_s^T \int_{[\tau,T]} d\mu(t)\; h'(\tau)\,d\tau + h(s)\,[\varphi(T) - \varphi(s)] =$$

$$= \int_s^T [\varphi(T) - \varphi(\tau)]\, h'(\tau)\,d\tau + h(s)\,[\varphi(T) - \varphi(s)] =$$

$$= \varphi(T)[h(T) - h(s)] - \int_s^T \varphi(\tau) h'(\tau)\,d\tau + h(s)\,[\varphi(T) - \varphi(s)] =$$

$$= [\varphi(T)h(T) - \varphi(s)h(s)] - \int_s^T \varphi(\tau) h'(\tau)\,d\tau .$$

Poichè vale

$$(A \times B)^* \simeq A^* \times B^*,$$

lo spazio $[C([a,T], \mathbb{R}^m)]^*$ è costituito da m-ple di misure con segno μ, associate a m-ple di funzioni φ_i, che sono differenze di funzioni crescenti e continue a sinistra, salvo al più nel punto T.

Si avrà perciò

$$\mu([c,d[) = (\varphi_1(d) - \varphi_1(c), \ldots, \varphi_m(d) - \varphi_m(c)) \quad \text{se } d < T,$$

$$m([c,T[) = (\varphi_1(T^-) - \varphi_1(c), \ldots, \varphi_m(T^-) - \varphi_m(c)).$$

Terminata la digressione, andiamo a calcolare

$$[\Phi'(y_0,u_0)^* \mu](h,k)$$

per $(h,k) \in C([a,T], \mathbb{R}^n) \times C([a,T], \mathbb{R}^m)$. Si ha, integrando per parti,

$$[\Phi'(y_0,u_0)^*\mu](h,k) = \mu\left(\Phi'(y_0,u_0)(h,k)\right) = \int_{[0,T]} [\Phi'(y_0,u_0)(h,k)](t)\,d\mu(t) =$$

$$= \int_{[0,T]} \left[h(t) - \int_0^t \left[f_y(y_0(s),u_0(s)) \cdot h(s) + f_u(y_0(s),u_0(s)) \cdot k(s) \right] ds \right] \cdot d\mu(t) =$$

$$= \int_{[0,T]} h(t) \cdot d\mu(t) - \int_0^T \left[f_y(y_0(s),u_0(s)) \cdot h(s) + f_u(y_0(s),u_0(s)) \cdot k(s) \right] \cdot [\varphi(T) - \varphi(s)]\, ds.$$

Dunque la condizione $J'(y_0,u_0) + \Phi'(y_0,u_0)^*\mu = 0$ diventa:

$$\int_0^T \left[g_y(y_0(s),u_0(s)) \cdot h(s) + g_u(y_0(s),u_0(s)) \cdot k(s) \right] ds + DG(y_0(T)) \cdot h(T) +$$

$$+ \int_{[0,T]} h(t) \cdot d\mu(t) - \int_0^T \left[f_y(y_0(s),u_0(s)) \cdot h(s) + f_u(y_0(s),u_0(s)) \cdot k(s) \right] \cdot [\varphi(T) - \varphi(s)]\, ds = 0$$

per ogni $(h,k) \in C([0,T],\mathbb{R}^n) \times C([0,T],\mathbb{R}^m)$. Equivalentemente, posto $\psi(s) = \varphi(T) - \varphi(s)$, abbiamo

$$\int_0^T \Big[\left[g_y(y_0(s),u_0(s)) - \psi(s) \cdot f_y(y_0(s),u_0(s)) \right] \cdot h(s) +$$

$$+ \left[g_u(y_0(s),u_0(s)) - \psi(s) \cdot f_u(y_0(s),u_0(s)) \right] \cdot k(s) \Big] ds + DG(y_0(T)) \cdot h(T) + \int_{[0,T]} h(t)\,d\mu(t) = 0$$

per ogni $(h,k) \in C([0,T],\mathbb{R}^n) \times C([0,T],\mathbb{R}^m)$.

Sfruttiamo adesso l'arbitrarietà di h e k. Scegliendo $h \equiv 0$,

$$\int_0^T \left[g_u(y_0(s),u_0(s)) - \psi(s) f_u(y_0(s),u_0(s)) \right] \cdot k(s)\, ds = 0 \quad \forall k \in C([0,T],\mathbb{R}^m)$$

e per il primo dei lemmi fondamentali del calcolo delle variazioni (in $L^1(0,T)$) si ottiene

$$g_u(y_0(s),u_0(s)) - \psi(s)\cdot f_u(y_0(s),u_0(s)) = 0 \quad \text{q.o. in } [0,T]$$

Scegliendo poi $h\in C_0^1([0,T],\mathbb{R}^n)$, ricaviamo

$$\int_0^T [g_y(y_0(s),u_0(s)) - \psi(s) f_y(y_0(s),u_0(s))]\cdot h(s)\,ds + \int_{[0,T]} h(t)\,d\mu(t) = 0;$$

dunque, essendo

$$\int_{[0,T]} h(t)\cdot d\mu(t) = -\int_0^T h'(t)\cdot \varphi(t)\,dt,$$

otteniamo per ogni $h\in C^1([0,T],\mathbb{R}^n)$

$$\int_0^T \left([g_y(y_0(s),u_0(s)) - \psi(s)\cdot f_y(y_0(s),u_0(s))]\cdot h(s) - \varphi(s)\cdot h'(s)\right) ds = 0$$

da cui, per definizione di derivata distribuzionale,

$$\exists\, \psi'(s) = -\varphi'(s) = -\psi(s)\cdot f_y(y_0(s),u_0(s)) + g_y(y_0(s),u_0(s)) \quad \text{q.o. in } [0,T].$$

Il secondo membro è una funzione sicuramente misurabile e limitata, quindi ψ è assolutamente continua in $[0,T]$. In particolare

$$\psi(T) = \lim_{t\to T} \psi(t) = \lim_{t\to T^-} [\varphi(T)-\varphi(t)] = \varphi(T) - \varphi(T^-).$$

Scegliamo infine, per $v\in\mathbb{R}^n$ fissato e $k\in\mathbb{N}^+$,

$$h(t) = \begin{cases} 0 & \text{se } t\in[0,T-\frac{1}{k}], \\ (k(t-T)+1)v & \text{se } t\in[T-\frac{1}{k},T]. \end{cases}$$

Allora si ha

$$\int_{T-\frac{1}{k}}^T [g_y(y_0(s),u_0(s)) - \psi(s)\cdot f_y(y_0(s),u_0(s))]\cdot h(s)\,ds + DG(y_0(T))\cdot v + \int_{[T-\frac{1}{k},T]} h(t)\cdot d\mu(t) = 0;$$

Il primo addendo tende a 0 per $k\to\infty$, per convergenza dominata: infatti $|h(t)|\le|v|$ per ogni $t\in[0,T]$. L'ultimo addendo, integrando per parti, vale

$$\int_{[T-\frac{1}{k},T]} h(t)\cdot d\mu(t) = \varphi(T)\cdot v - \int_{T-\frac{1}{k}}^{T} k\, v\cdot\varphi(t)\,dt\,,$$

da cui per $k\to\infty$

$$DG(y_0(T))\cdot v + [\varphi(T)-\varphi(T^-)]\, v = 0 \quad \forall v\in\mathbb{R}^n;$$

dunque

$$DG(y_0(T)) = -[\varphi(T)-\varphi(T^-)].$$

Si ottiene in definitiva il seguente sistema di condizioni necessarie:

$$\begin{cases} g_u(y_0(s),u_0(s)) - \psi(s)\cdot f_u(y_0(s),u_0(s)) = 0 & \text{q.o. in } [0,T] \\ \psi'(s) = -\psi(s)\cdot f_y(y_0(s),u_0(s)) + g_y(y_0(s),u_0(s)) & \text{q.o. in } [0,T] \\ \psi(T) + DG(y_0(T)) = 0 \end{cases}$$

In molti casi concreti, queste condizioni riescono a selezionare un'unica (candidata) coppia ottimale. Naturalmente, però, l'esistenza della coppia ottimale va dimostrata separatamente.