Appunti di Metodi Variazionali

Matteo Scarcella

Maggio 2024

Indice

1 Introduzione

Questi appunti sono soltanto un riordionamento dei risultati esposti nel libro di ”Metodi variazionali per il controllo ottimo” di Bruni, Di Pillo.

2 Spazi normati e Funzionali

2.1 Introduzione

Sia $\mathcal{Z} = \mathbb{R}^\nu$ lo spazio ambiente, con $z \in \mathcal{Z}$. Sia $\mathcal{D} \subset \mathcal{Z}$ un sotto-insieme dello spazio ambiente detto insieme ammissibile, se necessario descritto da $\mu < \nu$[1] vincoli di uguaglianza $h(z) = 0$ e σ vincoli di disuguaglianza $g(z) \leq 0$, dove con σ_a e $g_a(z)$ si fa riferimento ai soli vincoli di disuguaglianza attivi, ovvero verificati all'uguaglianza. Sia $[t_i, t_f]$ un intervallo di tempo. Siano

$$\|z\| = \sup_{t\in[t_i,t_f]} \|z(t)\| \tag{1}$$

$$\|z\| = \sup_{t\in[t_i,t_f]} \|z(t)\| + \sup_t \|\dot{z}(t)\| \tag{2}$$

Rispettivamente norma forte e norma debole, scelte a priori sullo spazio ambiente. Sia $J : \mathcal{Z} \to \mathbb{R}$ un funzionale di costo. L'obiettivo è trovare il controllo che permetta di minimizzare il valore del funzionale J, ovvero risolvere il problema

$$\begin{cases} \min J(z) \\ z \in \mathcal{D} \end{cases} \tag{3}$$

$\square$

Definizione 1.29 - z^* è un punto di minimo locale (forte o debole, in base alla norma scelta) se vale

$$J(z^*) \leq J(z) \quad \forall z \in \mathcal{D} \cap \mathcal{S}(z^*, \varepsilon) \tag{4}$$

$\square$

Definizione 2.5 - Definita la matrice Jacobiana dei vincoli attivi in un punto ammissibile $\bar{z}$

$$\left.\frac{\partial(h, g_a)}{\partial z}\right|_{\bar{z}} = \begin{pmatrix} \frac{dh}{dz} \\ \\ \frac{dg_a}{dz} \end{pmatrix}_{\bar{z}} \tag{5}$$

Allora $\bar{z}$ si dice punto di regolarità dei vincoli se la matrice Jacobiana dei vincoli attivi ha rango pieno, ovvero se

$$rank\left\{ \left.\frac{\partial(h, g_a)}{\partial z}\right|_{\bar{z}} \right\} = \mu + \sigma_a \tag{6}$$

$\square$

Definizione 2.6 - Si definisce la funzione Lagrangiana

$$L(z, \lambda_0, \lambda, \eta) = \lambda_0 J(z) + \lambda^T h(z) + \eta^T g(z) \tag{7}$$

Con λ_0, λ, η moltiplicatori opportuni $\square$

Teorema 2.7 - (Condizioni necessarie di minimo) In riferimento al problema di minimo vincolato con opportuni vincoli $h(z)$ e $g(z)$, sia z^* un punto di minimo locale. Allora esistono moltiplicatori $\lambda_0^*, \lambda^*, \eta^*$ non tutti simultaneamente nulli tali che:

$$\begin{cases} \left.\frac{\partial L}{\partial z}\right|^* = 0^T \\ \eta_i^* g_i(z^*) = 0 \quad i = 1, 2, \ldots, \sigma_a \\ \lambda_0^* \geq 0 \\ \eta_i^* \geq 0 \quad i = 1, 2, \ldots, \sigma_a \end{cases} \tag{8}$$

$\square$

[1] Nel caso $\mu = \nu$ si avrebbe un insieme di vincoli finito e verrebbe meno il concetto di minimizzazione. Ad esempio se $\mu = \nu = 2$, i due vincoli potrebbero essere due rette che si intersecano in un punto, e tale punto risulterebbe l'unico punto ammissibile

Teorema 2.8 - (Condizioni di Kuhn-Tucker) In riferimento al problema di minimo vincolato con opportuni vincoli $h(z)$ e $g(z)$, sia z^* un punto di minimo locale e di regolarità dei vincoli. Allora valgono le stesse condizioni del teorema (2.7), con l'aggiunta di

$$\lambda_0^* = 1 \tag{9}$$

Dimostrazione - Si supponga per assurdo $\lambda_0^* = 0$. Allora nella funzione lagrangiana sparirebbe il termine dipendente dal funzionale di costo $J(z)$ e la prima condizione delle (8) diventerebbe

$$0^T = \frac{\partial L}{\partial z}\Big|^* = 0 + \lambda^{*T}\frac{\partial h(z)}{\partial z} + \eta^{*T}\frac{\partial g(z)}{\partial z} \tag{10}$$

Che in forma vettoriale diventa

$$0^T = \begin{pmatrix} \lambda^{*T} & \eta^{*T} \end{pmatrix} \begin{pmatrix} \frac{\partial h(z)}{\partial z} \\ \frac{\partial g(z)}{\partial z} \end{pmatrix} \tag{11}$$

Dove l'ultimo vettore è proprio la matrice jacobiana dei vincoli attivi. Per ipotesi z^* è un punto di regolarità dei vincoli e quindi tale matrice ha rango pieno. Dunque l'unica possibilità per cui il prodotto con i moltiplicatori valga 0^T è che i moltiplicatori siano tutti nulli, il che è assurdo perchè contraddice l'ipotesi. Pertanto, siccome $\lambda_0^* \neq 0$, sarà sempre possibile dividere tutti i moltiplicatori per λ_0^*, ottenendo così dei nuovi moltiplicatori $\bar{\lambda}_0^*, \bar{\lambda}^*, \bar{\eta}^*$, con $\bar{\lambda}_0^* = 1$ □

Le due condizioni sono dei precedenti teoremi sono necessarie e quindi non danno direttamente luogo a dei punti di minimo, ma a dei punti candidati ad essere di minimo, che vengono definiti estremali. Gli estremali possono essere

- normali, se $\lambda_0^* \neq 0$ (e quindi conseguentemente $\lambda_0^* = 1$)
- non normali, se $\lambda_0^* = 0$

Il teorema (2.7) offre delle condizioni meno stringenti rispetto al (2.8), che fissando $\lambda_0^* = 1$ esclude gli estremali non normali, e pertanto fornisce un numero maggiore di candidati.

Esempio 2.1 □

Le condizioni necessarie appena viste diventano anche sufficienti nel caso in cui il funzionale di costo $J(z)$ e l'insieme ammissibile $\mathcal{D}$ siano convessi. La particolarizzazione dell'insieme di ammissibilità al caso convesso può essere effettuata sulla base del seguente lemma

Lemma 2.13 - Sia $\mathcal{D}$ un insieme definito da vincoli di uguaglianza $h(z) = 0$ e di disuguaglianza $g(z) \leq 0$. Se

- $g(z)$ è dato da tutte funzioni convesse
- $h(z)$ è dato da funzioni lineari (o affini) del tipo $h_j(z) = c_j^T z_j + b_j$, $j = 1, 2, \ldots, \mu$

Allora $\mathcal{D}$ è convesso. □

Teorema 2.14 - In riferimento al problema di minimo vincolato con un funzionale di costo $J(z)$ convesso e con opportuni vincoli $h(z)$ affini e $g(z)$ convessi, di modo che l'insieme $\mathcal{D}$ sia convesso. Se z^0 è un punto ammissibile ed esistono moltiplicatori λ^0 e η^0 tali che valgano

$$\begin{cases} \frac{dJ}{dz} + \lambda^{0T}\frac{dh}{dz} + \eta^{0T}\frac{dg}{gz} = 0^T \\ \eta_i^0 g_i(z^0) = 0 \quad i = 1, 2, \ldots, \sigma_a \\ \eta_i^0 \geq 0 \quad i = 1, 2, \ldots, \sigma_a \end{cases} \tag{12}$$

Allora z^0 è un punto di minimo globale. Inoltre, se J è strettamente convessa, z^0 è l'unico punto di minimo globale.

Dimostrazione - Posto z un qualunque punto ammissibile, allora valgono i vincoli e quindi si può scrivere

$$J(z) \geq J(z) + \underbrace{\lambda^{0T}h(z)}_{=0} + \underbrace{\eta^{0T}g(z)}_{\leq 0} \tag{13}$$

Siccome per ipotesi le funzioni J e g sono convesse e le funzioni h sono affini, allora si ha

$$J(z) \geq J(z^0) + \frac{dJ}{dz}(z - z^0) + \lambda^{0T}\left[h(z^0) + \frac{dh}{dz}(z - z^0)\right] + \eta^{0T}\left[g(z^0) + \frac{dg}{dz}(z - z^0)\right] \tag{14}$$

Per le ipotesi del teorema, nel punto z^0 si ha $\eta^{0T} g(z^0) = 0$, inoltre poichè, sempre per ipotesi, z^0 è un punto ammissibile, allora valgono i vincoli, ovvero $h(z^0) = 0$. Quindi raccogliendo $(z - z^0)$ si ottiene

$$J(z) \geq J(z^0) + \underbrace{\left[\frac{dJ}{dz} + \lambda^{0T}\frac{dh}{dz} + \eta^{0T}\frac{dg}{dz}\right]}_{=0^T}(z - z^0) \tag{15}$$

E di nuovo, per le ipotesi (la prima delle (12)), si può scrivere

$$J(z) \geq J(z^0) \tag{16}$$

Ovvero z^0 è un punto di minimo globale. Se inoltre J fosse strettamente convessa, allora le stesse relazioni varrebbero strettamente, giungendo quindi a

$$J(z) > J(z^0) \tag{17}$$

Che implica l'unicità del minimo. □

2.2 Spazi infinito-dimensionali

Dato un insieme $\mathcal{Z}$, si dice che è uno spazio lineare se è chiuso rispetto all'operazione di somma e all'operazione di moltiplicazione. Ad esempio lo spazio Euclideo $\mathbb{R}^\nu$ e lo spazio delle matrici quadrate di dimensione $n \times n$ sono spazi lineari. Un controesempio è dato dallo spazio delle matrici non-singolari, nel quale è possibile trovare due matrici che sommate restituiscano una matrice di rango non pieno, che non appartiene allo spazio di partenza. Affianco alla definizione di spazio lineare vi è quella di sottospazio lineare, che è un sotto-insieme dello spazio lineare che a sua volta soddisfa le condizioni di chiusura rispetto alle due operazioni fondamentali.

Dato un qualunque spazio lineare $\mathcal{Z}$ e presi alcuni suoi elementi $z_k \in \mathcal{Z}$, allora essi si dicono linearmente indipendenti se vale la condizione

$$\sum_k \alpha_k z_k = 0 \iff \alpha_k = 0 \forall k \tag{18}$$

Questa condizione permette di introdurre il concetto di sottospazio generato, che è difatti il sottospazio definito da tutti gli elementi z che sono il risultato della combinazione lineare fra dei vettori z_k fissati e dei pesi α_k arbitrari, ovvero

$$\left\{ z \in \mathcal{Z} : z = \sum_k \alpha_k z_k \right\} \tag{19}$$

Se la scelta dei vettori è tale per cui il sottospazio generato coincide con lo spazio di partenza $\mathcal{Z}$, allora tali vettori vengono definiti base per $\mathcal{Z}$. Tipicamente il concetto di base non è univoco, nel senso che posso esistere più basi in uno spazio. Al contrario, è univoco il concetto di dimensione di una base, che è il numero di elementi della base linearmente indipendenti. Ad esempio, in $\mathbb{R}^3$, la dimensione è 3, poichè è sufficiente considerare 3 vettori non nulli e linearmente indipendenti per poter ricostruire tutto $\mathbb{R}^3$ tramite opportune combinazioni lineari. Ci sono però dei casi in cui il numero di elementi in una base non è un numero finito, e in tali casi si parla di spazi infinito-dimensionali. Alcuni esempi sono $\bar{\mathcal{C}}^k$ e $\mathcal{C}^k$, ovvero rispettivamente gli insiemi delle funzioni continuamente derivabili k volte con e senza punti di discontinuità in numero finito.

2.3 Norme

***Definizione 1.6* -** Una norma è una funzione $\|\cdot\|$ che porta elementi di un qualsiasi spazio lineare in elementi non negativi di $\mathbb{R}$ e soddisfa le seguenti proprietà:

- $\|z\| = 0 \iff z = 0$
- $\|z' + z''\| \leq \|z'\| + \|z''\|$ (Disuguaglianza triangolare)
- $\|\alpha z\| = |\alpha| \cdot \|z\|$

Uno spazio lineare è detto normato se su di esso è definita una norma. □

Alcuni esempi di norme in spazi finito-dimensionali sono:

$$\|z\|_1 = \left(\sum_{i=1}^{n} |z_i| \right) \tag{20}$$

$$\|z\|_2 = \left(\sum_{i=1}^{n} z_i^2 \right)^{1/2} \tag{21}$$

$$\|z\|_k = \left(\sum_{i=1}^{n} z_i^k \right)^{1/k} \tag{22}$$

$$\|z\|_\infty = \max_i \left\{ \left| z_i \right| \right\} \tag{23}$$

Spostandosi su spazi infinito-dimensionali, il concetto di norma dev'essere esteso poichè gli elementi dell'insieme sono funzioni, e non è più sufficiente applicare una somma su tutte le componenti della funzione in un certo istante. Occorre piuttosto integrare su tutto l'intervallo di supporto dello spazio stesso. Ad esempio in $\mathcal{C}^k[t_i, t_f]$, per k fissato, si ha:

$$\|z\| = \int_{t_i}^{t_f} \|z(t)\| \, dt \tag{24}$$

$$\|z\| = \left(\int_{t_i}^{t_f} \|z(t)\|^2 \, dt \right)^{1/2} \tag{25}$$

$$\|z\| = \sup_{t \in [t_i, t_f]} \|z(t)\| \tag{26}$$

$$\|z\| = \sup_{t \in [t_i, t_f]} \|z(t)\| + \sup_{t \in [t_i, t_f]} \|\dot{z}(t)\| \tag{27}$$

Dove occorre prestare attenzione al fatto che il simbolo di norma che compare all'esterno è diverso da quello che compare all'interno: quello all'interno è da intendersi come una qualsiasi "classica" norma di uno spazio finito-dimensionale su tutte le componenti della funzione in un dato istante t, mentre quella all'esterno è una ri-definizione estesa a tutti i valori che la funzione assume in un dato intervallo, attraverso l'operazione di integrale. Inoltre, le ultime due norme assumono un carattere particolare poichè generalizzano rispettivamente i concetti di norma forte e norma debole. Una domanda lecita potrebbe essere quale norma scegliere fra tutti i tipi di norma esistenti. La risposta è semplice nel caso finito-dimensionale: infatti, in questo tipo di spazi, ogni norma è equivalente a meno di costanti moltiplicative opportune. Al contrario, la scelta della norma in spazio infinito-dimensionali è fondamentale poichè tali equivalenze non valgono più. Per questo motivo, qualunque eventuale risultato dev'essere affiancato dall'assunzione di utilizzare una specifica norma.

Definizione 1.9 - Dato uno spazio normato $\mathcal{Z}$, una successione $\{z_k\}$ è convergente a $\bar{z}$ se

$$\lim_{k \to \infty} \|z_k - \bar{z}\| = 0 \tag{28}$$

□

Esempio 1.16 □

Esempio 1.18 □

Esempio 1.19 □

Esempio 1.21 □

2.4 Funzionali

Definizione 1.18 - Dato uno spazio lineare $\mathbb{Z}$ e un sotto-insieme ammissibile $\mathcal{D}$, si definisce funzionale una qualsiasi legge del tipo $J : \mathcal{D} \to \mathbb{R}$. □ La definizione sembrerebbe non dissimile da quella di funzione, e difatti una funzione è un caso particolare di funzionale. La differenza sostanziale risiede sostanzialmente nell'insieme ammissibile $\mathcal{D}$ (che deriva le sue proprietà dallo spazio $\mathcal{Z}$), che in generale può essere definito come insieme di funzioni. Nel caso di funzione quindi, tale insieme si riduce a un insieme di numeri.

Fra tutti i possibili funzionali, si è particolarmente interessati al funzionale di Lagrange, ovvero la legge $J : \bar{\mathcal{C}}^1[t_i, t_f] \to \mathbb{R}$ tale che

$$J(z, \dot{z}, t) = \int_{t_i}^{t_f} L(z, \dot{z}, t) \, dt \tag{29}$$

Con $L : \mathbb{R}^\nu \times \mathbb{R}^\nu \times \mathbb{R} \to \mathbb{R}$ derivabile con continuità 2 volte.

Esempio 1.28 **-** Sia J il funzionale di Lagrange particolarizzato al caso in cui la funzione L non dipende esplicitamente dal tempo. Allora si ha:

$$|J(z) - J(\bar{z})| \leq \int_{t_i}^{t_f} |L(z, \dot{z}) - L(\bar{z}, \dot{\bar{z}})| \, dt \tag{30}$$

Dove il segno è dipeso dal fatto che la norma di un integrale è minore uguale dell'integrale della norma. Allora, per il teorema del valore medio esiste un punto $\omega(t)$ tale da poter riscrivere l'integrale come

$$\int_{t_i}^{t_f} |\frac{\partial L}{\partial z(t)}\bigg|_{\omega(t)} (z(t) - \bar{z}(t)) + \frac{\partial L}{\partial \dot{z}(t)}\bigg|_{\omega(t)} (\dot{z}(t) - \dot{\bar{z}}(t))| \, dt \tag{31}$$

E applicando la disuguaglianza di Cauchy-Schwartz si ottiene

$$\int_{t_i}^{t_f} \left\| \frac{\partial L}{\partial z(t)}\bigg|_{\omega(t)} \right\| \|z(t) - \bar{z}(t)\| \, dt + \int_{t_i}^{t_f} \left\| \frac{\partial L}{\partial \dot{z}(t)}\bigg|_{\omega(t)} \right\| \|\dot{z}(t) - \dot{\bar{z}}(t)\| \, dt \tag{32}$$

Poichè la funzione L è stata supposta appartenente a $\mathcal{C}^1$, la sua derivata è ancora una funzione continua, e per il teorema di Weierstrass ammette un massimo. Allora si introduce una costante C come un upper-bound di tale funzione, ovvero

$$\max_{t \in [t_i, t_f]} \left\{ \frac{\partial L}{\partial z(t)}|_{\omega(t)}, \frac{\partial L}{\partial \dot{z}(t)}|_{\omega(t)} \right\} \leq C \tag{33}$$

Pertanto si può minorare la somma dei due integrali con

$$C \cdot \int_{t_i}^{t_f} \|z(t) - \bar{z}(t)\| \, dt + \int_{t_i}^{t_f} \|\dot{z}(t) - \dot{\bar{z}}(t)\| \, dt \tag{34}$$

Ma il termine a destra è proprio la norma debole di $z - \bar{z}$, che è indipendente dal tempo, e quindi si ottiene complessivamente (svolgendo l'integrale)

$$|J(z) - J(\bar{z})| \leq C(t_f - t_i)\|z - \bar{z}\| \tag{35}$$

Notare nella norma l'assenza della dipendenza da t. Per la continuità si richiede che $\forall \ \varepsilon > 0, \ \exists \ \delta_\varepsilon(\bar{z}) > 0$ tale che

$$|J(z) - J(\bar{z})| < \varepsilon \quad \forall z \in \mathcal{S}(\bar{z}, \delta_\varepsilon(\bar{z})) \tag{36}$$

Allora nel caso in questione si ha

$$|J(z) - J(\bar{z})| \leq C(t_f - t_i)\|z - \bar{z}\| < \varepsilon \tag{37}$$

Assumendo per ipotesi che z sia in un intorno di $\bar{z}$, allora si può dire $\|z - \bar{z}\| < \delta_\varepsilon(\bar{z})$[2], e quindi

$$\delta_\varepsilon(\bar{z}) < \frac{\varepsilon}{C(t_f - t_i)} \tag{38}$$

E dal momento che tale δ esiste, allora si conclude che il funzionale J è continuo. Inoltre, notando che il δ ricavato non dipende da $\bar{z}$, si può aggiungere che J è uniformemente continuo. □

[2] Capire bene il passaggio logico