Università degli Studi di Pisa

Facoltà di Scienze Matematiche, Fisiche e Naturali

Corso di Laurea in Matematica

Tesi di Laurea

Funzioni gap e algoritmi risolutivi per problemi di equilibrio

Candidato

Donatella Lazzaro

RELATORE

Prof. Giandomenico Mastroeni

CORRELATORE

Prof. Paolo Acquistapace

CONTRORELATORE

Prof. Carlo Luigi Berselli

Anno Accademico 2014/2015

Alla mia famiglia
e a mio nonno

If you're trying to achieve, there will be roadblocks. I've had them; everybody has had them. But obstacles don't have to stop you. If you run into a wall, don't turn around and give up. Figure out how to climb it, go through it, or work around it.

Michael Jordan

Indice

Introduzione

Negli ultimi cinquant'anni, molti problemi di equilibrio noti in fisica e economia, e diversi e importanti problemi ingegneristici sono stati riformulati mediante le disequazioni variazionali, introdotte da Hartman e Stampacchia nel 1966. La disequazione variazionale è un caso particolare del problema di equilibrio: data una funzione in due variabili, un problema di equilibrio consiste nel cercare un valore ammissibile per la prima variabile tale che la funzione sia sempre non negativa sulla regione ammissibile per ogni valore della seconda variabile. Questo problema, così semplice da porre, permette di riformulare sia problemi teorici, come i problemi del punto fisso, che problemi concernenti le applicazioni come i problemi di equilibrio di Nash e gli ottimi di Pareto.

Più recente è la teoria sulle funzioni gap, inizialmente utilizzate in letteratura per le disequazioni variazionali. Una funzione gap è una funzione che risulta essere maggiore o uguale di zero sull'insieme su cui è definito il problema di equilibrio ed è nulla solo e soltanto in corrispondenza di una soluzione. Pertanto la minimizzazione della funzione gap è equivalente al problema dato.

Gli algoritmi risolutivi sono molti, ognuno con le sue peculiarità, che richiedono condizioni particolari sulla funzione gap per essere applicati. È quindi importante studiare le proprietà delle funzioni gap, in particolare la differenziabilità. Per ottenere la differenziabilità delle funzioni gap è utile introdurre problemi di equilibrio ausiliari, ottenuti regolarizzando in modo opportuno la funzione che definisce il problema di equilibrio.

Con l'introduzione del problema ausiliario e della relativa funzione gap è possibile trasformare il problema di equilibrio iniziale in un problema di minimizzazione di una funzione differenziabile. Per risolvere questo problema di minimizzazione, tra gli algoritmi risolutivi, abbiamo scelto i metodi di discesa. Ad ogni passo dell'algoritmo si sceglie una direzione di discesa, in cui la funzione gap decresce, e ci si sposta lungo quella direzione. Gli algoritmi di discesa si dividono in due gruppi, quelli con ricerca esatta e quelli con ricerca inesatta. Nel primo caso, ad ogni passo si cerca il minimo della funzione nella direzione di discesa; nel secondo, invece, ci si sposta lungo la direzione di discesa di un certo valore (detto "passo"), opportunamente scelto.

Particolare importanza assume quindi il Rate of convergence e l'Error bound associato ai problemi di equilibrio, ossia rispettivamente la velocità di convergenza e le limitazioni del valore della distanza di un generico punto dalla soluzione ottima del problema.

La tesi è strutturata nel modo seguente.

Nel primo capitolo introdurremo i problemi di equilibrio, mediante i quali riformuleremo una vasta gamma di problemi, come due problemi classici, ottimizzazione e ottimizzazione differenziabile convessa, per poi proseguire con alcuni casi speciali: ottimo di Pareto, punti di sella, ottimizzazione inversa, disuguaglianze variazionali per funzioni a più valori, problemi di equilibrio di Nash, problemi di complementarità, problemi del punto fisso e disuguaglianze variazionali. Daremo anche la dimostrazione di alcuni risultati di esistenza, sotto opportune ipotesi sulla funzione relativa al problema di equilibrio.

Nel secondo capitolo introdurremo le funzioni gap, e approfondiremo le loro proprietà di differenziabilità, necessarie per poter applicare alcuni algoritmi di ottimizzazione per la minimizzazione delle funzioni gap. E infine accenneremo alle funzioni D-gap.

Nel terzo capitolo descriveremo alcuni algoritmi che possono essere applicati alle funzioni gap per trovare una soluzione di ottimo per il problema di equilibrio. In particolare, gli algoritmi risolutivi per la minimizzazione delle funzioni gap che considereremo saranno basati su metodi di ricerca esatta o inesatta lungo una direzione di discesa. Studieremo l'Error bound e il Rate of convergence per alcuni problemi di equilibrio. Descriveremo inoltre le applicazioni alle diseguaglianze variazionali e ai problemi di ottimizzazione. Concluderemo infine con un'applicazione a un problema di flusso su reti.

Nella prima appendice richiameremo alcune definizioni di concetti base che non sono stati definiti nella tesi. Nella seconda appendice invece descriveremo il metodo del gradiente, che è alla base degli algoritmi descritti nel terzo capitolo. Nella terza appendice infine riportiamo il codice usato per trovare il minimo della funzione gap per l'applicazione al problema di flusso su reti.

Capitolo 1

Preliminari ed esistenza

In questo capitolo, introdurremo il problema di equilibrio. Riformuleremo mediante problemi di equilibrio alcuni problemi classici come quello di ottimizzazione differenziabile e non.

Nella prima sezione daremo alcune definizioni importanti per la comprensione di quello che segue. Nella seconda parte illustreremo come molti altri problemi sono equivalenti al problema di equilibrio e infine nella terza e ultima parte del capitolo parleremo dell' esistenza della soluzione per il problema di equilibrio. In particolare considereremo il classico Teorema di Ky-Fan [11] e le sue generalizzazioni.

Un **problema di equilibrio** consiste nel cercare $\bar{x} \in \mathbb{R}^n$ tale che

$$\bar{x} \in K, \quad f(\bar{x}, y) \geq 0, \ \forall y \in K, \tag{EP}$$

dove $K \subseteq \mathbb{R}^n$ è un insieme convesso, e $f\colon\ K \times K \longrightarrow \mathbb{R}$ è una funzione che soddisfa $f(x, x) = 0$ per ogni $x \in K$.

I primi a essere riformulati con il problema di equilibrio (EP) sono due problemi classici della matematica: il problema di ottimizzazione e il problema di ottimizzazione differenziabile convessa. Successivamente ne analizzeremo anche altri più recenti.

Ottimizzazione

Sia $\varphi : \mathbb{R}^n \longrightarrow \mathbb{R}$ una funzione. Un **problema di minimo vincolato** é il seguente:

$$\min_{x \in K} \varphi(x). \tag{1.1}$$

Esso può essere riformulato come un problema di equilibrio (EP) considerando la funzione

$$f(x, y) := \varphi(y) - \varphi(x)$$

e il problema (EP) relativo a questa funzione. Le soluzioni di questo problema (EP) sono le stesse del problema di ottimizzazione (1.1), infatti $f(x, y) \geq 0$ se e solo se $\varphi(x) \leq \varphi(y)$.

Ottimizzazione differenziabile convessa Il problema (1.1) e il corrispondente problema (EP) sono equivalenti perché gli ottimi globali coincidono con quelli locali. Analizziamo la sottile connessione che esiste tra il caso convesso e quello differenziabile.

Sia $\varphi\colon \mathbb{R}^n \longrightarrow \mathbb{R}$ una funzione convessa e differenziabile secondo Gâteaux (Definizione A.0.10), con differenziale di Gâteaux $D\varphi(x)$ in x. Supponiamo che K sia convesso. Consideriamo il problema:

$$\min\{\varphi(x) | x \in K\}. \tag{1.2}$$

Osservazione 1.0.1. *Dall'analisi convessa sappiamo che $\bar{x}$ è una soluzione del problema* (1.2) *se e solo se soddisfa*

$$\bar{x} \in K, \quad \langle D\varphi(\bar{x}), y - \bar{x} \rangle \geq 0, \ \forall y \in K.$$

Ponendo $f(x, y) := \langle D\varphi(x), y - x \rangle$ questo diventa un problema di equilibrio (EP)*.*

1.1 Definizioni preliminari

Richiamiamo nel seguito alcune definizioni basilari sulla convessità e sulla monotonia che utilizzeremo nel corso della trattazione.

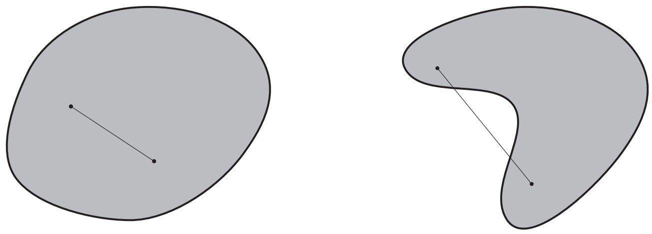

Figura 1.1: Un insieme convesso (a sinistra) e uno non convesso (a destra).

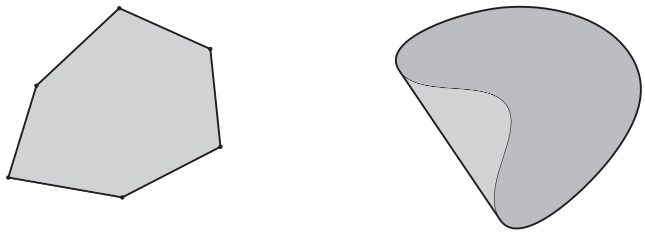

Figura 1.2: L'inviluppo convesso di sei punti (a sinistra) e dell'insieme non convesso della Figura 1.1-destra (a destra).

Concavità e convessità

Definizione 1.1.1. *Un sottoinsieme K di $\mathbb{R}^n$ si dice* **convesso** *se per ogni coppia di punti x e y in K il segmento con estremi x e y è contenuto in K, ossia*

$$\lambda x + (1-\lambda)y \in K, \quad \forall \lambda \in [0,1]$$

(Figura 1.1).

Definizione 1.1.2. *L'***inviluppo convesso** *di un qualsiasi insieme K è l'intersezione di tutti gli insiemi convessi che contengono K (Figura 1.2).*

Osservazione 1.1.3. *Dalla definizione segue che l'inviluppo convesso di un insieme è il più piccolo convesso che contiene l'insieme.*

Definizione 1.1.4. *Una* **combinazione convessa** *di n punti $x_1, \ldots, x_n \in K$ è la somma*

$$\sum_{i=1}^{n} \mu_i y_i \quad con \quad \mu_i \geq 0, \ \forall i \in \{1, \ldots, n\} \quad e \quad \sum_{i=1}^{n} \mu_i = 1,$$

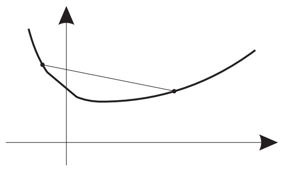

Figura 1.3: Il grafico di una funzione convessa.

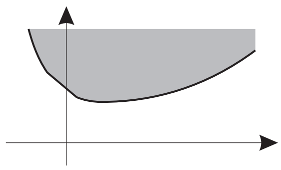

Figura 1.4: L'epigrafico di una funzione convessa.

Osservazione 1.1.5. *L'inviluppo convesso di un insieme K è formato da tutti i punti che sono combinazioni convesse di punti di K.*

Definizione 1.1.6. *Sia K un sottoinsieme convesso di $\mathbb{R}^n$. Una funzione $f\colon K \to \mathbb{R}$ si dice* **convessa** *se per ogni $x, y \in K$ e $\lambda \in [0,1]$ si ha*

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y)$$

(Figura 1.3). Se l'uguaglianza vale solo nei casi $x = y$, $\lambda = 0$ o $\lambda = 1$ allora la funzione f di dice **strettamente convessa**.

La convessità di una funzione è legata alla convessità del suo epigrafico.

Proposizione 1.1.7. *Sia K un sottoinsieme convesso di $\mathbb{R}^n$. Una funzione $f : K \longrightarrow \mathbb{R}$ è convessa se e solo se il suo epigrafico è convesso in $\mathbb{R}^{n+1}$ (Figura 1.4).*

Osservazione 1.1.8. *Un punto di minimo locale di una funzione convessa è anche un minimo globale. Una funzione strettamente convessa ha al più un punto di minimo globale. (Figura 1.5.)*

Definizione 1.1.9. *Una funzione $f : K \longrightarrow \mathbb{R}$ con $K \subseteq \mathbb{R}^n$ si dice* **concava** *se $-f$ è convessa. Essa si dice si dice* **strettamente concava** *se $-f$ è strettamente convessa.*

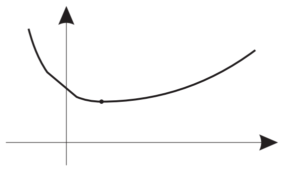

Figura 1.5: Un minimo locale di una funzione convessa è anche globale.

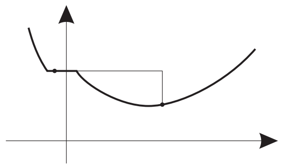

Figura 1.6: Il grafico di una funzione quasi-convessa.

Definizione 1.1.10. *Una funzione* $f : K \longrightarrow \mathbb{R}$ *con* $K \subseteq \mathbb{R}^n$ *si dice* **quasi-convessa** *nel suo dominio* K *se per ogni* $x, y \in K$ *e* $\lambda \in [0,1]$ *si ha*

$$f(\lambda x + (1-\lambda)y) \leq \max\{f(x), f(y)\}$$

(Figura 1.6). Se l'uguaglianza vale solo nei casi $x = y$, $\lambda = 0$ *o* $\lambda = 1$ *allora la funzione* f *di dice* **strettamente quasi-convessa**.

La funzione f *si dice* **quasi-concava** *se* $-f$ *è quasi-convessa. Se l'uguaglianza vale solo nei casi* $x = y$, $\lambda = 0$ *o* $\lambda = 1$ *allora la funzione* f *di dice* **strettamente quasi-concava**.

Definizione 1.1.11. *Una funzione* $f\colon \mathbb{R}^n \to \mathbb{R}$ *si dice* **fortemente convessa** *di modulo* $a > 0$ *su un convesso* $K \subseteq \mathbb{R}^n$ *se per ogni* $x, y \in K$ *e* $\lambda \in [0,1]$ *si ha*

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y) - a\frac{\lambda(1-\lambda)}{2} \parallel x - y \parallel^2 .$$

Monotonia

Definizione 1.1.12. *Una funzione* $f\colon K \times K \to \mathbb{R}$ *si dice* **monotona** *(***strettamente***) se*

$$f(x,y) + f(y,x) \leq 0 \quad \textit{per ogni} \quad x, y \in K$$

$$(f(x,y) + f(y,x) < 0 \quad \textit{per ogni} \quad x \neq y \in K).$$

Una funzione $F\colon K \longrightarrow \mathbb{R}^n$ *si dice* **monotona** *su* K *se lo è la funzione* $f(x, y) = \langle F(x), y - x\rangle$*, ossia se*

$$\langle F(x) - F(y), y - x\rangle \leq 0,\ \forall\ x, y \in K.$$

Per chiarire la definizione precedente, consideriamo una funzione $\varphi\colon\ K \to \mathbb{R}$. Essa è convessa se e solo se $D\varphi$ è monotona, ossia

$$\langle D\varphi(x) - D\varphi(y), x - y\rangle \geq 0, \quad \forall x, y \in K. \tag{1.3}$$

Se $f = \langle D\varphi(x), y - x\rangle$, per la linearità del prodotto scalare, il primo membro della (1.3) diventa

$$\langle D\varphi(x), x - y\rangle - \langle D\varphi(y), x - y\rangle = -f(x, y) - f(y, x)$$

e la (1.3) equivale a

$$-f(x, y) - f(y, x) \geq 0, \quad \forall x, y \in K; \tag{1.4}$$

quindi la monotonia di f deve essere definita tramite la (1.4).

Definizione 1.1.13. *Una funzione* $f : K \times K \longrightarrow \mathbb{R}$ *si dice* **fortemente monotona** *su* $K \subseteq \mathbb{R}^n$*, con modulo* $a > 0$*, se e solo se*

$$f(x, y) + f(y, x) \leq -a\|y - x\|^2$$

Una funzione $F\colon\ K \longrightarrow \mathbb{R}^n$ *si dice* **fortemente monotona, con modulo** $a > 0$**,** *se lo è la funzione* $f(x, y) = \langle F(x), y - x\rangle$*, ossia se*

$$\langle F(x) - F(y), y - x\rangle \leq -a\|y - x\|^2,\ \forall\ x, y \in K.$$

Osservazione 1.1.14. *Se* f *è differenziabile rispetto a* x *(rispettivamente* y*) denotiamo con* f'_x *il gradiente rispetto a* x *(rispettivamente* f'_y *il gradiente rispetto a* y*).*

Esempio 1.1.15. *La funzione* f*, definita per il problema di ottimizzazione, è monotona. Infatti abbiamo*

$$f(x, y) + f(y, x) = \big(\varphi(y) - \varphi(x)\big) + \big(\varphi(x) - \varphi(y)\big) = 0.$$

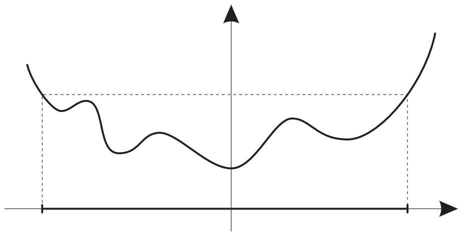

Figura 1.7: Il grafico di una funzione coerciva.

Esempio 1.1.16. *La funzione f, analizzata per il caso dell'ottimizzazione convessa, è monotona poiché l'applicazione $x \longmapsto D\varphi(x)$ è monotona, ossia*

$$\langle D\varphi(x) - D\varphi(y),\ y - x\rangle \leq 0,\ \forall\ x, y \in K.$$

Definizione 1.1.17. *Una funzione $f\colon\ K \times K \to \mathbb{R}$ si dice* **∇-monotona** *se*

$$\langle f'_x(x, y) + f'_y(x, y), y - x\rangle \geq 0,\ \forall\ x, y \in K.$$

Si dice **strettamente ∇-monotona** *se*

$$\langle f'_x(x, y) + f'_y(x, y), y - x\rangle > 0,\ \forall\ x, y \in K\ \text{con}\ x \neq y.$$

Coercività

Definizione 1.1.18. *Una funzione $f\colon\ X \to \mathbb{R}$ con $X \subseteq \mathbb{R}^n$ si dice* **coerciva** *se per ogni $t \in \mathbb{R}$ esiste un sottoinsieme compatto e chiuso K_t di X tale che*

$$\{x \in X | f(x) \leq t\} \subseteq K_t$$

(Figura 1.7).

Diamo una definizione alternativa per la funzione coerciva:

Definizione 1.1.19. *Una funzione $f\colon\ \mathbb{R}^n \to \mathbb{R}$ si dice* **coerciva** *su K se*

$$\lim_{x \in K, \|x\| \longrightarrow +\infty} f(x) = +\infty.$$

Centro di un insieme

Definizione 1.1.20. *Siano K e C sottoinsiemi di $\mathbb{R}^n$ con $C \subset K$. Il* **centro** *di C relativo a K è indicato con $\operatorname{core}_K C$ ed è definito tramite*

$$a \in \operatorname{core}_K C \quad :\Leftrightarrow \quad a \in C \text{ e } C \cap (a, y] \neq \emptyset \ \forall y \in K \setminus C.$$

Un punto appartiene al centro di C in K se è di accumulazione su ogni semiretta che esce da esso e che interseca K in un altro punto.

Osservazione 1.1.21. *Abbiamo $\operatorname{core}_K K = K$.*

1.2 Problemi riconducibili al problema (EP)

Descriviamo altri problemi che possono essere ricondotti al problema (EP) definendo una opportuna funzione f [4].

1.2.1 Ottimo di Pareto

Definizione 1.2.1. *Siano date*

- *m funzioni a valori reali $\psi_1, \ldots, \psi_m\colon\ \mathbb{R}^n \to \mathbb{R}$,*
- *la funzione $\psi = (\psi_1, \ldots, \psi_m)\colon\ \mathbb{R}^n \to \mathbb{R}^m$,*
- *un insieme chiuso $K \subseteq \mathbb{R}^n$.*

Un **minimo paretiano debole** *per la funzione $\psi = (\psi_1, \ldots, \psi_m)$ sull'insieme K è un $\bar{x} \in K$ tale che per ogni $y \in K$ esiste un indice i tale che*

$$\psi_i(y) \geq \psi_i(\bar{x}).$$

Trovare un minimo paretiano debole equivale a risolvere il problema (EP) con la funzione

$$f(x, y) = \max_{i=1,\ldots,m} \bigl\{\psi_i(y) - \psi_i(x)\bigr\}.$$

Infatti, il fatto che $f(\bar{x}, y) \geq 0$ per ogni $y \in K$ è equivalente al fatto che per ogni y esiste i tale che $\psi_i(y) \geq \psi_i(\bar{x})$.

Osservazione 1.2.2. *In questo caso f non è necessariamente monotona.*

1.2.2 Punti di sella

Dati due chiusi $K_1 \subset \mathbb{R}^{n_1}$ e $K_2 \subset \mathbb{R}^{n_2}$, e data $\varphi\colon K_1 \times K_2 \longrightarrow \mathbb{R}$, il punto $(\bar{x}_1, \bar{x}_2) \in \mathbb{R}^{n_1+n_2}$ è chiamato **punto di sella** per φ se

$$(\bar{x}_1, \bar{x}_2) \in K_1 \times K_2, \quad \varphi(\bar{x}_1, y_2) \leq \varphi(y_1, \bar{x}_2), \ \forall\ (y_1, y_2) \in K_1 \times K_2,$$

o, equivalentemente,

$$(\bar{x}_1, \bar{x}_2) \in K_1 \times K_2, \quad \varphi(\bar{x}_1, y_2) \leq \varphi(\bar{x}_1, \bar{x}_2) \leq \varphi(y_1, \bar{x}_2)\ \forall (y_1, y_2) \in K_1 \times K_2.$$

Sia $K := K_1 \times K_2$ e sia $f\colon\ K \times K \longrightarrow \mathbb{R}$ definita da

$$f\big((x_1, x_2), (y_1, y_2)\big) = \varphi(y_1, x_2) - \varphi(x_1, y_2).$$

Il problema (EP) è equivalente alla ricerca di punti di sella, infatti abbiamo $f\big((x_1, x_2), (y_1, y_2)\big) \geq 0$ se e solo se $\varphi(x_1, y_2) \leq \varphi(y_1, x_2)$.

Osservazione 1.2.3. *Anche in questo caso la funzione f è monotona. Infatti abbiamo*

$$\begin{aligned} & f\big((x_1, x_2), (y_1, y_2)\big) + f\big((y_1, y_2), (x_1, x_2)\big) = \\ & \qquad = \big(\varphi(y_1, x_2) - \varphi(x_1, y_2)\big) + \big(\varphi(x_1, y_2) - \varphi(y_1, x_2)\big) = 0. \end{aligned}$$

1.2.3 Equilibrio di Nash in giochi non cooperativi

Sia I un insieme finito di indici che rappresenta l'insieme dei giocatori. Per ogni $i \in I$ sia dato l'insieme $K_i \subseteq \mathbb{R}^m$ delle strategie dell'i-esimo giocatore e sia $K = \prod_{i \in I} K_i$. Per ogni $i \in I$ sia data la funzione $f_i\colon\ K \to \mathbb{R}$ di perdita dell'i-esimo giocatore, dipendente dalle strategie di tutti i giocatori. Per $x = (x_i)_{i \in I} \in K$ definiamo $x^j := (x_i)_{i \in I, i \neq j}$.

Definizione 1.2.4. *Il punto $\bar{x} = (\bar{x}_i)_{i \in I} \in K$ è chiamato* **equilibrio di Nash** *se per ogni $i \in I$ vale*

$$f_i(\bar{x}) \leq f_i\big(x_1, ..., y_i, ..., x_n\big) \quad \forall y_i \in K_i, \tag{1.5}$$

Osservazione 1.2.5. *In un equilibrio di Nash, ogni giocatore non può ridurre la sua perdita variando solo la sua strategia.*

Trovare un equilibrio di Nash è equivalente a risolvere il problema (EP) *con la funzione* $f\colon K \times K \longrightarrow \mathbb{R}$ *data da*

$$f(x,y) := \Sigma_{i\in I}\Big(f_i\big(x_1,...,y_i,...,x_n\big) - f_i(x)\Big),$$

detta **funzione di Nikaido-Isoda** *[24].*

Proposizione 1.2.6. *Un punto* $\bar{x} \in K$ *è un equilibrio di Nash se e solo se* $\bar{x}$ *è soluzione del problema* (EP) *relativo alla funzione di Nikaido-Isoda.*

Dimostrazione. Se la disuguaglianza (1.5) vale per tutti gli $i \in I$, allora è ovvio che $\bar{x}$ è soluzione di (EP). Viceversa, fissato $i \in I$ scegliamo $y \in K$ in modo che $\bar{x}^i = y^i$, si ha

$$f(\bar{x},y) = f_i\big(x_1,...,\bar{x},...,x_n\big) - f_i(\bar{x}).$$

Quindi (EP) implica (1.5) per ogni $i \in I$. □

Osservazione 1.2.7. *In questo caso* f *non è necessariamente monotona.*

1.2.4 Punti fissi

Sia $K \subseteq \mathbb{R}^n$ un chiuso e convesso. Sia data una funzione $T\colon K \longrightarrow K$. Un punto $\bar{x} \in K$ è chiamato **punto fisso** per T se:

$$\bar{x} = T(\bar{x}). \tag{1.6}$$

Trovare un punto fisso per T significa risolvere il problema (EP) con

$$f(x,y) := \langle x - T(x), y - x\rangle.$$

Proposizione 1.2.8. *Il punto* $\bar{x}$ *risolve il problema* (EP) *relativo a* $f(x,y)$ *se e solo se* $\bar{x}$ *è una soluzione di* (1.6).

Dimostrazione. Dimostriamo che un punto fisso è una soluzione del problema (EP). Se $\bar{x} = T(\bar{x})$ abbiamo

$$f(\bar{x},y) = \langle \bar{x} - \bar{x}, y - \bar{x}\rangle = \langle 0, y - \bar{x}\rangle = 0, \quad \forall y \in K.$$

Viceversa, se $\bar{x}$ è soluzione del problema (EP), allora scelto $y := T(\bar{x})$ si ha

$$0 \leq f\big(\bar{x}, T(\bar{x})\big) = \langle \bar{x} - T(\bar{x}), T(\bar{x}) - \bar{x} \rangle = -\|\bar{x} - T(\bar{x})\|,$$

ossia $\bar{x} = T(\bar{x})$. Quindi $\bar{x}$ è un punto fisso per T. □

Osservazione 1.2.9. *In questo caso f è monotona se e solo se*

$$\langle T(x) - T(y), x - y \rangle \leq \|x - y\|^2 \quad \forall x, y \in K.$$

Infatti abbiamo

$$\begin{aligned} f(x,y) + f(y,x) &= \langle x - T(x), y - x \rangle + \langle y - T(y), x - y \rangle = \\ &= \langle T(x) - T(y), x - y \rangle - \|x - y\|^2. \end{aligned}$$

Osservazione 1.2.10. *Se $T\colon\ K \rightrightarrows K$ è una funzione a più valori con valori compatti allora trovare $\bar{x} \in K$ tale che $\bar{x} \in T(\bar{x})$ significa risolvere il problema* (EP) *con $f(x,y) := \max_{u \in T(x)} \langle x - u, y - x \rangle$. Infatti, se $\bar{x} \in T(\bar{x})$ si ha $f(\bar{x}, y) \geq \langle \bar{x} - \bar{x}, y - \bar{x} \rangle = 0$ per ogni $y \in K$.*

1.2.5 Disequazioni variazionali

Definizione 1.2.11. *Sia $K \subseteq \mathbb{R}^n$ un insieme e $F : K \longrightarrow \mathbb{R}^n$ una funzione. Risolvere una* **disequazione variazionale** *consiste nel*

$$\textit{trovare} \quad \bar{x} \in K \quad \textit{tale che} \quad \langle F(\bar{x}), y - \bar{x} \rangle \geq 0, \quad \forall y \in K. \tag{VI}$$

Osservazione 1.2.12. *Ovviamente, un punto $\bar{x}$ è una soluzione della disequazione variazionale* (VI) *se e solo se è una soluzione del problema* (EP) *relativo alla funzione*

$$f(x,y) := \langle F(x), y - x \rangle.$$

Interpretazione geometrica della soluzione di una disequazione variazionale Dal punto di vista geometrico, risolvere la disequazione variazionale (VI) consiste nel trovare un punto $\bar{x}$ tale che $F(\bar{x})$ forma un angolo non ottuso con ogni vettore $y - \bar{x}$ al variare di y in K.

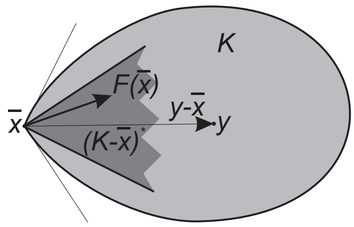

Figura 1.8: Interpretazione geometrica di una soluzione di una disequazione variazionale. L'insieme K è colorato in grigio chiaro, mentre il cono duale di $K - \bar{x}$ è colorato in grigio scuro.

Identificando gli elementi del duale di $\mathbb{R}^n$ con gli elementi di $\mathbb{R}^n$ stesso, tramite l'applicazione $\mathbb{R}^n \ni v \mapsto \langle v, \cdot \rangle \in \mathbb{R}^n$ abbiamo che la disequazione variazionale (VI) è equivalente al problema

$$\text{trovare } \bar{x} \in K \text{ tale che } F(\bar{x}) \in (K - \bar{x})^*,$$

dove K^* è il cono polare di K (Definizione A.0.3). Nella Figura 1.8 abbiamo mostrato un esempio in $\mathbb{R}^2$.

Disequazione variazionale per funzioni a più valori Analizziamo le disequazioni variazionali nel caso in cui F sia una funzione a più valori.

Sia $F\colon K \rightrightarrows \mathbb{R}^n$ una funzione a più valori, con $F(x)$ compatto, convesso e non vuoto per tutti gli $x \in \mathbb{R}^n$. Il problema consiste nel trovare

$$\bar{x} \in K, \quad \bar{\xi} \in F(\bar{x}), \ \textit{tale che} \ \langle \bar{\xi}, y - \bar{x} \rangle \geq 0 \ \forall y \in K. \tag{1.7}$$

Consideriamo la funzione

$$f(x, y) := \max_{\zeta \in F(x)} \langle \zeta, y - x \rangle.$$

Proposizione 1.2.13. *Il punto $\bar{x}$ risolve il problema* (EP) *relativo alla funzione f se e solo se esiste $\bar{\xi}$ tale che $\bar{x}$ e $\bar{\xi}$ risolvono il problema* (1.7).

Per dimostrare la proposizione abbiamo bisogno di un lemma.

Lemma 1.2.14. *Sia D un insieme convesso e compatto. Sia K un insieme convesso. Sia $p\colon D \times K \longrightarrow \mathbb{R}$ un funzione concava, semicontinua superiormente nel primo argomento, e convessa nel secondo argomento. Assumiamo* *che*

$$\max_{\xi \in D} p(\xi, y) \geq 0, \quad \forall y \in K.$$

Allora esiste $\bar{\xi} \in D$ *tale che*

$$p(\bar{\xi}, y) \geq 0, \quad \forall y \in K.$$

Dimostrazione. Supponiamo per assurdo che la tesi sia falsa. Allora per ogni $\zeta \in D$ esistono $y \in K$ e $\varepsilon \geq 0$ tali che $p(\zeta, y) < -\varepsilon$.

Sia

$$S(y, \varepsilon) := \{\zeta \in D | p(\zeta, y) < -\varepsilon\} \qquad (y \in K, \varepsilon > 0)$$

un ricoprimento aperto dell'insieme D. Per la compattezza di D esiste un sottoricoprimento finito

$$S(y_i, \varepsilon_i) := \{\zeta \in D \mid p(\zeta, y_i) < -\varepsilon_i\} \quad i = 1, \ldots, n.$$

Sia $\varepsilon := \min_i \varepsilon_i$; allora da $D \subset \bigcup_i S(y_i, \varepsilon)$ segue che

$$\min_i p(\zeta, y_i) \leq -\varepsilon \quad \forall y \in D.$$

Visto che le funzioni $p(\cdot, y_i)$ sono concave, segue da un risultato standard di analisi convessa che esistono numeri reali

$$\mu_i \geq 0 \ (i = 1, ..., m) \text{ tali che } \sum_i \mu_i p(\zeta_i, y_i) \leq -\varepsilon \ \forall \zeta \in D.$$

Allora si ha $\max_{\zeta \in D} p(\zeta, \bar{y}) < 0$, contraddicendo l'ipotesi. □

Ora possiamo dimostrare la proposizione.

Dimostrazione della Proposizione 1.2.13. Ovviamente, se $\bar{x}$ e $\bar{\xi}$ risolvono la disequazione variazionale (1.7), allora abbiamo $\max_{\zeta \in F(x)} \langle \zeta, y - x \rangle \geq 0$ e quindi $\bar{x}$ è soluzione del problema (EP).

Viceversa, se $\bar{x}$ è soluzione del problema (EP), applichiamo il Lemma 1.2.14 con

$$D := F(\bar{x}) \quad \text{e} \quad p(x, y) := \langle \xi, y - \bar{x} \rangle,$$

ottenendo che esiste $\bar{\xi} \in F(\bar{x})$ tale che $\langle \bar{\xi}, y - \bar{x} \rangle \geq 0$ per ogni $y \in K$, e quindi $\bar{x}$ e $\bar{\xi}$ risolvono la disequazione variazionale (1.7). □

Proposizione 1.2.15. *La funzione f è monotona se e solo se la mappa F è monotona, nel senso che*

$$\langle \eta - \xi, y - x \rangle \geq 0 \quad \forall x \in K,\ y \in K,\ \xi \in F(x),\ \eta \in F(y).$$

Dimostrazione. Abbiamo

$$\begin{aligned} f(x,y) + f(y,x) &= \max_{\xi \in F(x)} \langle \xi, y - x \rangle + \max_{\eta \in F(x)} \langle \eta, x - y \rangle = \\ &= \max_{\xi \in F(x), \eta \in F(y)} \langle \xi - \eta, y - x \rangle \leq \\ &\leq 0, \end{aligned}$$

che equivale a

$$\langle \eta - \xi, y - x \rangle \geq 0 \quad \forall x \in K,\ y \in K,\ \xi \in F(x),\ \eta \in F(y).$$

□

1.2.6 Problemi di complementarità

L'esempio seguente è un caso speciale di disequazione variazionale. Sia K un cono convesso chiuso. Sia $F\colon K \to \mathbb{R}^n$ una mappa data. Risolvere un problema di complementarità consiste nel cercare $\bar{x} \in \mathbb{R}^n$ tale che

$$\bar{x} \in K, \quad F(\bar{x}) \in K^*, \quad \langle F(\bar{x}), \bar{x} \rangle = 0, \tag{1.8}$$

dove K^* è il cono polare di K (Definizione A.0.3).

Proposizione 1.2.16. *Il problema di complementarità* (1.8) *è equivalente alla disequazione variazionale* (VI)*.*

Prima di dimostrare la proposizione notiamo che la disequazione variazionale (VI) è equivalente a un problema (EP), come abbiamo descritto nell'Osservazione 1.2.12, quindi anche il problema di complementarità (1.8) è equivalente a un problema (EP).

Dimostrazione della Proposizione 1.2.16. Supponiamo che $\bar{x}$ sia soluzione della disequazione variazionale (VI). Scegliendo nella (VI) $y := 2\bar{x}$ si ottiene $\langle F(\bar{x}), \bar{x} \rangle \geq 0$, mentre scegliendo $y := 0$ si ottiene $\langle F(\bar{x}), \bar{x} \rangle \leq 0$; quindi si ha

$\langle F(\bar{x}), \bar{x} \rangle = 0$. Sempre dalla (VI) si ottiene $\langle F(\bar{x}), y \rangle \geq 0$ per ogni $y \in K$, cioè $F(\bar{x}) \in K^*$. Quindi $\bar{x}$ è soluzione del problema (1.8).

Viceversa, se $\bar{x}$ è soluzione del problema (1.8) otteniamo

$$\langle F(\bar{x}), y - \bar{x} \rangle = \langle F(\bar{x}), y \rangle + \langle F(\bar{x}), \bar{x} \rangle = \langle F(\bar{x}), y \rangle,$$

che è non negativo perché $F(\bar{x}) \in K^*$, e quindi $\bar{x}$ è soluzione della disequazione variazionale (VI). $\square$

1.2.7 Problemi di ottimizzazione inversa

Siano dati un insieme chiuso $K \subset \mathbb{R}^n$, m funzioni $f_i \colon \mathbb{R}^n \to \mathbb{R}$, con $i = 1, \ldots, m$ e p funzioni $g_j \colon \mathbb{R}^n \to \mathbb{R}$, con $j = 1, \ldots, p$. Il problema di ottimizzazione inversa richiede di determinare un parametro $\lambda^* \in \mathbb{R}^m_+$ tale che almeno una soluzione ottima $\bar{x}$ del problema di minimo

$$\min_{x \in K} \left\{ \sum_{i=1}^{m} \lambda_i^* f_i(x) \right\}$$

soddisfi i vincoli

$$g_j(\bar{x}) \leq 0 \quad \forall j = 1, \ldots, p.$$

Ciò è equivalente al problema di Nash con tre giocatori in cui

- il primo giocatore controlla le variabili x e ha l'obiettivo di risolvere

$$\min_{x \in K} \left\{ \sum_{i=1}^{m} \lambda_i f_i(x) \right\},$$

- il secondo giocatore controlla le variabili ausiliari y e si pone l'obiettivo di risolvere

$$\max_{y \geq 0} \left\{ \sum_{j=1}^{p} g_j(x) y_i \right\},$$

- il terzo giocatore sceglie un vettore $\lambda \in \mathbb{R}^m_-$ o equivalentemente minimizza una funzione obiettivo costante su $\mathbb{R}^m_+$.

Per questo motivo, un problema di ottimizzazione inversa può essere scritto come un problema (EP) con la funzione di Nikaido-Isoda (Proposizione 1.2.6).

1.3 Risultati di esistenza

1.3.1 I risultati classici

In questo paragrafo enunceremo il teorema di Ky-Fan [11] (di cui non daremo la dimostrazione), che garantisce l'esistenza della soluzione per il problema di equilibrio e ci sarà utile per i risultati che si avranno nel secondo e nel terzo capitolo. Prima di enunciare il teorema di Ky-Fan riportiamo il seguente lemma propedeutico di Knaster, Kuratowski e Mazurkiewicz [17] .

Lemma 1.3.1 (KKM)**.** *Sia X un sottoinsieme di $\mathbb{R}^n$. Per ogni $y \in Y$ sia dato un sottoinsieme $S(y)$ di X tale che $S(y)$ sia compatto per almeno un $\bar{y} \in Y$. Se l'inviluppo convesso di ogni sottoinsieme finito $\{y_1, y_2, \dots, y_n\}$ di Y è contenuto nella corrispondente unione $\bigcup_{i=1}^{n} S(y_i)$, allora l'intersezione $\bigcap_{y \in Y} S(y)$ è non-vuota.*

Il Lemma KKM (Lemma 1.3.1) è molto utile per garantire una condizione sufficiente per l'esistenza di soluzioni del problema (EP). Infatti, se consideriamo la multifunzione

$$S\colon X \rightrightarrows \mathbb{R}^n$$

data da

$$S(y) = \{x \in X | f(x, y) \geq 0\},$$

l'insieme delle soluzioni del problema (EP), ossia degli $x \in X$ tali che $f(x, y) \geq 0$ per ogni $x \in X$, coincide con l'intersezione

$$S = \bigcap_{y \in X} S(y).$$

Ci siamo quindi ridotti a cercare ipotesi su f e su X che siano sufficienti per applicare il Lemma KKM (Lemma 1.3.1). Seguendo questa idea possiamo dimostrare il seguente teorema.

Teorema 1.3.2. *Sia X un sottoinsieme di $\mathbb{R}^n$ tale che*

- *X è convesso,*
- *X è limitato,*

- $f(x, \cdot)$ *è convessa per ogni* $x \in X$,
- $f(\cdot, y)$ *è continua per ogni* $y \in X$.

Allora S *è non vuoto.*

Dimostrazione. Visto che X è limitato e $f(\cdot, y)$ è continua per ogni $y \in X$, abbiamo che $S(y)$ è compatto per ogni $y \in X$ fissato. Da ciò deduciamo che $S(y)$ è chiuso per ogni $y \in X$ e che esiste $\bar{y} \in X$ tale che $S(\bar{y})$ è compatto.

Dal fatto che X è convesso e che $f(x, \cdot)$ è convessa per ogni $x \in X$ deduciamo che l'inviluppo convesso di ogni sottoinsieme finito $\{y_1, y_2, \ldots, y_n\}$ di X è contenuto nella corrispondente unione $\bigcup_{i=1}^n S(y_i)$.

Le ipotesi del Lemma KKM (Lemma 1.3.1) sono soddisfatte quindi S non è vuoto. □

Grazie al Lemma 1.3.1 abbiamo delle ipotesi su f e su X che ci permettono di scrivere il Teorema di Ky-Fan per il problema di equilibrio.

Teorema 1.3.3 (Teorema di min-max di Ky Fan). *Sia* $K \subset \mathbb{R}^n$ *convesso e compatto e sia* $f\colon K \times K \to \mathbb{R}$ *tale che*

- *per ogni* $x \in K$ *la funzione* $f(x, \cdot)$ *è semicontinua superiormente e quasi-concava,*
- *per ogni* $y \in K$ *la funzione* $f(\cdot, y)$ *è semicontinua inferiormente e quasi-convessa.*

Allora esiste un punto di sella per f, *ossia esiste* $(\bar{x}, \bar{y}) \in K \times K$ *tale che*

$$f(\bar{x}, \bar{y}) = \min_{x \in K} \max_{y \in K} f(x, y) = \max_{y \in K} \min_{x \in K} f(x, y).$$

1.3.2 Un secondo risultato di esistenza

In questa sezione, dimostriamo l'esistenza di un risultato base per i problemi di equilibrio (EP) nel caso in cui

$$f(x, y) = g(x, y) + h(x, y)$$

dove g è una funzione monotona e soddisfa una condizione di semicontinuità superiore nel primo argomento, mentre h non è necessariamente monotona, ma deve soddisfare una condizione di semicontinuità superiore nel secondo argomento [4]. Se $g = 0$, allora il risultato diventa una variante del Teorema di min-max di Ky Fan [11].

Enunciamo il teorema.

Teorema 1.3.4. *Supponiamo che le seguenti assunzioni siano soddisfatte:*

(i) $K \subset \mathbb{R}^n$ *è un insieme chiuso, convesso e non vuoto;*

(ii) $g\colon K \times K \longrightarrow \mathbb{R}$ *gode delle seguenti proprietà:*

- *(a)* $g(x,x) = 0$ *per ogni* $x \in K$*,*
- *(b)* $g(x,y) + g(y,x) \leq 0$ *per ogni* $x, y \in K$ *(monotona),*
- *(c) per tutti gli* $x, y \in K$ *la funzione*

$$[0,1] \ni t \longmapsto g(ty + (1-t)x, y) \in \mathbb{R}$$

  è semicontinua superiormente in $t = 0$*,*
- *(d)* g *è convessa e semicontinua inferiormente nel secondo argomento;*

(iii) $h\colon K \times K \longrightarrow \mathbb{R}$ *ha le seguenti proprietà:*

- *(a)* $h(x,x) = 0$ *per ogni* $x \in K$*,*
- *(b)* h *è semicontinua superiormente nel primo argomento,*
- *(c)* h *è convessa nel secondo argomento;*

(iv) esiste $C \subset K$ *compatto, convesso e non vuoto tale che per ogni* $x \in C \setminus \operatorname{core}_K C$ *esiste* $a \in \operatorname{core}_K C$ *tale che* $g(x,a) + h(x,a) \leq 0$*.*

Allora esiste $\bar{x} \in C$ *tale che* $0 \leq g(\bar{x}, y) + h(\bar{x}, y)$ *per ogni* $y \in K$*.*

Dimostriamo il seguente lemma che ha le stesse ipotesi del teorema precedente.

Lemma 1.3.5. *Supponiamo che le assunzioni del Teorema 1.3.4 siano soddisfatte. Allora esiste* $\bar{x} \in C$ *tale che* $g(y, \bar{x}) \leq h(\bar{x}, y)$ *per ogni* $y \in C$*.*

Dimostrazione. Consideriamo gli insiemi chiusi $S(y) := \{x \in C | g(y,x) \leq h(x,y)\}$ al variare di $y \in C$ e dimostriamo che $\cap_{y \in C} S(y) \neq \emptyset$. Siano y_i, con $i \in N$, un sottoinsieme finito di C. Sia I un sottoinsieme non vuoto di N. Sia ζ un elemento dell'inviluppo convesso $\mathrm{conv}\{y_i \mid i \in I\}$. Allora si ha

$$\zeta = \sum_{i \in I} \mu_i y_i \quad \text{con} \quad \mu_i \geq 0, \ \forall i \in I \quad \text{e} \quad \sum_{i \in I} \mu_i = 1.$$

Assumiamo, per assurdo, che

$$g(y_i, \zeta) > h(\zeta, y_i) \quad \forall i \in I. \tag{1.9}$$

Da ciò segue

$$\sum_{i \in I} \mu_i g(y_i, \zeta) > \sum_{i \in I} \mu_i h(\zeta, y_i) \tag{1.10}$$

visto che i μ_i non possono essere tutti contemporaneamente nulli. D'altra parte, dalla convessità e dalla monotonia di g segue

$$\begin{aligned} \sum_{i \in I} \mu_i g(y_i, \zeta) &\leq \sum_{i \in I} \sum_{j \in I} \mu_i \mu_j g(y_i y_j) = \\ &= \frac{1}{2} \sum_{i,j \in I} \mu_i \mu_j \left(g(y_i, y_j) + g(y_j, y_i) \right) \leq \\ &\leq 0 \end{aligned}$$

e dalla convessità e dalla (iii-a) segue:

$$0 = h(\zeta, \zeta) \leq \sum_{i \in I} \mu_i h(\zeta, y_i).$$

Allora si ha

$$\sum_{i \in I} \mu_i g(y_i, \zeta) \leq 0 \leq \sum_{i \in I} \mu_i h(\zeta, y_i),$$

che contraddice la (1.10). Allora la (1.9) non può essere vera, e quindi abbiamo $g(y_i, \zeta) \leq h(\zeta, y_i)$ per qualche $i \in I$ e quindi ζ appartiene a $S(y_i)$ per qualche $i \in I$.

Essendo $\zeta \in conv\{y_i | i \in I\}$ arbitrario, segue che

$$conv\{y_i | i \in I\} \subset \bigcup_{i \in I} \{S(y_i)\}.$$

Visto che questo è vero per tutti gli insiemi non vuoti $I \subset N$ e che gli insiemi $S(y_i)$ sono chiusi, segue dalla versione standard del Lemma KKM (Lemma 1.3.1) che $\bigcap_{i \in N} S(y_i) \neq \varnothing$.

Abbiamo quindi dimostrato che ogni sottofamiglia finita della famiglia $\{S(y)\}_{y \in C}$ ha intersezione non vuota. Visto che questi insiemi sono sottoinsiemi chiusi dell'insieme compatto C, segue che la famiglia intera ha intersezione non vuota. Allora

$$\bigcap_{y \in C} S(y) \neq \emptyset.$$

$\square$

Lemma 1.3.6. *Dato $\bar{x} \in C$ e g monotona, le seguenti asserzioni sono equivalenti:*

(a) per ogni $y \in C$ si ha $g(y, \bar{x}) \leq h(\bar{x}, y)$,

(b) per ogni $y \in C$ si ha $0 \leq g(\bar{x}, y) + h(\bar{x}, y)$.

Dimostrazione. Supponiamo che (b) sia vera e dimostriamo (a). Poiché g è monotona per ipotesi, abbiamo $g(\bar{x}, y) \leq -g(y, \bar{x})$. Da ciò segue ovviamente (a).

Supponiamo vera (a) e dimostriamo (b). Sia $y \in C$ arbitrario e sia

$$x_t := ty + (1-t)\bar{x} \quad 0 \leq t \leq 1.$$

Allora $x_t \in C$ e allora da (a) si ottiene

$$g(x_{t,\bar{x}}) \leq h(\bar{x}, x_t).$$

Dalle proprietà di g e h segue che, per ogni t,

$$\begin{aligned}
0 = g(x_t, x_t) &\leq \\
&\leq tg(x_t, y) + (1-t)g(x_t, \bar{x}) \leq \\
&\leq tg(x_t, y) + (1-t)h(\bar{x}, x_t) \leq \\
&\leq tg(x_t, y) + (1-t)\left(th(\bar{x}, y) + (1-t)h(\bar{x}, \bar{x})\right) = \\
&= tg(x_t, y) + (1-t)th(\bar{x}, y).
\end{aligned}$$

Dividendo per t otteniamo

$$0 \leq g(x_t, y) + (1-t)h(\bar{x}, y).$$

Per $t \to 0$ si ha $x \to \bar{x}$; usando la monotonia di g e considerando il limite per $t \to 0$, otteniamo

$$0 \leq g(\bar{x}, y) + h(\bar{x}, y),$$

ossia (b). □

Lemma 1.3.7. *Assumiamo che $\psi\colon K \longrightarrow \mathbb{R}$ è convessa. Sia $x_0 \in \operatorname{core}_K C$ tale che $\psi(x_0) \leq 0$ e supponiamo che $\psi(y) \geq 0$ per ogni $y \in C$. Allora si ha che $\psi(y) \geq 0$ per ogni $y \in K$.*

Dimostrazione. Assumiamo per assurdo che esista $y \in K \setminus C$ tale che $\psi(y) < 0$. Allora $\psi(\eta) < 0$ per ogni $\eta \in (x_0, y]$ e quindi esiste $\eta \in C$ con $\psi(\eta) < 0$: ciò contraddice l'ipotesi. □

Ora possiamo dimostrare il Teorema 1.3.4.

Dimostrazione del Teorema 1.3.4. Dal Lemma 1.3.5 otteniamo $\bar{x} \in C$ con $g(y, \bar{x}) \leq h(\bar{x}, y)$ per ogni $y \in C$. Dal Lemma 1.3.6 segue

$$0 \leq g(\bar{x}, y) + h(\bar{x}, y) \quad \forall y \in C.$$

Definiamo

$$\psi(\cdot) := g(\bar{x}, \cdot) + h(\bar{x}, \cdot).$$

Allora $\psi(\cdot)$ è convessa e $\psi(y) \geq 0$ per ogni $y \in C$. Se $\bar{x} \in \operatorname{core}_K C$, allora definiamo $x_0 := \bar{x}$. Se $\bar{x} \in C \setminus \operatorname{core}_K C$, allora per l'ipotesi (iv) esiste $x_0 \in \operatorname{core}_K C$ tale che $g(\bar{x}, x_0) + h(\bar{x}, x_0) \leq 0$. In entrambi i casi si ha

$$x_0 \in \operatorname{core}_K C \quad \text{e} \quad \psi(x_0) \leq 0.$$

Allora segue del Lemma 1.3.7 che $\psi(y) \geq 0$ per ogni $y \in K$, cioè

$$g(\bar{x}, y) + h(\bar{x}, y) \geq 0 \quad \forall y \in K.$$

□

Osservazione 1.3.8. *Se K è compatto, allora la condizione di coercività (iv) è banalmente soddisfatta con $C = K$, infatti in tal caso si ha*

$$C \setminus \operatorname{core}_K C = \emptyset.$$

Indichiamo ora alcune condizioni sufficienti affinché sia soddisfatta l'ipotesi (iv) del Teorema 1.3.4. Tutte sono soddisfatte se K è chiuso e limitato. Sia $a \in K$ un elemento fissato.

Proposizione 1.3.9. *Assumiamo*

$$g(x,a) + h(x,a) \longrightarrow -\infty \quad \textit{se} \quad \|x-a\| \longrightarrow +\infty \ (x \in K). \tag{1.11}$$

Allora l'ipotesi (iv) del Teorema 1.3.4 è soddisfatta.

Prima di dimostrare la proposizione notiamo che l'ipotesi della proposizione è verificata se g è fortemente monotona e h è identicamente nulla.

Dimostrazione. Siano $r > 0$ sufficientemente grande e $x \in K$ tali che

$$\|x-a\| = r \text{ e } g(x,a) + h(x,a) \leq 0.$$

Sia $C := \{x \in K| \quad \|x-a\| \leq r\}$. Allora C è compatto e $a \in \operatorname{core}_K C$. Inoltre abbiamo

$$x \in C \setminus \operatorname{core}_K C \quad \Longrightarrow \quad \|x-a\| = r.$$

Allora abbiamo

$$g(x,a) + h(x,a) \leq 0 \quad \forall x \in C \setminus \operatorname{core}_K C$$

e l'ipotesi (iv) del Teorema 1.3.4 è soddisfatta. □

Proposizione 1.3.10. *Assumiamo*

$$\begin{cases} h(x,a) \leq M\|x-a\| & \forall x \in K \ \textit{con} \ \|x-a\| \geq r \\ \frac{g(x,a)}{\|x-a\|} \longrightarrow -\infty & \textit{se} \quad \|x-a\| \longrightarrow +\infty \ (x \in K). \end{cases} \tag{1.12}$$

Allora l'ipotesi (iv) del Teorema 1.3.4 è soddisfatta.

Dimostrazione. Dalla (1.12) segue

$$\frac{g(x,a)+h(x,a)}{\|x-a\|} \longrightarrow -\infty \quad \text{se} \quad \|x-a\| \longrightarrow +\infty \ (x \in K)$$

e questo implica la (1.11), ossia la tesi per la Proposizione 1.3.9. □

Proposizione 1.3.11. *Assumiamo*

$$\frac{h(x,a)}{\|x-a\|} \longrightarrow -\infty \quad \textit{se} \quad \|x-a\| \longrightarrow +\infty \ (x \in K). \tag{1.13}$$

Allora l'ipotesi (iv) del Teorema 1.3.4 è soddisfatta.

Dimostrazione. Sia

$$B := \{x \in K | \|x-a\| \leq 1\}.$$

L'insieme B è compatto e $g(x,\cdot)$ è semicontinua inferiormente; allora esiste un numero reale M tale che $M \leq g(a,x)$ per ogni $x \in B$. Da $g(a,a)=0$ e dalla convessità di $g(a,\cdot)$ segue

$$M \leq \frac{g(a,x)}{\|x-a\|} \quad \forall x \in K \text{ con } \|x-a\| \geq 1. \tag{1.14}$$

Inoltre dalla (1.13) segue

$$\frac{-g(a,x)+h(x,a)}{\|x-a\|} \longrightarrow -\infty \quad \text{se} \quad \|x-a\| \longrightarrow +\infty \ (x \in K).$$

Visto che g è monotona si ha

$$g(x,a)+h(x,a) \leq -g(a,x)+h(x,a).$$

Ciò implica la (1.11), ossia la tesi per la Proposizione 1.3.9. □

Proposizione 1.3.12. *Assumiamo*

$$\begin{cases} h(x,a) \leq M\|x-a\| \quad \forall x \in K \ \textit{con} \ \|x-a\| \geq r \\ \frac{g(x,a)+g(a,x)}{\|x-a\|} \longrightarrow -\infty \quad \textit{se} \quad \|x-a\| \longrightarrow +\infty \ (x \in K). \end{cases} \tag{1.15}$$

Allora l'ipotesi (iv) del Teorema 1.3.4 è soddisfatta.

Prima di dimostrare la proposizione notiamo che l'ipotesi della proposizione è verificata se g è fortemente monotona e h è identicamente nulla.

Dimostrazione. La (1.14), con la (1.15), implica

$$\frac{g(x,a)}{\|x-a\|} \longrightarrow -\infty \quad \text{se} \quad \|x-a\| \longrightarrow +\infty \quad x \in K.$$

Con la prima condizione della (1.15) questo implica la (1.12), ossia la tesi per la Proposizione 1.3.10. □

Sottolineiamo il fatto che la condizione di coercività (iv) può essere sostituita, ossia vale il seguente risultato.

Teorema 1.3.13. *Supponiamo che siano soddisfatte le assunzioni (i), (ii) e (iii) del Teorema 1.3.4. Supponiamo che la seguente assunzione sia soddisfatta:*

(iv') Esiste $B \subset K$ compatto, convesso e diverso dal vuoto, tale che per ogni $x \in K \setminus B$ esiste $a \in B$ con

$$g(x,a) + h(x,a) < 0. \tag{1.16}$$

Allora esiste $\bar{x} \in B$ tale che $0 \leq g(y,\bar{x}) + h(\bar{x},y)$ per ogni $y \in B$.

Dimostrazione. Sia $\{y_i\}_{i\in N}$ un sottoinsieme finito di K. Sia $C := \mathrm{conv}\{B, \bigcup_{i\in N} y_i\}$ con C convesso e compatto. Per i risultati del Teorema 1.3.4 esiste $x \in C$ tale che $0 \leq g(x,y) + h(x,y)$ per ogni $y \in C$. Per l'ipotesi (iv') otteniamo $x \in B$. Dalla monotonia di g segue $g(x,y) \leq h(x,y)$ per ogni $y \in C$ e, in particolare, $g(y_i,x) \leq h(x,y_i)$ per ogni $i \in N$. Quindi ogni sottofamiglia finita della famiglia scelta degli insiemi chiusi

$$S(y) := \{x \in B | g(y,x) \leq h(x,y)\} \quad (y \in K)$$

ha intersezione non vuota ed, essendo B compatto, si ha

$$\bigcap_{y\in K} S(y) \neq \varnothing.$$

Ossia esiste $\bar{x} \in B$ tale the

$$g(y,\bar{x}) \leq h(\bar{x},y) \quad \forall y \in K.$$

Dal Lemma 1.3.6 (dove C può essere un qualsiasi sottoinsieme convesso di K) segue

$$0 \leq g(\bar{x},y) + h(\bar{x},y) \quad \forall y \in K.$$

Ossia la conclusione del teorema. □

Corollari per il problema (EP) Enunciamo due corollari dei risultati appena ottenuti.

Corollario 1.3.14. *Il problema di equilibrio* (EP) *relativo a una funzione f fortemente monotona sull'insieme K ammette una soluzione.*

Dimostrazione. La funzione f può essere scritta come $g+h$ con $g=f$ fortemente monotona e h identicamente nulla. La Proposizione 1.3.12 implica che il problema di equilibrio (EP) relativo alla funzione $g+h$ ammette un'unica soluzione, infatti le due ipotesi sono soddisfatte:

- la condizione

  $$h(x,a) \leq M\|x-a\| \quad \forall x \in K \text{ con } \|x-a\| \geq r$$

  è verificata perché h è identicamente nulla;

- la condizione

  $$\frac{g(x,a)+g(a,x)}{\|x-a\|} \longrightarrow -\infty \quad \text{se} \quad \|x-a\| \longrightarrow +\infty \ (x \in K)$$

  è verificata perché dalla monotonia di f, $f(x,a)+f(a,x) \leq -b\|a-x\|^2$ con $b>0$, si deduce

  $$\frac{g(x,a)+g(a,x)}{\|x-a\|} \leq -b\|a-x\|$$

  e perché $-b\|x-a\| \longrightarrow -\infty \quad$ se $\quad \|x-a\| \longrightarrow +\infty$ $(x \in K)$.

□

Corollario 1.3.15. *Il problema di equilibrio* (EP) *relativo a una funzione f su un compatto K in $\mathbb{R}^n$ ammette soluzione.*

Dimostrazione. Visto che l'insieme K è compatto, l'ipotesi (iv') del Teorema 1.3.13 è verificata scegliendo $B=K$, quindi si ottiene la tesi. □

Capitolo 2

Funzioni gap

In questo capitolo introdurremo le funzioni gap, che permettono di studiare il problema di equilibrio. Per quanto detto nel capitolo precedente le funzioni gap ci consentono di studiare in modo alternativo anche tutti i problemi equivalenti a quello di equilibrio. Con l'introduzione non solo della funzione gap, ma anche di un problema ausiliario e della sua formulazione minimax è possibile esprimere il problema di equilibrio (EP) tramite la minimizzazione di una funzione gap differenziabile. La differenziabilità della funzione gap è necessaria per poter applicare gli algoritmi di discesa analizzati nel prossimo capitolo. Descriveremo anche la relazione tra le funzioni gap e i principi variazionali. Infine per completezza, daremo la definizione delle funzioni D-gap, che permettono di formulare il problema di equilibrio con problemi di minimo non vincolato.

L'insieme K su cui sono definite le funzioni è supposto compatto. Per poter considerare funzioni differenziabili su K, avremmo bisogno di un dominio aperto. Per ovviare a questo problema, ogni volta che considereremo una funzione differenziabile f con dominio compatto K (risp. $K \times K$), supporremo (senza dirlo esplicitamente) che la funzione f è definita su V (risp. $V \times V$), dove V è un aperto che contiene K.

2.1 Definizione ed esempi

Definizione 2.1.1. *Sia $K \subseteq \mathbb{R}^n$. La funzione $p\colon \mathbb{R}^n \to \mathbb{R}$ è una* **funzione gap** *per il problema* (EP)

$$\bar{x} \in K, \quad f(\bar{x}, y) \geq 0 \ \forall y \in K \tag{EP}$$

se valgono le due condizioni seguenti:

i) $p(x) \geq 0$ *per ogni* $x \in K$,

ii) $p(x) = 0$ *e* $x \in K$ *se e solo se* x *è una soluzione del problema* (EP).

Osservazione 2.1.2. *Un punto è una soluzione di un problema* (EP) *se e solo se è un valore del minimo per una (qualsiasi) funzione gap e il corrispondente valore minimo è uguale a zero.*

Prima di considerare un esempio di funzione gap dimostriamo l'equivalenza tra il problema di equilibrio (EP) e un problema di minimax attraverso il seguente lemma tratto da [20].

Il problema di minimax che consideriamo è

$$\min_{x \in K} \sup_{y \in K} [-f(x, y)], \tag{2.1}$$

e dedurremo che il problema di equilibrio (EP) ammette soluzione se e solo se il valore dell'ottimo di (2.1) è zero. Questo risultato ci permetterà di considerare come funzione gap

$$p(x) = \sup_{y \in K} [-f(x, y)] \tag{2.2}$$

la cui minimizzazione sull'insieme K coincide con il problema (2.1).

Lemma 2.1.3. *Supponiamo che $f(x, x) = 0$ per ogni $x \in K$. Allora le affermazioni seguenti sono equivalenti:*

i) esiste $\bar{x} \in K$ tale che $f(\bar{x}, y) \geq 0$ per ogni $y \in K$;

ii) esiste $\bar{x} \in K$ soluzione del problema

$$\min_{y \in K} f(\bar{x}, y); \tag{2.3}$$

iii) si ha $\min_{x\in K}\sup_{y\in K}[-f(x,y)] = 0$.

Dimostrazione. Dimostriamo l'equivalenza tra la i) e la ii).

Abbiamo che $\min_{y\in K} f(\bar{x},y) \geq 0$. Visto che $f(\bar{x},\bar{x}) = 0$, abbiamo che $\min_{y\in K} f(\bar{x},y) = 0$ e quindi che $\bar{x} \in K$ è una soluzione del problema (2.3).

Viceversa, se vale la condizione ii), abbiamo che $\min_{y\in K} f(\bar{x},y) = f(\bar{x},\bar{x})$. Visto che $f(\bar{x},\bar{x}) = 0$, abbiamo che $\min_{y\in K} f(\bar{x},y) = 0$ e quindi che $f(\bar{x},y) \geq 0$ per ogni $y \in K$.

Dimostriamo l'equivalenza tra la i) e la iii).

$f(x,x) = 0$ implica che $\sup_{y\in K} f(x,y) \leq 0$ per ogni $x \in K$, e quindi $\min_{x\in K}\sup_{y\in K} f(x,y) \leq 0$. La condizione i) implica che $\min_{x\in K}\sup_{y\in K} f(x,y) \geq 0$, e quindi $\min_{x\in K}\sup_{y\in K} f(x,y) = 0$.

Viceversa, se vale la condizione iii), allora esiste $\bar{x}$ tale che $\sup_{y\in K} f(\bar{x},y) = 0$, e ciò equivale alla condizione i), visto che abbiamo $\min_{x\in K}\sup_{y\in K} f(x,y) \leq 0$, per ogni $y \in K$. □

Osservazione 2.1.4. *Per il lemma precedente, la funzione:*

$$p(x) := \sup_{y\in K}[-f(x,y)]$$

è una funzione gap per il problema (EP). *Questo approccio generalizza quello di Auslender [1] che introdusse la funzione gap per le disuguaglianze variazionali*

$$\sup_{y\in K}\langle F(x), y-x\rangle. \tag{2.4}$$

2.2 Il problema della differenziabilità delle funzioni gap

Le funzioni gap per un certo problema di equilibrio possono essere differenziabili o no. A seconda del metodo utilizzato per risolvere il problema attraverso funzioni gap, la differenziabilità di quest'ultima può essere determinante.

In questa sezione studiamo il problema della differenziabilità delle funzioni gap.

Se assumiamo che, per ogni $\bar{x} \in K$, il problema (2.3) ha un'unica soluzione, possiamo utilizzare il seguente metodo iterativo.

Algoritmo generico

i) Fissiamo $k = 0$ e $x^0 \in K$;

ii) sia x^{k+1} la soluzione del problema

$$\max_{y \in K} \{-f(x^k, y)\}; \tag{2.5}$$

iii) se $\|x^{k+1} - x^k\| < \mu$, per qualche fissato $\mu > 0$, allora STOP, altrimenti poniamo $k = k + 1$ e andiamo al passo ii).

In alcuni casi non è possibile utilizzare l'algoritmo precedente per trovare la soluzione del problema (EP). Per garantire la differenziabilità della funzione gap furono introdotti problemi di equilibrio ausiliari da Fukushima [12], il cui approccio in seguito fu generalizzato. Nel seguito verranno date condizioni sufficienti che garantiranno la differenziabilità della funzione gap (2.2).

Adesso introduciamo il seguente teorema propedeutico di cui non daremo la dimostrazione che può essere trovata in [2].

Teorema 2.2.1. *Supponiamo che le seguenti condizioni siano soddisfatte:*

i) $K = \mathbb{R}^n$;

ii) $M : \Lambda \rightrightarrows \mathbb{R}^n$ *è una funzione semicontinua inferiormente secondo Berge in* $y_0 \in \Lambda$ *(cfr. A.0.19), dove* Λ *è un sottoinsieme di* $\mathbb{R}^n$ *aperto, non vuoto e convesso;*

iii) $\psi(y_0)$ *è non vuoto e limitato;*

iv) f *è semicontinua inferiormente su* $\mathbb{R}^n \times \{y_0\}$ *ed esiste un punto* $x_0 \in \psi(y_0)$ *tale che* f *sia semicontinua superiormente in* (x_0, y_0)*;*

v) $f(\cdot, y)$ *è quasi-convessa su* $\mathbb{R}^n$ *per ogni fissato* $y \in \Lambda$*;*

vi) $M(y_0)$ *è convesso e chiuso;*

vii) tutti gli insiemi $M(y)$*, con* $y \in \Lambda$*, sono convessi;*

viii) la mappa M *è chiusa (cfr. Definizione A.0.18) in* y_0*;*

dove

$$\phi\colon\ K \ni y \longrightarrow \inf_{x\in M(y)}\{f(x,y)\} \in \mathbb{R}\cup\{-\infty\}$$

e

$$\psi\colon\ K \ni y \rightrightarrows \big\{x \in K | f(x,y) = \inf_{x\in M(y)}\{f(x,y)\}\big\} \in \mathbb{R}\cup\{-\infty\}.$$

Allora ϕ è continua in y_0 e ψ è semicontinua superiormente secondo Berge in y_0.

Teorema 2.2.2. *Sia $\varphi : \mathbb{R}^m \times K \longrightarrow \mathbb{R}$ una funzione continua e $K \subset \mathbb{R}^n$. Si denoti con $\psi(x) = \inf\{\varphi(x,y) \mid y \in K\}$, $M(x) = \{y \in K \mid \varphi(x,y) = \psi(x)\}$. Si supponga che per ogni $x \in \mathbb{R}^m$, $M(x) \neq \emptyset$ e che M è limitata nell'intorno di x. Sia Ω un aperto di $\mathbb{R}^m$ tale che φ'_x (gradiente rispetto a x) esiste su $\Omega \times K$ ed è continua su $\Omega \times K$. Allora per ogni $x \in \Omega$, si ha*

$$\psi'(x,d) = \inf\{(\varphi'_x(x,y),d) \mid y \in M(x)\},\ \forall\ d \in \mathbb{R}^m.$$

Se in un punto $x \in \Omega$, $M(x)$ è uguale a un solo punto $y(x)$, allora ψ è derivabile nel senso di Gâteaux in quel punto e si ha

$$\psi^{'}(x) = \varphi^{'}_x(x,y(x)).$$

In generale la funzione (2.2) non è differenziabile. Per la prima volta Fukushima in [12] definì una funzione gap differenziabile con continuità per le disequazioni variazionali. Nella seguente proposizione verranno date condizioni sufficienti che garantiranno la differenziabilità della funzione (2.2). Introdurremo infine un problema di equilibrio ausiliario che servirà per regolarizzare il problema (EP).

Proposizione 2.2.3. *Assumiamo che le seguenti condizioni siano soddisfatte:*

i) $f(x,y)$ *è una funzione strettamente convessa rispetto a* y *per ogni* $x \in K$;

ii) f *è differenziabile rispetto a* x *per ogni* $y \in K$ *e* f'_x *è continua su* $K \times K$;

iii) l'estremo superiore in

$$\sup_{y\in K}[-f(x,y)]$$

è assunto per ogni $x \in K$.

Allora

$$p(x) := \sup_{y \in K} [-f(x, y)]$$

è una funzione gap differenziabile per (EP) *e il suo gradiente è dato da*

$$p'(x) = -f'_x(x, y(x)) \tag{2.6}$$

dove

$$y(x) = \operatorname{argmin}_{y \in K} f(x, y) \tag{2.7}$$

(cfr. Definizione A.0.22).

Dimostrazione. Per ogni $x \in K$, visto che $f(x, y)$ è una funzione strettamente convessa rispetto a y e che l'estremo superiore in (2.1) è assunto per ogni $x \in K$, esiste un unico punto di minimo $y(x)$ del problema $\min_{y \in K} f(x, y)$. Applicando il Teorema 2.2.1, otteniamo che $y(x)$ è semicontinua superiormente secondo Berge in x e, assumendo che $y(x)$ consiste di un solo valore per ogni x, si ha che essa è continua in x.

Visto che f'_x è continua, dal Teorema 2.2.2 segue che

$$p'(x) = -f'_x(x, y(x)).$$

Dalla continuità di f'_x e $y(x)$ segue che $p'(x)$ è continua in x, da cui la differenziabilità di p. □

Osservazione 2.2.4. *La condizione iii) della proposizione precedente è soddisfatta se $f(x, y)$ è semicontinua inferiormente rispetto a y e l'insieme K è compatto o se $f(x, \cdot)$ è fortemente convessa. La Proposizione 2.2.3 è vera anche se $f(x, \cdot)$ è strettamente quasi-convessa e K è compatto [27].*

L'ipotesi della stretta convessità della funzione $f(x, \cdot)$ non è sempre soddisfatta: per esempio, per la disequazione variazionale (VI) la funzione è lineare. Questo problema può essere superato introducendo un problema di equilibrio ausiliario, aggiungendo alla funzione f un termine $H(x, \cdot)$ strettamente convesso.

Sia $H(x, y)\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ una funzione differenziabile rispetto a y su un insieme convesso K tale che:

$$H(x, y) \geq 0, \qquad \forall (x, y) \in K \times K; \tag{2.8}$$

$$H(x, x) = 0, \qquad \forall x \in K; \tag{2.9}$$

$$H'_y(x, x) = 0, \qquad \forall x \in K. \tag{2.10}$$

Il **problema di equilibrio ausiliario** consiste nel cercare $\bar{x} \in \mathbb{R}^n$ tale che

$$\bar{x} \in K, \quad f(\bar{x}, y) + H(\bar{x}, y) \geq 0, \ \forall y \in K. \tag{AEP}$$

Proposizione 2.2.5. *Sia $f(x, y)$ una funzione convessa, differenziabile rispetto a y in $x = \bar{x}$. Sia $\varepsilon > 0$. Sia $H(x, y) : K \times K \longrightarrow \mathbb{R}$ una funzione strettamente convessa su un insieme convesso $K \subseteq \mathbb{R}^n$ rispetto a y per cui valgono le condizioni* (2.8)*,* (2.9) *e* (2.10) *allora $\bar{x}$ è una soluzione del problema* (EP) *se e solo se è una soluzione del problema di equilibrio ausiliario:*

$$\bar{x} \in K, \quad \varepsilon f(\bar{x}, y) + H(\bar{x}, y) \geq 0, \ \forall y \in K. \tag{AEP$_\varepsilon$}$$

Dimostrazione. La condizione (2.8) implica ovviamente che se $\bar{x}$ è una soluzione del problema (EP) allora $\bar{x}$ è una soluzione del problema ausiliario (AEP$_\varepsilon$).

Viceversa, sia $\bar{x}$ una soluzione del problema ausiliario (AEP$_\varepsilon$). Allora $\bar{x}$ è un punto di minimo del problema

$$\min_{y \in K} [\varepsilon f(\bar{x}, y) + H(\bar{x}, y)]. \tag{2.11}$$

Poiché K è un insieme convesso allora $\bar{x}$ è una soluzione di ottimo per (2.11) se e solo se

$$\langle \varepsilon f'_y(\bar{x}, \bar{x}) + H'_y(\bar{x}, \bar{x}), y - \bar{x} \rangle \geq 0, \quad \forall y \in K.$$

Per la condizione (2.10) abbiamo

$$\langle \varepsilon f'_y(\bar{x}, \bar{x}), y - \bar{x} \rangle \geq 0, \quad \forall y \in K,$$

che, dividendo per ε, diventa

$$\langle f'_y(\bar{x}, \bar{x}), y - \bar{x} \rangle \geq 0, \quad \forall y \in K.$$

Per la convessità di $f(\bar{x}, \cdot)$, otteniamo

$$f(\bar{x}, y) \geq f(\bar{x}, \bar{x}), \quad \forall y \in K,$$

e quindi, visto che $f(\bar{x}, \bar{x}) = 0$,

$$f(\bar{x}, y) \geq 0, \quad \forall y \in K,$$

ossia $\bar{x}$ è una soluzione del problema (EP). $\square$

Corollario 2.2.6. *Un punto $\bar{x}$ è una soluzione del problema* (EP) *se solo se $\bar{x}$ è una soluzione di ottimo per il problema*

$$\min_{y \in K} [\varepsilon f(\bar{x}, y) + H(\bar{x}, y)].$$

Teorema 2.2.7. *Supponiamo che $f(x, y)$ soddisfi le seguenti proprietà:*

i) f sia differenziabile, semicontinua inferiormente e convessa rispetto a y, per ogni $x \in K$;

ii) f sia differenziabile rispetto a x, per ogni $y \in K$;

iii) f'_x sia continua su $K \times K$.

Sia $H(x, y)\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ una funzione differenziabile con continuità su $K \times K$, fortemente convessa rispetto a y, per ogni $x \in K$, e tale che le condizioni (2.8), (2.9) *e* (2.10) *siano soddisfatte.*

Allora

$$h(x) := \max_{y \in K} [-f(x, y) - H(x, y)] \tag{2.12}$$

è una funzione gap differenziabile per (EP) *e il suo gradiente è dato da*

$$h'(x) = -f'_x(x, y(x)) - H'_x(x, y(x)),$$

dove $y(x) := \operatorname{argmin}_{y \in K} [f(x, y) + H(x, y)]$.

Dimostrazione. Applicando la Proposizione 2.2.5 con $\varepsilon = 1$, otteniamo che il problema (EP) è equivalente al problema ausiliario (AEP). Applicando la Proposizione 2.2.3 al problema ausiliario (AEP), completiamo la dimostrazione. $\square$

Osservazione 2.2.8. *Quanto fatto finora è una generalizzazione di risultati di Zhu e Marcotte [36], che hanno considerato il caso particolare in cui $f(x, y) := \langle F(x), y - x \rangle$.*

2.3 Funzioni gap e principi variazionali

Vediamo ora un esempio di principio variazionale che può essere visto come una funzione gap. Diamo la definizione generale di principio variazionale.

Definizione 2.3.1. *Diciamo che un* **principio variazionale** *vale per il problema di equilibrio* (EP) *se e solo se esiste una funzione* $W\colon K \longrightarrow \mathbb{R} \cup \{-\infty\}$ *tale che l'insieme delle soluzioni di* (EP) *coincida con l'insieme delle soluzioni del problema di ottimizzazione*

$$\max\{W(x) | x \in K\}. \tag{2.13}$$

Esempio 2.3.2. *Se* $f(x,y) := \langle D\varphi(x), y-x\rangle$, *dove* $\varphi\colon \mathbb{R}^n \longrightarrow \mathbb{R}$ *è convessa e differenziabile secondo Gâteaux con differenziale* $D\varphi(x)$ *in* x, *allora* $\bar{x}$ *è soluzione di* (EP) *se e solo se* $\bar{x}$ *è soluzione di*

$$\min\{\varphi(x) | x \in K\},$$

per l'Osservazione 1.0.1. Allora, ponendo $W := -\varphi$, *si ottiene un principio variazionale per* (EP).

Osservazione 2.3.3. *Nei principi variazionali non è richiesto che il massimo della funzione sia uguale a zero, mentre per una funzione gap il minimo è uguale a zero.*

Nel caso che analizzeremo la funzione gap sarà $p(x) = -W(x)$ perché $\max\{W(x)|x \in K\}$ sarà uguale a zero.

Consideriamo il caso in cui

$$f(x,y) = g(x,y) + h(x,y),$$

e consideriamo il problema (EP)

$$\bar{x} \in K, \quad g(\bar{x},y) + h(\bar{x},y) \geq 0, \quad \forall y \in K. \tag{2.14}$$

Assumiamo che l'insieme $K \subset \mathbb{R}^n$ e le funzioni $g, h\colon K \times K \longrightarrow \mathbb{R}$ soddisfino le Assunzioni (i), (ii) e (iii) del Teorema 1.3.4, ossia

(i) $K \subset \mathbb{R}^n$ sia un insieme chiuso, convesso e non vuoto;

(ii) $g\colon K \times K \longrightarrow \mathbb{R}$ goda delle seguenti proprietà:

(a) $g(x, x) = 0$ per ogni $x \in \mathbb{R}^n$,

(b) $g(x, y) + g(y, x) \leq 0$ per ogni $x, y \in K$ (monotona),

(c) per tutti gli $x, y \in K$ la funzione

$$[0, 1] \ni t \longmapsto g(ty + (1-t)x, y) \in \mathbb{R}$$

sia semicontinua superiormente in $t = 0$,

(d) g sia convessa e semicontinua inferiormente nel secondo argomento;

(iii) $h\colon K \times K \longrightarrow \mathbb{R}$ abbia le seguenti proprietà:

(a) $h(x, x) = 0$ per ogni $x \in K$,

(b) h sia semicontinua superiormente nel primo argomento,

(c) h sia convessa nel secondo argomento.

Sia $\pi\colon K \times K \longrightarrow \mathbb{R} \cup \{+\infty\}$ una funzione che soddisfa

$$\pi(x, y) \geq 0, \quad \pi(x, x) = 0, \quad \pi(x, ty + (1-t)x) = o(t) \ (0 \leq t \leq 1). \tag{2.15}$$

Definiamo $W\colon K \longrightarrow \mathbb{R} \cup \{-\infty\}$ come

$$W(x) := \inf_{y \in K} \{-g(y, x) + h(x, y) + \pi(x, y)\} \tag{2.16}$$

e associamo al problema (2.14) il problema di ottimizzazione

$$\max\{W(x) | x \in K\}. \tag{2.17}$$

Osservazione 2.3.4. *Visto che abbiamo*

$$g(x, x) + h(x, x) + \pi(x, x) = 0,$$

otteniamo

$$W(x) \leq 0 \quad \forall x \in K. \tag{2.18}$$

Il principio variazionale che vogliamo ottenere è basato sul lemma seguente.

Vista l'osservazione precedente se poniamo $p = -W$ abbiamo la proprietà i) delle funzioni gap (Definizione 2.1.1).

Lemma 2.3.5. *Il problema* (2.14) *è equivalente a cercare* $\bar{x}$ *tale che*

$$\bar{x} \in K, \ W(\bar{x}) = 0. \tag{2.19}$$

Dimostrazione. Assumiamo vera la (2.14). Poiché g è monotona (ossia $g(x,y)+g(y,x) \leq 0$ per ogni $x, y \in K$), abbiamo che

$$g(\bar{x}, y) \leq -g(y, \bar{x}) \quad \forall y \in K,$$

quindi abbiamo che

$$0 \leq g(\bar{x}, y) + h(\bar{x}, y) \leq -g(y, \bar{x}) + h(\bar{x}, y) \quad \forall y \in K.$$

Dal fatto che $\pi \geq 0$, otteniamo

$$0 \leq -g(y, \bar{x}) + h(\bar{x}, y) + \pi(\bar{x}, y) \quad \forall y \in K.$$

Dalla definizione di W segue $W(\bar{x}) \geq 0$. Quindi la (2.19) è verificata.

Viceversa, supponiamo vera la (2.19). Allora dalla definizione di W segue

$$0 \leq -g(\eta, \bar{x}) + h(\bar{x}, \eta) + \pi(\bar{x}, \eta) \quad \forall \eta \in K.$$

Fissiamo $y \in K$ arbitrariamente. Sia

$$x_t := ty + (1-t)\bar{x} \quad 0 < t \leq 1.$$

Allora si ha $x_t \in K$ e dalle disuguaglianze precedenti segue

$$g(x_t, \bar{x}) \leq h(\bar{x}, x_t) + \pi(\bar{x}, x_t).$$

Da questa, usando

$$h(\bar{x}, x_t) \leq th(\bar{x}, y) + (1-t)h(\bar{x}, \bar{x}) = th(\bar{x}, y),$$

otteniamo

$$\begin{aligned}
0 = g(x_t, x_t) \leq \\
&\leq tg(x_t, y) + (1-t)g(x_t, \bar{x}) \leq \\
&\leq tg(x_t, y) + (1-t)\big(h(\bar{x}, x_t) + \pi(\bar{x}, x_t)\big) \leq \\
&\leq tg(x_t, y) + (1-t)\big(th(\bar{x}, y) + o(t)\big).
\end{aligned}$$

Dividendo per t otteniamo

$$0 \leq g(x_t, y) + (1-t)\left(h(\bar{x}, y) + \frac{o(t)}{t}\right);$$

passando al limite per $t \longrightarrow 0$ (quindi $x_t \longrightarrow \bar{x}$) ed usando la semicontinuità di g abbiamo

$$0 \leq g(\bar{x}, y) + h(\bar{x}, y).$$

Questo è vero per ogni $y \in K$, quindi abbiamo dimostrato che la (2.14) è verificata. $\square$

È ora possibile enunciare il teorema relativo al principio variazionale.

Teorema 2.3.6. *Un elemento $\bar{x}$ è una soluzione del problema* (2.14) *se e solo se $W(\bar{x}) = 0$. Se l'insieme delle soluzioni del problema* (2.14) *è non vuoto, allora l'insieme delle soluzioni del problema* (2.14) *coincide con quelle del problema di ottimizzazione* (2.17)*.*

Dimostrazione. Per l'Osservazione 2.3.4 abbiamo $W(x) \leq 0$ per ogni $x \in K$ quindi il massimo del problema di ottimizzazione (2.17)

$$\max\{W(x) | x \in K\}$$

è minore o uguale a 0. Inoltre per il Lemma 2.3.5 abbiamo che il problema (2.14) è equivalente a cercare $\bar{x}$ tale che $\bar{x} \in K$ e $W(\bar{x}) \geq 0$. Quindi il problema è equivalente a cercare $\bar{x}$ tale che $\bar{x} \in K$ e $W(\bar{x}) = 0$, ossia la prima parte dell'enunciato.

Per dimostrare la seconda parte dell'enunciato, notiamo che, per quanto abbiamo appena detto, se il problema (2.14) ha una soluzione $\bar{x}$, abbiamo che $W(\bar{x}) = \max\{W(x) | x \in K\} = 0$, e viceversa. $\square$

Osservazione 2.3.7. *Notiamo che, sempre per il Lemma 2.3.5, se il problema* (2.14) *non ha soluzione, non esistono $\bar{x}$ tali che $\bar{x} \in K$ e $W(\bar{x}) \geq 0$, e quindi abbiamo* $\sup\{W(x) | x \in K\} < 0$.

Osservazione 2.3.8. *Auchmuty considerò il caso in cui $g(x, y)$ è identicamente nulla e*

$$h(x, y) := \langle Tx, y - x \rangle \quad \textit{dove} \quad T \colon K \longrightarrow \mathbb{R}^n. \tag{2.20}$$

Esempio 2.3.9. *Sia* $\varphi\colon \mathbb{R}^n \longrightarrow \mathbb{R}$ *convessa e differenziabile secondo Gâteaux con differenziale di Gâteaux* $D\varphi(x) \in \mathbb{R}^n$*; definiamo*

$$\pi(x,y) := \varphi(y) - \varphi(x) - \langle D\varphi(x), y - x\rangle. \tag{2.21}$$

Le condizioni (2.15) per π sono soddisfatte, infatti abbiamo:

- Usando la definizione di $D\varphi(x)$, ossia

$$\langle D\varphi(x), v\rangle = \lim_{t\to 0} \frac{\varphi\big(x + tv\big) - \varphi(x)}{t}$$

e la convessità di φ, ossia

$$\varphi\big((1-t)x + ty\big) \leq (1-t)\varphi(x) + t\varphi(y),$$

otteniamo

$$\begin{aligned}
\pi(x,y) &= \varphi(y) - \varphi(x) - \langle D\varphi(x), y - x\rangle = \\
&= \varphi(y) - \varphi(x) - \lim_{t\to 0} \frac{\varphi\big(x + t(y-x)\big) - \varphi(x)}{t} = \\
&= \varphi(y) - \varphi(x) - \lim_{t\to 0} \frac{\varphi\big((1-t)x + ty\big) - \varphi(x)}{t} \geq \\
&\geq \varphi(y) - \varphi(x) - \lim_{t\to 0} \frac{(1-t)\varphi(x) + t\varphi(y) - \varphi(x)}{t} = \\
&= \varphi(y) - \varphi(x) - \lim_{t\to 0} \frac{-t\varphi(x) + t\varphi(y)}{t} = \\
&= 0.
\end{aligned}$$

- Abbiamo $\pi(x,x) = \varphi(x) - \varphi(x) - \langle D\varphi(x), x - x\rangle = 0$.

- Abbiamo

$$\begin{aligned}
\pi(x, ty + (1-t)x) &= \varphi(ty + (1-t)x) - \varphi(x) \\
&\quad - \langle D\varphi(x), \big(ty + (1-t)x\big) - x\rangle = \\
&= \varphi(ty + (1-t)x) - \varphi(x) - t\langle D\varphi(x), y - x\rangle,
\end{aligned}$$

quindi, usando la definizione di $D\varphi(x)$, otteniamo

$$\begin{aligned}\lim_{t\to 0^+} \frac{\pi(x, ty + (1-t)x)}{t} &= \lim_{t\to 0^+} \frac{\varphi(ty + (1-t)x) - \varphi(x)}{t} \\ &\quad - \langle D\varphi(x), y - x\rangle = \\ &= \lim_{t\to 0^+} \frac{\varphi(x + t(y-x)) - \varphi(x)}{t} \\ &\quad - \langle D\varphi(x), y - x\rangle = \\ &= 0.\end{aligned}$$

Corollario 2.3.10. *La funzione* $p = -W$ *definita dalla* (2.16) *è una funzione gap per il problema* (2.14).

D-gap

Abbiamo analizzato le funzioni gap, nel caso dei problemi di minimo vincolato. Per completezza, definiamo le funzioni D-gap, che vengono applicate nel caso dei problemi di minimo non vincolato, anche se computazionalmente non sono utilizzate.

In generale la funzione gap (2.2) non è finita, non è convessa e non è differenziabile. Per queste ragioni, fu introdotta la funzione gap regolarizzata

$$p_\alpha(x) \doteq \max_{y\in K} \left\{ -f(x,y) - \frac{\alpha}{2}\|y - x\|^2 \right\}. \tag{2.22}$$

Ciò permette di riformulare il problema di equilibrio (EP) come problema di minimizzazione della funzione p_α su K, che è differenziabile su K se lo è f. Si ha

$$\nabla p_\alpha(x) = -\nabla_x f(x, y_\alpha(x)) - \alpha[x - y_\alpha(x)],$$

dove $y_\alpha(x)$ è l'unico valore massimo del problema (2.22). Il termine $\|y - x\|^2$ è chiamato termine di regolarizzazione per la funzione f.

Definizione 2.3.11. *La funzione* $p_{\alpha\beta}\colon \mathbb{R}^n \longrightarrow \mathbb{R}$ *data da*

$$p_{\alpha\beta}(x) = p_\alpha(x) - p_\beta(x),$$

dove α *e* β *sono parametri arbitrari non negativi su* $\mathbb{R}^n$ *tali che* $0 < \alpha < \beta$, *è detta* **D-gap**.

Osservazione 2.3.12. *Si hanno le seguenti proprietà:*

- $p_{\alpha\beta}(x) \geq 0$ *per ogni* $x \in \mathbb{R}^n$,
- $p_{\alpha\beta}(\bar{x}) = 0$ *se e solo se* $\bar{x}$ *risolve il problema di equilibrio.*

Il minimo globale di $p_{\alpha\beta}$ su $\mathbb{R}^n$ coincide con le soluzioni del problema di equilibrio (EP) [18, 35].

Capitolo 3

Algoritmi

In questo capitolo analizzeremo alcuni algoritmi che utilizzando le funzioni gap trovano una soluzione di ottimo per il problema di equilibrio (EP). Con l'introduzione del problema ausiliario (AEP) e della relativa funzione gap (2.12) è possibile esprimere il problema di equilibrio (EP) tramite la minimizzazione di una funzione differenziabile.

In particolare, considereremo algoritmi risolutivi per la minimizzazione della funzione gap basati su metodi di ricerca esatta o inesatta lungo una direzione di discesa.

Inoltre, descriveremo applicazioni alle diseguaglianze variazionali e ai problemi di ottimizzazione, e studieremo Error bound e Rate of convergence.

Infine, applicheremo tutta la teoria a un esempio di flusso su reti.

Preliminari Faremo le seguenti assunzioni basilari. In questo capitolo, se non differentemente specificato (oltre alle ipotesi fatte nel primo e secondo capitolo) assumeremo le seguenti ipotesi:

- la funzione considerata è $f : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}$;
- $f(x, y)$ è una funzione differenziabile convessa rispetto a y, per ogni $x \in K$;
- $f(x, y)$ è differenziabile su K rispetto a x, per ogni $y \in K$;
- f'_x è continua su $K \times K$.

3.1 Metodi di ricerca esatta

In questa sezione descriveremo alcuni metodi di ricerca esatta. Prima di analizzare un algoritmo sul metodo di ricerca esatta lungo una direzione di discesa, definiamo una direzione di discesa per una funzione.

Definizione 3.1.1. *Sia p una funzione gap. Una direzione d è detta* **ammissibile** *in $\bar{x}$ se $\bar{x} + td$ appartiene a K per ogni $t \in [0, \delta]$ con $\delta > 0$.*

Definizione 3.1.2. *Sia p una funzione gap. Una direzione d è una* **direzione di discesa** *per p in $\bar{x}$ se*

- *d è ammissibile,*
- *risulta*
$$\langle \nabla p(\bar{x}), d \rangle < 0.$$

Algoritmo 3.1.3.

Passo 1: *Siano $k = 0$ e $x_0 \in K$.*

Passo 2: *Sia $x_{k+1} \doteq x_k + t_k d_k$ dove $d_k \doteq y(x_k) - x_k$ e $y(x_k)$ è una soluzione del problema:*
$$\min_{y \in K} f(x_k, y)$$
e t_k è la soluzione del problema
$$\min_{0 \leq t \leq 1} p(\bar{x} + t_k d_k).$$

Passo 3: *Se $\|x_{k+1} - x_k\| < \mu$, per qualche fissato $\mu > 0$, allora STOP, altrimenti poni $k = k + 1$ e vai al Passo 2.*

Si deve provare che d_k è una direzione di discesa per p nel punto x_k. Per raggiungere questo obiettivo è necessario aggiungere anche le ipotesi di stretta ∇-monotonia di f:

$$\langle f_x'(x, y) + f_y'(x, y), y - x \rangle > 0, \quad \forall (x, y) \in K \times K \text{ con } x \neq y. \tag{3.1}$$

Per questo enunciamo la seguente proposizione:

Proposizione 3.1.4. *Supponiamo vere le assunzioni della Proposizione 2.2.3 e la* (3.1)*. Allora se $y(x) \neq x$, $d(x) \doteq y(x) - x$ è una direzione di discesa per p in $x \in K$.*

Dimostrazione. Osserviamo che $\bar{x}$ è una soluzione per il problema (EP) se e solo se $y(\bar{x}) = \bar{x}$. Poiché $y(x) \doteq \text{argmin}_{y \in K} f(x, y)$ e $f(x, \cdot)$ è strettamente convessa, vale la seguente disuguaglianza variazionale:

$$\langle f_y^{'}(x, y(x)), z - y(x) \rangle \geq 0, \quad \forall z \in K. \tag{3.2}$$

Posto $z \doteq x$ otteniamo $\langle f_y^{'}(x, y(x)), y(x) - x \rangle \leq 0$. Considerando la (3.1) si ha:

$$0 \geq \langle f_y^{'}(x, y(x)), y(x) - x \rangle > -\langle f_x^{'}(x, y(x)), y(x) - x \rangle$$

Dalla (2.6) otteniamo $\langle p^{'}(x), y(x) - x \rangle < 0$. $\square$

Enunciamo adesso il teorema di Zangwill che ci servirà nel seguito. Supponiamo che $T : X \rightrightarrows X$ sia una mappa che rappresenta un algoritmo, X sia un insieme e M sia un insieme di punti di X che soddisfa un'opportuna condizione necessaria di ottimalità. Diamo la seguente definizione.

Definizione 3.1.5. *$z : X \longrightarrow \mathbb{R}$ è una* **funzione di discesa** *(relativa all'algoritmo T) se è continua ed ha le seguenti proprietà:*

(1) $x \notin M$ implica $z(y) < z(x)$ $\forall y \in T(x)$,

(2) $x \in M$ implica $z(y) \leq z(x)$ $\forall y \in T(x)$.

Ora possiamo enunciare il teorema di Zangwill [34] come riportato in [23].

Teorema 3.1.6. *Sia $\{x_k\}$ una successione generata dall'algoritmo ossia tale che soddisfi $x_{k+1} \in T(x^k)$. Se le seguenti tre condizioni sono soddisfatte:*

(i) Tutti i punti x_k sono contenuti nell'insieme compatto $K \subset X$;

(ii) Esiste una funzione di discesa $z : X \longrightarrow \mathbb{R}$;

(iii) La funzione T è chiusa (cfr. Definizione A.0.18) su $X \setminus M$ e $\forall x \in X \setminus M$, $T(x) \neq \emptyset$.

Allora ogni punto limite appartiene a M.

Teorema 3.1.7. *Supponiamo che K sia un insieme compatto in X, che l'assunzione* (3.1) *sia soddisfatta e che $f(x, y)$ sia una funzione strettamente convessa rispetto a y, per ogni $x \in K$. Allora, per ogni $x_0 \in K$ la successione $\{x_k\}$ definita dall'Algoritmo 3.1.3, appartiene all'insieme K e ogni punto di accumulazione di $\{x_k\}$ è la soluzione di* (EP).

Dimostrazione. La convessità di K implica che la successione $\{x_k\}$ è contenuta in K poiché $t_k \in [0, 1]$. Poiché $y(x)$ è continua anche la funzione $d(x) \doteq y(x) - x$ è continua su K. La mappa:

$$U(x, d) \doteq \{y : y = x + t_k d,\ p(x + t_k d) = \min_{0 \le t \le 1} p(y + td)\}$$

è chiusa (cfr. Definizione A.0.18) ogni volta che p è una funzione continua. Pertanto la funzione algoritmo $x_{k+1} = U(x_k, d(x_k))$ è chiusa (cfr. Teorema 3.3 in [23] e [19]). Il teorema di convergenza di Zangwill [34] implica che qualsiasi punto di accumulazione della successione $\{x_k\}$ è una soluzione del problema di equilibrio (EP). □

Osservazione 3.1.8. *Se al Passo 2 dell'Algoritmo 3.1.3 scegliamo sempre $t_k = 1$ otteniamo:*

$$y_{k+1} = y_k + d_k = y_k + (y_k) - y_k = x(y_k).$$

L'algoritmo quindi diventa un metodo del punto fisso.

Applicando l'Algoritmo 3.1.3 al problema di equilibrio ausiliario (AEP) si ha il seguente:

Algoritmo 3.1.9. *Sia h la funzione gap, definita in* (2.12)*, associata al problema ausiliario* (AEP).

Passo 1: *Siano $k = 0$ e $x_0 \in K$.*

Passo 2: *Sia $x_{k+1} \doteq x_k + t_k d_k$, $k = 1, \ldots,$ dove $d_k \doteq y(x_k) - x_k$, $y(x_k)$ è una soluzione del problema:*

$$\min_{y \in K} [f(x_k, y) + H(x_k, y)],$$

e t_k è una soluzione del problema

$$\min_{0 \le t \le 1} h(x_k + td_k).$$

Passo 3: *Se* $\|x_{k+1} - x_k\| < \mu$ *per qualche fissato* $\mu > 0$*, allora STOP, altrimenti poni* $k = k + 1$ *e vai al Passo 2.*

Per applicare l'Algoritmo 3.1.9 sostituiamo la (3.1) con la seguente condizione:

$$\langle f_x^{'}(x,y) + H_x^{'}(x,y) + f_y^{'}(x,y) + H_y^{'}(x,y), x - y\rangle > 0, \quad \textit{per ogni } (x,y) \in K \times K \textit{ con } x \neq y. \tag{3.3}$$

Se facciamo la seguente assunzione (cfr. [20] e [33]):

$$H_x^{'}(x,y) + H_y^{'}(x,y) = 0, \ \forall \ (x,y) \in K \times K \tag{3.4}$$

allora la (3.3) ovviamente coinciderà con la (3.1). Osserviamo che la (3.4) è soddisfatta nell'algoritmo di ricerca esatta lungo una direzione di discesa proposta da Fukushima [12] che si ottiene ponendo:

$$H(x,y) = \frac{1}{2}\langle M(x-y), x-y\rangle,$$

dove M è una matrice definita positiva di ordine n.

Osservazione 3.1.10. *Il Teorema 3.1.7 garantisce la convergenza dell'Algoritmo 3.1.9.*

Teorema 3.1.11. *Sia* $H(x,y) : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}$ *una funzione differenziabile su* $K \times K$*, fortemente convessa rispetto a* x*, per ogni valore di* $y \in K$*, e tale che siano soddisfatte anche le condizioni* (2.8)*,* (2.9) *e* (2.10)*. Supponiamo che* K *sia un insieme compatto di* $\mathbb{R}^n$ *e che la condizione* (3.3) *sia soddisfatta allora, per ogni* $x_0 \in K$*, la successione* $\{x_k\}$*, definita dall'Algoritmo 3.1.9, appartiene all'insieme* K *e ogni punto di accumulazione della successione* $\{x_k\}$ *è una soluzione del problema* (EP)*.*

Dimostrazione. Poichè il problema di equilibrio (EP) è equivalente al problema ausiliario (AEP) per la Proposizione 2.2.3 possiamo applicare il Teorema 3.1.7 al problema ausiliario e quindi abbiamo dimostrato la nostra tesi. □

3.2 Metodi di ricerca inesatta

Abbiamo precedentemente illustrato l'Algoritmo 3.1.3 basato sul metodo di ricerca esatta lungo una direzione di discesa, ora considereremo quello con ricerca inesatta. Generalizziamo la condizione (3.1) che ci servirà per descrivere il prossimo algoritmo:

$$\langle f'_x(x,y) + f'_y(x,y), y - x \rangle \geq \mu \|y - x\|^2, \quad \forall\ (x,y) \in K \times K \tag{3.5}$$

dove $\mu > 0$ é un'opportuna costante.

Osservazione 3.2.1. *La disequazione* (3.5) *è soddisfatta se la funzione è fortemente convessa rispetto alla variabile y, con modulo 2μ, per ogni $x \in K$, e concava rispetto a x, per ogni $y \in K$.*

Proposizione 3.2.2. *Sia K un insieme compatto in $\mathbb{R}^n$ e assumiamo che la condizione* (3.5) *sia soddisfatta per qualche $\mu > 0$ allora*

$$\langle p'(x), d(x) \rangle \leq -\mu \|d(x)\|^2$$

dove $d(x) \doteq y(x) - x$, p è definita in (2.2) *e $y(x)$ nella* (2.7).

Dimostrazione. Considerando la (3.1) otteniamo:

$$0 \geq \langle f'_x(y(x), y), y(x) - x \rangle \geq -\langle f'_y(y(x), y), y(x) - x \rangle + \mu \|y(x) - y\|^2.$$

Dalla (2.6) otteniamo

$$\langle p'(x), d(x) \rangle \leq -\mu \|d(x)\|^2.$$

□

Algoritmo 3.2.3.

Passo 1: *Sia $x_0 \in K$, sia ϵ un fattore tolleranza e siano β, σ parametri dell'intervallo aperto* $(0,1)$. *Sia $k = 0$.*

Passo 2: *Se $p(x_k) = 0$, allora STOP, altrimenti vai al passo 3.*

Passo 3: *Sia* $d_k(x_k) \doteq y(x_k) - x_k$. *Sia* m *il più piccolo intero non negativo tale che*

$$p(x_k) - p(x_k + \beta^m d_k) \geq \sigma\beta^m \|d_k\|^2,$$

e siano fissati $\alpha_k = \beta^m$ *e* $x_{k+1} = x_k + \alpha_k d_k$*; se* $\|x_{k+1} - x_k\| < \epsilon$*, allora STOP, altrimenti sia* $k = k+1$ *e torna al passo 2.*

Teorema 3.2.4. *Sia* $\{x_k\}$ *la successione definita nell'Algoritmo 3.2.3, sia* K *un insieme compatto in* $\mathbb{R}^n$*, sia* $f(x,y)$ *una funzione strettamente convessa rispetto a* y *per ogni* $x \in K$*, sia soddisfatta l'asserzione* (3.5) *per un opportuno* $\mu > 0$*, e sia* $\sigma < \dfrac{\mu}{2}$*. Allora, per ogni* $x_0 \in K$ *la successione* $\{x_k\}$ *è contenuta in* K *e qualsiasi punto di accumulazione di* $\{x_k\}$ *è una soluzione del problema* (EP)*.*

Dimostrazione. La convessità di K implica che la successione $\{x_k\}$ è contenuta in K, poichè $\alpha_k \in [0,1]$.

La compattezza di K garantisce che $\{x_k\}$ abbia almeno un punto di accumulazione. Sia $\{\tilde{x}_k\}$ una sottosuccessione convergente di $\{x_k\}$ e sia $\bar{x}$ il suo punto limite. Dimostreremo che $\bar{x}$ è un punto fisso per $y(x)$, $y(\bar{x}) = \bar{x}$, cosicché $\bar{x}$ è una soluzione per il problema (EP).

Sia $d(x) \doteq y(x) - x$, una direzione di discesa; poiché la funzione $y(x)$ è continua per la Proposizione 2.2.3, ne segue che $d(x)$ è continua, pertanto otteniamo che $d(\tilde{x}_k) \to d(\bar{x}) \doteq d^*$ e $p(\tilde{x}_k) \to p(\bar{x}) = p^*$. Dal metodo di ricerca esatta abbiamo:

$$p(x_k) - p(x_{k+1}) \geq \sigma\alpha_k \|d_k\|^2,$$

e questa relazione continua ad essere valida per la sottosuccessione $\{\tilde{x}_k\}$. Pertanto,

$$\tilde{\alpha}_k \|d(\tilde{x}_k)\|^2 \to 0$$

per una sottosuccessione $\{\tilde{\alpha}_k\} \subseteq \{\alpha_k\}$.

Se $\tilde{\alpha}_k > \gamma > 0$ con $\gamma \in \mathbb{R}$, per ogni $k \in \mathbb{N}$, allora $\|d(\tilde{x}_k)\| \to 0$ e $y(\bar{x}) = \bar{x}$. Altrimenti, esiste una successione $\{\alpha_{k'}\} \subseteq \{\tilde{\alpha}_k\}$ con $\alpha_{k'} \to 0$. Dal metodo di ricerca esatta lungo una direzione abbiamo

$$\frac{p(\tilde{x}_{k'}) - p(\tilde{x}_{k'} + \bar{\alpha}_{k'} d(\tilde{x}_{k'}))}{\bar{\alpha}_{k'}} < \sigma \|d(\tilde{x}_{k'})\|^2, \tag{3.6}$$

dove $\bar{\alpha}_{k'} = \frac{\alpha_{k'}}{\beta}$. Prendendo il limite di (3.6) per $k' \to \infty$, da $\bar{\alpha}_{k'} \to 0$ e essendo p differenziabile, otteniamo

$$-\langle p'(\bar{x}), d^* \rangle \leq \sigma \|d^*\|^2. \tag{3.7}$$

Poichè $\sigma < \frac{\mu}{2}$, deve essere $\|d^*\| = 0$, ciò implica che $y(\bar{x}) = \bar{x}$. $\square$

L'Algoritmo 3.2.3 può essere utilizzato anche per la minimizzazione con ricerca inesatta della funzione gap h, definita dalla (2.12), purché l'asserzione (3.5) sia sostituita dalla seguente:

$$\langle f'_x(x,y) + H'_x(x,y) + f'_y(x,y) + H'_y(x,y), y - x \rangle \geq \mu \|x - y\|^2, \tag{3.8}$$

per ogni $(x,y) \in K \times K$.

Nel caso in cui $f(x,\cdot)$ è una funzione convessa, ma non necessariamente fortemente convessa si ha il seguente algoritmo:

Algoritmo 3.2.5.

Passo 1: *Sia $x_0 \in K$, sia ϵ un fattore di tolleranza e siano β, σ parametri nell'intervallo $(0,1)$ e $k = 0$.*

Passo 2: *Se $h(x_k) = 0$, allora STOP, altrimenti vai al passo 3.*

Passo 3: *Sia $d_k(x_k) \doteq y(x_k) - x_k$. Sia m il più piccolo intero non negativo tale che*

$$h(x_k) - h(x_k + \beta^m d_k) \geq \sigma \beta^m \|d_k\|^2,$$

fissati $\alpha_k = \beta^m$ e $y_{k+1} = y_k + \alpha_k d_k$, se $\|x_{k+1} - x_k\| < \epsilon$, allora STOP, altrimenti sia $k = k + 1$ e vai al passo 2.

Corollario 3.2.6. *Sia $\{x_k\}$ una successione definita dall'Algoritmo 3.2.5. Supponiamo che K sia un compatto in $\mathbb{R}^n$, che l'asserzione* (3.8) *sia soddisfatta per un opportuno $\mu > 0$, e che $\sigma < \frac{\mu}{2}$. Allora, per ogni $x_0 \in K$ la successione $\{x_k\}$ è contenuta in K e ogni punto di accumulazione di $\{x_k\}$ è soluzione del problema di equilibrio* (EP).

3.3 Error bound e Rate of convergence

Nel prossimo paragrafo parleremo dell'Error bound che stima quanto si è lontani dalla soluzione ottima del problema. Analizzeremo il caso in cui K non è necessariamente un insieme compatto, f è una funzione fortemente monotona e $H_y'(x,\cdot)$ una funzione lipschitziana.

Se f è una funzione fortemente monotona vedremo che la funzione p, definita nella (2.2), e la funzione h, definita nella (2.12) forniscono un error bound globale.

Proposizione 3.3.1. *Sia f una funzione fortemente monotona su K, con modulo b. Allora*

$$p(x) \geq b\|x-\bar{x}\|^2, \quad \forall x \in K \tag{3.9}$$

dove $\bar{x}$ è soluzione del problema (EP).

Dimostrazione. Osserviamo che, essendo f fortemente monotona su K, (EP) ammette una e una sola soluzione $\bar{x}$ su K. Poiché per ogni $y \in K$, si ha $p(x) \geq -f(x,y)$ allora

$$\begin{aligned} p(x) &\geq -f(\bar{x},x) - f(x,\bar{x}) + f(\bar{x},x) \geq \\ &\geq b\|x-\bar{x}\|^2 + f(\bar{x},x) \geq \\ &\geq b\|x-\bar{x}\|^2. \end{aligned}$$

□

Per estendere il risultato precedente alla funzione gap h assumiamo che H_y' sia una funzione lipschitziana (cfr. Definizione A.0.8).

Proposizione 3.3.2. *Sia f una funzione fortemente monotona su K, con modulo b, $H(x,\cdot)$ sia convessa e $H_y'(x,\cdot)$ lipschitziana con modulo $L < 2b$, per ogni $x \in K$. Allora*

$$h(x) \geq (b - \frac{L}{2})\|x-\bar{x}\|^2, \textit{ per ogni } y \in K, \tag{3.10}$$

dove $\bar{x}$ è una soluzione per il problema (EP).

Dimostrazione. Per ogni $x, y \in K$, abbiamo

$$h(x) \geq -f(x,y) - H(x,y).$$

Pertanto, per $y = \bar{x}$,

$$\begin{aligned} h(x) &\geq -f(\bar{x},x) - H(x,\bar{x}) - f(x,\bar{x}) + f(\bar{x},x) \geq \\ &\geq b\|x-\bar{x}\|^2 + f(\bar{x},x) - H(x,\bar{x}). \end{aligned}$$

Otteniamo così

$$h(x) \geq b\|x-\bar{x}\|^2 - H(x,\bar{x}). \tag{3.11}$$

Poiché $H_y^{'}(x,\cdot)$ è lipschitziana vale la seguente disuguaglianza:

$$H(x,\bar{x}) = H(x,\bar{x}) - H(x,x) \leq \frac{L}{2}\|\bar{x}-x\|^2, \ \forall x \in K.$$

Combinando la precedente disuguaglianza con la (3.11), otteniamo:

$$h(x) \geq \left(b - \frac{L}{2}\right)\|x-\bar{x}\|^2.$$

□

Osservazione 3.3.3. *Poiché $\{p(x_k)\}$ è una successione strettamente decrescente, le disequazioni (3.9) e (3.10) garantiscono che la successione $\{x_k\}$, generata da tutti gli algoritmi descritti precedentemente, sia contenuta in un insieme compatto.*

Rate of convergence È importante anche studiare la velocità di convergenza. Ci limitiamo a dare la definizione e ad enunciare un risultato.

Definizione 3.3.4. *Una successione $\{x^k\}$ si dice che converge a $\bar{x}$ con tasso di convergenza uguale a r se*

$$\limsup_{k\longrightarrow+\infty} \frac{\|x^{k+1}-\bar{x}\|}{\|x^k-\bar{x}\|^r} = \gamma \ \in (0,+\infty).$$

Se $r=1$ e $\gamma \in (0,1)$, allora la convergenza è detta lineare, se $r>1$, allora la convergenza è detta superlineare, e, in particolare, se $r=2$, la convergenza si dice quadratica.

Consideriamo l'Algoritmo 3.1.9 con $H(x,y) = \frac{1}{2}\|x-y\|^2$ e dimostriamo che converge linearmente [29]. Prima di enunciare il teorema in cui viene dimostrato che la convergenza è lineare assumiamo che per ogni $x \in K$, la funzione f è convessa rispetto a y, semicontinua inferiormente rispetto a y e sottodifferenziabile su K.

Per ogni $x \in K$, definiamo la mappa

$$S(x) \doteq \text{argmin}_{y\in K}\left\{f(x,y) + \frac{1}{2}\|y-x\|^2\right\}. \tag{3.12}$$

Poiché la funzione è fortemente convessa nella seconda variabile, (3.12) ammette un'unica soluzione. La mappa è ben definita ed è a valori singoli.

Osservazione 3.3.5. *$S(x) = y(x)$ dove $H(x,y) = \frac{1}{2}\|x-y\|^2$ dove $t_k = 1$.*

Per risolvere il problema di equilibrio (EP) consideriamo l'algoritmo iterativo $x^{k+1} = S(x^k)$. In [21] è stato dimostrato che la successione $\{x^k\}$ converge fortemente all'unica soluzione del problema di equilibrio (EP) se f è fortemente monotona e soddisfa la seguente condizione di lipschitzianità:

esistono costanti $L_1 > 0$ e $L_2 > 0$ tali che

$$f(x,y)+f(y,z) \geq f(x,z)-L_1\|x-y\|^2-L_2\|y-z\|^2,\ \forall\, x,y,z \in K. \tag{3.13}$$

Ponendo $x = z$ si ha

$$f(x,y) + f(y,x) \geq -(L_1+L_2)\|x-y\|^2,\ \forall x,y \in K.$$

Così, se inoltre f è fortemente monotona su K con modulo τ, allora $\tau \leq L_1 + L_2$. Nel seguito chiameremo L_1 e L_2 le costanti di Lipschitz per f.

Ora possiamo enunciare il teorema che dimostra che la successione $\{x^k\}$ definita da $x^{k+1} = S(x^k)$ converge all'unica soluzione del problema (EP).

Teorema 3.3.6. *Supponiamo che f sia fortemente monotona su K con modulo τ e soddisfi la condizione di Lipschitz* (3.13). *Allora, per ogni punto iniziale $x_0 \in K$ la successione $\{x^k\}$ definita da*

$$x^{k+1} \doteq \text{argmin}_{y\in K}\left\{f(x^k,y) + \frac{1}{2}\|y-x^k\|^2\right\}, \tag{3.14}$$

soddisfa

$$\|x^{k+1} - \bar{x}\|^2 \leq \alpha \|x^k - \bar{x}\|^2, \ \forall k \geq 0, \tag{3.15}$$

purché $L_2 \leq \dfrac{1}{2}$, *dove* $L_1, L_2 > 0$ *sono le costanti di Lipschitz per* f, $\bar{x}$ *è l'unica soluzione del problema* (EP) *e* $\alpha \doteq 1 - 2(\tau - L_1)$.

Dimostrazione. Per ogni $k \geq 0$ sia

$$f_k(x) \doteq f(x^k, x) + \frac{1}{2}\|x - x^k\|^2.$$

Dalla convessità di $f(x^k, \cdot)$ deduciamo che la funzione f_k è fortemente convessa su K con modulo 1. Ciò implica

$$f_k(x^{k+1}) + \langle w^k, x - x^{k+1} \rangle + \frac{1}{2}\|x - x^{k+1}\|^2 \leq f_k(x), \ \forall x \in K, \tag{3.16}$$

dove $w^k = f_k(x^{k+1})$. Poiché x^{k+1} è la soluzione del problema (3.14), abbiamo che $\langle w^k, x - x^{k+1} \rangle \geq 0$ per ogni $x \in K$. Allora, dalla (3.16), segue che

$$f_k(x^{k+1}) + \frac{1}{2}\|x - x^{k+1}\|^2 \leq f_k(x), \ \forall x \in K. \tag{3.17}$$

Applicando la disuguaglianza (3.17) con $x = \bar{x}$ e usando la definizione di f_k, otteniamo

$$\begin{aligned} &\|x^{k+1} - \bar{x}\|^2 \leq \\ &\qquad \leq 2\big(f(x^k, \bar{x}) - f(x^k, x^{k+1})\big) + \|x^k - \bar{x}\|^2 - \|x^{k+1} - x^k\|^2. \end{aligned} \tag{3.18}$$

Poichè f è fortemente monotona su K con modulo τ, otteniamo la disuguaglianza

$$f(x^k, \bar{x}) \leq -f(\bar{x}, x^k) - \tau\|x^k - \bar{x}\|^2.$$

Sostituendo questa diseguaglianza nella (3.18), abbiamo

$$\begin{aligned} &\|x^{k+1} - \bar{x}\|^2 \leq \\ &\leq (1 - 2\tau)\|x^k - \bar{x}\|^2 + 2\big(-f(\bar{x}, x^k) - f(x^k, x^{k+1})\big) - \|x^{k+1} - x^k\|^2. \end{aligned} \tag{3.19}$$

Ora, applicando la condizione di Lipschitz (3.13) con $x = \bar{x}$, $y = \bar{x}$, e $z = x^{k+1}$, otteniamo

$$\begin{aligned} -f(x^k, x^{k+1}) - f(\bar{x}, x^k) &\leq -f(\bar{x}, x^{k+1}) + L_1\|\bar{x} - x^k\|^2 + L_2\|x^k - x^{k+1}\|^2 \leq \\ &\leq L_1\|\bar{x} - x^k\|^2 + L_2\|x^k - x^{k+1}\|. \end{aligned} \tag{3.20}$$

L'ultima diseguaglianza nella (3.20) segue dal fatto che $f(\bar{x}, x^{k+1}) \geq 0$, poiché $\bar{x}$ è la soluzione del problema di equilibrio (EP). Sostituendo nella (3.19), otteniamo

$$\|x^{k+1} - \bar{x}\|^2 \leq \big(1 - 2(\tau - L_1)\big)\|x^k - \bar{x}\|^2 - (1 - 2L_2)\|x^{k+1} - x^k\|^2. \quad (3.21)$$

Poiché $L_2 \leq \frac{1}{2}$, segue la disuguaglianza

$$\|x^{k+1} - \bar{x}\|^2 \leq [1 - 2(\tau - L_1)]\|x^k - \bar{x}\|^2, \quad (3.22)$$

ossia la tesi. □

Il seguente corollario segue immediatamente dal teorema precedente.

Corollario 3.3.7. *Siano L_1, L_2 le costanti di Lipschitz per f tali che $L_1 < \tau$ e $L_2 \leq \frac{1}{2}$. Allora*

$$\|x^{k+1} - \bar{x}\| \leq r\|x^k - \bar{x}\|, \ \forall k \geq 0,$$

dove $0 < r \doteq \sqrt{1 - 2(\tau - L_1)} < 1$.

Osservazione 3.3.8. *Poichè $\tau \leq L_1 + L_2$ e $L_2 \leq \frac{1}{2}$, si ha che $2(\tau - L_1) < 1$. Allora r raggiunge il valore minimo in $L_2 = \frac{1}{2}$.*

3.4 Applicazioni alle disuguaglianze variazionali e ai problemi di ottimizzazione

In questa sezione, come detto all'inizio del capitolo sono analizzati casi particolari di problemi di equilibrio che sono risolvibili con algoritmi che utilizzano le funzioni gap. Nel caso particolare di una disequazione variazionale generalizzata, vedremo che la condizione di forte monotonia per la funzione F è sufficiente per trovare la soluzione ottima applicando gli algoritmi con ricerca inesatta 3.2.3 e 3.2.5.

Consideriamo la diseguaglianza:

trovare $\bar{x} \in K$ tale che

$$\langle F(\bar{x}), y - \bar{x}\rangle \geq \phi(\bar{x}) - \phi(y), \ \forall\, y \in K, \quad (VI(F, \phi))$$

dove $F : \mathbb{R}^n \longrightarrow \mathbb{R}^n$, K è un insieme convesso di $\mathbb{R}^n$, $\phi : \mathbb{R}^n \longrightarrow \mathbb{R}$.

Se definiamo $f(x, y) \doteq \langle F(x), y-x\rangle - \phi(x) + \phi(y)$ allora il problema (EP) è equivalente al problema $(VI(F, \phi))$. Quando la funzione F è monotona e ϕ è una funzione convessa, $(VI(F, \phi))$ è equivalente alla seguente disuguaglianza variazionale:

trovare $\bar{x} \in K$ tale che

$$\langle F(y), y - \bar{x}\rangle \geq \phi(\bar{x}) - \phi(y), \ \forall y \in K. \qquad (MVI(F, \phi))$$

Proposizione 3.4.1. *Sia $\phi : \mathbb{R}^n \longrightarrow \mathbb{R}$ una funzione convessa su K e F sia una funzione continua monotona su K. Allora $\bar{x}$ è una soluzione per $(VI(F, \phi))$ se e solo se è una soluzione di $(MVI(F, \phi))$.*

Proposizione 3.4.2. *Sia F differenziabile con continuità su K, sia ϕ una funzione convessa differenziabile su K e sia*

i) $f(x, y) \doteq \langle F(y), y - x\rangle + \phi(y) - \phi(x)$ *oppure*

ii) $f(x, y) \doteq \langle F(x), y - x\rangle + \phi(y) - \phi(x)$.

Allora

1) se F è strettamente monotona su K allora vale (3.1),

2) se F è fortemente monotona su K, con modulo μ, allora vale (3.5).

Dimostrazione. Nel caso (i), noi abbiamo

$$f_x(x, y) + f_y(x, y) = \nabla F(y)(y - x) + \phi'(y) - \phi'(x),$$

mentre nel caso (ii)

$$f_x(x, y) + f_y(x, y) = \nabla F(x)(y - x) + \phi'(y) - \phi'(x).$$

E' noto [25] che F è strettamente monotona su K se e solo se $\nabla F(y)$ è una matrice definita positiva, $\forall y \in K$, e che F è fortemente monotona su K, con modulo μ, se e solo se

$$\langle \nabla F(x)d, d\rangle \geq \mu \|d\|^2,$$

per ogni $d \in \mathbb{R}^n$ e per ogni $x \in K$. Ponendo $d \doteq y - x$ e ricordando che ϕ è convessa e differenziabile, si ha la tesi. □

Osservazione 3.4.3. *Se F è una funzione differenziabile con continuità e monotona (risp. fortemente monotona) su K e ϕ è una funzione differenziabile strettamente convessa su K, allora l'Algoritmo 3.1.3 (risp. 3.2.3) può essere applicato per risolvere la $VI(F, \phi)$. Per la VI, ottenuta ponendo $\phi(y) = 0$, per ogni $y \in K$, possiamo applicare l'Algoritmo 3.1.9: visto che $H(x, \cdot)$ è una funzione strettamente convessa, per ogni $x \in K$, allora nel Teorema 3.1.7, la funzione $f(x, \cdot)$ può essere supposta convessa. Circa la condizione* (3.3), *noi osserviamo che è soddisfatta quando F è strettamente monotona su K provato che:*

$$\langle H_x^{'}(x, y) + H_y^{'}(x, y), y - x\rangle \geq 0, \ \forall (x, y) \in K \times K.$$

Come già osservato, se H soddisfa l'asserzione:

$$H_x^{'}(x, y) + H_y^{'}(x, y) = 0, \quad \forall (x, y) \in K \times K \tag{3.23}$$

allora è sufficiente supporre che F sia strettamente monotona su K.

In modo simile, gli algoritmi generali possono essere applicati a $MVI(F, \phi)$; l'Algoritmo 3.1.3 (risp. 3.2.3) può essere applicato, nelle ipotesi della stretta monotonia (risp. forte monotonia) dell'operatore F, con l'ulteriore condizione:

> $f(x, y) \doteq \langle F(y), y - x\rangle - \phi(x) + \phi(y)$ è una funzione strettamente convessa su K, rispetto a y, per ogni $x \in K$.

Gli Algoritmi 3.1.9 e 3.2.5 richiedono la stretta convessità (rispetto a y) della funzione:

$$f(x, y) \doteq \langle F(y), y - x\rangle + H(x, y) - \phi(x) + \phi(y),$$

così che possano essere applicate nel caso in cui la funzione

$$\langle F(y), y - x\rangle - \phi(x) + \phi(y)$$

è convessa, ma non necessariamente strettamente convessa. Ciò succede, ad esempio, se $F(y) = Ay$ con A una matrice definita positiva di ordine n e $\phi(y) = 0$, per $y \in K$.

Per quanto riguarda le applicazioni ai problemi di ottimizzazione, è noto (si veda ad esempio [7]) che $(VI(\nabla\psi, \phi))$ rappresenta la condizione di ottimalità al primo ordine del problema di estremo vincolato

$$\min_{x \in K}\{\psi(x) + \phi(x)\} \tag{3.24}$$

dove

- $\psi\colon \mathbb{R}^n \longrightarrow \mathbb{R}$ è una funzione convessa differenziabile con continuità sull'insieme convesso K,
- $\phi\colon \mathbb{R}^n \longrightarrow \mathbb{R}$ è una funzione continua fortemente convessa su K con derivata direzionale finita $\phi'(y - z, y)$ per ogni $y, z \in K$.

Per completezza enunciamo la proposizione che dimostra l'equivalenza tra il problema (3.24) e le corrispondenti disuguaglianze variazionali.

Proposizione 3.4.4. *$\bar{x}$ è una soluzione del problema* (3.24) *se e solo se è una soluzione del problema $(VI(\nabla\psi, \phi))$, o equivalentemente del problema $(MVI(\nabla\psi, \phi))$.*

Dimostrazione. Cominciamo dimostrando che il problema $(VI(\nabla\psi, \phi))$ è equivalente al problema $(MVI(\nabla\psi, \phi))$. Ciò è vero per la Proposizione 3.4.1, infatti la funzione ϕ è convessa e la funzione $\nabla\psi$ è monotona perché ψ è convessa (si veda quanto detto nella Sezione 1.1).

Dimostriamo ora che una soluzione del problema (3.24) è anche una soluzione del problema $(VI(\nabla\psi, \phi))$. Supponiamo che $\bar{x}$ sia un punto di minimo per il problema (3.24), ossia

$$\psi(\bar{x}) + \phi(\bar{x}) \leq \psi(y) + \phi(y), \ \forall y \in K.$$

Consideriamo il punto $y_t = (1 - t)\bar{x} + ty$ al variare di t nell'intervallo aperto $(0, 1)$ e di y nell'insieme K. Dal fatto che $\bar{x}$ è un punto di minimo per il problema (3.24) e dalla convessità di ϕ deduciamo

$$\begin{aligned}\psi(\bar{x}) + \phi(\bar{x}) \leq \psi(y_t) + \phi(y_t) &\leq \\ &\leq \psi(y_t) + (1 - t)\phi(\bar{x}) + t\phi(y),\end{aligned}$$

e quindi

$$\psi(\bar{x}) \leq t(\phi(y) - \phi(\bar{x})) + \psi(y_t),$$

ossia

$$\frac{\psi(y_t) - \psi(\bar{x})}{t} + \phi(y) - \phi(\bar{x}) \geq 0.$$

Passando al limite e utilizzando la differenziabilità di ψ otteniamo

$$\langle \nabla\psi(\bar{x}), y - \bar{x} \rangle + \phi(y) - \phi(\bar{x}) \geq 0, \ \forall y \in K,$$

ossia $\bar{x}$ è una soluzione del problema $(VI(\nabla\psi, \phi))$.

Dimostriamo infine il viceversa. Supponiamo che $\bar{x}$ sia una soluzione del problema $(VI(\nabla\psi, \phi))$, ossia

$$\langle \nabla\psi(\bar{x}), y - \bar{x} \rangle \geq \phi(\bar{x}) - \phi(y), \ \forall y \in K.$$

Dalla convessità di ψ e dalla disuguaglianza precedente deduciamo che

$$\begin{aligned} 0 &\leq \langle \nabla\psi(\bar{x}), y - \bar{x} \rangle - \phi(\bar{x}) + \phi(y) \leq \\ &\leq \psi(y) - \psi(\bar{x}) - \phi(\bar{x}) + \phi(y), \end{aligned}$$

per ogni $y \in K$. Quindi abbiamo

$$\psi(y) - \psi(\bar{x}) - \phi(\bar{x}) + \phi(y) \geq 0, \ \forall y \in K,$$

ossia $\bar{x}$ è una soluzione del problema (3.24). □

La proposizione precedente permette di applicare gli algoritmi proposti al problema $(VI(\nabla\psi, \phi))$ e al problema $(MVI(\nabla\psi, \phi))$ per risolvere il problema di ottimizzazione (3.24), se assumiamo che ϕ è differenziabile su K.

3.5 Flusso su reti

In questa sezione vediamo una applicazione delle funzioni gap e D-gap a un problema di flusso su reti.

Una **rete** di flusso consiste in un insieme di **nodi** $\mathcal{N}$, un insieme di **archi** $\mathcal{A} \subset \mathcal{N} \times \mathcal{N}$, e un insieme di **coppie origine/destinazione** $\mathcal{W} \subset \mathcal{N} \times \mathcal{N}$. Per ogni coppia w di origini/destinazioni il problema consiste nel distribuire attraverso i nodi che connettono w una richiesta di flusso d_w. Denotiamo con $\mathcal{P}_w$ l'insieme di tutti i cammini che connettono w, con x_p il flusso sul cammino $p \in \mathcal{P}_w$, e con $(x_p)_{p\in\mathcal{P}_w, w\in\mathcal{W}}$ il vettore di tutti i flussi sui cammini. L'insieme dei flussi sui cammini ammissibili è dato da

$$X = \left\{ x \geq 0 \,\middle|\, \sum_{p\in\mathcal{P}_w} x_p = d_w, \ \forall w \in \mathcal{W} \right\}.$$

Il flusso f_a su ogni arco $a \in \mathcal{A}$ è la somma di tutti i flussi lungo i cammini a cui gli archi appartengono, quindi il vettore dei flussi sugli archi $f = (f_a)_{a\in\mathcal{A}}$ può essere scritto come $f = \Delta x$, dove Δ è la matrice di incidenza arco-cammino:

$$\Delta_{a,p} = \begin{cases} 1 & \text{se } a \in p, \\ 0 & \text{altrimenti.} \end{cases}$$

Per ogni arco a consideriamo una funzione di costo non negativa $t_a(f)$, che rappresenta il tempo di viaggio associato all'arco a e dipende dal vettore dei flussi sugli archi f. La corrispondente funzione di costo sui cammini è supposta additiva, ossia il tempo di viaggio $T_p(x)$ sul cammino p è la somma dei tempi di viaggio degli archi che appartengono al cammino p:

$$T_p(x) = \sum_{a\in p} t_a(\Delta x).$$

In accordo con il principio di equilibrio di Wardrop [32], un flusso su tutti i possibili cammini $\bar{x} \in X$ è detto *equilibrio di rete* se è positivo solo su cammini di costo minimo, ossia si ha l'implicazione

$$\bar{x}_p > 0 \qquad \Longrightarrow \qquad T_p(\bar{x}) = \min_{q\in\mathcal{P}_w} \{T_q(\bar{x})\}$$

per ogni coppia origine/destinazione $w \in \mathcal{W}$ e per ogni cammino $p \in \mathcal{P}_w$.

Il problema di trovare un equilibrio di rete è equivalente a risolvere la seguente disuguaglianza variazionale [8]:

$$\text{trovare} \quad \bar{x} \in X \quad \text{tale che} \quad \langle T(\bar{x}), y - \bar{x} \rangle \geq 0, \quad \forall y \in X. \tag{3.25}$$

Consideriamo la rete mostrata nella Figura 3.1, con due coppie origine/destinazione:

- $w_1 = (1, 4)$ con richiesta $d_1 = 4$,

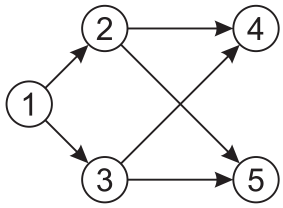

Figura 3.1: Rete dell'esempio.

- $w_2 = (1,5)$ con richiesta $d_2 = 6$.

Ogni coppia è connessa con due cammini:

- $\mathcal{P}_{w_1} = \{(1,2),(2,4);\ (1,3),(3,4)\}$,
- $\mathcal{P}_{w_2} = \{(1,2),(2,5);\ (1,3),(3,5)\}$.

Denotiamo il flusso sui cammini con x_1, x_2, x_3, x_4, rispettivamente. Quindi l'insieme dei flussi sui cammini ammissibili è dato da

$$X = \left\{ x \in \mathbb{R}^4_+ \middle| x_1 + x_2 = 4,\ x_3 + x_4 = 6 \right\}.$$

Assumiamo che le funzioni di costo sugli archi siano date come segue:

$$\begin{cases} t_{12} \doteq f_{12} + 1 = x_1 + x_3 + 1, \\ t_{13} \doteq 3f_{13} + 2 = 3(x_2 + x_4) + 2, \\ t_{24} \doteq 2f_{24} + f_{34} + 1 = 2x_1 + x_2 + 1, \\ t_{25} \doteq 2f_{25} + f_{35} + 3 = 2x_3 + x_4 + 3, \\ t_{34} \doteq f_{34} + 2 = x_2 + 2, \\ t_{35} \doteq 4f_{35} + 1 = 4x_4 + 1. \end{cases}$$

Quindi i corrispondenti costi sui cammini sono

$$\begin{cases} T_1 = t_{12} + t_{24} = 3x_1 + x_2 + x_3 + 2, \\ T_2 = t_{13} + t_{34} = 4x_2 + 3x_4 + 4, \\ T_3 = t_{12} + t_{25} = x_1 + 3x_3 + x_4 + 4, \\ T_4 = t_{13} + t_{35} = 3x_2 + 7x_4 + 3, \end{cases}$$

ossia l'operatore della disequazione variazionale (3.25) è $T(x) = Ax + B$ con

$$A = \begin{pmatrix} 3 & 1 & 1 & 0 \\ 0 & 4 & 0 & 3 \\ 1 & 0 & 3 & 1 \\ 0 & 3 & 0 & 7 \end{pmatrix} \qquad \text{e} \qquad B = \begin{pmatrix} 2 \\ 4 \\ 4 \\ 3 \end{pmatrix}.$$

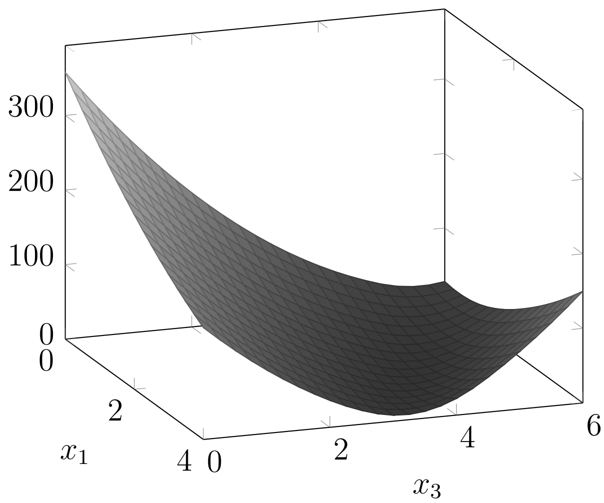

Figura 3.2: Il grafico della funzione gap regolarizzata p^F nell'esempio.

Notiamo che la matrice A è definita positiva, quindi la mappa T è fortemente monotona e esiste una soluzione unica della disequazione variazionale (3.25), ossia un solo equilibrio di rete.

Consideriamo ora la funzione gap regolarizzata di Fukushima, ossia

$$p^F = \max_{y \in X} \left\{ \langle T(x), x - y) \rangle - \frac{1}{2} \|x - y\|^2 \right\}.$$

Questa funzione è differenziabile con continuità e fortemente convessa, perché la matrice $A + A^T - I$ è definita positiva. Nella Figura 3.2 mostriamo il grafico di p^F relativamente alla coppia di variabili (x_1, x_3), con $x_1 \in [0, 4]$ e $x_3 \in [0, 6]$. Non è necessario considerare anche le variabili x_2 e x_4, perché i vincoli relativi alla richiesta di flusso permettono di scrivere le variabili x_2 e x_4 in funzione delle variabili x_1 e x_3.

L'algoritmo di discesa 3.1.3 applicato a p^F può essere utilizzato per calcolare l'equilibrio di rete. Abbiamo implementato l'algoritmo utilizzando il programma Sage (il codice e l'output possono essere trovati nell'Appendice C). La Tabella 3.1 riporta le prime quattro iterazioni dell'algoritmo partendo dal flusso ammissibile $(4, 0, 6, 0)$. I risultati confermano quelli trovati in [3].

Notiamo che i costi dei cammini che corrispondono alla soluzione di

Iterazione	x_1	x_2	x_3	x_4	$p^F(x)$
0	4	0	6	0	$1.5000 \cdot 10^2$
1	2.679622	1.320378	4.019434	1.980566	$8.8840 \cdot 10^{-3}$
2	2.631471	1.368529	4.052452	1.947548	$1.5156 \cdot 10^{-6}$
3	2.631587	1.368413	4.052626	1.947374	$2.5905 \cdot 10^{-10}$
4	2.631579	1.368421	4.052632	1.947368	$1.1450 \cdot 10^{-14}$

Tabella 3.1: Risultati numerici dell'algoritmo di discesa applicato all'esempio.

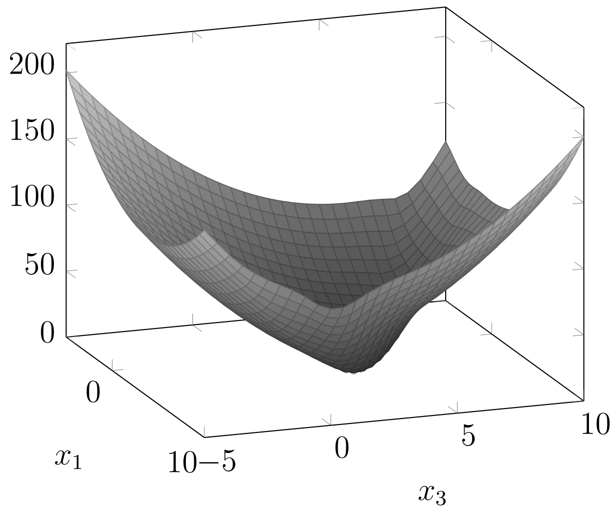

Figura 3.3: Il grafico della funzione D-gap $p_{\alpha\beta}$, con $\alpha = 1$, $\beta = 2$ e $H(x, y) = \frac{\|x-y\|^2}{2}$ nell'esempio.

equilibrio

$$\bar{x} = (2.6316,\ 1.3684,\ 4.0526,\ 1.9474)$$

sono

$$T(\bar{x}) = (15.3158,\ 15.3158,\ 20.7368,\ 20.7368),$$

ossia i due cammini che connettono ogni coppia origine/destinazione hanno lo stesso costo.

La Figura 3.3 mostra il grafico della funzione D-gap $p_{\alpha\beta}$, con $\alpha = 1$, $\beta = 2$ e $H(x, y) = \frac{\|x-y\|^2}{2}$, definita sullo spazio (x_1, x_3). Notiamo che questa funzione è sempre non negativa, anche nei punti non ammissibili, e il suo minimo globale è $\bar{x}$.

Appendice A

Richiami sulle definizioni utilizzate nella tesi

In questa appendice richiamiamo alcune definizioni base che abbiamo usato nella tesi.

Prodotto scalare

Definizione A.0.1. *Il* **prodotto scalare** *tra due vettori di* $\mathbb{R}^n$ *è indicato con*

$$\langle \cdot, \cdot \rangle : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}.$$

Dualità Definiamo il duale di $\mathbb{R}^n$.

Definizione A.0.2. *Il* **duale** *di* $\mathbb{R}^n$ *è lo spazio* $(\mathbb{R}^n)^* := \{f \colon \mathbb{R}^n \to \mathbb{R} \text{ lineare}\}$.

Polarità Definiamo il cono polare di un insieme.

Definizione A.0.3. *Dato* $K \subseteq \mathbb{R}^n$ *l'insieme*

$$K^* := \{x^* \in (\mathbb{R}^n)^* | \langle x^*, y \rangle \geq 0 \ \forall y \in K\}$$

è chiamato **cono polare** *di* K.

(a) Semicontinua inferiormente (b) Semicontinua superiormente

Figura A.1: Grafici di funzioni semicontinue.

Continuità Diamo la definizione di **limsup** e **liminf**.

Definizione A.0.4. *Sia X un aperto di $\mathbb{R}^n$ e sia $f : X \longrightarrow \mathbb{R}$ una funzione. Definiamo per un punto di accumulazione a di E*

$$\limsup_{x \to a} f(x) = \lim_{\varepsilon > 0} \big(\sup\{ f(x) | x \in X \cap B(a, \varepsilon) - \{a\}\} \big)$$

e

$$\liminf_{x \to a} f(x) = \lim_{\varepsilon > 0} \big(\inf\{ f(x) | x \in X \cap B(a, \varepsilon) - \{a\}\} \big),$$

dove $B(a, \varepsilon)$ è la palla di centro a e raggio ε.

Definizione A.0.5. *Sia X un aperto di $\mathbb{R}^n$ e sia $f : X \longrightarrow \mathbb{R}$ una funzione. Diciamo che f è* **semicontinua inferiormente nel punto** *$x_0 \in X$ se si ha*

$$f(x_0) \leq \liminf_{x \to x_0} f(x)$$

(Figura A.1(a)). Diciamo che f è **semicontinua superiormente nel punto** *$x_0 \in X$ se si ha*

$$f(x_0) \geq \limsup_{x \to x_0} f(x)$$

(Figura A.1(b)). La funzione f si dice **semicontinua inferiormente**, *o* **superiormente**, **in** *X se lo è in ogni punto di X.*

Definizione A.0.6. *Se $X \subseteq \mathbb{R}^n$, l'***epigrafico** *di una funzione $f : X \longrightarrow \mathbb{R}$ è l'insieme*

$$\{(x_1, \ldots, x_n, y) \in \mathbb{R}^{n+1} \mid y \geq f(x_1, \ldots, x_n)\}$$

(Figura A.2).

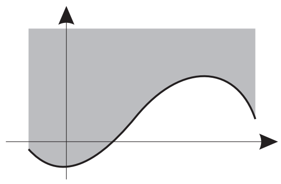

Figura A.2: Epigrafico di una funzione da $\mathbb{R}$ in $\mathbb{R}$.

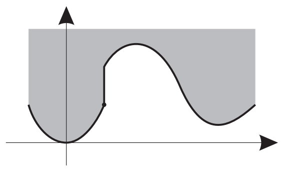

Figura A.3: L'epigrafico di una funzione semicontinua inferiormente.

Osservazione A.0.7. *Una funzione è semicontinua inferiormente se e solo se l'epigrafico è chiuso (Figura A.3); è semicontinua superiormente se e solo se l'epigrafico di* $-f$ *è chiuso.*

Definizione A.0.8. *Sia* $X \subseteq \mathbb{R}^n$ *e sia* $F\colon X \to \mathbb{R}^m$*. Diciamo che* F *è* **Lipschitziana***, con modulo* $L > 0$*, se si ha*

$$\|F(x) - F(y)\| \leq L\|x - y\| \quad \textit{per ogni} \quad x, y \in X$$

(Figura A.4).

Osservazione A.0.9. *Se una funzione* $f\colon X \to \mathbb{R}$ *è differenziabile con continuità e se* ∇f *è Lipschitziana con modulo* L*, abbiamo*

$$f(x) - f(y) \leq \langle \nabla f(y), x - y \rangle + \frac{L}{2}\|x - y\|^2 \quad \textit{per ogni} \quad x, y \in X.$$

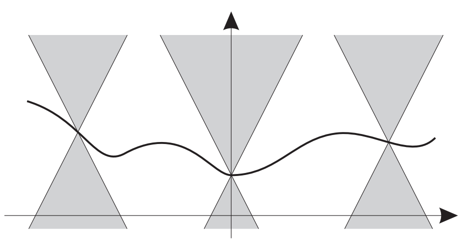

Figura A.4: Il grafico di una funzione Lipschitziana. Le regioni colorate sono esempi di zone del piano che non devono contenere pezzi del grafico.

Differenziabilità secondo Gâteaux Da noi verrà utilizzata la differenziabilità secondo Gâteaux la quale è diversa dalla definizione di differenziabilità usuale perché il limite della seconda è fatto su tutto un intorno mentre il limite della prima è fatto su rette.

Definizione A.0.10. *Sia $\varphi\colon X \to \mathbb{R}$ con $X \subseteq \mathbb{R}^n$ aperto. Il* **differenziale secondo Gâteaux di f in** $x \in X$ **nella direzione** $v \in \mathbb{R}^n$ *è definito come*

$$Df(x;v) := \lim_{t\to 0} \frac{f(x+tv) - f(x)}{t}$$

se il limite esiste. Se il limite esiste per ogni $v \in \mathbb{R}^n$ e la funzione $Df(x;\cdot)$ è lineare, la funzione f si dice **differenziabile secondo Gâteaux in** x.

La definizione seguente è equivalente a quella di funzione differenziabile secondo Gâteaux.

Definizione A.0.11. *Sia X un aperto di $\mathbb{R}^n$. Una funzione $f : X \longrightarrow \mathbb{R}^m$ si dice* **di Gâteaux** *nel punto $x_0 \in X$ (o più brevemente* **G-differenziabile** *in x_0) se esiste un operatore lineare $A \in L(\mathbb{R}^n, \mathbb{R}^m)$ tale che*

$$\lim_{t\to 0} \left\| \frac{f(x_0+tv) - f(x_0) - Av}{t} \right\| = 0, \ \forall\ v \in \mathbb{R}^n.$$

A è detto **operatore differenziale di Gâteaux** *di f in x_0.*

Osservazione A.0.12. *In questa definizione si richiede non solo che esista il differenziale di Gâteaux di f in x_0 secondo la direzione $v \in X$, ma anche che la dipendenza di tale derivata da v sia lineare; ciò non è sempre vero.*

Definizione A.0.13. *Una funzione $f\colon X \to \mathbb{R}^m$ con $X \subseteq \mathbb{R}^n$ aperto è* **differenziabile** *in un punto x_0 del dominio se esiste un'applicazione lineare $L\colon \mathbb{R}^n \to \mathbb{R}^m$ tale che*

$$\lim_{\|h\|\to 0} \frac{f(x_0+h) - f(x_0) - L(h)}{\|h\|} = 0.$$

Sottogradiente Definiamo il sottogradiente.

Definizione A.0.14. *Dati un insieme convesso* $C \subseteq \mathbb{R}^n$ *e una funzione* $f : C \longrightarrow \mathbb{R}$ *convessa su* C*, un vettore* $\gamma \in \mathbb{R}^n$ *è un* **sottogradiente** *di* f *in un punto* $\bar{x} \in C$ *se*

$$f(x) \geq f(\bar{x}) + \gamma^t \cdot (x - \bar{x}), \quad \forall x \in C.$$

L'iperpiano

$$H = \{(x, y) | y = f(\bar{x}) + \gamma^t \cdot (x - \bar{x})\}$$

è detto **di supporto** *all'epigrafico di* f*. L'insieme dei sottogradienti di* f *in* $\bar{x}$ *è indicato con* $\nabla f(\bar{x})$*.*

Osservazione A.0.15. *Una funzione* $f : C \longrightarrow \mathbb{R}$ *convessa possiede almeno un sottogradiente in ogni punto interno* $\bar{x} \in C$*. Ossia se* $\bar{x} \in intC$ *allora esiste* $\gamma \in \mathbb{R}^n$ *tale che l'iperpiano:*

$$H = \{(x, y) | y = f(\bar{x}) + \gamma^t \cdot (x - \bar{x})\}$$

è di supporto all'epigrafico di f *in* $(\bar{x}, f(\bar{x}))$*.*

Osservazione A.0.16. *Sia data* $f : C \longrightarrow \mathbb{R}$*.*

1) L'esistenza di almeno un sottogradiente in ogni punto di $intC$*, con* C *convesso, è una condizione necessaria e sufficiente affinché* f *sia convessa su* $intC$*.*

2) Se f *è convessa e* $x \in C$*,* $\nabla f(x)$ *è un insieme non vuoto, convesso e limitato.*

3) Un punto $\bar{x}$ *è un ottimo (globale) di* f *se e solo se* $0 \in \nabla f(\bar{x})$*.*

Multifunzioni

Definizione A.0.17. *Siano* X *e* Y *due insiemi e sia* $\mathcal{P}(Y)$ *l'insieme delle parti di* Y*. Una* **multifunzione** f *da* X *in* Y *è una funzione*

$$f\colon\ X \longrightarrow \mathcal{P}(Y).$$

Essa viene anche indicata con

$$f\colon\ X \rightrightarrows Y.$$

Quando parliamo di multifunzioni chiameremo quelle classiche **a valori singoli**.

Definizione A.0.18. *Sia* $X \subseteq \mathbb{R}^n$*. Una multifunzione* $f\colon X \rightrightarrows \mathbb{R}^m$ *si dice* **chiusa in un punto** $x \in X$ *se*

$$x_k \longrightarrow x,\ y_k \longrightarrow y,\ \text{con } y_k \in f(x_k)\ \forall k \quad \Longrightarrow \quad y \in f(x)$$

La funzione f *si dice* **chiusa in un sottoinsieme** $S \subset X$ *se è chiusa in ogni punto di* S*.*

Definizione A.0.19. *Sia* $X \subseteq \mathbb{R}^n$*. Una multifunzione* $f : X \rightrightarrows \mathbb{R}^m$ *è* **semicontinua superiormente secondo Berge** *nel punto* $\bar{x} \in X$ *se, per ogni aperto* U *contenente* $f(\bar{x})$*, esiste un intorno* V *di* $\bar{x}$ *tale che*

$$f(x) \subset U, \quad \forall x \in V.$$

Definizione A.0.20. *Sia* $X \subseteq \mathbb{R}^n$*. Una multifunzione* $f : X \rightrightarrows \mathbb{R}^m$ *è* **semicontinua inferiormente secondo Berge** *nel punto* $\bar{x} \in X$ *se, per ogni aperto* U *che interseca* $f(\bar{x})$*, esiste un intorno* V *di* $\bar{x}$ *tale che*

$$f(x) \cap U \neq \emptyset, \quad \forall x \in V.$$

Definizione A.0.21. *Sia* $X \subseteq \mathbb{R}^n$*. Una multifunzione* $f : X \rightrightarrows \mathbb{R}^m$ *è* **semicontinua secondo Berge** *se è semicontinua inferiormente e superiormente secondo Berge.*

Definizione A.0.22. *Dati* $X \subseteq \mathbb{R}^n$ *e una funzione* $f : X \longrightarrow \mathbb{R}$*. La funzione* **argmin** *rappresenta l'argomento del minimo, cioè l'insieme dei punti per i quali la funzione* f *raggiunge i suoi valori minimi:*

$$\operatorname{argmin}_{x \in X} f(x) \doteq \{x \in X \mid f(y) \geq f(x)\ \forall y \in X\}.$$

Appendice B

Metodi del gradiente

Dopo aver analizzato i metodi di ricerca esatta e inesatta che ci permettono di trovare un minimo locale per problemi vincolati con un insieme convesso, in questa appendice analizzeremo il Metodo del gradiente usato per problemi non vincolati.

Sia $f : \mathbb{R}^n \longrightarrow \mathbb{R}^n$, differenziabile e d una direzione di discesa allora

$$f(x_k + d) = f(x_k) + \langle \nabla f(x), d \rangle + \beta_1(x_k, d)$$

con $\lim_{\|d\| \to 0} \dfrac{\beta_1(x_k, d)}{\|d\|} = 0$.

Vogliamo approssimare la funzione $f(x_k + d)$ con la funzione $\psi_k(d) \doteq f(x_k) + \langle \nabla f(x_k), d \rangle$ scegliendo come direzione quella che minimizza la funzione $\psi_k(d)$ nella sfera di raggio unitario ovvero d è la soluzione di

$$\min_{\|d=1\|} \psi_k(d)$$

che equivale a

$$\min_{\|d=1\|} \langle \nabla f(x_k), d \rangle.$$

Si ha quindi la soluzione:

$$\bar{d} = -\frac{\nabla f(x_k)}{\|\nabla f(x_k)\|}$$

e quindi abbiamo

$$x_{k+1} = x_k - \tilde{\alpha}_k \frac{\nabla f(x_k)}{\|\nabla f(x_k)\|}.$$

Metodo del sottogradiente Il metodo del sottogradiente è un metodo di discesa che risolve il seguente problema:

$$\min_{x \in \mathbb{R}^n} f(x)$$

(i) con $f : \mathbb{R}^n \longrightarrow \mathbb{R}$,

(ii) f di classe C^1 su $\mathbb{R}^n$,

(iii) $f(x) \geq 0, \forall x \in \mathbb{R}^n$,

(iv) $\gamma_k \in \nabla f(x_k)$ con $\gamma_k \neq 0$.

Sia $-\dfrac{\gamma_k}{\|\gamma_k\|}$ una direzione di discesa.

Ora possiamo descrivere l'algoritmo:

Algoritmo B.0.23. *Sia f una funzione tale che valgano le proprietà (i),...,(iv) di sopra*

Passo 1: *Scegliamo $x_0 \in X$, $\epsilon > 0$ e $k = 0$.*

Passo 2: *Se $\|\gamma_0\| < \epsilon$* **STOP** .

Passo 3: *Sia $d_k = -\dfrac{\gamma_k}{\|\gamma_k\|}$ e sia α_k la soluzione ottima del problema di minimizzare $f(x^k + \alpha)d_k)$ con vincolo $\alpha \geq 0$. Sia $x^{k+1} = x^k + \alpha_k d_k$. Poniamo $k = k + 1$.*

Passo 4: *Se $\|\gamma_k(x^k)\| < \epsilon$* **STOP**. *Altrimenti vai al* **Passo 2**.

Non si effettua la ricerca in una direzione perché nel caso delle funzioni non ovunque differenziabili $-\gamma$ con $\gamma \in \partial f(x)$ non è necessariamente una direzione di discesa. Valgono i seguenti risultati di cui non daremo la dimostrazione.

Lemma B.0.24. *Sia x_k una soluzione non ottima e $\bar{x}$ una qualsiasi soluzione ottima, allora*

$$0 < \alpha_k < 2\frac{f(x_k) - f(\bar{x})}{\|\gamma_k\|}$$

implica

$$\|x_{k+1} - \bar{x}\| < \|x_k - \bar{x}\|.$$

Teorema B.0.25. *Se f è convessa, $\lim_{\|x\|\to\infty} f(x) = +\infty$, $\lim_{k\to\infty} \alpha_k = 0$ e $\sum_{k=0}^{\infty} \alpha_k = \infty$, allora il metodo del sottogradiente termina dopo un numero finito di iterazioni con una soluzione ottima $\bar{x}$ o genera una successione infinita di punti x_k che ammette una sottosuccessione che converge verso $\bar{x}$.*

Selezione del passo: Sia $\{\alpha_k\}$ tale che $\lim_{k\to\infty} \alpha_k = 0$ e $\sum_{k=0}^{\infty} \alpha_k = \infty$ e $0 < \alpha_k < 2\dfrac{f(x_k) - f(\bar{x})}{\|\gamma_k\|}$ per ogni k (con stima di $f(\bar{x})$ se il valore non è noto).

Criterio di arresto: Si effettua un numero massimo di iterazioni perché anche se $0 \in \nabla f(x_k)$ non è detto che in x_k si adoperi quel sottogradiente.

Osservazione B.0.26. *Per estendere il metodo, al caso con vincolo insiemistico (limiti superiori e/o inferiori sulle variabili), basta effettuare ad ogni iterazione una proiezione.*

Appendice C

Codice

In questa appendice mostriamo il codice usato per applicare il metodo di discesa con ricerca esatta alla funzione gap p^F dell'esempio del flusso su reti (Sezione 3.5). Abbiamo usato il programma Sage.

```
# variabili
x1, x2, x3, x4 = var('x1, x2, x3, x4')
y1, y2, y3, y4 = var('y1, y2, y3, y4')
X = vector([x1,x2,x3,x4])
Y = vector([y1,y2,y3,y4])
alfa = var('alfa') # coefficiente per funzione gap

# vincoli
x2 = 4 - x1
x4 = 6 - x3
y2 = 4 - y1
y4 = 6 - y3

# funzione dei costi sui cammini
A = matrix([[3,1,1,0],[0,4,0,3],[1,0,3,1],[0,3,0,7]])
B = vector([2,4,4,3])
T = A*X + B
T

# funzione da massimizzare nella funzione gap
f = ( T.row()*(X-Y) - alfa/2*(X-Y).row()*(X-Y) )[0]
f.expand() # per visualizzare la forma esplicita
```

```
# Metodo di discesa per trovare il punto di minimo della funzione gap
n = 10 # intervalli
esponentePrecisione = 7 # livello di precisione: ordine di grandezza
Xstar = vector([4,6]) # punto iniziale
for iterazione in range(1,5):
    print ''
    print 'Iterazione', iterazione
    Xstart = Xstar
    g(y1,y3) = -f(x1=Xstart[0],x3=Xstart[1],alfa=1)
    Ydiscesa = minimize_constrained(g,[[0,4],[0,6]],[2,3])
    pXstart = -g(y1=Ydiscesa[0], y3=Ydiscesa[1])
    pXstar = pXstart
    DirezioneDiscesa = Ydiscesa - Xstart
    print 'Xstart =', Xstart
    print 'pXstart =', pXstart
    print 'Ydiscesa =', Ydiscesa
    print 'DirezioneDiscesa =', DirezioneDiscesa
    g(y1,y3) = -f(x1=Ydiscesa[0],x3=Ydiscesa[1],alfa=1)
    minYdiscesa = minimize_constrained(g,[[0,4],[0,6]],[2,3])
    pYdiscesa = -g(y1=minYdiscesa[0], y3=minYdiscesa[1])
    tXstar = 0 # t corrispondente al punto con funzione gap minore
    for j in range(1,esponentePrecisione+1):
        # al passo j ci muoviamo vicino a tPassoj
        # (che e' il tXstar trovato al passo j precedente)
        tPassoj = tXstar
        if tXstar > 0: # se possiamo muoverci a sinistra di tXstar
            for i in range(0,n+1):
                t = tPassoj - i/(n**j)
                XstarParziale = Xstart+t*DirezioneDiscesa
                g(y1,y3) = -f(x1=XstarParziale[0],x3=XstarParziale[1],alfa=1)
                minXstarParziale = minimize_constrained(g,[[0,4],[0,6]],[2,3])
                pXstarParziale = -g(y1=minXstarParziale[0], y3=minXstarParziale[1])
                if pXstarParziale < pXstar:
                    # aggiorniamo punto con funzione gap minore
                    Xstar = XstarParziale
                    # aggiorniamo minimo della funzione gap
                    pXstar = pXstarParziale
                    # aggiorniamo t corrispondente al punto con funzione gap minore
                    tXstar = t
        if tXstar < 1: # se possiamo muoverci a destra di tXstar
            for i in range(0,n+1):
```

```
                t = tPassoj + i/(n**j)
                XstarParziale = Xstart+t*DirezioneDiscesa
                g(y1,y3) = -f(x1=XstarParziale[0],x3=XstarParziale[1],alfa=1)
                minXstarParziale = minimize_constrained(g,[[0,4],[0,6]],[2,3])
                pXstarParziale = -g(y1=minXstarParziale[0], y3=minXstarParziale[1])
                if pXstarParziale < pXstar:
                    # aggiorniamo punto con funzione gap minore
                    Xstar = XstarParziale
                    # aggiorniamo minimo della funzione gap
                    pXstar = pXstarParziale
                    # aggiorniamo t corrispondente al punto con funzione gap minore
                    tXstar = t
    print 'Xstar =', Xstar
    print 'pXstar =', pXstar
```

Il codice di sopra produce come output il seguente (cfr. Tabella 3.1).

```
Iterazione 1
Xstart = (4, 6)
pXstart = 150.0
Ydiscesa = (0.0, 0.0)
DirezioneDiscesa = (-4.0, -6.0)
Xstar = (2.6796224, 4.0194336)
pXstar = 0.00888395335935

Iterazione 2
Xstart = (2.6796224, 4.0194336)
pXstart = 0.00888395335935
Ydiscesa = (2.60188799793, 4.07273760052)
DirezioneDiscesa = (-0.0777344020735, 0.0533040005171)
Xstar = (2.63147111503, 4.05245187827)
pXstar = 1.51557543072e-06

Iterazione 3
Xstart = (2.63147111503, 4.05245187827)
pXstart = 1.51557543072e-06
Ydiscesa = (2.63215337848, 4.05347513813)
DirezioneDiscesa = (0.000682263455797, 0.00102325985811)
Xstar = (2.63158714962, 4.05262590715)
pXstar = 2.59046559106e-10
```

```
Iterazione 4
Xstart = (2.63158714962, 4.05262590715)
pXstart = 2.59046559106e-10
Ydiscesa = (2.63157388575, 4.05263502601)
DirezioneDiscesa = (-1.32638724835e-05, 9.11886622479e-06)
Xstar = (2.63157894948, 4.05263154471)
pXstar = 1.14499851285e-14
```

Bibliografia

[1] A. Auslender, "Optimization, Methodes Numeriques", Masson, Paris, 178 pp.

[2] B. Bank – J. Guddat – D. Klatte – B. Kummer – K. Tammer, "Nonlinear Parametric Optimization", Birkhauser Verlag (1983).

[3] G. Bigi – M. Castellani – M. Pappalardo – M. Passacantando, *Existence and solution methods for equilibria*, European Journal of Operational Research **227** (2013), 1–11.

[4] E. Blum – W. Oettli, *From optimization and variational inequalities to equilibrium problems*, The Mathematics Student **63**, n. 1–4 (1993), 1–23.

[5] V. G. Boltyanski – H. Martini – P. Soltan, *Excursions into combinatorial geometry*, New York, Springer (1997), 418+xi pp.

[6] M. Castellani – M. Giuli, *On equivalent equilibrium problems*, Journal of Optimization Theory and Applications **147** (2010), 157–168.

[7] G. Cohen, *Auxiliary problem principle extended to variational inequalities*, Journal of Optimization Theory and Applications **59** (1988), 325–333.

[8] S. Dafermos, *Traffic equilibrium and variational inequalities*, Transportation Science **14** (1980), 42–54.

[9] Q. T. Dinh – L. D. Muu – V. H. Nguyen, *Extragradient methods extended to equilibrium problems*, Optimization **57** (2008), 749–776.

[10] F. Facchinei – J.-S. Pang, *Finite-Dimensional Variational Inequalities and Complementarity Problems. Vol. I,* Springer Verlag, New York, 2003, xxxiv+624+I69 pp.

[11] K. Fan, *A minimax inequality and applications*, In: Inequalities III, ed. by O. Shisha, Accademic Press, New York (1972), pp 103–113.

[12] M. Fukushima, *Equivalent differentiable optimization problems and discendent methods for asymmetric variational inequality problems*, Mathematical Programming **53** (1992), 99–110.

[13] F. Giannessi, *On minty variational principle*, In: News Trends in Mathematical Programming, F: Giannessi, S. Komlosi e T.Rapcsak, (eds.), Kluwer Academic Publishers, Dordrecht (1998), 93–99.

[14] C. J. Goh – X. Q. Yang, *Duality in optimization and variational inequalities*, London; New York: Taylor e Francis (2002), xvi+313 pp.

[15] M. S. Gowda, *Complementarity problems over locally compact cones*, SIAM J. Control Optimization **27** (1989), 836–841.

[16] S. Karamardian, *An existence theorem for the complementary problem*, Journal of Optimization Theory and Applications **18**, 445–456.

[17] R. Knaster – C. Kuratowski – S. Mazurkiewicz, *Ein Beweis des Fixpunktsatzes für n-dimensionale Simplexe*, Fund. Math. **14** (1929), 132–137.

[18] I. V. Konnov – O. V. Pinyagina, *D-gap functions for a class of equilibrium problems in Banach spaces*, Comput Methods Appl Math **3** (2003), 274–286.

[19] D. G. Luemberge, *Introduction to Linear and Nonlinear Programming*, Addison-Wesley.

[20] G. Mastroeni, *Minimax and extremum problems associated to a Variational Inequality*, Supplemento ai Rendiconti del Circolo Matematico di Palemo **58** (1999), 185–196.

[21] G. Mastroeni, *On auxiliary principle for equilibrium problems*, Technical Report of the Department of Mathematics of Pisa University, Italy 3.244.1258.

[22] G. Mastroeni, *Gap Functions for Equilibrium Problems*, Journal of Global Optimization **27** (2003), 411–426.

[23] M. Minoux, *Mathematical Programming Theory and Algorithms*, New York, John Wiley, (1986).

[24] H. Nikaido – K. Isoda, *Note on noncooperative convex games*, Pacific Journal of Mathematics **5** (1955), 807–815.

[25] J. M. Ortega – W. C. Rheinboldt, *Iterative Solutions of Nonlinear Equations in Several Variables*, Academic Press, New York (1970), xx+572 pp.

[26] M. Pappalardo – G. Mastroeni – M. Passacantando, *Merit fuctions: a bridge between optimization and equilibria*, Springer-Verlag Berlin Heidelberg 2014.

[27] J. Ponstein, *Seven kind of comvexity*, SIAM Rev **9** (1967), 115–119.

[28] T. D. Quoc – L. D. Muu, *Iterative methods for solving monotone equilibrium problems via dual gap fuctions*, Computational Optimization and Applications **51** (2012), 709–728.

[29] T. D. Quoc – L. D. Muu, *Regularization Algorithms for Solving Monotone Ky Fan Inequalities with Application to a Nash-Cournot Equilibrium Model*, J. Optim. Theory App. **142** (2009), 185–204.

[30] A. G. Ramm – P. N. Shivakumar – A. V. Strauss editors, *Operator theory and its applications*, Providence, R.I.: American Mathematical Society (2000), 574 pp.

[31] M. V. Solodov, *Merit functions and error bounds for generalized variational inequalities*, Journal of Mathematical Analysis and Applications **287** (2003), 405–414.

[32] J. G. WARDROP, *Some theoretical aspects of road traffic research*, Proceedings of the Institute of Civil Engeneers, Vol. I (1952), 325–378.

[33] N. YAMASHITA – K. TAJI – M. FUKUSHIMA, *Unconstrained optimization reformulations of variational inequality problems*, Journal of Optimization Theory and Applications **92** (1997), 439–456.

[34] W. I. ZANGWILL, *Nonlinear Programming: a unified approach*, Prentice-Hall, Englewood Cliffs, NY (1969), 356 pp.

[35] L. P. ZHANG – J. Y. HAN, *Unconstrained optimization reformulations of equilibrium problems*, Acta Math Sin **25** (2009), 343–354.

[36] D. L. ZHU – P. MARCOTTE, *An extended descent framework for variational inequalities*, Journal of Optimization Theory and Applications, **80** (1994), 349–366.

Ringraziamenti

Chi mi conosce sa che non sono abituata alle smancerie, ma in questo caso mi sembrano doverose, e anche piacevoli. Sono arrivata qui non solo con le mie forze, ma anche con l'aiuto e il sostegno di tantissime persone, che voglio ringraziare.

In primis i miei genitori, che nonostante la strada in salita hanno sempre creduto che ce l'avrei fatta.

Le mie sorelle, inseparabili compagne di viaggio. Senza il loro supporto non sarei qui oggi.

I piccolini, che quando ero triste mi facevano tornare il sorriso chiedendomi quanto mancasse per venire a Pisa.

Tutti gli zii, i cugini, i cognati e gli amici di famiglia, che sono tantissimi, che hanno avuto un peso più o meno determinante al conseguimento di questo risultato.

Chi mi ha dato la forza di riprendere i libri in mano per terminare quello che iniziai.

La mia "segretaria personale": senza di lei la burocrazia, e non solo, sarebbe stata un inferno. E suo "marito", colui che ascolta.

Chi mi ha regalato la "Pacientina" e mi ha sostenuto passo–passo incoraggiandomi e sopportandomi.

Decky e la Fra', che hanno percorso con me la salita degli ultimi esami, non solo compagne di studio, ma anche amiche.

Il buon Nicola, che quando il computer mi abbandonò, pur stanco ebbe la pazienza di smontarlo.

Vania, la mia compagna di fatto.

Laura, l'amica di cicchetti, che mi ha cucinato quando ero stanca.

Mia "figlia" Valentina, che ha creduto più di me nell'ultimo esame e mi ha aspettato per brindare, e non solo.

Tutti coloro che sono stati presenti in questi anni pisani, i miei compagni di studio, di gioia e di dolore, le mie amiche e i miei amici: Marta, Maria B., Natascia, Roberto, Annalisa, Fausto, Luca, Dinora, Michela, Gerardo, Cristian, Sara, Laura M., Tommaso (in ordine volutamente casuale).

I miei "grilli parlanti" Giacomina e Alessandro, che mi hanno sostenuto moralmente.

I miei alunni, che senza saperlo mi hanno spronato ad arrivare fin qui.

Uno dei migliori uomini che esistano per umanità e umiltà, lui che mi è stato vicino durante i miei "scleri", e anche adirandosi non mi ha mai lasciata sola, spronandomi.

Tutti coloro che hanno contribuito matematicamente alla realizzazione di questa tesi, in particolare il paziente Prof. Mastroeni (che mi ha portato a convergere) e il disponibile Prof. Acquistapace.

E, infine, l'ultimo nell'elenco ma tra i più importanti, la persona più buona che abbia mai conosciuto, che non c'è e non può vedermi in questo giorno, ma è stato sempre presente.

Donatella