

Università degli Studi di Trieste – sede amministrativa del
dottorato

Università Commerciale “L. Bocconi” – sede consorziata

Università degli Studi di Torino – sede consorziata

Università degli Studi di Venezia “Ca’Foscari” – sede consorziata

Università degli Studi di Udine – sede consorziata

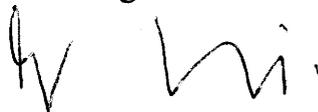
Dottorato in “**Matematica per le decisioni economiche**”
XIII Ciclo

*Un approccio dinamico ai problemi di ottimizzazione
vettoriale*

Tesi di dottorato di **Enrico Miglierina** 16P

CS

Relatore: chiar.mo prof. **Angelo Guerraggio**
Università degli Studi dell’Insubria



Correlatore: chiar.mo prof. **Erio Castagnoli**
Università Commerciale “L.Bocconi”



Coordinatore: chiar.mo prof. **Marco Zecchin**
Università degli Studi di Trieste



*A
Giovanna
e
Guido*

*“La voce della realtà è nel senso dei simboli”
R. Thom*

Indice

Introduzione	iii
1 Ottimizzazione vettoriale	1
1.1 Spazi parzialmente ordinati	2
1.1.1 Definizioni	2
1.1.2 Ordini e coni	3
1.1.3 Esempi	5
1.2 Nozioni di soluzione	7
1.2.1 Elementi minimali di un sottoinsieme di uno spazio vettoriale ordinato	8
1.2.2 Soluzioni di un problema di ottimizzazione vettoriale	11
1.3 Scalarizzazione	13
1.3.1 La scalarizzazione lineare	14
1.3.2 Scalarizzazioni basate su “distanze”	16
1.4 Applicazioni	20
1.4.1 I teoremi del benessere	20
1.4.2 La selezione del portafoglio	23
1.4.3 I giochi cooperativi	24
1.4.4 Applicazioni all’ingegneria	25
2 Inclusioni differenziali	26
2.1 Mappe punto-insieme	26
2.1.1 Generalità e notazioni	26
2.1.2 Continuità per mappe punto-insieme	28
2.1.3 Selezioni di mappe punto-insieme	31
2.2 Teoremi di esistenza per le inclusioni differenziali	37
2.2.1 Inclusioni differenziali a valori convessi	40
2.2.2 Inclusioni differenziali a valori non convessi	43
2.2.3 Mappe monotone massimali	45
3 Approcci “dinamici” ai problemi vettoriali	47
3.1 I lavori di Smale	48
3.2 Il caso convesso: Aubin e Cellina	58

3.3	L'approccio di Malivert	60
3.4	Alcuni lavori recenti	62
4	Un nuovo approccio dinamico	65
4.1	Punti critici per una funzione vettoriale	66
4.2	Lo pseudogradiante generalizzato	69
4.3	L'inclusione differenziale e le sue traiettorie	73
4.3.1	Esistenza delle soluzioni	74
4.3.2	Proprietà asintotiche delle traiettorie	78
4.4	Linee di ricerca future	79
	Appendice 1 - Il teorema della migliore approssimazione	81
	Appendice 2 - Lo pseudogradiante (scalare)	84

Introduzione

Le idee che stanno alla base della presente tesi scaturiscono sostanzialmente da due ordini di osservazioni differenti:

- in un problema di ottimizzazione vettoriale o multiobiettivo risulta utile lavorare nello spazio di definizione della funzione obiettivo di tale problema; la presente osservazione risulta strettamente legata alle applicazioni di natura economica (si pensi ad esempio alla massimizzazione simultanea delle funzioni d'utilità di più agenti oppure alla massimizzazione del valore atteso di un portafoglio ed alla contemporanea minimizzazione della varianza) in quanto i decisori possono agire solo sulle variabili indipendenti per perseguire i propri obiettivi. Questa semplice riflessione suggerisce che l'approccio ai problemi di ottimizzazione vettoriale mediante lo studio di condizioni nello spazio immagine, particolarmente seguita nella letteratura contemporanea (si veda ad esempio [25]), risulta, seppur molto interessante dal punto di vista matematico, relativamente lontana dalle applicazioni; da qui l'esigenza di studiare un nuovo metodo per l'analisi di questi problemi che lavori nello spazio di definizione della funzione obiettivo;
- dal punto di vista teorico risulta poi particolarmente interessante introdurre un elemento dinamico nello studio dei problemi di ottimizzazione vettoriale che usualmente vengono studiati da un punto di vista statico; questo tipo di considerazioni permette infatti di evidenziare legami con la teoria dei sistemi dinamici e sembra essere una strada promettente per studiare la stabilità nell'ambito di questa classe di problemi.

Pertanto lo scopo di questo lavoro consiste nel costruire un possibile approccio dinamico ad un problema di ottimizzazione vettoriale con un occhio di riguardo (almeno negli sviluppi futuri) alla nozione di stabilità ed alle applicazioni in campo economico finanziario.

Alcuni tentativi in questo senso, seppur non numerosissimi, sono noti in letteratura. A questo proposito prima di tutto occorre ricordare i pionieristici lavori, esplicitamente motivati dalle applicazioni economiche, di Stephen Smale [57] -

[66]; in questi lavori viene sviluppato un approccio nel quale i problemi di ottimizzazione vettoriale vengono affrontati alla luce dei risultati, allora appena formulati, riguardanti la teoria delle singolarità di mappe e dei sistemi dinamici, questo tipo di metodi è abbastanza differente da quello impiegato nella presente tesi, tuttavia di questa serie di lavori viene recuperato, in questa sede, lo spirito di fondo che consiste nella sottolineatura degli aspetti "dinamici" del problema. Altro lavoro molto importante in questo filone di ricerche è poi [40], nel quale per la prima volta vengono utilizzate le inclusioni differenziali. Recentemente si segnalano i lavori [24], [27] e [18]; in particolare i primi due sono dedicati anche ad aspetti algoritmici mentre il terzo, di carattere più teorico, è incentrato sullo studio di una teoria dei punti critici per funzioni a valori vettoriali mirata a risultati analoghi al teorema del passo di montagna valido nel caso scalare. Il lavoro presentato in questa tesi si caratterizza, all'interno di questo panorama, per l'utilizzo di idee mutuata dalla teoria dei punti critici moderna (cfr. ad esempio [45]) e si differenzia rispetto ai lavori precedenti per due motivi principali: viene formulata esplicitamente la legge che definisce la dinamica associata al problema di ottimizzazione vettoriale (contrariamente a quanto fatto in [24] e [27]) e simultaneamente viene esplicitata una particolare classe di traiettorie che sono soluzioni di questa legge (mentre in [40] ne veniva garantita solo l'esistenza).

Di seguito riassumiamo brevemente le linee di sviluppo della parte originale di questa tesi, che, come già detto si propone di sviluppare un approccio all'ottimizzazione vettoriale basto sull'introduzione di aspetti dinamici nel problema.

Siano Ω un aperto di \mathbb{R}^n , $f : \Omega \rightarrow \mathbb{R}^m$ una funzione tale che $f \in C^1(\Omega)$. Si introduce allora l'insieme K_f dei punti critici per f nel modo seguente:

$$K_f = \left\{ x \in \Omega : \left[f'(x) \right] (\mathbb{R}^n) \cap -\text{int}\mathbb{R}^m = \emptyset \right\}$$

dove $f'(x) : \mathbb{R}^n \rightarrow \mathbb{R}^m$ è la trasformazione lineare che ha come matrice rappresentativa la matrice Jacobiana di f in x e \mathbb{R}_+^m indica l'insieme dei punti di \mathbb{R}^m che hanno tutte le componenti maggiori o uguali di 0. Si osservi che i punti di K_f giocano il ruolo dei punti stazionari nel caso scalare.

Si dimostra che ogni punto efficiente di f è un elemento di K_f (non vale il viceversa), tuttavia l'insieme K_f è molto più "trattabile" rispetto all'insieme dei punti efficienti di f in quanto come vedremo si può dimostrare che traiettorie di un opportuno "sistema dinamico" convergono ad un punto di K_f . Inoltre è utile sottolineare che questo approccio permette anche lo studio di punti di equilibrio che non sono punti efficienti. Introduciamo poi il seguente strumento (radicale cambiamento di uno strumento introdotto nella teoria dei punti critici moderna nel caso scalare, cfr [45]).

Definiamo prima di tutto la seguente funzione $s_f : \mathbb{R}^n \rightarrow \mathbb{R}_+$

Definizione 1 $s_f(x) := \min \left\{ \left\| \sum_{i=1}^m \lambda_i \nabla f_i(x) \right\| : \lambda_i \in \mathbb{R}, \lambda_i \geq 0, \sum_{i=1}^m \lambda_i = 1 \right\}.$

Definizione 2 Un vettore $v \in \mathbb{R}^n$ si dice pseudogradiante generalizzato per f in x se:

- (1) $\|v\| \leq 2s_f(x)$,
 (2) $[f'(x)]v \in -[s_f(x)]^2 e - \mathbb{R}_+^m$
 dove $e = [1 \ \dots \ 1]^T \in \text{int}\mathbb{R}_+^m$.

(Lo pseudogradiante, nel caso scalare, è uno strumento chiave nei lemmi di deformazione ed individua le direzioni di "discesa" per la funzione f ; si osservi che nel caso scalare $-\nabla f(x)$ è uno pseudogradiante di f in x).

A questo punto ad ogni $x \in \Omega$ è possibile associare un insieme $V_f(x)$ così definito:

$$V_f(x) = \{v \in \mathbb{R}^n : v \text{ è uno pseudogradiante di } f \text{ in } x\}$$

Consideriamo ora il seguente sistema dinamico

$$(*) \quad x'(t) \in V_f(x) \quad \text{con } t \in [0, +\infty)$$

dove al posto di un'equazione differenziale standard vi è un'inclusione differenziale.

Il primo passo consiste nel dimostrare l'esistenza di uno pseudogradiante per ogni punto dello spazio, in seguito si studiano le proprietà della mappa punto insieme V_f in modo tale da poter arrivare ad un teorema d'esistenza per l'inclusione differenziale (*) (utilizzando teoremi sulle inclusioni differenziali, cfr. [3]). Il passo seguente è basato poi sullo studio delle proprietà delle traiettorie utilizzando in particolare la nozione di soluzione per (*) "slow". Si osservi come in questo contesto venga esplicitamente costruita una traiettoria che risulta soluzione dell'inclusione differenziale considerata.

Successivamente si dimostra che alcune traiettorie, soluzione di (*), "convergono" (nel senso che la loro distanza dall'insieme dei punti critici tende a zero) ad un punto appartenente a K .

E' interessante sottolineare come questa tematica (cioè lo studio della dinamica definita da (*)) possa risultare rilevante per tre ordini di motivi (fornendo così anche spunti per un proseguimento delle ricerche):

1) di natura teorica, per la miglior comprensione, o meglio per la comprensione da un punto di vista diverso e nuovo, dei problemi di ottimizzazione vettoriale; inoltre potrebbe essere una base di partenza per una teoria dei punti critici nel caso vettoriale che permetta di studiare e classificare le varie tipologie dei punti critici tenendo conto anche dell'ordine dato nello spazio immagine, fatto questo che non è tenuto in considerazione nei classici studi delle singolarità di mappe differenziabili (cfr. [26]).

2) di natura applicativa, poichè lo studio di problemi di ottimizzazione vettoriale risulta di grande rilevanza in molti campi (teoria economica, teoria delle

decisioni, teoria dei giochi, problemi finanziari); questo approccio serve ad introdurre un elemento dinamico in questi settori che potrebbe migliorare la comprensione e l'analisi dei modelli studiati, ed in particolare permette di individuare traiettorie che portino ai punti di equilibrio;

3) di natura computazionale, la conoscenza delle traiettorie che portano ai punti critici potrebbe portare, almeno in casi particolari (ad esempio in ipotesi di convessità), alla costruzione di algoritmi per individuare punti efficienti basati sulla conoscenza delle traiettorie (in analogia con il caso scalare) che si potrebbero implementare discretizzando le traiettorie o più in generale il sistema dinamico (*).

A conclusione di questa breve introduzione diamo qualche dettaglio sulla struttura della tesi. Un primo capitolo, a carattere introduttivo, è dedicato all'ottimizzazione vettoriale ed alle sue applicazioni; si segnala che nell'ambito di questo capitolo sono sintetizzati alcuni risultati originali conseguiti dall'autore (in parte in collaborazione con E. Molho, cfr. [49] e [48]) riguardanti le scalarizzazioni di un problema di ottimizzazione vettoriale. Il secondo capitolo è dedicato ad alcuni cenni sulle inclusioni differenziali con lo scopo di introdurre gli strumenti tecnici indispensabili nel seguito. Il terzo capitolo fornisce una panoramica, arricchita con alcune osservazioni originali, sulla letteratura nota nel campo di ricerca in cui si inserisce questa tesi. Il quarto ed ultimo capitolo costituisce la parte in cui vengono forniti i risultati completamente originali che danno corpo all'approccio dinamico all'ottimizzazione vettoriale prima riassunto. Chiudono il lavoro due appendici di carattere tecnico che richiamo alcuni risultati molto usati nella tesi.

L'autore ringrazia il prof. Angelo Guerraggio per l'assistenza prestata, oltre che durante lo svolgimento della presente tesi, anche in tutti gli anni del Dottorato; si ringrazia inoltre il prof. Marco Degiovanni dell'Università Cattolica del Sacro Cuore di Brescia per la grande disponibilità dimostrata in numerosi ed utilissimi colloqui sull'argomento di questa tesi.

L'autore infine si scusa infine per alcune imprecisioni ortografiche verificatesi nello spezzare le parole andando a capo; purtroppo il programma di scrittura (o forse meglio l'imperizia dello scrivente con questo programma) non ha permesso altrimenti.

Capitolo 1

Ottimizzazione vettoriale

In questo capitolo ci proponiamo di fornire una breve e sintetica introduzione alla teoria dell'ottimizzazione vettoriale ed alle sue possibili applicazioni. Per ottimizzazione vettoriale si intende la minimizzazione o la massimizzazione di una funzione a valori vettoriali (o equivalentemente di più funzioni simultaneamente); naturalmente ciò richiede di definire cosa si intende per minimo o massimo di una funzione a valori vettoriali, come vedremo nella sezione 1.2 di questo capitolo. Ancor più in generale si può dire che per ottimizzazione vettoriale si intende ottimizzazione rispetto ad un ordine parziale dato nello spazio immagine della funzione obiettivo del problema in esame.

La nascita di studi riguardanti problematiche ora note come problemi di ottimizzazione vettoriale ha sostanzialmente due fonti: da una parte di natura economica legate agli studi di Edgeworth ([21]) e Pareto ([53]) sull'equilibrio, dall'altra di natura matematica connessa agli studi sugli spazi ordinati di Cantor e Hausdorff. Questi primi studi si collocano a cavallo tra la fine dell'ottocento ed i primi anni del novecento. In seguito si registrano in questo settore contributi sia di matematici (legati alla teoria dei giochi) come Borel ([10]) e von Neumann ([68]) sia di economisti come Koopmans ([34]), questi lavori tuttavia non hanno carattere organico rispetto all'ottimizzazione vettoriale e sono incentrati spesso su altre tematiche. Solo con gli articoli di Kuhn e Tucker ([35]) sulle condizioni necessarie e sufficienti per l'efficienza e con il lavoro di Debreu ([17]) sui rapporti tra equilibri e punti Pareto ottimi, negli anni cinquanta, questa area di ricerca guadagna una sua autonomia ed un suo status. A partire poi dagli anni settanta si assiste ad un notevole sviluppo di attività in questo settore sia sul versante delle applicazioni sia dal punto di vista teorico, come testimonia l'imponente massa di letteratura apparsa in questo campo; per una panoramica seppure parziale dei contributi più recenti si vedano le monografie [37], [46] e [22]. Questi brevissimi cenni storici mettono bene in luce come l'ottimizzazione vettoriale sia un settore di studio fortemente interdisciplinare, connotato dallo stretto e costante rapporto tra teoria ed applicazioni.

Articoleremo la nostra esposizione in quattro sezioni: la prima sarà dedicata

alle relazioni d'ordine, strumento indispensabile per poter parlare di ottimizzazione vettoriale, la seconda sarà incentrata sulle varie nozioni di soluzioni di un problema di ottimizzazione vettoriale, nella terza si tratterà la riduzione di un problema di ottimizzazione vettoriale ad uno di programmazione scalare ed infine l'ultima sezione di questo capitolo sarà dedicata all'indicazione di alcune tematiche in cui l'ottimizzazione vettoriale trova, da un lato, parte delle motivazioni che spingono allo studio della sua teoria e, dall'altro, utili applicazioni. Per ulteriori approfondimenti riguardanti i diversi aspetti concernenti l'ottimizzazione vettoriale si rimanda a [31], [37], [46] e [22], si segnala inoltre il volume [73] che presenta tematiche inerenti alle "decisioni multicriteria" (ovvero processi decisionali in cui c'è la necessità di considerare contemporaneamente più criteri) e ne evidenzia le connessioni con l'ottimizzazione vettoriale.

1.1 Spazi parzialmente ordinati

1.1.1 Definizioni

Nell'ambito dell'ottimizzazione vettoriale, ed in particolare per la definizione di possibili nozioni di punti ottimali per una funzione $f : X \rightarrow Y$ con Y insieme qualunque (in generale diverso da \mathbb{R} , che possiede già un suo ordine naturale), risulta indispensabile introdurre una relazione d'ordine in Y ; infatti in assenza di una struttura di questo tipo su Y , non si potrebbe neppure parlare del problema di determinare punti che siano, in un qualche senso che preciseremo meglio più avanti, ottimali per la funzione f . In questa sezione introdurremo quindi alcune nozioni riguardanti la possibilità di dotare un qualsiasi insieme di una relazione d'ordine, perlomeno parziale. Per un'approfondita trattazione degli spazi parzialmente ordinati si rimanda ad esempio alla monografia [33].

Per far questo dobbiamo richiamare brevemente il concetto di relazione binaria.

Dato un insieme arbitrario Y una *relazione binaria* \mathcal{R} in Y è determinata da un qualunque sottoinsieme R non vuoto del prodotto cartesiano $Y \times Y$; inoltre se la coppia ordinata (y_1, y_2) appartiene ad R si scrive che $y_1 \mathcal{R} y_2$ e si dice che y_1 è in relazione con y_2 secondo la relazione \mathcal{R} . Le relazioni binarie possono godere di molte proprietà; di seguito elenchiamo le principali. Una relazione \mathcal{R} in Y si dice:

- *riflessiva* se $y \mathcal{R} y$, per ogni $y \in Y$;
- *simmetrica* se $y_1 \mathcal{R} y_2$ implica che $y_2 \mathcal{R} y_1$, per ogni $y_1, y_2 \in Y$;
- *antisimmetrica* se $y_1 \mathcal{R} y_2$ implica che non si abbia $y_2 \mathcal{R} y_1$, per ogni $y_1, y_2 \in Y$, $y_1 \neq y_2$;
- *transitiva* se $y_1 \mathcal{R} y_2$ e $y_2 \mathcal{R} y_3$ implicano $y_1 \mathcal{R} y_3$, per ogni $y_1, y_2, y_3 \in Y$;

- *completa* se o $y_1 \mathcal{R} y_2$ oppure $y_2 \mathcal{R} y_1$, per ogni $y_1, y_2 \in Y$, $y_1 \neq y_2$.

Se si suppone poi che l'insieme Y sia dotato della struttura di spazio vettoriale topologico una relazione \mathcal{R} si dice:

- *lineare* se $y_1 \mathcal{R} y_2$ implica che $(\alpha y_1 + y_3) \mathcal{R} (\alpha y_2 + y_3)$, per ogni $y_1, y_2, y_3 \in Y$ e per ogni $\alpha \in \mathbb{R}$, $\alpha > 0$,
- *chiusa* se l'insieme R che la determina la relazione \mathcal{R} è un chiuso in $Y \times Y$.

A questo punto possiamo dare la definizione di ordine.

Definizione 3 Una relazione binaria \succeq è un ordine parziale in Y se \succeq è riflessiva e transitiva.

Se lo spazio Y che si sta considerando è uno spazio vettoriale (sul campo reale) usualmente si richiede che l'ordine \succeq sia compatibile con la struttura lineare dello spazio e quindi si richiede che la relazione \succeq sia lineare.

E' significativo, anche se immediato, osservare che l'esistenza di un ordine parziale \succeq in Y non garantisce affatto che due qualsiasi elementi y_1, y_2 di Y siano sempre fra loro confrontabili; ciò significa che non è detto che, dati $y_1, y_2 \in Y$, valga almeno una delle due seguenti affermazioni

$$y_1 \succeq y_2 \text{ oppure } y_2 \succeq y_1.$$

Questa osservazione giustifica l'aggettivo parziale che accompagna il sostantivo ordine; una relazione che risulta essere simultaneamente un ordine parziale e completa prende il nome di *ordine completo*.

1.1.2 Ordini e coni

Assumiamo ora di considerare l'insieme Y dotato della struttura di spazio vettoriale; in questo caso esiste allora una caratterizzazione degli ordini parziali in Y , mediante i coni convessi, molto importante e naturale. Prima di esplicitare questa relazione tra coni ed ordini ci sembra utile richiamare la definizione di cono ed alcune loro proprietà elementari.

Definizione 4 Un sottoinsieme K dello spazio Y si dice cono quando

$$k \in K, \quad \lambda \geq 0 \implies \lambda k \in K.$$

Particolare rilevanza nell'ambito della caratterizzazione di un ordine mediante un cono hanno i coni convessi, che altro non sono che sottoinsiemi di Y che, oltre ad essere coni, risultano anche convessi. Essi sono caratterizzati dalla seguente proprietà.

Proposizione 5 *Un cono $K \subset Y$ è convesso se e solo se per ogni $k_1, k_2 \in K$ si ha che $k_1 + k_2 \in K$.*

Risulta inoltre necessario introdurre una particolare classe di coni definita nel modo seguente:

Definizione 6 *Un cono K è puntato quando*

$$K \cap (-K) = \{0_Y\},$$

dove 0_Y indica il vettore nullo di Y .

Le principali relazioni tra ordine e coni in uno spazio vettoriale sono contenute nel seguente risultato.

Teorema 7 *Sia Y uno spazio vettoriale.*

1. *Se \succeq è un ordine parziale in Y , allora l'insieme*

$$K := \{y \in Y : y \succeq 0_Y\}$$

risulta essere un cono convesso. Inoltre, se la relazione \succeq gode della proprietà antisimmetrica, allora K è un cono puntato.

2. *Se $K \subset Y$ è un cono convesso, allora la relazione binaria \succeq_K , definita nel modo seguente:*

$$y_1 \succeq_K y_2 \iff y_1 - y_2 \in K,$$

è un ordine parziale in Y . Inoltre, se K è puntato, la relazione \succeq_K gode della proprietà antisimmetrica.

Alla luce del teorema precedente risulta chiaro il motivo per cui un cono convesso che caratterizza un ordine nello spazio Y viene usualmente chiamato *cono d'ordine* o anche *cono dei positivi*.

Nel seguito considereremo solo ordini parziali definiti mediante coni convessi. A tale proposito introduciamo le seguenti notazioni.

Notazione 8 *Sia $K \subset Y$ un cono convesso, allora per ogni $y_1, y_2 \in Y$*

$$y_1 \geq_K y_2 \iff y_1 - y_2 \in K,$$

$$y_1 >_K y_2 \iff y_1 \geq_K y_2 \text{ ma non } y_2 \geq_K y_1,$$

inoltre se $\text{int}(K) \neq \emptyset$

$$y_1 \gg_K y_2 \iff y_1 >_{K'} y_2 \text{ con } K' = \{0\} \cup \text{int}(K).$$

1.1.3 Esempi

Scopo di questo paragrafo è fornire alcuni esempi di coni d'ordine; ci si propone inoltre di mostrare come, accanto ai coni che inducono ordini che si possono ritenere in un certo senso naturali, esistano anche coni che caratterizzano ordini ben più complessi. Incominceremo con l'espone alcuni esempi che riguardano spazi vettoriali di dimensione finita per passare poi a citare alcuni coni d'ordine in spazi più generali.

1. L'esempio più semplice, ma allo stesso tempo di importanza fondamentale, è dato dall'ordine componente per componente definito nell'ordinario spazio euclideo, noto anche come *ordine paretiano*. In questo caso $Y = \mathbb{R}^m$ e

$$K = \mathbb{R}_+^m = \{y = (y^1, \dots, y^m) \in \mathbb{R}^m : y^i \geq 0 \text{ per } i = 1, \dots, m\}.$$

Si ottiene allora il ben noto ordine parziale in \mathbb{R}^m , rispetto al quale, le notazioni introdotte in precedenza, prendono i seguenti significati particolari:

$$\begin{aligned} y_1 \geq_K y_2 &\iff y_1^i \geq y_2^i \quad \forall i = 1, \dots, m; \\ y_1 >_K y_2 &\iff y_1^i \geq y_2^i \quad \forall i = 1, \dots, m, \quad y_1 \neq y_2; \end{aligned}$$

inoltre poichè $\text{int}(\mathbb{R}_+^m) = \{y = (y^1, \dots, y^m) \in \mathbb{R}^m : y^i > 0 \text{ per } i = 1, \dots, m\} \neq \emptyset$, si ha

$$y_1 \gg_K y_2 \iff y_1^i > y_2^i \quad \forall i = 1, \dots, m.$$

E' immediato verificare che l'ordine associato al cono considerato gode della proprietà antisimmetrica (quindi il cono \mathbb{R}_+^m è puntato), è lineare e chiuso; inoltre si osserva che non è un ordine completo.

2. Nello spazio euclideo m -dimensionale si possono definire anche altri coni d'ordine che, naturalmente, portano relazioni d'ordine diverse da quella considerata nell'esempio precedente. Si considerino ad esempio i seguenti coni:

$$K_1 = \left\{ \begin{array}{l} y = (y^1, \dots, y^m) \in \mathbb{R}^m : y^i \geq 0 \text{ per } i = 1, \dots, n, \\ y^j = 0 \text{ per } j = n + 1, \dots, m \end{array} \right\};$$

$$K_2 = \{y = (y^1, \dots, y^m) \in \mathbb{R}^m : y^i = 0 \text{ per } i = 1, \dots, n, \text{ con } n < m\}.$$

In quest'ultimo caso si vede che $y_1 \geq_K y_2$ se e solo se le prime n componenti dei due vettori in esame coincidono; inoltre si verifica facilmente che tale ordine gode della proprietà simmetrica, è lineare e chiuso. Si osservi poi che mentre il cono K_1 è puntato, il cono K_2 non gode di questa proprietà.

3. In ogni spazio vettoriale Y esistono due coni d'ordine, che possiamo chiamare *banali*; essi sono lo stesso spazio Y e l'insieme $\{0_Y\}$ costituito dal solo vettore nullo di Y .

4. Consideriamo lo spazio vettoriale Ω di tutte le successioni $y = \{y^n\}_{n \in \mathbb{N}}$ di numeri reali. Possiamo considerare come cono d'ordine in Ω il seguente cono

$$K = \{y = \{y^n\} \in \Omega : y^n \geq 0 \forall n \in \mathbb{N}\}$$

che risulta essere puntato.

5. Sia Θ il sottospazio vettoriale dello spazio Ω , considerato nell'esempio precedente, composto dalle successioni con elementi tutti nulli eccetto un numero finito. Consideriamo ora l'insieme K costituito dalla successione nulla e dalle successioni il cui ultimo termine non nullo è positivo. Questo insieme risulta essere un cono convesso e puntato ed inoltre è possibile dimostrare che il sottospazio lineare generato da K è l'intero Θ .

6. Si consideri lo spazio vettoriale normato

$$\ell^p := \left\{ y = \{y^n\} \in \Omega : \|y\|_p := \left(\sum_{n \in \mathbb{N}} |y^n|^p \right)^{\frac{1}{p}} < \infty \right\} \text{ con } 1 \leq p < \infty.$$

Il cono che induce l'ordine cosiddetto naturale in ℓ^p è il seguente

$$K_{\ell^p} = \{y = \{y^n\} \in \ell^p : y^n \geq 0 \forall n \in \mathbb{N}\}.$$

Nello spazio ℓ^p è tuttavia possibile definire altre relazioni d'ordine, si pensi ad esempio a quella indotta dal cono K (che prende il nome di cono lessicografico) che è definito nel modo seguente: K è l'insieme delle successioni di ℓ^p tali che il primo termine non nullo sia positivo.

7. Si consideri lo spazio vettoriale normato

$$\ell^\infty := \left\{ y = \{y^n\} \in \Omega : \|y\|_\infty := \sup_{n \in \mathbb{N}} |y^n| < \infty \right\}.$$

Il cono che definisce l'ordine naturale in ℓ^∞ è dato da:

$$K_{\ell^\infty} = \{y = \{y^n\} \in \ell^\infty : y^n \geq 0 \forall n \in \mathbb{N}\}.$$

8. Sia ora H un qualunque spazio di Hausdorff compatto. L'insieme $C(H)$ delle funzioni a valori reali, continue e definite su H costituisce uno spazio normato rispetto alla norma data da

$$\|f\|_{C(H)} := \sup_{x \in H} \{f(x)\} \text{ per ogni } f \in C(H).$$

Il cono che induce l'ordine naturale in $C(H)$ è il seguente

$$K_{C(H)} = \{f \in C(H) : f(x) \geq 0 \forall x \in H\}.$$

Si osservi che questo cono è chiuso ed ha interno non vuoto nella topologia di $C(H)$ indotta dalla norma.

9. In analogia con quanto fatto nell'esempio 6, si può considerare lo spazio vettoriale normato $L^p(U)$ ($p \in [1, \infty)$) costituito dalle (classi d'equivalenza di) funzioni $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ la cui potenza p -esima è Lebesgue-integrabile, con norma data da:

$$\|f\|_{L^p(U)} := \left(\int_U |f(x)|^p \right)^{\frac{1}{p}} \text{ per ogni } f \in L^p(U).$$

Il cono che induce l'ordine naturale in $L^p(U)$ è dato da

$$K_{L^p(U)} = \{f \in L^p(U) : f(x) \geq 0 \text{ per quasi ogni } x \in U\}.$$

Si osservi che il cono considerato ha interno vuoto (nella topologia indotta dalla norma). Per inciso osserviamo che nello stesso modo si possono rendere spazi vettoriali ordinati anche gli spazi $L^p(U)$ con $p \in (0, 1)$, che pur non risultano essere spazi normati (cfr. [37], pag. 3).

10. Analogamente a quanto detto nell'esempio 7, si consideri ora lo spazio vettoriale normato $L^\infty(U)$ costituito dalle (classi d'equivalenza di) funzioni $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ essenzialmente limitate (ovvero limitate tranne che su insieme di punti aventi misura di Lebesgue nulla), con norma data da:

$$\|f\|_{L^\infty(U)} := \left(\int_U |f(x)|^p \right)^{\frac{1}{p}} \text{ per ogni } f \in L^p(U).$$

L'ordine naturale nello spazio $L^\infty(U)$ risulta essere quello associato al cono

$$K_{L^\infty(U)} = \{f \in L^\infty(U) : f(x) \geq 0 \text{ per quasi ogni } x \in U\}$$

che si prova avere interno non vuoto rispetto alla topologia indotta dalla norma.

1.2 Nozioni di soluzione

In questa sezione esporremo le principali nozioni di soluzione per un problema di ottimizzazione vettoriale. La sezione si articola in due paragrafi: il primo dedicato alla definizione e alle prime proprietà degli elementi minimali di un insieme mentre il secondo, mediante l'uso delle definizioni date nel primo paragrafo, tratterà le possibili nozioni di soluzione per un problema di ottimizzazione vettoriale. Nel presente lavoro ci limiteremo prevalentemente a trattare spazi ordinati mediante un cono convesso e puntato, da un lato, per non appesantire con un eccessivo formalismo la trattazione, dall'altro, perchè in realtà i coni puntati godono di grande importanza; si segnala comunque che la maggior parte del materiale qui esposto si può estendere anche al caso di coni d'ordine non puntati (cfr. [37]).

1.2.1 Elementi minimali di un sottoinsieme di uno spazio vettoriale ordinato

Sia Y uno spazio vettoriale topologico dotato di una relazione d'ordine indicata con \geq_K , generata da un cono K convesso e puntato di Y .

Definizione 9 *Sia S un sottoinsieme non vuoto di Y ; diciamo allora che:*

1. $y_0 \in S$ è un elemento minimale ideale (o efficiente ideale) di S rispetto a K quando $y \geq_K y_0$ per ogni $y \in S$; l'insieme degli elementi minimali ideali di S si denota con $\text{IMin}(S, K)$;
2. $y_0 \in S$ è un elemento minimale (o efficiente) di S rispetto a K quando $y_0 \geq_K y$, per qualche $y \in S$, implica che $y \geq_K y_0$; l'insieme degli elementi minimali di S si denota con $\text{Min}(S, K)$.

In modo del tutto naturale si definiscono poi gli elementi massimali ideali e gli elementi massimali di S (i cui insiemi sono rispettivamente indicati con $\text{IMax}(S, K)$ e $\text{Max}(S, K)$). Si osservi poi che, mentre le precedenti definizioni non presentano alcun problema anche nel caso di coni non puntati, l'osservazione successiva (in particolare il punto 2.) non vale per coni non puntati.

Osservazione 10 *Grazie alla struttura di spazio vettoriale posseduta da Y , le definizioni precedenti si possono riformulare, privilegiando un approccio di natura insiemistico-geometrica, nel modo seguente:*

1. $y_0 \in \text{IMin}(S, K)$ se e solo se $S \subseteq y_0 + K$;
2. $y_0 \in \text{Min}(S, K)$ se e solo se $(\{y_0\} - K) \cap S = \{y_0\}$.

Oltre a queste nozioni ne esistono altre che si configurano come concetti più deboli (elementi debolmente minimali) o come rafforzamenti dei precedenti (elementi propriamente minimali). Incominciamo con l'esaminare il concetto di elemento debolmente minimale; per far questo occorre sottolineare che la definizione seguente ha significato solo nel caso in cui il cono d'ordine K ha interno non vuoto.

Definizione 11 *Siano K un cono convesso tale che $\text{int}(K) \neq \emptyset$ ed S un sottoinsieme non vuoto di Y ; diciamo allora che $y_0 \in S$ è un elemento debolmente minimale di S rispetto a K quando $y_0 \in \text{Min}(S, \{0\} \cup \text{int}K)$; l'insieme degli elementi debolmente minimali di S si denota con $\text{WMin}(S, K)$.*

Analogamente a quanto detto in precedenza, anche in questo caso si può parlare anche di elementi debolmente massimali adattando la definizione precedente in modo del tutto naturale; l'insieme dei punti debolmente massimali si indicherà

allora con $\text{WMax}(S, K)$. Si osservi inoltre che la definizione precedente si può riformulare nel modo seguente (a patto di supporre che il cono K non coincida con l'intero spazio): $y_0 \in \text{WMin}(S, K)$ se e solo se $(\{y_0\} - \text{int}K) \cap S = \emptyset$, o equivalentemente se e solo se non esiste alcun elemento $y \in S$ tale che $y_0 \gg_K y$.

Introduciamo ora il concetto di elementi propriamente minimali. Esistono svariate e differenti nozioni di propria minimalità, tuttavia tutte sono accomunate dall'esigenza di escludere dall'insieme $\text{Min}(S, K)$ degli elementi minimali di S alcuni punti che risultano "cattivi" elementi minimali in quanto non soddisfano requisiti che si possono classificare come "robustezza" della soluzione. Per avere propria minimalità si richiede infatti che un elemento minimale soddisfi proprietà aggiuntive (a carattere locale o globale) come ad esempio l'essere minimale rispetto ad un cono più ampio di K oppure l'essere elemento minimale anche di opportune approssimazioni coniche dell'insieme S . per una rassegna completa delle varie nozioni di propria minimalità e per i rapporti che intercorrono tra loro si rimanda a [28], mentre in questa sede ci limiteremo ad enunciare due particolari definizioni di elementi propriamente minimali, a nostro modo di vedere particolarmente significative.

Definizione 12 *Siano S un sottoinsieme non vuoto di Y e K un cono convesso puntato di Y ; diciamo allora che $y_0 \in S$ è un elemento propriamente minimale (secondo Henig) di S rispetto a K quando esiste un cono K' convesso, diverso dall'intero spazio Y e soddisfacente la relazione $(K \setminus \{0\}) \subset \text{int} K'$, tale che $y_0 \in \text{Min}(S, K')$. L'insieme degli elementi propriamente minimali (secondo Henig) di S si indica con $\text{PrMin}_H(S, K)$.*

Al fine di poter enunciare la definizione del secondo concetto di propria minimalità, occorre ricordare la seguente notazione che definisce una approssimazione conica di un sottoinsieme S non vuoto dello spazio Y :

$$\text{cone}(S) := \{y \in Y : y = \lambda s, \lambda > 0, s \in S\},$$

questo insieme prende il nome di *cono generato* da S .

Definizione 13 *Siano S un sottoinsieme non vuoto di Y e K un cono convesso puntato di Y ; diciamo allora che $y_0 \in S$ è un elemento propriamente minimale (secondo Borwein) di S rispetto a K quando $y_0 \in \text{Min}(\text{cl}(\text{cone}(S - y_0)), K)$. L'insieme degli elementi propriamente minimali (secondo Borwein) di S si indica con $\text{PrMin}_B(S, K)$.*

Come è facile immaginare, anche per queste definizioni, risulta agevole la riformulazione nel caso di elementi propriamente massimali, ottenendo così gli insiemi $\text{PrMax}_H(S, K)$ e $\text{PrMax}_B(S, K)$.

Una prima osservazione riguardo questi due concetti di propria minimalità motiva l'affermazione fatta in precedenza che attribuiva particolare rilievo a queste

due definizioni; la prima definizione sottolinea infatti una sorta di stabilità degli elementi propriamente minimali secondo Henig rispetto a perturbazioni dell'ordine, mentre la seconda evidenzia come un elemento propriamente minimale secondo Borwein sia stabile rispetto ad opportune perturbazioni della regione ammissibile S . Queste due definizioni mettono cioè in chiara luce come la nozione di propria minimalità sia ricollegabile a richieste di "stabilità" dell'elemento minimale rispetto a perturbazioni che toccano i due dati fondamentali del problema: l'insieme S da una parte ed il cono d'ordine K dall'altra. Risulta inoltre interessante segnalare che questi due differenti nozioni di propria minimalità coincidono se lo spazio Y è uno spazio vettoriale di dimensione finita (cfr. [28]).

Esempio 14 Forniamo ora alcuni semplici esempi per meglio illustrare il significato delle definizioni precedenti.

1. Sia $Y = \mathbb{R}^2$ e si considerino gli insiemi

$$\begin{aligned} S_1 &= \left\{ (y^1, y^2) \in \mathbb{R}^2 : (y^1)^2 + (y^2)^2 \leq 1, y^2 \leq 0 \right\} \cup \\ &\quad \left\{ (y^1, y^2) \in \mathbb{R}^2 : y^1 \geq 0, -1 \leq y^2 \leq 0 \right\}, \\ S_2 &= S_1 \cup \{(-2, -2)\}. \end{aligned}$$

Supponiamo di ordinare lo spazio \mathbb{R}^2 rispetto al cono \mathbb{R}_+^2 allora

$$\begin{aligned} \text{IMin}(S_2, \mathbb{R}_+^2) &= \text{PrMin}_H(S_2, \mathbb{R}_+^2) = \text{Min}(S_2, \mathbb{R}_+^2) = \text{WMin}(S_2, \mathbb{R}_+^2) = \\ &= \{(-2, -2)\}, \\ \text{IMin}(S_1, \mathbb{R}_+^2) &= \emptyset, \\ \text{PrMin}_H(S_1, \mathbb{R}_+^2) &= \left\{ (y^1, y^2) \in \mathbb{R}^2 : (y^1)^2 + (y^2)^2 = 1, y^1 < 0, y^2 < 0 \right\}, \\ \text{Min}(S_1, \mathbb{R}_+^2) &= \text{PrMin}_H(S_1, \mathbb{R}_+^2) \cup (0, -1) \cup (-1, 0), \\ \text{WMin}(S_1, \mathbb{R}_+^2) &= \text{Min}(S_1, \mathbb{R}_+^2) \cup \left\{ (y^1, y^2) \in \mathbb{R}^2 : y^2 = -1, y^1 \geq 0 \right\}. \end{aligned}$$

Supponiamo poi di dotare lo spazio \mathbb{R}^2 dell'ordine indotto dal cono

$$K = \{(y^1, y^2) \in \mathbb{R}^2 : y^2 = 0\};$$

i vari insiemi degli elementi minimali diventano allora

$$\begin{aligned} \text{IMin}(S_2, K) &= \emptyset, \\ \text{PrMin}_H(S_2, K) &= \text{Min}(S_2, K) = \text{WMin}(S_2, K) = S_2, \\ \text{IMin}(S_1, K) &= \emptyset, \\ \text{PrMin}_H(S_1, K) &= \text{Min}(S_1, K) = \text{WMin}(S_1, K) = S_1. \end{aligned}$$

Si osservi che il secondo cono qui considerato non è puntato.

2. Consideriamo ora lo spazio Ω_1 definito nell'esempio 5 della sezione precedente, si dimostra che questo spazio può essere dotato di una norma nel modo seguente

$$\|y\|_{\Omega_1} := \max \{|y^n| : n \in \mathbb{N}\} \quad \forall y = \{y^n\}_{n \in \mathbb{N}}.$$

Consideriamo poi il seguente insieme

$$S = \{y \in \Omega_1 : \|y\|_{\Omega_1} \leq 1\};$$

se l'ordine in è quello indotto dal cono K definito nell'esempio 4, si ha che

$$\text{Min}(S, K) = \{y \in \Omega_1 : \|y\|_{\Omega_1} = 1\} \cap (-K)$$

mentre se il cono d'ordine è dato dal cono K dell'esempio 5, si ha che

$$\text{Min}(S, K) = \emptyset.$$

Il risultato seguente riassume le relazioni che intercorrono tra le varie categorie di elementi ottimali.

Proposizione 15 *Siano, Y uno spazio vettoriale, S un suo sottoinsieme non vuoto e K un cono convesso e puntato con interno non vuoto. Valgono allora le seguenti relazioni*

$$(\text{PrMin}_B(S, K)) \quad \text{PrMin}_H(S, K) \subseteq \text{Min}(S, K) \subseteq \text{WMin}(S, K);$$

inoltre se $\text{IMin}(S, K) \neq \emptyset$, allora

$$\text{IMin}(S, K) = \text{Min}(S, K)$$

ed inoltre $\text{IMin}(S, K)$ si riduce ad essere un singolo.

Osservazione 16 *Nel teorema precedente l'ipotesi riguardante l'interno del cono d'ordine serve solo per poter parlare di elementi debolmente minimali. Inoltre l'ipotesi di K cono puntato è fondamentale solo per provare che $\text{IMin}(S, K)$ si riduce ad un singolo.*

1.2.2 Soluzioni di un problema di ottimizzazione vettoriale

Incominciamo con il definire esattamente cosa si intende per problema di ottimizzazione vettoriale.

Consideriamo:

- uno spazio vettoriale topologico X ed un suo sottoinsieme Q non vuoto che prende il nome di *regione ammissibile*,
- uno spazio vettoriale Y ordinato mediante un cono $K \subset Y$ convesso, chiuso e puntato e con interno non vuoto; con \geq_K indichiamo l'ordine associato al cono K ,
- una funzione $f : Q \rightarrow Y$ che prende il nome di *funzione obiettivo*.

Per problema di ottimizzazione vettoriale (di minimizzazione) si intende la ricerca di punti $x \in Q$ tali che $f(x)$ sia elemento (debolmente, propriamente, ideale) minimale dell'insieme $f(Q)$, immagine di Q mediante f , rispetto al cono d'ordine K . Un punto di questo tipo prende il nome di *punto minimale* di f in Q (rispetto al cono K) o di soluzione minimale del problema di ottimizzazione vettoriale che usualmente viene indicato nel seguente mod

$$\min_{x \in Q} f(x) \quad (VP).$$

L'insieme dei *punti minimali* di f in Q si indica con

$$\text{SMin}(f, Q) := \{x \in X : f(x) \in \text{Min}(f(Q), K)\};$$

analogamente si definiscono l'insieme dei *punti minimali ideali* di f in Q :

$$\text{SIMin}(f, Q) := \{x \in X : f(x) \in \text{IMin}(f(Q), K)\},$$

dei *punti debolmente minimali* di f in Q :

$$\text{SWMin}(f, Q) := \{x \in X : f(x) \in \text{WMin}(f(Q), K)\}$$

e dei *punti propriamente minimali* di f in Q :

$$\text{SPrMin}(f, Q) := \{x \in X : f(x) \in \text{PrMin}(f(Q), K)\}.$$

L'insieme $\text{Min}(f(Q), K)$ (rispettivamente $\text{WMin}(f(Q), K)$, $\text{PrMin}(f(Q), K)$) prende il nome di insieme dei *valori minimali* (rispettivamente *debolmente*, *propriamente*) di f in Q .

Ricordiamo ancora una volta che la richiesta che l'interno del cono d'ordine sia non vuoto è indispensabile solo nel caso in si parli di debole minimalità.

Nel caso in cui non venga esplicitamente indicato un insieme ammissibile Q , nel quale ricercare i punti minimali, si assume che la regione ammissibile del problema in esame coincida con l'intero spazio X ed in questo caso si parla di problema di ottimizzazione vettoriale libera.

Risulta del tutto evidente che cosa si intenda per problema vettoriale di massimizzazione e quali siano le notazioni in questo caso. A titolo d'esempio consideriamo l'insieme dei punti massimali di f in Q che costituisce l'insieme delle soluzioni del problema

$$\max_{x \in Q} f(x)$$

e si indica con $S\text{Max}(f, Q) := \{x \in X : f(x) \in \text{Max}(f(Q), K)\}$.

Segnaliamo inoltre che fra i due tipi di problemi intercorre una naturale relazione che fa sì che sia lecito indagare uno solo dei due, dal momento che l'altro si può in ogni caso ricondurre al primo senza difficoltà. E' infatti immediato verificare che

$$S\text{Max}(f, Q) = S\text{Min}(-f, Q)$$

ed una relazione dello stesso tipo vale per tutte le diverse nozioni di soluzione definite in precedenza.

Consideriamo ora una situazione particolare ma di grande importanza da molti punti di vista ed in particolare per l'ultimo capitolo di questa tesi.

Esempio 17 (Pareto Ottimalità) *Siano $X = \mathbb{R}^n$, Q sottoinsieme non vuoto di \mathbb{R}^n , $Y = \mathbb{R}^m$, $K = \mathbb{R}_+^m = \{(y^1, \dots, y^m) \in \mathbb{R}^m : y^i \geq 0 \forall i = 1, \dots, m\}$ cono d'ordine in Y . Si ricorda che quindi stiamo considerando nello spazio euclideo ad m dimensioni l'ordine paretiano definito nell'esempio 1 della sezione precedente. Consideriamo poi come funzione obiettivo una funzione $f = (f_1, \dots, f_m) : Q \rightarrow \mathbb{R}$ dove ogni componente f_i ($i = 1, \dots, m$) è una funzione definita in Q a valori in \mathbb{R} . In questo contesto è interessante esplicitare, in termini delle componenti della funzione obiettivo, la natura dei punti minimali di f in Q rispetto all'ordine paretiano: essi sono quei punti $\tilde{x} \in Q$ tali per cui non esiste alcun punto $x \in Q$ tale che*

$$\begin{aligned} f_i(x) &\leq f_i(\tilde{x}) \quad \forall i = 1, \dots, m \\ &\text{e contemporaneamente} \\ \exists j &\in \{1, \dots, m\} : f_j(x) < f_j(\tilde{x}). \end{aligned}$$

Questo, da un punto di vista semplicemente intuitivo, significa che non ci si può spostare da un punto minimale di f senza che il valore di almeno una delle componenti della funzione f cresca. In questa situazione particolare un punto minimale prende il nome di punto Pareto minimale di f . In modo del tutto simile si possono definire i punti Pareto massimali di f .

Concludiamo questa sezione segnalando che è possibile costruire problemi di ottimizzazione vettoriale ancora più generali, considerando al posto della funzione obiettivo $f : Q \rightarrow Y$ una mappa punto-insieme $F : Q \rightrightarrows Y$. Per questo tema si veda ad esempio [37], cap.2, sez.5.

1.3 Scalarizzazione

Questa sezione è dedicata alla scalarizzazione di un problema del tipo (VP) definito nella sezione precedente. In generale per scalarizzazione di un problema di ottimizzazione vettoriale si intende il sostituire il problema vettoriale con

un problema di ottimizzazione scalare, le cui soluzioni siano in relazione con le soluzioni del problema vettoriale di partenza. Questa tematica risulta avere un'importanza centrale nell'ambito dell'ottimizzazione vettoriale perchè da un lato costituisce una delle vie per caratterizzare le soluzioni, fornendo così condizioni necessarie e/o sufficienti affinché un dato punto sia soluzione di un problema (VP), dall'altro permette di costruire esplicitamente algoritmi per il calcolo delle soluzioni. In letteratura esistono molti e differenti metodi di scalarizzazione di un problema vettoriale; in questa sede noi esamineremo solo due classi di particolare rilevanza: quella delle scalarizzazioni lineari ed una classe molto ampia, definita in modo assiomatico, che comprende anche le scalarizzazioni basate sulla minimizzazione di una particolare "distanza" da un fissato obiettivo di riferimento e su considerazioni di natura geometrica. Chiaramente il problema della scalarizzazione può essere semplificato considerando, al posto del problema (VP), il solo problema di individuare gli elementi ottimali di un sottoinsieme dello spazio Y ; dai risultati riguardanti questo caso è poi immediato ricondursi alla considerazione del problema (VP), è infatti sufficiente comporre la funzione scalarizzante (a valori reali) con la funzione obiettivo del problema (VP). Sottolineiamo infine che in tutta questa sezione parleremo di soluzioni minimali del problema (VP) e di conseguente problema di minimizzazione scalare, tuttavia, a patto di ovvie modifiche, tutto quello che diremo varrà anche nel caso di soluzioni massimali.

1.3.1 La scalarizzazione lineare

Nel presente paragrafo si intende esporre per sommi capi il metodo di scalarizzazione lineare che consiste nel associare al problema di ottimizzazione vettoriale (VP) un problema di natura scalare la cui funzione obiettivo è data dalla composizione tra la funzione $f : Q \rightarrow Y$ obiettivo del problema (VP) ed un opportuno funzionale lineare ℓ definito sullo spazio Y .

Valgano ora le ipotesi della seconda parte della sezione precedente; risulta inoltre indispensabile introdurre le seguenti notazioni:

- indichiamo con Y^* lo spazio duale (topologico) dello spazio Y ;
- $K^* := \{\ell \in Y^* : \ell(k) \geq 0 \forall k \in K\}$ prende il nome di *cono polare* del cono d'ordine K ;
- $K^\# := \{\ell \in Y^* : \ell(k) > 0 \forall k \in K \setminus \{0\}\}$ prende il nome di *cono polare stretto* del cono d'ordine K .

E' immediato constatare che $K^\# \subset K^*$; si ricordi inoltre che l'insieme $K^\#$ si dimostra essere non vuoto quando su Y si fa l'ulteriori ipotesi che sia uno spazio normato separabile.

La risoluzione del problema vettoriale (VP) si riconduce in questo caso alla minimizzazione della funzione scalare $\ell \circ f : X \rightarrow \mathbb{R}$ dove ℓ è un opportuno elemento di K^* (od eventualmente di $K^\#$). I principali risultati riguardanti questa

classe di scalarizzazioni sono contenute nei teoremi citati di seguito, per le dimostrazioni dei quali si rimanda ad esempio a [32]; incominciamo con l'esaminare il caso delle soluzioni debolmente minimali.

Teorema 18 *Siano soddisfatte le ipotesi sopra elencate riguardo Y e K .*

1. *Se esistono $x_0 \in Q$ ed $\ell \in K^* \setminus \{0\}$ tali che*

$$\ell(f(x_0)) \leq \ell(f(x)) \quad \forall x \in Q$$

allora $x_0 \in \text{SWMin}(f, Q)$.

2. *Si supponga inoltre che l'insieme $f(Q) + K$ sia convesso. Se $x_0 \in \text{SWMin}(f, Q)$ allora esiste $\ell \in K^* \setminus \{0\}$ tale che*

$$\ell(f(x_0)) \leq \ell(f(x)) \quad \forall x \in Q.$$

La completa caratterizzazione delle soluzioni debolmente minimali fornita dal teorema precedente costituisce proprio uno dei principali motivi di interesse della nozione di soluzioni debolmente minimale di un problema di ottimizzazione vettoriale. Per le soluzioni minimali non è possibile dimostrare un teorema di caratterizzazione così completo, si ottiene infatti, come miglior risultato, il seguente teorema.

Teorema 19 *Siano soddisfatte le ipotesi sopra elencate riguardo Y e K .*

1. *Se esistono $x_0 \in Q$ ed $\ell \in K^\#$ tali che*

$$\ell(f(x_0)) \leq \ell(f(x)) \quad \forall x \in Q$$

allora $x_0 \in \text{SMin}(f, Q)$.

2. *Si supponga inoltre che l'insieme $f(Q) + K$ sia convesso. Se $x_0 \in \text{SMin}(f, Q)$ allora esiste $\ell \in K^* \setminus \{0\}$ tale che*

$$\ell(f(x_0)) \leq \ell(f(x)) \quad \forall x \in Q.$$

Passando ora al caso delle soluzioni propriamente minimali si possono completamente caratterizzare queste soluzioni mediante la scalarizzazione lineare con elementi del cono polare stretto. In questo frangente ci restringiamo al caso che lo spazio Y sia di dimensione finita, ipotesi questa non indispensabile ma che semplifica il discorso in quanto le nozioni di propria minimalità, con questa ipotesi aggiuntiva, vengono a coincidere.

Teorema 20 *Siano soddisfatte le ipotesi sopra elencate riguardo Y e K ed inoltre Y sia di dimensione finita.*

1. Se esistono $x_0 \in Q$ ed $\ell \in K^\#$ tali che

$$\ell(f(x_0)) \leq \ell(f(x)) \quad \forall x \in Q$$

allora $x_0 \in \text{SPrMin}(f, Q)$.

2. Si supponga inoltre che l'insieme $f(Q) + K$ sia convesso. Se $x_0 \in \text{SPrSMin}(f, Q)$ allora esiste $\ell \in K^* \setminus \{0\}$ tale che

$$\ell(f(x_0)) \leq \ell(f(x)) \quad \forall x \in Q.$$

Esempio 21 Consideriamo ora il caso paretiano (cfr. gli esempi 1, 17), allora $K^* = \mathbb{R}_+^m$ e $K^\# = \text{int}(\mathbb{R}_+^m)$. ne segue allora che in questo caso il problema scalare consiste nel minimizzare la funzione scalare

$$\sum_{i=1}^m \lambda^i f_i(x) \quad \text{con } \lambda^i \geq 0 \ (\lambda^i > 0) \text{ per ogni } i = 1, \dots, m.$$

A conclusione a questo paragrafo vogliamo osservare che le condizioni necessarie, formulate mediante la scalarizzazione lineare, affinché un punto sia soluzione del problema di ottimizzazione vettoriale valgono solo in presenza di ipotesi di convessità. A tale proposito si sottolinea che la condizione $f(Q) + K$ convesso è verificata da tutte le funzioni K -convesse (cfr. [37], pag. 29), cioè da quelle funzioni $f : Q \rightarrow Y$ che soddisfano (una volta che Q sia supposto convesso)

$$f(\alpha x_1 + (1 - \alpha)x_2) - \alpha f(x_1) - (1 - \alpha)f(x_2) \in -K \quad \forall x_1, x_2 \in Q, \forall \alpha \in [0, 1];$$

tuttavia esistono anche funzioni f non K -convesse tali che $f(Q) + K$ sia convesso; basti pensare alla funzione $f : [0, \infty[\rightarrow \mathbb{R}^2$ definita da

$$f(x) := \begin{bmatrix} x \\ \sin x \end{bmatrix}$$

che non risulta \mathbb{R}_+^2 -convessa, mentre $f([0, \infty[) + \mathbb{R}_+^2$ è un sottoinsieme convesso di \mathbb{R}^2 .

1.3.2 Scalarizzazioni basate su “distanze”

In questo paragrafo si intende studiare una famiglia di scalarizzazioni definita in modo assiomatico. In questa classe si vedrà poi come rientrino molte scalarizzazioni note in letteratura e basate sulla minimizzazione di un’opportuna “distanza” da un prefissato obiettivo di riferimento (si vedano ad esempio [51], [31], [22] sec. 4.4). Questa assiomatizzazione e lo studio delle sue proprietà sono sviluppate ad opera dell’autore della presente tesi, in collaborazione con E. Molho, nel lavoro [48] a cui si rimanda per una esposizione completa e dettagliata.

Consideriamo una funzione $f : X \rightarrow Y$ dove Y è uno spazio vettoriale normato dotato di una struttura d'ordine indotta da un cono K chiuso, convesso, puntato e con interno non vuoto e studiamo il problema di ottimizzazione vettoriale (VP) associato alla funzione f ed al sottoinsieme Q di X . A tal fine consideriamo una funzione scalare $s : Y \rightarrow \mathbb{R}$, dipendente da un *arbitrario* parametro $\bar{y} \in Y$ e così definita:

$$s : y \mapsto s(y, \bar{y})$$

che soddisfa le seguenti proprietà:

$$\mathbf{P1} \quad \forall y_1, y_2 \in A : y_2 - y_1 \in K \implies s(y_1, \bar{y}) \leq s(y_2, \bar{y});$$

$$\mathbf{P2} \quad \forall y_1, y_2 \in A : y_2 - y_1 \in \text{int}K \implies s(y_1, \bar{y}) < s(y_2, \bar{y});$$

$$\mathbf{R1} \quad S_0 := \{y \in Y : s(y, \bar{y}) \geq 0\} \subseteq \bar{y} + K; \quad s(\bar{y}, \bar{y}) = 0;$$

$$\mathbf{R2} \quad \hat{S}_0 := \{y \in Y : s(y, \bar{y}) > 0\} \subseteq \bar{y} + \text{int}K.$$

Si considera allora come scalarizzazione del problema di ottimizzazione vettoriale, il seguente problema di ottimizzazione scalare

$$\min_{x \in Q} s(f(x), \bar{y})$$

le cui soluzioni, come vedremo, sono strettamente correlate con le soluzioni del problema (PV).

Una funzione che soddisfa le proprietà P1 e P2 si dice che *preserva strettamente l'ordine* mentre quando soddisfa R1 ed R2 si dice che *rappresenta strettamente l'ordine*. Le prime due proprietà costituiscono richieste naturali di monotonia che vengono usualmente adottate come proprietà di una funzione scalarizzante (cfr. ad es. [37]) in modo tale che attraverso la scalarizzazione si possano trovare condizioni sufficienti affinché un punto sia soluzione ottimale del problema (VP). Le altre due proprietà invece sono indispensabili per ottenere, attraverso la scalarizzazione, condizioni necessarie per l'ottimalità in assenza di ipotesi di convessità. In questa sezione enunceremo una serie di risultati che mostrano come, attraverso una funzione che preservi strettamente e simultaneamente rappresenti strettamente l'ordine, sia possibile fornire una completa caratterizzazione, mediante scalarizzazione, delle soluzioni di un problema di ottimizzazione vettoriale in assenza di qualsiasi ipotesi di convessità, contrariamente a quanto visto nel caso di scalarizzazioni lineari. Si sottolinea inoltre come questo tipo di scalarizzazione privilegi gli aspetti di struttura e geometrici del problema. Incominciamo con l'espore il risultato che riguarda le soluzioni debolmente minimali.

Proposizione 22 *Siano Y, K, Q, f ed s come sopra. Valgono le seguenti affermazioni:*

1. se esiste $\bar{y} \in Y$ tale che x_0 sia punto di minimo (globale) per la funzione $s(f(x), \bar{y})$ in Q allora $x_0 \in \text{WSMin}(f, Q)$;
2. se $x_0 \in \text{WSMin}(f, Q)$ allora esiste $\bar{y} \in Y$ tale che x_0 sia punto di minimo (globale) per la funzione $s(f(x), \bar{y})$ in Q .

Analogo risultato vale per le soluzioni minimali del problema.

Proposizione 23 *Siano Y, K, Q, f ed s come sopra. Valgono le seguenti affermazioni:*

1. se esiste $\bar{y} \in Y$ tale che x_0 sia punto di minimo stretto (globale) per la funzione $s(f(x), y)$ in Q allora $x_0 \in \text{SMin}(f, Q)$;
2. se $x_0 \in \text{SMin}(f, Q)$ allora esiste $\bar{y} \in Y$ tale che x_0 sia punto di minimo stretto (globale) per la funzione $s(f(x), \bar{y})$ in Q .

La caratterizzazione della propria minimalità richiede in questo caso l'introduzione di alcune nozioni alquanto tecniche; in particolare si legano le soluzioni propriamente minimali (nel senso di Borwein) del problema (VP) a proprietà di stabilità del problema scalarizzato espresse in termini di opportune convergenze d'insiemi (per questa nozione si veda ad es. [38]). Poichè questi argomenti risultano vicini ma tangenziali rispetto al filone principale di questa tesi riteniamo opportuno non addentrarci nella loro esposizione dettagliata, per la quale preferiamo rimandare al citato lavoro [48]. Intendiamo però sottolineare due importanti aspetti di questa famiglia di scalarizzazioni. Il primo consiste nel fatto che la funzione scalarizzante, con cui si compone la funzione obiettivo del problema (VP) , è sempre la stessa per tutti le differenti classi di soluzioni, (cosa questa che non avveniva nel caso della scalarizzazione lineare) solo vengono considerate differenti nozioni di soluzioni di un medesimo problema scalare (minimo, minimo stretto, soluzione "stabile"). Viene inoltre messa in luce un interessante trama di corrispondenze che, a nozioni via via più stringenti del problema vettoriale, associa concetti sempre più restrittivi di soluzioni del problema scalarizzato, come ben evidenziato dal seguente schema:

Problema vettoriale	\iff	Scalarizzazione
<i>debole minimalità</i>	\iff	<i>punto di minimo</i>
<i>minimalità</i>	\iff	<i>punto di minimo stretto</i>
<i>propria minimalità</i>	\iff	<i>stabilità</i>

Il secondo punto su cui ci soffermiamo mette in evidenza come la classe di funzioni scalarizzanti considerata in questo paragrafo comprenda interessanti scalarizzazioni proposte in letteratura. A titolo d'esempio ne citeremo due tra le più significative rimandando a [71] e a [22] per altri casi.

La prima è l'importante scalarizzazione introdotta da J. Jahn (si veda ad es. [32]) e basata sull'utilizzo di opportune norme. Per illustrare questo approccio al problema della scalarizzazione è indispensabile ricordare la nozione di *intervallo d'ordine* in uno spazio parzialmente ordinato mediante un cono; siano $y_1, y_2 \in Y$ allora

$$[y_1, y_2] = (y_1 + K) \cap (y_2 - K).$$

Ora, per ogni $a \in \text{int}K$ con $\|\cdot\|_a$ denotiamo il funzionale di Minkowski associato all'intervallo d'ordine, cioè

$$\|y\|_a := \inf \left\{ \lambda > 0 : \frac{1}{\lambda}y \in [-a, a] \right\} \quad \text{per ogni } y \in Y$$

Si prova facilmente che al variare di a in $\text{int}K$ $\|\cdot\|_a$ costituisce una famiglia parametrica di norme. Imponiamo poi la validità della seguente ipotesi aggiuntiva: esista un elemento $\hat{y} \in Y$ tale che

$$f(Q) \subseteq \hat{y} - \text{int}K.$$

Definendo allora la funzione s_J in questo modo

$$s_J(y, \bar{y}) = 1 - \|y - \hat{y}\|_{\bar{y} - \hat{y}}.$$

si prova che la scalarizzazione di Jahn rientra nella classe di scalarizzazioni considerata in questo paragrafo (per la dimostrazione di questo fatto si veda [48]).

La seconda scalarizzazione che intendiamo ricordare è stata introdotta per la prima volta in [15] (in un contesto completamente differente dal presente), e la funzione scalarizzante è data da

$$s(f(x), \bar{y}) := \Delta_{-K}(f(x) - \bar{y})$$

dove la funzione $\Delta_{-K} : Y \rightarrow \mathbb{R}$ è definita nel modo seguente (si ricordi che dato $H \subset Y$ si ha che $d(y) := \inf_{h \in H} \{\|h - y\|\}$ per ogni $y \in Y$)

$$\Delta_{-K}(y) := d_{-K}(y) - d_{Y \setminus -K}(y).$$

Questa particolare funzione è stata studiata approfonditamente, in relazione alla scalarizzazione di problemi di ottimizzazione vettoriale, dall'autore della presente tesi nel lavoro [47] nel quale, oltre a mostrare la validità in questo caso specifico dello schema sopra esposto, le soluzioni propriamente minimali vengono caratterizzate mediante una differente richiesta di "stabilità" del problema scalarizzato basata sulla nozione di buona posizione secondo Tyhkonov (per la definizione di questo concetto si veda ad es. [38]).

1.4 Applicazioni

Negli ultimi due decenni le tecniche di ottimizzazione multiobiettivo, oltre ad essere state approfonditamente studiate dal punto di vista teorico, hanno trovato numerosissime applicazioni sia in altri settori della matematica sia negli ambiti scientifico - tecnologici più svariati. In questa sezione ci proponiamo quindi di esporre sinteticamente alcuni esempi di applicazioni dell'ottimizzazione multiobiettivo, con lo scopo di illustrare l'utilità dei risultati e delle tecniche di ottimizzazione vettoriale in campi diversi come la teoria economica, la finanza, la teoria dei giochi. Naturalmente non abbiamo nessuna pretesa di esaurire con questi brevi cenni un campo tanto vasto come quello considerato; per ulteriori approfondimenti si segnalano gli ormai classici volumi [51], [73], [31] e i più recenti [46] e [22] e si rimanda inoltre alle ampie bibliografie contenute in queste opere per una panoramica della letteratura riguardante le applicazioni dell'ottimizzazione multiobiettivo.

1.4.1 I teoremi del benessere

In questa sezione delineamo per sommi capi l'importanza della nozione di Pareto-ottimalità (cfr. esempio 17) nell'ambito della Teoria economica con particolare riguardo ai cosiddetti due Teoremi fondamentali del benessere. Questo esempio mostra come problemi di ottimizzazione vettoriale nascano naturalmente nello studio di argomenti tipici delle scienze economiche. Per questi argomenti facciamo riferimento al testo [9], si può inoltre consultare per eventuali approfondimenti il classico volume [67].

Consideriamo un'economia con m consumatori ed n beni; con x_j^k denotiamo la quantità dell' j -esimo bene consumata dal k -esimo consumatore, supponiamo poi che il per il k -esimo consumatore sia definita una funzione d'utilità $u^k : \mathbb{R}_+^n \rightarrow \mathbb{R}$ di classe C^1 e che esso possieda una *dotazione iniziale* $h^k \in \mathbb{R}_+^n$; infine supponiamo di ambientare il discorso in un'economia chiusa nella quale l'ammontare totale di ogni bene rimane invariato durante le interazioni fra tutti i consumatori presenti. A questo punto si ha quindi che, indicato con $b = \sum_{k=1}^m h^k$, lo spazio degli stati del problema è dato da:

$$\Omega := \left\{ \mathbf{z} = (z^1, \dots, z^m) \in (\mathbb{R}_+^n)^m : \sum_{k=1}^m z^k = b \right\}.$$

Un elemento di Ω è detto *allocazione dei beni*; ogni funzione d'utilità u^k si può rileggere come una funzione \mathbf{u}^k definita in Ω ponendo

$$\mathbf{u}^k(z^1, \dots, z^k) := u^k(z^k) \quad \text{per ogni } k = 1, \dots, m,$$

in modo tale che l'insieme Ω diventi il dominio comune a tutte le m funzioni d'utilità considerate. Infine raggruppiamo le funzioni d'utilità in un'unica *mappa*

d'utilità

$$\mathbf{u} := (\mathbf{u}^1, \dots, \mathbf{u}^m) : \Omega \longrightarrow \mathbb{R}^m$$

e ordiniamo lo spazio \mathbb{R}^m rispetto all'ordine associato al cono \mathbb{R}_+^m , ovvero dotiamolo dell'ordine paretiano (vd. esempio 1). Sotto queste ipotesi un'economia è completamente specificata da una terna $(\mathbf{u}, (h^1, \dots, h^m), p) \in C^1(\Omega, \mathbb{R}^m) \times \Omega \times \mathbb{R}_+^n$ dove $p = (p_1, \dots, p_n)$ indica un sistema di prezzi (p_j è il prezzo unitario del j -esimo bene).

Un'allocazione Pareto ottima è un punto ideale che il mercato dovrebbe raggiungere per allocare al meglio le risorse disponibili, tuttavia i punti che il mercato effettivamente raggiunge sono i cosiddetti *equilibri competitivi*, che verranno definiti nel seguito. La questione fondamentale che a cui dobbiamo rispondere è la seguente: "Come il mercato può allocare le risorse disponibili nel miglior modo possibile?". Risulta chiaro, in base a quanto detto poco sopra, che questa domanda risulta essere equivalente a "Quali sono le relazioni tra le allocazioni Pareto-ottime e gli equilibri competitivi?". Alla risposta a questa domanda dobbiamo premettere alcune osservazioni e definizioni.

Come prima cosa si osservi che le allocazioni Pareto ottime non sono nient'altro che le soluzioni del seguente problema di ottimizzazione vettoriale

$$\max_{\mathbf{z} \in \Omega} \mathbf{u}(\mathbf{z}),$$

le quali, con le notazioni introdotte in precedenza, risultano essere gli elementi dell'insieme $\text{SMax}(\mathbf{u}, \Omega)$.

Definiamo poi il concetto di equilibrio competitivo. Per far questo dobbiamo considerare una dotazione iniziale globale $\mathbf{h} = (h^1, \dots, h^m) \in \Omega$. Se nel mercato vige il sistema di prezzi p , il consumatore che ha la dotazione iniziale h^k , possiede la ricchezza $w^k = \sum_{j=1}^n p_j h_j^k$. In questo caso l'agente considerato potrà possedere in futuro solo dotazioni di beni x^k appartenenti al seguente insieme

$$B(p, h^k) := \left\{ x^k \in \mathbb{R}_+^n : \sum_{j=1}^n p_j x_j^k \leq w^k \right\}$$

individuato da *vincoli di bilancio*. Naturalmente il k -esimo consumatore mira a massimizzare la sua funzione d'utilità u^k , restando però nel suo insieme di bilancio $B(p, h^k)$; indichiamo con $s^k(p, h^k)$ la soluzione di questo problema. Tale soluzione non è detto che sia unica, abbiamo quindi una mappa punto insieme $s^k : \mathbb{R}_+^n \times \mathbb{R}_+^n \rightrightarrows B(p, h^k)$ che associa ad ogni coppia di condizioni iniziali (p, h^k) il paniere $s^k(p, h^k)$; noi per semplicità supporremo che tale mappa punto-insieme sia una vera e propria funzione che prende il nome di *funzione di domanda*. La maggior parte di quanto si vedrà in seguito continua comunque a valere anche nel caso di una mappa punto-insieme.

La *domanda aggregata* degli m consumatori considerati è data dal vettore $\sum_{k=1}^m s^k(p, h^k)$, ed inoltre sia

$$\mathbf{z}(p, \mathbf{h}) := (s^1(p, h^1), \dots, s^m(p, h^m)) \in (\mathbb{R}_+^n)^m$$

l'allocazione dei beni domandati.

Definizione 24 Se $\mathbf{z}(p, \mathbf{h})$ si dice che:

1. p è un prezzo d'equilibrio per h ,
2. $\mathbf{z}(p, \mathbf{h})$ è un'allocazione d'equilibrio competitivo, e
3. $(p, \mathbf{z}(p, \mathbf{h}))$ è un equilibrio competitivo (relativo all'allocazione iniziale \mathbf{h})

quando vale la seguente relazione

$$\sum_{k=1}^m s^k(p, h^k) = \sum_{k=1}^m h^k = b.$$

Si osservi che quando il mercato si trova in una situazione di equilibrio competitivo, ogni consumatore massimizza la sua utilità rispettando i propri vincoli di bilancio e contemporaneamente la domanda aggregata è bilanciata dalla quantità totale di beni disponibili b .

A questo punto possiamo rispondere alla domanda formulata in precedenza affermando, che sotto opportune ipotesi, esiste sempre un sistema di prezzi p tale che un'allocazione Pareto-ottima possa essere vista come un equilibrio competitivo e, viceversa, ogni equilibrio competitivo è una allocazione Pareto-ottima. Quanto detto è contenuto nei due teoremi riportati di seguito, che sono noti come i due *teoremi fondamentali del benessere*.

Teorema 25 Supponiamo che ogni funzione d'utilità u^k sia quasiconcava e che soddisfi la seguente ipotesi di monotonia:

$$\frac{\partial u^k}{\partial x_j^k}(h^k) > 0 \quad \forall k = 1, \dots, m; j = 1, \dots, n.$$

Sia $\mathbf{h} = (h^1, \dots, h^m)$ la dotazione iniziale globale. Se (p, \mathbf{z}) è un equilibrio competitivo relativo ad \mathbf{h} ed inoltre $\mathbf{z} \neq 0$, allora \mathbf{z} è un punto Pareto massimale per la funzione \mathbf{u} ed inoltre $\mathbf{z} \geq_{\mathbb{R}_+^2} \mathbf{h}$.

Teorema 26 Supponiamo che valgano le ipotesi del teorema precedente. Se $\mathbf{z} = (z^1, \dots, z^m)$ è un punto Pareto massimale per la funzione \mathbf{u} , tale che $z_j^k > 0$ per ogni $k = 1, \dots, m$ e per ogni $j = 1, \dots, n$, allora esiste un sistema di prezzi $p \in \mathbb{R}_+^n$ ($p \neq 0$) tale che (p, \mathbf{z}) sia un equilibrio competitivo in Ω .

1.4.2 La selezione del portafoglio

Il problema della selezione del portafoglio costituisce una delle questioni centrali della matematica finanziaria moderna e consiste nel determinare la miglior ripartizione (secondo opportuni criteri) di un capitale tra più investimenti a carattere aleatorio (usualmente si considerano titoli di vario genere). Per una prima introduzione a questa tematica si rimanda a [11], [16] ed alle bibliografie ivi contenute; si segnalano inoltre i lavori [8] e [70], dove le tecniche di ottimizzazione multiobiettivo sono centrali; in particolare nel primo si utilizzano tecniche che sostanzialmente rientrano nelle scalarizzazioni, viste in precedenza, basate su opportune "distanze" da punti di riferimento prefissati.

Incominciamo con il supporre che il capitale di cui si dispone ai fini dell'investimento sia unitario; si osservi che questa ipotesi non è una vera restrizione in quanto ciò che si vuole ottenere come soluzione di questo problema è una determinazione delle quote percentuali di capitale da investire nei singoli titoli presi in esame; supponiamo poi di considerare n titoli candidati alla scelta ed indichiamo con x^k ($k = 1, \dots, n$) la quota da investire nel k -esimo titolo, naturalmente dovrà essere soddisfatto il vincolo $\sum_{k=1}^n x^k = 1$ ed inoltre $x^k \geq 0$ per ogni $k = 1, \dots, n$. Quest'ultima condizione potrebbe essere tralasciata se si ammette che il decisore possa anche vendere ($x^k < 0$) i titoli oltre che comprarli. Indichiamo con Z_k la variabile aleatoria "rendimento del k -esimo titolo" e con Z la variabile aleatoria "rendimento del portafoglio" composto dalle quote x^1, \dots, x^n dei vari titoli, che sarà naturalmente dato da

$$Z = \sum_{k=1}^n x^k Z_k.$$

Nell'affrontare il problema della selezione del portafoglio le quantità che vengono considerate sono il valor medio di Z

$$E(Z) = \sum_{k=1}^n x^k E(Z_k)$$

che naturalmente ha il significato di rendimento medio del portafoglio nel suo complesso e la varianza di Z

$$Var(Z) = E(Z - E(Z))^2 = \sum_{k=1}^n \sum_{s=1}^n x^k x^s Cov(Z_k, Z_s)$$

che è intesa come "misura del rischio" del portafoglio in esame.

L'obiettivo che solitamente un investitore si prefigge è quello di costruire un portafoglio con rendimento medio "abbastanza alto" e di rischio "abbastanza basso". L'idea è allora quella di individuare quei portafogli, detti *portafogli efficienti*, per i quali non è possibile simultaneamente migliorare il rendimento e abbassare il rischio. Risulta chiaro a questo punto che ci si trova di fronte ad un

problema di ottimizzazione multiobiettivo. Ci proponiamo ora di formalizzare precisamente questo problema: definiamo le due funzioni scalari (della variabile vettoriale $x = (x^1, \dots, x^n) \in \mathbb{R}^n$)

$$\begin{aligned} f_1(x) &: = E(Z) \\ f_2(x) &: = Var(Z). \end{aligned}$$

Siamo quindi interessati a massimizzare f_1 e simultaneamente a minimizzare f_2 . Ricordando ora che in generale vale la relazione

$$\max_x f_1(x) = - \min_x (-f_1(x)),$$

definendo la funzione $f : Q \rightarrow \mathbb{R}^2$ come $f(x) := (-f_1(x), f_2(x))$, posto $Q = \{x \in \mathbb{R}^n : \sum_{k=1}^n x^k = 1\}$ e considerato in \mathbb{R}^2 l'ordine Paretiano, si constata immediatamente che determinare i portafogli efficienti equivale a trovare le soluzioni minimali del seguente problema di ottimizzazione vettoriale

$$\min_{x \in Q} f(x).$$

Il punto chiave della risoluzione del problema di selezione del portafoglio consiste proprio nella determinazione dei portafogli efficienti. Bisogna tuttavia osservare che in generale l'insieme $SMin(f, Q)$, e quindi l'insieme dei portafogli efficienti, non si riduce ad un singolo. Questo fa sì che, per risolvere completamente il problema di selezione del portafoglio, occorra individuare un solo portafoglio ottimo fra quelli efficienti; esistono diversi metodi per effettuare questa scelta, per la loro esposizione si rimanda a [16].

1.4.3 I giochi cooperativi

Consideriamo un *gioco cooperativo con n giocatori*. Siano:

- X, Y_1, \dots, Y_n spazi vettoriali,
- S un sottoinsieme di X non vuoto,
- C_1, \dots, C_n coni convessi contenuti rispettivamente in Y_1, \dots, Y_n .

Per ogni giocatore si definisca poi una funzione obiettivo $f_k : S \rightarrow Y_k$ ($k = 1, \dots, n$). Il k -esimo giocatore cercherà di massimizzare, rispetto all'ordine indotto in Y_k dal cono C_k , la sua funzione obiettivo f_k . Dal momento che i giocatori interagiscono in modo esclusivamente cooperativo (a differenza del concetto introdotto da Von Neumann), essi non possono danneggiarsi a vicenda (per approfondimenti circa questo aspetto vd. [31], cap.10). Al fine di introdurre in

questo contesto una qualche nozione di ottimalità risulta conveniente definire lo spazio prodotto

$$Y := \prod_{k=1}^n Y_k$$

ed ordinarlo mediante la relazione d'ordine (ordine prodotto) associata al cono prodotto così definito

$$C := \prod_{k=1}^n C_k.$$

L'ordine prodotto ben si adatta a descrivere il concetto di cooperazione considerato, infatti un elemento $\bar{x} \in S$ è "preferito" quando è "preferito" da ogni singolo giocatore. Risulta allora evidente che è rilevante determinare l'insieme $S\text{Max}(f, S)$ delle soluzioni del problema di ottimizzazione vettoriale

$$\max_{x \in S} f(x)$$

dove f è la seguente funzione

$$f := (f_1, \dots, f_k) : S \longrightarrow Y.$$

Si conclude quindi che un gioco cooperativo con n giocatori del tipo considerato si può ricondurre ad un problema astratto di ottimizzazione vettoriale. Per ulteriori approfondimenti si rimanda al volume [31].

1.4.4 Applicazioni all'ingegneria

L'ottimizzazione vettoriale ha trovato numerose applicazioni in ambito ingegneristico, in particolare nella progettazione di strutture complesse. A tale proposito, in questa sede, ci limitiamo a fornire alcune citazioni bibliografiche: in [32] vengono trattati due esempi particolari (progettazione di una sbarra speciale e di un reattore chimico) mentre per una trattazione più estesa si può far riferimento al volume [23]; un esempio concreto e molto interessante delle applicazioni di tecniche di ottimizzazione multiobiettivo è poi fornito da alcune tematiche sviluppate in un progetto sviluppato dal Centro Ricerche FIAT in collaborazione con diversi partner europei (SNECMA-Francia, Intrasoftware-Grecia, Università di Warwick, CNRS- Francia e Politecnico di Torino) che sono poi state impiegate tra l'altro nella progettazione della parte frontale dell'automobile FIAT Multipla (vd. [6]).

Capitolo 2

Inclusioni differenziali

In questo capitolo ci proponiamo di illustrare alcuni aspetti della teoria riguardante le inclusioni differenziali, fornendo nel contempo un sintetico inquadramento generale di questo argomento. Non tutti i risultati e le considerazioni esposti saranno utilizzati nel seguito, tuttavia per agevolare la comprensione delle applicazioni all'ottimizzazione vettoriale di questi strumenti e per rendere più comprensibile l'intera presentazione, ci è parso opportuno includere in questa breve esposizione alcuni argomenti di grande rilevanza, anche se non strettamente necessari per gli sviluppi successivi trattati in questa tesi.

I principali testi a cui faremo riferimento sono le monografie di Aubin-Cellina ([3]) e di Deimling ([20]) a cui si rimanda per ulteriori approfondimenti.

2.1 Mappe punto-insieme

In questa sezione intendiamo fornire una breve introduzione alla cosiddetta set-valued analysis, in particolare sottolineeremo alcuni risultati riguardanti le mappe punto insieme che risultano strettamente interconnessi con la trattazione delle inclusioni differenziali. Per questo argomento oltre ai testi sopra citati si può anche fare ricorso a [4] ed a [19] (cap. 8). Altri testi trattano questi argomenti in modo lievemente diverso dalla presentazione che abbiamo preferito seguire in questa sede; ricordiamo, fra gli altri, [54] (cap.5) e [2] (cap.14).

Notazione 27 *Nel seguito le mappe punto-insieme saranno indicate con lettere latine maiuscole (F, \dots) mentre le funzioni (single-valued) saranno denotate con lettere latine minuscole (f, \dots).*

2.1.1 Generalità e notazioni

Incominciamo con il formalizzare precisamente il concetto di mappa punto-insieme.

Siano X ed Y due insiemi. Una *mappa punto-insieme* F da X a Y è una mappa che associa ad ogni elemento $x \in X$ un sottoinsieme $F(x)$ di Y e nel

seguito sarà indicata con la notazione seguente

$$F : X \rightrightarrows Y$$

L'insieme $F(x)$ prende il nome di *immagine* (o *valore*) di x mediante F . L'insieme

$$\text{Dom}(F) := \{x \in X : F(x) \neq \emptyset\}$$

è chiamato *dominio* di F . Quando $\text{Dom}(F) = X$ la mappa F si dice *stretta*. Una mappa, il cui dominio è non vuoto, prende il nome di mappa *propria*. E' utile osservare che una qualunque mappa punto-insieme, definita solamente su di un sottoinsieme proprio K di X con immagini non vuote, può essere estesa in modo naturale ad una mappa propria F_K definita su tutto X :

$$F_K := \begin{cases} F(x) & \text{se } x \in K \\ \emptyset & \text{se } x \notin K \end{cases};$$

si osserva poi immediatamente che $\text{Dom}(F_K) = K$. Nel seguito supporremo sempre, salvo diverso avviso, che una mappa punto-insieme F sia tale che $F(x) \neq \emptyset$ per ogni $x \in X$.

Spesso, per descrivere una mappa punto-insieme F , è molto utile ricorrere alla nozione di *grafico*:

$$\text{Graph}(F) := \{(x, y) \in X \times Y : y \in F(x)\}.$$

Si associa così ad una mappa punto-insieme un sottoinsieme dell'insieme dato dal prodotto cartesiano tra X ed Y ; in modo del tutto simmetrico è possibile associare ad un sottoinsieme \mathcal{F} non vuoto di $X \times Y$ una mappa punto-insieme F ponendo

$$F(x) := \{y \in Y : (x, y) \in \mathcal{F}\},$$

si ha allora che $\text{Graph}(F) = \mathcal{F}$.

Definiamo ora un sottoinsieme di Y che prende il nome di *range* di F :

$$R(F) := \bigcup_{x \in X} F(x).$$

Possiamo allora definire l'*inversa* F^{-1} di una mappa punto-insieme F nel seguente modo

$$x \in F^{-1}(y) \iff y \in F(x).$$

Si osservi poi che la *controimmagine* $F^{-1}(A)$ di un qualunque sottoinsieme $A \subset Y$ risulta essere allora l'insieme

$$F^{-1}(A) = \{x \in X : F(x) \cap A \neq \emptyset\}.$$

Introduciamo infine anche il concetto di *core* di A mediante F

$$F^{+1}(A) = \{x \in X : F(x) \subset A\}$$

In base a quanto visto sopra, ad F è possibile associare un sottoinsieme non vuoto \mathcal{F} del prodotto cartesiano $X \times Y$; risulta allora che

$$F^{-1}(y) = \{x \in X : (x, y) \in \mathcal{F}\}.$$

Due osservazioni riguardanti i concetti sopra definiti sono immediate:

1. $\text{Dom}(F)$ è la proiezione di \mathcal{F} su X ,
2. $R(F) = \text{Dom}(F^{-1})$ è la proiezione di \mathcal{F} su Y .

Se lo spazio Y è dotato di struttura topologica si dice che una mappa punto-insieme F è *compatta (limitata)* se il suo range $R(F)$ è un sottoinsieme compatto (limitato) di Y . Analogamente si parla di mappa *localmente compatta (localmente limitata)* quando per ogni $x \in \text{Dom}(F)$ esiste un intorno di x nel quale F è compatta (limitata).

2.1.2 Continuità per mappe punto-insieme

Ci proponiamo ora di introdurre alcune definizioni che riguardano le proprietà di continuità delle mappe punto-insieme; per far questo dobbiamo però pensare agli insiemi X ed Y dotati perlomeno di una struttura topologica, in particolare, noi supporremo che X ed Y siano spazi metrici. Sia F una mappa punto insieme a valori non vuoti, incominciamo con il definire la nozione di semicontinuità superiore che fu introdotta nei primi anni trenta da Bouligand, Kuratowski e Wilson proprio nel contesto di spazi metrici.

Definizione 28 *La mappa punto-insieme F si dice semicontinua superiormente in $x_0 \in X$ se per ogni aperto N contenente $F(x_0)$ esiste un intorno M di x_0 tale che $F(M) \subset N$.*

Si dice poi che F è semicontinua superiormente se F è semicontinua superiormente in x , per ogni $x \in X$.

Grazie al fatto che abbiamo supposto, per ipotesi, che gli insiemi X ed Y siano spazi metrici possiamo caratterizzare la semicontinuità superiore mediante l'utilizzo di successioni: la mappa punto-insieme F è semicontinua superiormente in x_0 se e solo se dati una successione $\{x_n\} \subset X$ ed un sottoinsieme chiuso A di X , allora $x_n \rightarrow x_0$ e $F(x_n) \cap A \neq \emptyset$ per ogni $n \geq 1$ implicano che $F(x_0) \cap A \neq \emptyset$. Questa caratterizzazione della semicontinuità superiore è particolarmente significativa perchè ben si presta a fornire una spiegazione "intuitiva" di questo concetto: $F(x_0)$ deve essere almeno tanto grande quanto il "limite" di $\{F(x_n)\}$; in altre parole ciò significa che alle immagini di una mappa F semicontinua superiormente è permesso "saltare" solo verso l'alto (rispetto all'inclusione insiemistica). Chiaramente le osservazioni precedenti sottointendono che sia possibile parlare

di convergenze d'insiemi, tuttavia in questa sede non ci è possibile dilungarci su questi aspetti per cui si rimanda a [4].

Una altra possibile riformulazione della precedente definizione di natura esclusivamente topologica è data dalla seguente proposizione.

Proposizione 29 *La mappa punto-insieme F è semicontinua superiormente se e solo se $F^{-1}(A)$ è chiuso per ogni sottoinsieme A chiuso di Y .*

Si osservi che l'affermazione precedente è del tutto equivalente a richiedere che l'insieme $F^{+1}(V) = \{x \in X : F(x) \subset V\}$ sia aperto ogni qualvolta V è un sottoinsieme aperto di Y . In base alla proposizione precedente è agevole osservare che, se F è una normale funzione definita in X a valori in Y , la semicontinuità superiore coincide con la usuale nozione di continuità; questo fatto sgombra subito il campo da un equivoco che potrebbe sorgere, infatti la semicontinuità superiore di mappe punto-insieme non ha relazioni con la nozione di semicontinuità superiore per funzioni definite in X a valori in \mathbb{R} .

Diventa a questo punto spontaneo chiedersi perchè utilizzare per le mappe punto-insieme il termine semicontinuità superiore al posto di quello di continuità. Una delle possibili risposte a questa domanda risiede nel fatto che una delle più utili caratterizzazioni delle funzioni continue - una funzione f è continua se e solo se manda successioni convergenti ad x_0 in successioni convergenti ad $f(x_0)$ - non è più valida nell'ambito delle mappe punto-insieme semicontinue superiormente. Per far sì che questa proprietà continui ad essere valida anche nel caso di mappe punto-insieme bisogna introdurre un differente tipo di continuità.

Definizione 30 *Si dice che una mappa punto-insieme F è semicontinua inferiormente in $x_0 \in X$ quando per ogni $y_0 \in F(x_0)$ e per ogni intorno U di y_0 , esiste un intorno V di x_0 tale che*

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

Si dice poi che F è semicontinua inferiormente se F è semicontinua inferiormente in x , per ogni $x \in X$.

Ricordando ora che, dal momento che noi abbiamo supposto X ed Y essere spazi metrici, possiamo caratterizzare le mappe semicontinue superiormente da un punto di vista successionale (ricollegandoci così a quanto detto immediatamente prima della definizione e mostrando che per questa classe di mappe continua a valere l'analogo della proprietà citata per le funzioni continue): la mappa punto-insieme F è semicontinua inferiormente in x_0 se e solo se dati una successione $\{x_n\} \subset X$ tale che $x_n \rightarrow x_0$ ed un sottoinsieme aperto V di X , allora $F(x_0) \cap V \neq \emptyset$ implica che $F(x_n) \cap V \neq \emptyset$ definitivamente. In base a quanto detto è immediato constatare che nel caso di funzioni definite in X a valori in Y la nozione di semicontinuità inferiore per mappe punto-insieme coincide con

l'usuale continuità di funzioni; quindi anche in questo caso si vede che non ci sono relazioni sostanziali fra la semicontinuità inferiore per mappe punto-insieme e la semicontinuità inferiore per funzioni a valori reali.

Anche in questo caso è possibile fornire una caratterizzazione della semicontinuità inferiore basandoci sulle proprietà topologiche delle controimmagini di insiemi.

Proposizione 31 *La mappa punto-insieme F è semicontinua inferiormente se e solo se $F^{-1}(V)$ è aperto per ogni sottoinsieme V aperto di Y .*

Passando ai complementari si giunge immediatamente ad affermare che F è semicontinua inferiormente se e solo se l'insieme $F^{+1}(A)$ è chiuso ogni qualvolta A è un sottoinsieme chiuso di Y .

Come già evidenziato nelle considerazioni precedenti, si ha che le nozioni di semicontinuità superiore ed inferiore per una mappa punto-insieme $F : X \rightrightarrows Y$ coincidono entrambe con la continuità, nel caso particolare in cui F sia una usuale funzione (single-valued) definita in X a valori in Y . Tuttavia, sfortunatamente, in generale le due nozioni non coincidono, come mostra il seguente esempio in cui compaiono mappe punto-insieme che soddisfano una delle due proprietà ma non l'altra.

Esempio 32 *Consideriamo la mappa punto-insieme $F_1 : \mathbb{R} \rightrightarrows \mathbb{R}$ così definita*

$$F_1(x) = \begin{cases} [-1, +1] & \text{se } x \neq 0 \\ \{0\} & \text{se } x = 0 \end{cases} ,$$

essa risulta semicontinua inferiormente in zero ma non è superiormente semicontinua nello stesso punto.

D'altro canto si consideri la mappa punto-insieme $F_2 : \mathbb{R} \rightrightarrows \mathbb{R}$ definita da

$$F_2(x) = \begin{cases} \{0\} & \text{se } x \neq 0 \\ [-1, +1] & \text{se } x = 0 \end{cases} ,$$

si constata facilmente che F_2 è semicontinua superiormente in zero ma non è inferiormente semicontinua in zero.

Alla luce di quanto visto, anche con quest'ultimo esempio è sensato introdurre la seguente definizione.

Definizione 33 *Una mappa punto insieme $F : X \rightrightarrows Y$ si dice continua quando è simultaneamente inferiormente e superiormente semicontinua.*

Le mappe continue avranno grande rilevanza nella parte originale di questo lavoro.

Come ultima osservazione, nell'ambito di questa breve presentazione delle nozioni di continuità per mappe punto-insieme, sottolineiamo che esistono anche raffinamenti delle definizioni finora date (si pensi ad esempio alla ε - δ -semicontinuità inferiore e superiore, vedi [3] e [20]) che esulano però dai nostri interessi in questa sede; ricordiamo infine che, sempre nell'ambito di nozioni di regolarità per mappe punto-insieme, si può parlare di mappe lipschitziane e localmente lipschitziane (cfr. ad esempio [4]). E' infatti possibile formulare la seguente definizione.

Definizione 34 Sia $F : X \rightrightarrows Y$ una mappa punto-insieme. Si dice che F è localmente lipschitziana se per ogni $x_0 \in X$, esiste un intorno $U(x_0) \subset X$ ed una costante reale $L(x_0) \geq 0$ tale che

$$x, x^* \in U(x_0), \quad F(x) \subset B_{L(x_0)d(x,x^*)}(F(x^*)).$$

F si dice poi lipschitziana se esiste una costante reale L tale che

$$x, x^* \in X, \quad F(x) \subset B_{Ld(x,x^*)}(F(x^*)).$$

2.1.3 Selezioni di mappe punto-insieme

Uno dei principali temi dell'analisi delle mappe punto insieme riguarda i rapporti che si possono stabilire tra mappe punto-insieme e funzioni usuali. In questo ambito è cruciale il concetto di selezione di una mappa punto-insieme.

Definizione 35 Sia $F : X \rightrightarrows Y$ una mappa punto-insieme. Una funzione (single-valued) $f : X \rightarrow Y$ si dice selezione di F quando

$$\forall x \in X, \quad f(x) \in F(x).$$

In termini molto generali appare evidente che una selezione esiste sempre, qualunque mappa punto-insieme stretta si consideri; questo fatto è infatti immediata conseguenza della validità dell'assioma della scelta che ci permette di scegliere un elemento $f(x)$ in tutti gli insiemi $F(x)$. Naturalmente, data una mappa punto-insieme non è certo garantita l'unicità della selezione. I problemi sorgono però quando si vuole che la selezione goda di prefissate proprietà di regolarità; è infatti naturale chiedersi se, data una mappa punto-insieme con certe proprietà di regolarità (ad esempio semicontinuità inferiore o superiore,...), esse vengono ereditate da una selezione della mappa considerata. In altre parole ci si pone il seguente problema: data una mappa F con certi requisiti di regolarità è possibile provare l'esistenza di una sua selezione f che soddisfi anch'essa delle condizioni di regolarità prefissate. La risposta a questo quesito non è affatto scontata, come mostra il seguente esempio.

Esempio 36 In questo esempio costruiremo una mappa punto-insieme F continua e definita sulla bolla unitaria chiusa di \mathbb{R}^2 (indicata con $\overline{B_1((0,0))}$) a valori nei sottoinsiemi compatti della stessa. che tuttavia non possiede alcuna selezione continua.

Sia $Q = \{(\rho, \theta) : 0 \leq \rho \leq 1, 0 \leq \theta < 2\pi\}$, definiamo allora la seguente funzione

$$\nu(\rho, \theta) := (\rho \cos \theta, \rho \sin \theta);$$

si poi osserva subito che ν è una biiezione da Q in $\overline{B_1((0,0))} \setminus \{(0,0)\}$. Definiamo ora su Q la seguente mappa punto-insieme a valori in \mathbb{R}

$$\Phi(\rho, \theta) := \{\omega \in \mathbb{R} : -\theta - 2\pi(1 - \rho) \leq \omega \leq -\theta + 2\pi(1 - \rho)\}.$$

In particolare si ha che $\Phi(1, \theta) = \{-\theta\}$.

Costruiamo poi la funzione $\tau : \mathbb{R} \rightarrow [0, 2\pi)$ ponendo $\tau(x) := x \bmod(2\pi)$ e definiamo $\Psi : Q \rightarrow [0, 2\pi)$ nel modo seguente:

$$\Psi(\rho, \theta) := \tau(\Phi(\rho, \theta))$$

Infine costruiamo la mappa punto-insieme $F : \overline{B_1((0,0))} \rightarrow \overline{B_1((0,0))}$ cercata:

$$F(x, y) := \begin{cases} \nu(1, \Psi(\nu^{-1}(x, y))) & \text{se } (x, y) \neq (0, 0) \\ \{(\xi, \eta) \in \mathbb{R}^2 : \xi^2 + \eta^2 = 1\} & \text{se } x = y = 0. \end{cases}$$

F è continua ed a valori compatti, inoltre $F(x, y) \subseteq F(0, 0)$ per ogni $(x, y) \in \overline{B_1((0,0))}$.

Supponiamo ora che F ammetta una selezione continua f . Allora f manda $\overline{B_1((0,0))}$ in se stessa e quindi, in forza del teorema di punto fisso di Brouwer (vd. [49], pp. 13-15), ha un punto fisso (x_0, y_0) . Naturalmente $(x_0, y_0) = f(x_0, y_0) \in F(x_0, y_0)$ e $\|(x_0, y_0)\| = 1$. Possiamo dunque concludere che $F(x_0, y_0) = \{-(x_0, y_0)\}$ e quest'ultima affermazione porta ad una contraddizione.

Accanto al problema sopra esposto ha poi grande importanza la costruzione esplicita, data la mappa F , di una sua selezione f .

In questa sezione daremo, seppure in estrema sintesi, una risposta ai problemi fino ad ora delineati. Incominciamo con l'affrontare il problema dell'esistenza di una selezione per mappe punto-insieme semicontinue inferiormente. Il risultato chiave riguardante questa classe di mappe è il teorema di Michael. Questo teorema assicura l'esistenza di una selezione continua in ipotesi di convessità per le immagini della mappa punto-insieme.

Teorema 37 (di Michael) Siano X uno spazio metrico ed Y uno spazio di Banach. Sia $F : X \rightarrow Y$ una mappa punto-insieme semicontinua inferiormente con immagini chiuse e convesse. Allora esiste una selezione f continua di F .

Questo teorema ha grande importanza perchè fissato un qualunque punto $(x_0, y_0) \in \text{Graph}(F)$ permette di trovare una selezione continua di F passante per il punto dato. Vale infatti il seguente corollario.

Corollario 38 *Siano F, X ed Y come nel teorema precedente. Siano poi $A \subset X$ un sottoinsieme e $\varphi : A \rightarrow Y$ una funzione continua tale che per ogni $x \in A$, $\varphi(x) \in F(x)$. Allora la funzione può essere estesa ad una selezione f continua (definita sull'intero spazio X) tale che la restrizione di f ad A coincida con φ .*

In particolare si ottiene che per ogni $y^ \in F(x^*)$ esiste una selezione continua f di F tale che $f(x^*) = y^*$.*

Passiamo ora ad affrontare il problema di trovare una selezione per mappe punto-insieme che siano semicontinue superiormente. Sfortunatamente la sola ipotesi di semicontinuità superiore non basta ad assicurare l'esistenza di una selezione continua, neppure se si adotta l'ipotesi aggiuntiva di immagini chiuse e convesse. Si consideri infatti il seguente esempio.

Esempio 39 *Consideriamo la mappa punto-insieme $F : \mathbb{R} \rightrightarrows \mathbb{R}$ definita da*

$$F := \begin{cases} \{-1\} & \text{se } x < 0 \\ [-1, +1] & \text{se } x = 0 \\ \{+1\} & \text{se } x > 0 \end{cases},$$

è immediato constatare che F è semicontinua superiormente ed è altrettanto chiaro che F non ammette selezioni continue.

Nel caso di mappe punto-insieme semicontinue superiormente è allora più appropriato parlare di *selezioni approssimate*; l'esistenza di questo tipo di selezioni, che godono di proprietà più deboli rispetto alle selezioni in senso proprio, è infatti comunque sufficiente ad assicurare la validità della maggior parte dei risultati interessanti (ad esempio esistenza di soluzioni per inclusioni differenziali). Il teorema chiave concernente le soluzioni approssimate è dovuto ad A. Cellina (cfr. [13] per il lavoro originale, altrimenti vd. [3] oppure [4]; si osservi inoltre come questo teorema sia stato dimostrato in modo alternativo da S. Smale in [66]). Tale risultato è noto come "Teorema della selezione approssimata".

Teorema 40 *Siano X uno spazio metrico, Y uno spazio di Banach, $F : X \rightrightarrows Y$ una mappa punto-insieme semicontinua superiormente ed a valori convessi. Allora per ogni $\varepsilon > 0$ esiste una funzione f_ε localmente lipschitziana tale che*

$$\text{Graph}(f_\varepsilon) \subset \text{Graph}(F) + B_\varepsilon(0)$$

ed inoltre per ogni $x \in X$, $f_\varepsilon(x)$ appartiene all'involucro convesso di $R(F)$.

E' particolarmente interessante osservare come da questo teorema si possa dedurre in modo agevole uno strumento fondamentale dell'analisi, spesso utilizzato anche in applicazioni alla teoria economica, quale il teorema di punto fisso di Kakutani. La dimostrazione di questo teorema mediante il risultato precedente è particolarmente elegante e breve anche se fa ricorso al teorema di Schauder (vd. [19] pag. 60) che altro non è che una generalizzazione a spazi infinito dimensionali del teorema di punto fisso di Brouwer (vd. [49] pp.13-15 oppure [19] pag.17).

Corollario 41 (Teorema di punto fisso di Kakutani) *Siano K un sottoinsieme compatto e convesso di X spazio di Banach ed $F : K \rightrightarrows K$ una mappa punto-insieme semicontinua superiormente ed a valori compatti e convessi. Allora esiste $x^* \in K$ tale che $x^* \in F(x^*)$.*

Dimostrazione. (cfr. [3] pag.85). In forza del teorema precedente possiamo costruire una successione $\{f_n\}_{n \in \mathbb{N}}$ di funzioni continue da K in K tali che

$$\text{Graph}(f_n) \subset \text{Graph}(F) + B_{\varepsilon_n}(0).$$

con $\{\varepsilon_n\}$ successione reale tale che $\varepsilon_n > 0$ per ogni n ed $\varepsilon_n \rightarrow 0$. Per il teorema di Schauder possiamo ora affermare che esiste una successione $\{x_n\}$ tale che $x_n = f(x_n)$; inoltre una sua opportuna sottosuccessione $\{x_{n_j}\}$ convergerà ad un elemento x^* di K . Perciò si ottiene che

$$d((x^*, x^*), \text{Graph}(F)) = \lim_{j \rightarrow +\infty} d((x_{n_j}, f(x_{n_j})), \text{Graph}(F)) = 0.$$

Dal momento che F è semicontinua superiormente $\text{Graph}(F)$ è chiuso (cfr. [3], pag.41), perciò $(x^*, x^*) \in \text{Graph}(F)$ ovvero $x^* \in F(x^*)$. ■

Ci proponiamo ora di costruire esplicitamente una selezione, la cosiddetta *selezione minimale* di una mappa punto-insieme $F : X \rightrightarrows Y$ prefissata. Questa selezione è costruita seguendo un processo molto naturale: si sceglie infatti di ogni insieme $F(x)$ l'elemento di norma minima (ovviamente le richieste sullo spazio Y devono essere tali perlomeno da garantire che abbia significato parlare di elemento di minima norma). Data la rilevanza che la selezione minimale avrà nel seguito forniremo anche le dimostrazioni dei risultati che la riguardano.

Siano X uno spazio metrico ed Y uno spazio di Banach, consideriamo poi una mappa punto-insieme $F : X \rightrightarrows Y$; possiamo allora dare la seguente definizione

Definizione 42 *Definiamo la mappa punto-insieme*

$$m(F(x)) := \left\{ u \in F(x) : \|u\| = \min_{z \in F(x)} \|z\| \right\}$$

che prende il nome di mappa minimale.

In base al teorema della migliore approssimazione (vd. Appendice 1) si prova che quando Y è uno spazio di Hilbert la mappa minimale diventa una funzione (single-valued) che prende il nome di selezione minimale di F . Più in generale questo fatto è vero quando Y è uno spazio di Banach strettamente convesso (si ricordi che uno spazio di Banach Y si dice strettamente convesso quando per ogni $x, y \in Y$ non collineari si ha che $\|x + y\| < \|x\| + \|y\|$).

A questo punto è naturale chiedersi quali proprietà di regolarità soddisfi la selezione minimale di una mappa F ed in particolare che rapporti ci sono tra la regolarità della mappa punto-insieme F e le caratteristiche della selezione. La risposta a questo quesito è fornita dal seguente teorema.

Teorema 43 *Siano X uno spazio metrico, Y uno spazio di Hilbert ed $F : X \rightrightarrows Y$ una mappa punto insieme a valori chiusi e convessi. Allora la funzione $x \mapsto m(F(x))$ è una selezione continua di F .*

Dimostrazione. (cfr. [3] pag.70) Incominciamo con il fissare $\varepsilon > 0$ ed $x \in X$. Per la continuità di F ogniqualvolta z è sufficientemente vicino ad x si ha che

$$\|m(F(x))\|^2 \leq \|m(F(z))\|^2 + \varepsilon.$$

La relazione precedente implica che $m(F(x)) \in B_{\sqrt{\varepsilon}}(0)$ quando $\|m(F(z))\| = 0$ e ciò prova il teorema in questo caso particolare.

Esaminiamo ora il caso in cui $\|m(F(z))\| > 0$. Per ogni x sufficientemente vicino a z esiste un $y_x \in F(z)$ tale che

$$\|m(F(x)) - y_x\| \leq \frac{\varepsilon}{\|m(F(z))\|}.$$

Possiamo perciò scrivere che

$$\begin{aligned} \langle m(F(z)), m(F(x)) \rangle &= & (2.1) \\ &= \langle m(F(z)), y_x \rangle + \langle m(F(z)), m(F(x)) - y_x \rangle \geq \\ &\geq \langle m(F(z)), y_x \rangle - \varepsilon \end{aligned}$$

Dal teorema della migliore approssimazione (in particolare dalla caratterizzazione variazionale dell'elemento di minima norma, vd. Appendice 1) segue ora che, per ogni $y \in F(z)$, si ha che

$$\langle m(F(z)), y \rangle \geq \langle m(F(z)), m(F(z)) \rangle$$

allora la (2.1) diventa

$$\langle m(F(z)), m(F(x)) - m(F(z)) \rangle \geq -\varepsilon \quad (2.2)$$

D'altro canto si ha anche che

$$\begin{aligned} \|m(F(x))\|^2 &= \|m(F(z))\|^2 + \|m(F(z)) - m(F(x))\|^2 \\ &\quad + 2 \langle m(F(z)), m(F(x)) - m(F(z)) \rangle. \end{aligned}$$

Questa relazione, insieme con la (2.2) portano a

$$\begin{aligned} \|m(F(z)) - m(F(x))\|^2 - 2\varepsilon &\leq \\ &\leq \|m(F(z)) - m(F(x))\|^2 \\ &\quad + 2 \langle m(F(z)), m(F(x)) - m(F(z)) \rangle \\ &\leq \varepsilon \end{aligned}$$

da cui

$$\|m(F(z)) - m(F(x))\|^2 \leq 3\varepsilon$$

Abbiamo così provato che la funzione $x \mapsto m(F(x))$ è continua. ■

Il teorema precedente rimane valido anche in spazi più generali con l'aggiunta di un'ipotesi di compattezza, a tale proposito e per ulteriori approfondimenti riguardanti la selezione minimale si rimanda a [4] pp. 360-364.

Con il teorema 43 abbiamo quindi visto che la selezione minimale eredita la proprietà di continuità dalla continuità della mappa punto-insieme a cui è associata. E' però interessante osservare come la sola semicontinuit , sia inferiore sia superiore, non basti a garantire la continuit  della selezione minimale. Valgono infatti i seguenti esempi.

Esempio 44 Consideriamo la mappa punto-insieme $F : \mathbb{R} \rightrightarrows \mathbb{R}$ cos  definita

$$F(x) := \begin{cases} \{2\} & \text{se } x \neq 0 \\ [1, 2] & \text{se } x = 0 \end{cases}.$$

F risulta essere semicontinua superiormente ed a valori compatti e convessi tuttavia la sua selezione minimale

$$m(F(x)) = \begin{cases} 2 & \text{se } x \neq 0 \\ 1 & \text{se } x = 0 \end{cases}$$

non   continua in 0.

Esempio 45 Consideriamo la mappa punto-insieme $F : \mathbb{R} \rightrightarrows \mathbb{R}$ cos  definita

$$F(x) := \begin{cases} [0, 1] & \text{se } x \neq 0 \\ \{1\} & \text{se } x = 0 \end{cases}.$$

F risulta essere semicontinua inferiormente ed a valori compatti e convessi tuttavia la sua selezione minimale

$$m(F(x)) = \begin{cases} 0 & \text{se } x \neq 0 \\ 1 & \text{se } x = 0 \end{cases}$$

non   continua in 0.

Si osserva che la selezione minimale non eredita la lipschitzianità; si può infatti costruire la selezione minimale di una opportuna mappa punto-insieme lipschitziana e dimostrare che essa non è una funzione lipschitziana (vd. [3], pp.70-72).

Esistono anche altre procedure che portano ad una costruzione esplicita di una selezione per una mappa punto-insieme (selezione di Chebishev, selezione baricentrica), tuttavia la loro esposizione esula dagli scopi di questo lavoro, perciò si rimanda ancora una volta al volume [3].

2.2 Teoremi di esistenza per le inclusioni differenziali

Uno dei problemi di maggiore rilevanza nella trattazione delle inclusioni differenziali, in analogia con la teoria delle equazioni differenziali, consiste nello studio di teoremi d'esistenza per le soluzioni di una data inclusione differenziale.

In questa sezione ci occuperemo proprio di passare in rassegna teoremi d'esistenza per le inclusioni differenziali della forma

$$x'(t) \in F(x(t)) \quad (2.3)$$

oppure

$$x'(t) \in F(t, x(t)) \quad (2.4)$$

In questa sezione seguiremo principalmente il volume di Aubin e Cellina ([3]) a cui si rimanda per una esposizione completa e dettagliata dell'argomento. Una trattazione dell'argomento parzialmente differente (distinzione tra il caso finito dimensionale e quello infinito dimensionale) si trova in [20].

Prima di affrontare direttamente i teoremi di esistenza è necessario fare un passo indietro e chiarire cosa si intenda per soluzione delle inclusioni differenziali (2.3) e (2.4).

Nel caso di equazioni differenziali ordinarie della forma $x'(t) = f(x(t))$ oppure $x'(t) = f(t, x(t))$, almeno nel caso in cui f sia supposta continua, la situazione è chiara: la soluzione è una funzione $t \mapsto x(t)$ differenziabile con continuità ovunque. Solo nel caso di equazioni differenziali autonome, qualora si decida di prendere in considerazione funzioni f che siano solo misurabili rispetto alla variabile t , si deve ricorrere ad un concetto di soluzione più debole, cioè si accettano come soluzioni funzioni derivabili tranne che in un insieme di misura nulla (soluzioni nel senso di Carathéodory; per la loro definizione cfr. [72]).

Quando si passa invece ad esaminare la situazione nel caso di inclusioni differenziali le cose si complicano. Consideriamo infatti la seguente inclusione differenziale

$$x'(t) \in \{-1, +1\}$$

e chiedamoci quali siano le sue soluzioni passanti per 0 all'istante $t = 0$, cioè tali che $x(0) = 0$. Se imponiamo che le soluzioni siano derivabili con derivata continua, troviamo solo due funzioni che rispettano le nostre condizioni: $x(t) = t$ e $x(t) = -t$. Sembra tuttavia del tutto sensato accettare anche altre soluzioni dell'inclusione differenziale considerata; potremmo ad esempio considerare soluzioni tali che la loro derivata non sia definita su di un insieme sufficientemente "piccolo" (ad esempio finito, oppure numerabile, o ancora di misura nulla). La scelta della classe di soluzioni che si vuole considerare dipende solitamente da due fattori, uno legato alle applicazioni, l'altro di carattere prettamente matematico; il primo infatti è legato a considerazioni esterne al campo matematico che fanno preferire una certa classe di soluzioni rispetto ad un'altra, il secondo invece è dato dall'esigenza di riuscire a provare un teorema d'esistenza per la classe di inclusioni considerata. Chiaramente, rispetto a questo secondo punto, è immediato osservare che più debole è il concetto di soluzione adottato, tanto sarà più facile provare un teorema d'esistenza per la classe considerata. Tuttavia è necessario imporre alcuni limitazioni affinché alcune proprietà vengano mantenute. Si pensi infatti al caso delle equazioni differenziali ordinarie dove per conservare l'importante proprietà che stabilisce l'equivalenza tra l'equazione differenziale considerata e la corrispondente equazione integrale

$$x(t) = x(0) + \int_0^t f(u, x(u)) du$$

si deve imporre che la soluzione $t \mapsto x(t)$, oltre che continua, sia assolutamente continua (per la definizione e le proprietà delle funzioni assolutamente continue si veda ad esempio [36] vol.3, pp. 260-284). Conveniamo perciò, anche nel caso di inclusioni differenziali, di adottare lo stesso tipo di approccio al problema. Giungiamo così alla seguente definizione.

Definizione 46 *Una funzione $t \mapsto x(t)$ è soluzione dell'inclusione differenziale (2.3) (rispettivamente (2.4)) quando è assolutamente continua e soddisfa la (2.3) (rispettivamente la (2.4)) quasi ovunque.*

Salvo diverso avviso nel seguito si intenderà quasi ovunque rispetto alla misura di Lebesgue.

Come si diceva sopra, il poter accettare come soluzioni anche funzioni non ovunque differenziabili è spesso motivato da esigenze di natura applicativa dove per meglio descrivere certi fenomeni è necessari utilizzare funzioni anche non di classe C^1 ; quindi la definizione precedente può essere messa in discussione in base a precise esigenze applicative. Tuttavia la richiesta dell'assoluta continuità delle soluzioni fa sì che vengano scartate situazioni "patologiche" (si pensi ad esempio al classico esempio di Vitali di una funzione continua la cui derivata esiste quasi ovunque ed è zero in tutti i punti in cui esiste, inoltre vale 0 in 0 ed 1 in 1; per la sua costruzione cfr. [36] vol.3, pp.235-237). Dal momento

che abbiamo stabilito che le soluzioni di un'inclusione differenziale si assumono essere perlomeno assolutamente continue può essere interessante stabilire sotto quali ulteriori ipotesi esistano soluzioni con un grado di regolarità più alto.

Nello sviluppo di teoremi di esistenza per le soluzioni di inclusioni differenziali si impongono richieste sulla mappa punto-insieme F ; esse sono generalmente di due tipi diversi:

- di *regolarità* sulla mappa punto-insieme F (ad esempio semicontinuità o continuità),
- di *natura topologico-geometrica* (compattezza e convessità) sulle immagini della mappa F .

Naturalmente è possibile combinare in diversi modi queste ipotesi ottenendo così differenti tipologie di risultati.

L'approccio più immediato e naturale al problema della dimostrazione di teoremi di esistenza per inclusioni differenziali consiste nel ricondursi a considerare il corrispondente problema per equazioni differenziali ordinarie. Gli strumenti che si utilizzano per operare in questo senso sono costituiti dalle selezioni di mappe punto-insieme esposte in 2.1.3; si deve infatti dimostrare che dentro un'inclusione differenziale si trova "nascosta" un'equazione differenziale, è sufficiente allora utilizzare un teorema che garantisca l'esistenza di una soluzione per questa equazione differenziale poichè, se esiste una selezione f continua della mappa punto-insieme F , è chiaro che le soluzioni dell'equazione differenziale ordinaria $x'(t) = f(t, x(t))$ sono soluzioni anche dell'inclusione differenziale $x'(t) \in F(t, x(t))$.

Passiamo ora ad esporre in dettaglio alcuni teoremi d'esistenza per le soluzioni di un'inclusione differenziale o, per essere più precisi, per le soluzioni di un problema di Cauchy associato all'inclusione differenziale (2.4) dove il problema di Cauchy per un'inclusione differenziale è definito in modo del tutto analogo a quello ben noto un'equazione differenziale ordinaria. Consideriamo pertanto il seguente problema di Cauchy

$$\begin{cases} x'(t) \in F(t, x(t)) \\ x(0) = x_0 \end{cases} \quad (2.5)$$

Distingueremo il caso delle mappe punto-insieme a valori convessi da quello delle mappe le cui immagini non godono di proprietà di convessità. Mentre la prima classe sarà trattata con un certo dettaglio dal momento che alcuni risultati saranno utilizzati nel seguito, per la seconda ci limiteremo ad alcuni cenni dato che esulano dai nostri interessi. Dedicheremo inoltre alcune osservazioni a parte al caso particolare e molto interessante in cui la mappa $-F$ sia monotona massimale.

2.2.1 Inclusioni differenziali a valori convessi

Se la mappa punto-insieme F è semicontinua inferiormente ed ha immagini chiuse e convesse utilizzando il teorema di selezione di Michael (vd. teorema 37) si arriva al seguente risultato, di cui forniremo anche la dimostrazione che risulta essere molto istruttiva ai fini della comprensione di quanto si andava dicendo sopra riguardo le tecniche dimostrative dei teoremi d'esistenza.

Teorema 47 *Siano $\Omega \subseteq \mathbb{R} \times \mathbb{R}^n$ un aperto non vuoto ed $F : \Omega \rightrightarrows \mathbb{R}^n$ una mappa punto-insieme semicontinua inferiormente a valori chiusi e convessi. Sia inoltre $(0, x_0) \in \Omega$. Allora esistono un intervallo $I = (\omega_-, \omega_+)$, $\omega_- < 0 < \omega_+$ ed almeno una funzione $x : I \rightarrow \mathbb{R}^n$ differenziabile con continuità su tutto I , che risulta essere soluzione del problema di Cauchy (2.5). Inoltre si ha che o $\omega_+ = +\infty$ oppure la soluzione $x(t)$ tende alla frontiera di quando $t \rightarrow \omega_+$, e analogamente avviene per ω_- .*

Dimostrazione. L'insieme $\Omega \subseteq \mathbb{R} \times \mathbb{R}^n$ è chiaramente uno spazio metrico; possiamo allora applicare il teorema 37 che garantisce l'esistenza di una selezione $f : \Omega \rightarrow \mathbb{R}^n$ continua di F . Possiamo perciò considerare il seguente problema di Cauchy per un'equazione differenziale ordinaria

$$\begin{cases} x'(t) = f(x(t)) \\ x(0) = x_0 \end{cases}$$

che in base al teorema di Peano (cfr. [36], vol.2, teor. 5.1.6) ha almeno una soluzione definita su di un intervallo massimale I . Le informazioni sul comportamento delle soluzioni ai punti estremi di questo intervallo si dimostrano mediante un classico risultato riguardante la prolungabilità di una soluzione di un problema di Cauchy per equazioni differenziali ordinarie (vd. [36], vol.2, teor.5.3.3). ■

Osservazione 48 *Si osservi che il fatto che $x(t)$ tende alla frontiera di quando $t \rightarrow \omega_+$ non implica affatto che esista $\lim_{t \rightarrow \omega_+} x(t)$. L'affermazione precedente infatti significa che*

$$\lim_{t \rightarrow \omega_+} d(x(t), \Omega^c)$$

dove la scrittura Ω^c denota il complementare insiemistico di Ω .

Tralasciamo per il momento il caso in cui F è solamente semicontinua superiormente, che richiede alcune considerazioni a parte, e passiamo ad esaminare il caso in cui si assume che F sia continua. Si ricordi inoltre che $m(F(x(t)))$ indica l'elemento di minima norma dell'insieme $F(x(t))$ (per la definizione rigorosa vd. 42)

Teorema 49 (esistenza di soluzioni lente) *Siano $\Omega \subseteq \mathbb{R} \times \mathbb{R}^n$ un aperto non vuoto ed $F : \Omega \rightrightarrows \mathbb{R}^n$ una mappa punto-insieme continua ed a valori chiusi e convessi. Sia inoltre $(0, x_0) \in \Omega$. Allora esiste un intervallo $I = (\omega_-, \omega_+)$, $\omega_- < 0 < \omega_+$ nel quale è definita una sola soluzione $x : I \rightarrow \mathbb{R}^n$ tale che*

$$x'(t) = m(F(x(t))) \quad \text{per quasi ogni } t \in I.$$

Inoltre si ha che o $\omega_+ = +\infty$ oppure la soluzione $x(t)$ tende alla frontiera di quando $t \rightarrow \omega_+$, e analogamente avviene per ω_- .

La dimostrazione di questo teorema è del tutto simile a quella del teorema 47; l'unica differenza consiste nel fatto che, al posto del teorema di Michael, si ricorre al teorema 43 che garantisce l'esistenza della selezione minimale di F e la sua continuità. Le soluzioni individuate dal teorema precedente hanno grande importanza, perciò hanno diritto ad un nome particolare; si dà infatti, la seguente definizione.

Definizione 50 *Si dice che una soluzione dell'inclusione differenziale (2.5) definita nel intervallo $[0, T]$ è una soluzione lenta se per quasi ogni $t \in [0, T]$, $x'(t) = m(F(x(t)))$.*

Osservazione 51 *Il teorema 49 non garantisce che il problema di Cauchy (2.5) abbia una sola soluzione (possono esistere altre soluzioni che non siano lente); vedremo però in seguito che se $-F$ è una mappa monotona massimale la soluzione lenta esiste ancora ed è anche l'unica soluzione del problema di Cauchy (2.5).*

0A questo punto ci resta da esaminare il caso in cui F è semicontinua superiormente. La situazione qui è più complicata a causa della mancanza di un teorema, analogo al teorema di Michael per le mappe semicontinue inferiormente, che assicuri l'esistenza di una selezione continua di F . In questo caso compare un problema tipico delle inclusioni differenziali (compare anche nel caso di mappe a valori non convessi) che non si incontra nelle equazioni differenziali ordinarie; più precisamente bisogna provare la convergenza della derivate di soluzioni approssimate, fatto che, nel caso delle equazioni differenziali ordinarie è implicato dalla convergenza delle soluzioni approssimate. Una spiegazione approfondita dei metodi che si utilizzano per superare questo problema ci portebbe troppo lontano dai nostri scopi, per cui ci limiteremo a dei semplici cenni, per ulteriori approfondimenti si rimanda perciò a [3].

Enunciamo ora un teorema di esistenza locale (analogo al teorema di Peano per le equazioni differenziali ordinarie) valido nel caso in cui F sia semicontinua superiormente. Si ricordi che una funzione ϕ si dice localmente compatta quando per ogni punto del suo dominio esiste un intorno U tale che $\phi(U)$ è compatto.

Teorema 52 *Siano X uno spazio di Hilbert, $\Omega \subset \mathbb{R} \times X$ un insieme aperto e $(0, x_0)$. Sia F una mappa punto-insieme semicontinua superiormente ed a valori chiusi e convessi. Si supponga poi che la funzione $(t, x) \mapsto m(F(t, x))$ sia localmente compatta. Allora esistono un numero reale $T > 0$ ed una funzione $t \mapsto x(t)$ assolutamente continua definita in $[0, T]$, che risulta essere soluzione del problema di Cauchy (2.5).*

Esistono diverse dimostrazioni di questo teorema; noi qui ne citeremo tre, tutte basate sulla costruzione di soluzioni approssimate e sulla dimostrazione della loro convergenza insieme a quella delle loro derivate; il punto chiave di queste dimostrazioni, come già detto sopra, consiste proprio nel provare la convergenza delle soluzioni approssimate. Si noti come ci siano delle somiglianze tra questo tipo di approccio e la dimostrazione del teorema di Peano (vd. [36], vol.2, pp.76-78) dove la convergenza delle soluzioni approssimate è provata mediante il teorema di Ascoli-Arzelà (vd. [36], vol.2, teor.5.1.7).

1. La prima dimostrazione è basata sul teorema della selezione approssimata (teorema 40) attraverso il quale si costruisce una successione di soluzioni approssimate. La convergenza di questa successione si prova mediante l'uso di un argomento di compattezza ([3], teor.4 pag.13) basato sul teorema di Ascoli-Arzelà e questo teorema insieme con il teorema di convergenza ([3], teor.1, pag.60) prova anche la convergenza delle derivate.
2. La seconda dimostrazione sfrutta ancora il teorema 40 per costruire una successione di soluzioni approssimate; la convergenza delle derivate è invece dimostrata ricorrendo ad un lemma di rappresentazione integrale ([3], lemma 1, pag.99).
3. La terza dimostrazione segue un'idea dovuta ad Eulero e costruisce una successione di soluzioni approssimate mediante delle approssimazioni poligonali della soluzione, la cui convergenza è mostrata mediante il teorema di compattezza citato al punto 1.. In questo caso la convergenza delle derivate delle soluzioni approssimate è provata ricorrendo di nuovo al teorema di convergenza.

Si può osservare come la dimensione dell'intervallo $[0, T]$ dipenda dalla grandezza di Ω e dall'intorno che la mappa minimale m manda in un compatto. Nel caso in cui $\Omega = \mathbb{R} \times X$ e $m(F(t, x))$, per ogni $(t, x) \in \mathbb{R} \times X$, rimanga in un insieme compatto K di X si può ottenere il seguente teorema di esistenza globale.

Teorema 53 *Sia $F : [0, \infty) \times X \rightrightarrows X$ una mappa punto-insieme semicontinua superiormente ed a valori chiusi e convessi. Si supponga poi che $m(F(t, x)) \in K$ per ogni $(t, x) \in [0, \infty) \times X$, dove K è un sottoinsieme compatto di X . Allora, per ogni $x_0 \in X$, esiste una soluzione $t \mapsto x(t)$ del problema di Cauchy (2.5) definita in $[0, \infty)$.*

2.2.2 Inclusioni differenziali a valori non convessi

Trattiamo ora il problema dell'esistenza per il problema di Cauchy (2.5) dove la mappa punto-insieme F è supposta essere a valori non convessi. Come già detto in precedenza, anche in questo caso, il punto chiave delle dimostrazioni dei teoremi d'esistenza risiede nel provare la convergenza delle derivate delle soluzioni approssimate. Sfortunatamente qui non si può ricorrere al teorema di convergenza, infatti nella dimostrazione di questo risultato la convessità gioca un ruolo cruciale, in quanto permette di passare dalla convergenza rispetto alla topologia debole (provata mediante un semplice argomento di compattezza) a quella rispetto alla topologia forte, che è indispensabile per provare che il limite delle soluzioni approssimate è proprio una soluzione del problema di Cauchy. Questo fatto obbliga, nel caso di inclusioni differenziali in cui la mappa a secondo membro abbia immagini non convesse, ad utilizzare argomenti di compattezza direttamente nella topologia forte per provare la convergenza delle derivate ed in generale questo porta ad argomentazioni più complicate rispetto al caso di ipotesi di convessità. Vedremo infatti che l'ipotesi di convessità delle immagini sarà sostituita, in questo contesto, dalla compattezza delle stesse.

Teorema 54 *Siano $\Omega \subseteq \mathbb{R} \times \mathbb{R}$ un aperto contenente $(0, x_0)$ ed $F : \Omega \rightrightarrows \mathbb{R}$ una mappa punto-insieme a valori compatti. Allora esistono un numero reale $T > 0$ ed una funzione $t \mapsto x(t)$ assolutamente continua definita in $[0, T]$, che risulta essere soluzione del problema di Cauchy (2.5).*

La dimostrazione di questo teorema si basa su di un risultato di compattezza ambientato nello spazio delle funzioni limitate (cfr.[3], teor.5, pag.15).

Nel caso in cui si assumano proprietà di regolarità più forti per la mappa F si giunge a risultati che garantiscono l'esistenza di soluzioni di classe C^1 (cfr. [3], cap.2 sez. 3-4). Ricordiamo a titolo d'esempio, il seguente risultato riguardante inclusioni differenziali autonome con mappe lipschitziane.

Teorema 55 *Siano $\Omega \subseteq \mathbb{R}^n$ un aperto ed $F : \Omega \rightrightarrows \mathbb{R}^n$ una mappa punto-insieme a valori chiusi e lipschitziana con costante k ; siano fissati $x_0 \in \Omega$ e $v_0 \in F(x_0)$; inoltre sia $T > 0$ tale che*

$$x_0 + \left[\|v_0\| \frac{e^{kT} - 1}{k} \right] B_1(0) \subset \Omega.$$

Allora esiste una soluzione $t \mapsto x(t)$ di classe C^1 definita in $(-T, +T)$ del seguente problema

$$\begin{cases} x'(t) \in F(x(t)) \\ x(0) = x_0 \\ x'(0) = v_0 \end{cases} .$$

Si osservi come la lipschitzianità di F permetta di evitare la richiesta di compattezza delle immagini di F e come si consideri una versione modificata del problema di Cauchy.

Le mappe lipschitziane sono altresì molto importanti perchè permettono di legare inclusioni differenziali a valori non convessi a inclusioni differenziali a valori convessi. Si considerino infatti le due seguenti inclusioni differenziali

$$\begin{aligned}x'(t) &\in F(x(t)), \\x'(t) &\in \overline{\text{co}}F(x(t))\end{aligned}$$

dove F è una mappa punto-insieme a valori non convessi e $\overline{\text{co}}(F(x))$ indica la chiusura convessa di $F(x)$. Naturalmente le soluzioni della prima inclusione sono soluzioni anche della seconda; nasce quindi il problema di capire quali nuove soluzioni introduce la convessificazione del secondo membro ed in particolare ci si chiede quali siano le condizioni che devono valere affinché l'insieme delle soluzioni della prima inclusione sia denso nell'insieme delle soluzioni dell'inclusione convessificata. Poichè le soluzioni di quest'ultima inclusioni sono spesso chiamate *soluzioni rilassate* di $x'(t) \in F(x(t))$, questo problema prende il nome di *problema del rilassamento*; a tale riguardo vale il seguente importante teorema.

Teorema 56 (di rilassamento) *Si consideri il seguente sottoinsieme di \mathbb{R}^n*

$$Q := \{x \in \mathbb{R}^n : \|x - x_0\| \leq b\}$$

e sia $F : Q \rightrightarrows \mathbb{R}$ una mappa punto-insieme lipschitziana e a valori compatti. Fissato un intervallo $I = [-T, +T]$, sia $x : I \rightarrow Q$ una soluzione di

$$\begin{cases} x'(t) \in \overline{\text{co}}(F(x(t))) \\ x(0) = x_0 \end{cases}$$

tale che $\|x(t) - x_0\| < b$ per ogni $t \in I$. Allora per ogni $\varepsilon > 0$, esiste $y : I \rightarrow Q$ soluzione di

$$\begin{cases} x'(t) \in F(x(t)) \\ x(0) = x_0 \end{cases}$$

tale che $\|y(t) - x(t)\| \leq \varepsilon$ per ogni $t \in I$.

Concludiamo questa brevissima rassegna riguardante le inclusioni differenziali con mappe a valori non convessi citando un differente metodo di risoluzione del problema dell'esistenza per inclusioni differenziali a valori non convessi, che lo rilegge come un problema di punto fisso (in analogia al caso delle equazioni differenziali ordinarie) ed è basato su di un approccio al problema della selezione di mappe punto-insieme dovuto ad Antosiewicz e Cellina. Questi risultati, riguardanti l'esistenza di selezioni, costituiscono un analogo del teorema di Michael, nel caso di mappe a valori non convessi (e non necessariamente semicontinue inferiormente). Per questo approccio si rimanda ancora una volta a [3] (cap.2, sez.5-6).

2.2.3 Mappe monotone massimali

A conclusione della sezione dedicata ai teoremi d'esistenza dedichiamo un piccolo spazio ad una classe di inclusioni differenziali particolare per cui è possibile formulare un risultato, oltre che di esistenza, anche di unicità per le soluzioni del problema di Cauchy. Per far questo dobbiamo introdurre una nuova classe di mappe punto-insieme. In questo ambito X denoterà sempre uno spazio di Hilbert.

Definizione 57 Una mappa punto insieme $M : X \rightrightarrows X$ si dice monotona quando

$$\forall x_1, x_2 \in \text{Dom}(M), v_i M(x_i) \text{ per } i = 1, 2 \quad \langle v_1 - v_2, x_1 - x_2 \rangle \geq 0. \quad (2.6)$$

Una mappa punto-insieme M si dice poi monotona massimale quando M è monotona ed inoltre non esistono mappe-punto-insieme \widetilde{M} monotone tali che

$$\text{Graph}(M) \subset \text{Graph}(\widetilde{M}).$$

Nel caso in cui $X = \mathbb{R}$ ed M sia un'usuale funzione f , la nozione di monotonia vista sopra coincide con la monotonia standard per funzioni reali di variabile reale; cioè $x_1 \leq x_2$ implica $f(x_1) \leq f(x_2)$; nel caso in cui X abbia dimensione maggiore di 1, anche nel caso di funzioni, questa definizione ha poco a che vedere con nozioni di monotonia rispetto ad ordini parziali dati in X , ma è piuttosto una generalizzazione del concetto di operatore lineare semidefinito positivo, infatti la (2.6) è equivalente a

$$\langle Hx, x \rangle \geq 0 \quad \forall x \in X$$

nel caso in cui $M(x) = \{Hx\}$ ed H è lineare su tutto X .

E' possibile dimostrare che una mappa punto-insieme M monotona massimale gode di buone proprietà. Le sue immagini sono infatti chiuse e convesse, tuttavia la mappa M non è detto sia necessariamente semicontinua superiormente. Questa mancanza di continuità può però essere superata mediante una proprietà di approssimazione, analoga a quella formulata nel teorema 40 della selezione approssimata valido per le mappe semicontinue superiormente, nota con il nome di *approssimazione di Yosida* (vd. [3], pp. 144-146) che fornisce una approssimazione (in un senso opportuno), mediante funzioni lipschitziane, della mappa M . In base a questo risultato è allora possibile ricondursi a considerare una successione di equazioni differenziali ordinarie, le cui soluzioni si provano poi essere convergenti ad una soluzione dell'inclusione differenziale $x'(t) \in -M(x(t))$. Vale infatti il seguente importante teorema.

Teorema 58 *Sia $M : X \rightrightarrows X$ una mappa punto-insieme monotona massimale. Si consideri il problema di Cauchy*

$$\begin{cases} x'(t) \in -M(x(t)) \\ x(0) = x_0 \end{cases} \quad (2.7)$$

dove $x_0 \in \text{Dom}(M)$. Allora esiste una ed una sola soluzione $t \mapsto x(t)$ definita in $[0, \infty)$, che risulta essere la soluzione lenta.

Inoltre:

1. $t \mapsto x'(t)$ è decrescente;
2. siano $t \mapsto x(t)$ e $t \mapsto y(t)$ due soluzioni di (2.7) uscenti rispettivamente da x_0 ed y_0 , allora per $t \in [0, \infty)$ si ha

$$\|x(t) - y(t)\| \leq \|x_0 - y_0\|;$$

3. per ogni $t \in [0, \infty)$ si ha che

$$x'(t) = \lim_{h \rightarrow 0^+} \frac{x(t+h) - x(t)}{h}$$

e la funzione $t \mapsto x'(t)$ è continua da destra.

Consideriamo infine una particolare classe di inclusioni differenziali con mappe monotone massimali.

Osservazione 59 (inclusioni gradiente) *Si consideri il sottodifferenziale $\partial f(x)$ di una funzione f convessa propria e semicontinua inferiormente; si dimostra allora che la mappa punto insieme definita da*

$$x \mapsto \partial f(x)$$

è massimale monotona. Allora il seguente problema di Cauchy per l'inclusione differenziale gradiente

$$\begin{cases} x'(t) \in -\partial f(x(t)) \\ x(0) = x_0 \end{cases}$$

ha un'unica soluzione $t \mapsto x(t)$; si dimostra poi che la funzione $t \mapsto V(x(t))$ è convessa e decrescente. Inoltre se f raggiunge il suo valore minimo in qualche punto, la traiettoria $x(t)$ converge debolmente, per $t \rightarrow \infty$, ad un punto che minimizza f . (Per la definizione di sottodifferenziale e le sue proprietà vedi ad esempio [38] oppure [54]).

Capitolo 3

Approcci “dinamici” ai problemi vettoriali

Lo scopo di questo capitolo consiste nel delineare i tratti principali di alcuni lavori, per la verità relativamente poco numerosi anche se a mio giudizio particolarmente interessanti, che affrontano le tematiche dell’ottimizzazione vettoriale seguendo metodi e modalità eterodosi rispetto a quelli che appaiono nella maggior parte della letteratura sviluppatasi in questo campo. A questo filone di ricerche appartiene la parte completamente originale di questa tesi, che sarà esposta nel capitolo successivo. In particolare questi lavori, pur molto diversi fra loro, sono caratterizzati da alcuni elementi comuni che appartengono principalmente da almeno uno dei due seguenti punti fondamentali:

- attenzione agli aspetti dinamici del problema di ottimizzazione vettoriale e conseguente utilizzo di idee e strumenti tipici della teoria dei sistemi dinamici;
- Utilizzo di tecniche e risultati della teoria delle singolarità (o dei punti critici) di mappe e funzioni.

Nel corso dell’esposizione, facendo ricorso anche agli aspetti cronologici con cui queste tematiche si sono sviluppate, si vedrà come questo tipo di approccio ai problemi di ottimizzazione vettoriale sia riaffiorato diverse volte nel corso degli anni, ad opera anche di studiosi di grande rilievo; tuttavia questi articoli sono rimasti, almeno numericamente, minoritari all’interno della vasta produzione dedicata alle tematiche dell’ottimizzazione vettoriale o multiobiettivo. Senza alcuna pretesa di essere esauriente e in modo non approfondito, vorrei però accennare ad una breve e personale spiegazione di questo fatto. La motivazione principale per cui questo tipo di ricerche sono rimaste a carattere sporadico risiede, secondo me, nel fatto che le competenze e le conoscenze di carattere matematico richieste sono di natura assolutamente trasversale, dal momento che spaziano dalla teoria dell’ottimizzazione vettoriale alla teoria dei sistemi dinamici, delle inclusioni

differenziali, delle singolarità e ad alcuni aspetti della topologia differenziale. Mi sembra tuttavia che questo tipo di approccio indaghi in profondità la natura dei problemi connessi all'ottimizzazione di funzioni a valori vettoriali e possa portare a sviluppi significativi sia dal punto di vista teorico sia da quello applicativo.

3.1 I lavori di Smale

Nella prima metà degli anni settanta, Steve Smale, uno dei più importanti matematici viventi (vincitore della Medaglia Fields nel 1966 per la dimostrazione della congettura di Poincaré in dimensione maggiore uguale a cinque) pubblicò sul *Journal of Mathematical Economics* una serie di articoli nel campo dell'economia matematica e, più precisamente, riguardanti la teoria dell'equilibrio economico il cui titolo "collettivo" ("Global Analysis and Economics") è particolarmente significativo (cfr. [58],[59],[60],[61] e [64]). I contenuti di questi articoli sono poi sintetizzati dallo stesso Smale nel suo contributo all'opera "Handbook of Mathematical Economics" ([66]). In questi lavori Smale utilizzava alcune tecniche matematiche tipiche dell'analisi globale per provare, o meglio per ridimostrare in modo alternativo e con ipotesi diverse, alcuni fondamentali risultati nell'ambito della teoria economica; in particolare si proponeva di sostituire le usuali ipotesi di convessità con richieste di differenziabilità. Curiosamente si constata che la serie degli articoli apparsi sul *Journal of Mathematical Economics* incomincia con il numero II; in effetti qualche tempo prima era uscito un articolo dal titolo "Global Analysis and Economics I: Pareto Optimum and a Generalization of Morse Theory" ([57]) in cui ci si proponeva di investigare attraverso tecniche matematiche avanzate il problema di determinare i punti Pareto ottimali, ovvero di studiare il problema di ottimizzare più funzioni scalari simultaneamente in relazione alla nozione di punti Pareto ottimali introdotta nell'ambito della teoria economica che, come sottolineato nel lavoro, costituiva una delle principali motivazioni di questo studio. Proprio questo articolo costituisce il primo tentativo di affrontare un problema di ottimizzazione vettoriale (limitato al caso paretiano) mediante un approccio che si ricollega ai sistemi dinamici, infatti utilizzando le parole di Smale stesso "we obtain a natural notion of dynamics in this setting to generalize that of a gradient flow" (vd. [57]). E' interessante sottolineare che questo lavoro ha caratteristiche abbastanza particolari in quanto segnala interessanti problemi aperti e suggerisce linee d'attacco a queste questioni molto originali ed innovative (soprattutto nelle sezioni 5 e 6), tuttavia alcuni risultati in esso contenuti sono solo abbozzati e non vengono dimostrati rigorosamente. Oltre a questo lavoro Smale pubblicò poi altri due contributi dedicati alle tematiche dell'ottimizzazione vettoriale ([63], [62]) contraddistinti anch'essi da un approccio dinamico a questi problemi. Ai fini della nostra tesi risulta quindi chiaro che gli articoli di Smale di maggior importanza tra quelli citati sono proprio quelli dedicati direttamente

ai problemi di ottimizzazione vettoriale e cioè [57], [63] e [62]; in questa sezione intendiamo perciò esporre e commentare i punti più significativi di questi lavori.

Nel seguito si utilizzeranno le nozioni di varietà (differenziabile), di spazio tangente e cotangente e di differenziale di una mappa; per le loro definizioni e per una sintetica rassegna delle principali proprietà riguardanti questi concetti si rimanda a [49]; per una trattazione più ampia si può consultare, ad esempio, il volume [5].

Negli articoli citati si considerano un insieme W aperto dello spazio euclideo \mathbb{R}^n o, più in generale, una varietà differenziabile W e m funzioni lisce (cioè di classe C^∞ in W) $u_i : W \rightarrow \mathbb{R}$, $i = 1, \dots, m$ che abbiano tutte come dominio proprio l'insieme W , si suppone inoltre che $\dim W \geq m$. Si osservi, per inciso, che Smale sottolinea come queste assunzioni di carattere prettamente matematico modellizzino una tipica situazione economica: W rappresenta lo spazio degli stati di un'economia mentre le funzioni u_i si possono riguardare come le funzioni d'utilità di m consumatori che agiscono nell'ambito dell'economia considerata, infine anche l'ipotesi sulla dimensione di W è giustificata da un punto di vista economico (si veda [57], pp. 532-533). A questo punto si passa a considerare l'insieme delle soluzioni del seguente problema di ottimizzazione vettoriale

$$\max_{x \in W} u(x)$$

dove la funzione obiettivo è data da $u := (u_1, \dots, u_m) : W \rightarrow \mathbb{R}^m$ e l'ordine nello spazio \mathbb{R}^m è il tradizionale ordine paretiano associato all'ortante \mathbb{R}_+^m . Naturalmente Smale non usa queste stesse notazioni, che si sono affermate in letteratura più tardi, e chiama gli elementi dell'insieme $S\text{Max}(u, W)$, da noi definito nella sezione 1.2, *punti Pareto classici* per la funzione u in W . Il passo successivo consiste nel definire un insieme θ più grande di quello dei punti Pareto classici e che sia individuato da condizione di natura differenziale, che prende il nome di *insieme Pareto critico*. Lo scopo che si prefigge Smale in questi lavori (in particolare in [57] ed in [63]) è quello di studiare la struttura di questo insieme inizialmente con gli strumenti del calcolo differenziale, in analogia a quanto si fa nel caso di una funzione scalare per l'insieme dei suoi punti stazionari (o critici), ed a un livello più avanzato con gli strumenti, allora recentissimi, nati nell'ambito della teoria delle singularità di mappe (per questa interessante teoria si rimanda al volume [26]). Vediamo ora in dettaglio quali considerazioni portano alla costruzione dell'insieme θ sopra definito seguendo quanto fatto nei lavori citati.

Per prima cosa consideriamo una curva definita su di un intervallo reale, cioè una applicazione $\varphi : [a, b] \rightarrow W$, e diciamo che una curva siffatta è ammissibile quando

$$\frac{d}{dt} u_i(\varphi(t)) > 0 \quad \forall t \in (a, b), \forall i = 1, \dots, m.$$

E' immediato osservare che lungo curve ammissibili si incrementano i valori assunti da ogni singola componente della funzione u e quindi queste curve non danno origine ad alcun fenomeno di ricorrenza, cioè $\varphi(x_1) \neq \varphi(x_2)$ per ogni $x_1, x_2 \in [a, b]$.

Nell'ambientazione economica prima accennata le curve ammissibili possono essere pensate come sequenze di scambi tra gli agenti presenti nell'economia. A questo punto è possibile fornire una prima definizione dell'insieme Pareto critico θ : un punto x di W appartiene a θ quando non esistono curve ammissibili che passano per x . Nel caso $m = 1$ l'insieme Pareto critico per la funzione u coincide con l'insieme dei punti stazionari per la funzione u .

Le curve ammissibili possono essere meglio caratterizzate facendo ricorso al differenziale $Du(x)$ della funzione u in un punto x di W , che altro non è che una applicazione lineare tra lo spazio tangente $T_x(W)$ a W in x e lo spazio tangente a \mathbb{R}^m in $u(x)$ che coincide con \mathbb{R}^m stesso.

Definizione 60 Una curva $\varphi : [a, b] \rightarrow W$ si dice ammissibile quando

$$Du(\varphi(t))\varphi'(t) \in \text{int}\mathbb{R}_+^m \text{ per ogni } t \in (a, b)$$

dove $\varphi'(t) = \frac{d}{dt}\varphi(t) \in T_{\varphi(t)}(W)$ indica il vettore tangente in $\varphi(t)$ alla curva φ .

L'introduzione delle curve ammissibili porta quindi ad un elemento dinamico nello studio del problema di ottimizzazione vettoriale perchè diventa fondamentale studiare il comportamento di u quando ci si muove proprio lungo le curve ammissibili stesse.

Per poter fornire una definizione più precisa dell'insieme Pareto critico dobbiamo però far ricorso ad una descrizione delle curve ammissibili di natura più marcatamente geometrica. Per ogni $x \in W$ sia C_x il cono, contenuto nello spazio tangente $T_x(W)$, definito da

$$C_x := \{v \in T_x(W) : Du(x)v \in \text{int}\mathbb{R}_+^m\}.$$

Si osservi che $C_x = Du(x)^{-1}(\text{int}\mathbb{R}_+^m)$, perciò C_x risulta essere un cono aperto e convesso (eventualmente vuoto) in virtù della linearità di $Du(x)$; vale inoltre la seguente proposizione che evidenzia proprietà di continuità della mappa punto insieme che associa ad ogni $x \in W$ l'insieme C_x .

Proposizione 61 (vd. [57]) Siano $x_0 \in W$ e $v : W \rightarrow T(W)$ un campo vettoriale continuo tale che $v(x_0) \in C_{x_0}$. Allora esiste un intorno U di x_0 in W tale che $v(x) \in C_x$ per ogni $x \in U$.

Una volta introdotto il cono C_x , le curve ammissibili possono essere ridefinite nel modo seguente: una curva $\varphi : [a, b] \rightarrow W$ è ammissibile quando soddisfa l'inclusione

$$\varphi'(t) \in C_{\varphi(t)}$$

che in linea di principio identifica un sistema dinamico. Possiamo allora dare la seguente definizione che meglio precisa l'idea di insieme Pareto critico.

Definizione 62 *L'insieme Pareto critico per la funzione u è l'insieme*

$$\theta := \{x \in W : C_x = \emptyset\}.$$

E' immediato osservare che θ coincide con l'insieme

$$\{x \in W : Du(x) [T_x(W)] \cap \text{int}(\mathbb{R}_+^m) = \emptyset\}.$$

Smale osserva che ogni soluzione del problema di ottimizzazione vettoriale considerato, ovvero ogni punto Pareto-massimale di u in W è un elemento di θ mentre il viceversa non vale (per le dimostrazioni rigorose di queste asserzioni si rimanda all'ultimo capitolo della presente tesi, in quanto nei lavori citati esse sono omesse); inoltre sottolinea come, mediante l'uso di tecniche di calcolo differenziale ed analisi globale, l'insieme θ sia trattabile più facilmente rispetto all'insieme dei punti Pareto massimali. Nell'articolo [63], viene infatti enunciata una proposizione che caratterizza i punti appartenenti a θ attraverso l'uso dei differenziali delle singole componenti della funzione u .

Proposizione 63 ("first order proposition") *Siano $x \in W$ e $u : W \rightarrow \mathbb{R}^m$. Allora $x \in \theta$ se e solo se*

1. *i vettori $Du_i(x)$ non sono tutti contenuti in un medesimo semispazio aperto dello spazio cotangente $T_x^*(W)$;*
2. *$\exists \lambda_i \geq 0, i = 1, \dots, m$, non tutti nulli, tali che*

$$\sum_{i=1}^m \lambda_i Du_i(x) = 0.$$

Osservazione 64 *Nell'articolo citato la dimostrazione di questa proposizione è soltanto delineata a grandi linee (per una dimostrazione dettagliata si rimanda alla proposizione 72 dell'ultimo capitolo), inoltre si può osservare che la proprietà 1. è implicata dalla proprietà 2., ed è quindi, in un certo senso, ridondante. Per dimostrare questa osservazione ricordiamo il seguente fatto: dati i vettori $v_1, \dots, v_m \in V$ spazio vettoriale di dimensione n , essi appartengono ad un medesimo semispazio aperto di V se esiste un'applicazione lineare $\ell : V \rightarrow \mathbb{R}$ tale che $\ell(v_i) > 0$ per ogni $i = 1, \dots, m$. Valga ora la proprietà 2. della proposizione precedente e si supponga per assurdo che i vettori $Du_1(x), \dots, Du_m(x)$ appartengano tutti ad un medesimo sottospazio aperto di $T(W)$ (che è uno spazio vettoriale di dimensione n), allora, considerando gli scalari $\lambda_i \geq 0$ e non tutti nulli, la cui esistenza è garantita da 2., si ha che*

$$\sum_{i=1}^m \lambda_i \ell(Du_i(x)) > \ell(0) = 0.$$

Quest'ultima relazione costituisce un assurdo contro la validità del punto 2. della proposizione precedente.

Nei tre lavori di Smale presi in considerazione, oltre a quanto esposto sin qui, vengono affrontate anche altre tematiche, naturalmente sempre connesse al problema di ottimizzare la funzione u rispetto all'ordine paretiano in \mathbb{R}^m , che schematicamente possono essere suddivise in tre aree distinte.

- La prima area è incentrata principalmente sulla ricerca di condizioni necessarie e sufficienti affinché un punto sia un punto Pareto massimale (locale) ed è sviluppata principalmente in [62]. In questo filone di ricerca si inquadra anche il lavoro [69]. Il risultato principale in quest'area è costituito da un teorema in cui, accanto alla condizione necessaria del prim'ordine già nota in letteratura, viene enunciata una condizione sufficiente del second'ordine.

In questa sede riportiamo questo risultato seguendo [62] e ci restringiamo pertanto a considerare il caso di W aperto di \mathbb{R}^n ; per il caso più generale in cui W è una generica varietà si rimanda a [69]. Introduciamo poi la seguente notazione: $D^2u_i(x)$ indica l'Hessiana delle singole funzioni scalari u_i , pensata come forma bilineare simmetrica a valori reali. Si osservi inoltre che, essendo W un aperto di \mathbb{R}^n , vale, per ogni $v = (v, \dots, v) \in \mathbb{R}^n$, la seguente relazione

$$\langle v, \nabla u_i(x) \rangle = \sum_{j=1}^n v^j \frac{\partial u_i}{\partial x^j}(x) = Du_i(x)v$$

che lega il differenziale $Du_i(x)$ costituito da una applicazione lineare da \mathbb{R}^n in \mathbb{R} , al gradiente di u_i .

Teorema 65 (vd. [62]) *Sia $x \in W$ un punto Pareto massimale locale per la funzione $u : W \rightarrow \mathbb{R}^m$. Allora esistono $\lambda_1, \dots, \lambda_m \geq 0$ e non tutti nulli, tali che*

$$\sum_{i=1}^m \lambda_i Du_i(x) = 0$$

Siano ora fissati $\lambda_1, \dots, \lambda_m \geq 0$ non tutti nulli ed $x \in W$ in modo tale che sia soddisfatta la relazione precedente. Se la forma bilineare simmetrica così definita

$$\sum_{i=1}^m \lambda_i D^2u_i(x)$$

risulta definita negativa sul sottospazio vettoriale

$$\{v \in \mathbb{R}^n : \langle v, \lambda_i \nabla u_i(x) \rangle = 0, \forall i = 1, \dots, m\}$$

allora x è un punto Pareto massimale per u .

Da questo teorema risulta subito evidente che la condizione del prim'ordine che caratterizza l'insieme Pareto critico è anche condizione necessaria affinché un punto sia Pareto massimale; fatto questo ben noto e strettamente connesso alla scalarizzazione lineare di un problema di ottimizzazione vettoriale (cfr. 1.3.1). La condizione sufficiente del second'ordine invece presenta alcuni aspetti innovativi che, nell'enunciato da noi considerato, non appaiono del tutto evidenti, ma che lo diventano una volta che si consideri il caso più generale studiato da Wan (cfr. [69]) oppure l'articolo [63]; essi sono legati alla definizione intrinseca di una "Hessiana" per la funzione $u : W \rightarrow \mathbb{R}^m$ che sfrutta idee e metodi della teoria delle singolarità, in quegli anni appena formulati.

Infine segnaliamo che riguardo quest'area di ricerca bisogna citare anche il lavoro [61] appartenente alla serie "Global Analysis and Economics", che affronta il problema di determinare condizioni necessarie e sufficienti per il problema di ottimizzazione vettoriale in presenza di vincoli della forma $g_j(x) \geq 0$ con $g_j : W \rightarrow \mathbb{R}$ ($j = 1, \dots, l$); anche in questo caso vengono individuate condizioni necessarie del prim'ordine e condizioni sufficienti del second'ordine. Sempre in questo ambito tematico, ricordiamo poi l'articolo [42], che contiene, tra l'altro, risultati che forniscono condizioni necessarie e sufficienti affinché un punto sia punto massimale (locale) per una funzione $u : W \rightarrow \mathbb{R}^m$ (dove W è una varietà) rispetto ad un ordine in \mathbb{R}^m indotto da un cono chiuso, convesso e con interno non vuoto generico; in questo lavoro sono anche accennate condizioni di ordine superiori al secondo giungendo alla conclusione che esse non hanno molto significato e rilevanza.

- La seconda area approfondisce gli studi riguardanti la struttura dell'insieme Pareto critico θ , definito in precedenza, utilizzando strumenti che, all'epoca di pubblicazione dei contributi considerati, erano recentissimi; in particolare vengono delineati metodi di studio della struttura di questo insieme attraverso opportune modificazioni delle tecniche della teoria delle singolarità di mappe differenziabili. Entrando, seppur sinteticamente, nei dettagli si osserva che nell'articolo [57] si giunge a formulare l'ipotesi che se valgono opportune condizioni di trasversalità, ottenute a partire da condizioni di trasversalità di R. Thom, l'insieme θ possieda la struttura di sottovarietà della varietà W (per la definizione di sottovarietà si veda [5]); vengono inoltre studiate le proprietà della frontiera di θ . Sempre nel medesimo lavoro, nella sezione 6, viene proposta una possibile estensione della teoria di Morse al caso di funzioni a valori vettoriali e viene formulato un teorema che contiene una disuguaglianza di Morse generalizzata. Questi risultati, per quanto interessanti e affascinanti, sono solo enunciati e mancano di una vera dimostrazione. La struttura dell'insieme Pareto critico viene studiata anche in [59] ed in [64], dove questo studio è direttamente finalizzato ad applicazioni alla teoria economica (prezzi d'equilibrio, teorema

fondamentale dell'economia del benessere); in quest'ottica segnaliamo anche l'articolo [56]. Infine ricordiamo che vengono evidenziati legami molto forti tra le diverse classi di singolarità per mappe differenziabili ed i punti Pareto-ottimali (si veda [43]).

In conclusione si può affermare questo indirizzo di ricerca evidenzia molti ed interessanti legami tra l'ottimizzazione vettoriale e la teoria delle singolarità, sviluppatasi negli anni '50 e '60 nell'ambito della topologia differenziale; si nota però come i risultati concernenti questi aspetti siano spesso rimasti ad un livello poco più avanzato di quello di semplici congetture. A mio modo di vedere potrebbe essere molto proficuo, dal punto di vista della teoria dell'ottimizzazione vettoriale, proseguire ricerche di questo tipo basandosi, invece che sulla teoria della singolarità di mappe differenziabili, sui più recenti sviluppi della teoria dei punti critici per funzionali scalari, che tra l'altro gettano una nuova luce anche su tecniche più vecchie. Naturalmente questi strumenti dovranno essere opportunamente modificati tenendo presenti le specificità del caso vettoriale, ad esempio come fatto nell'ultimo capitolo di questa tesi.

- Nella terza area di ricerca viene affrontato il problema della stabilità dei punti dell'insieme θ , in analogia con quanto usualmente si fa nel caso dei sistemi dinamici, ed in particolare nel caso dei sistemi definiti dal gradiente di una funzione scalare, per i quali esistono definizioni standard di stabilità dovute principalmente a Liapunov (cfr. [30], [29] e [72]). Lo studio delle possibili nozioni di stabilità risulta di grande importanza dal punto di vista teorico; la sua rilevanza tuttavia è notevole anche dal punto di vista applicativo, perchè è evidente l'importanza di stabilire quando un punto critico, che rappresenta un equilibrio del sistema in esame, non è influenzato da perturbazioni delle variabili indipendenti, almeno nel caso in cui queste siano di piccola entità.

Nei pionieristici lavori di Smale il problema di determinare un sottoinsieme θ_S dell'insieme Pareto critico θ composto da punti in un qualche senso "stabili", viene subito visto come uno dei punti fondamentali del approccio ai problemi di ottimizzazione vettoriale proposto. Questo fatto è motivato da considerazioni economiche, oltre che da motivazioni interne al discorso puramente matematico; infatti Smale asserisce che punti di θ che non siano stabili hanno soltanto un significato indiretto in economia. Dobbiamo però osservare che la trattazione del problema della stabilità di un punto dell'insieme Pareto critico è ben lungi dall'essere conclusa in questi lavori; permangono infatti molti punti oscuri e non rigorosamente dimostrati e non c'è neppure accordo tra i vari lavori su di un'unica definizione di stabilità.

Nell'articolo [57] si procede nel modo seguente: supponiamo che $\varphi : [a, b] \rightarrow$

W sia una curva ammissibile e che

$$\lim_{t \rightarrow b} \varphi(t) = x;$$

si dice allora che φ parte da $\varphi(a) = w$ e termina in x . Possiamo allora formulare la seguente definizione di punto stabile per la funzione u .

Definizione 66 *Un punto $x \in \theta$ si dice stabile (nel primo senso) o $x \in \theta_S^1$, quando per ogni intorno U di x in θ esiste un intorno V di x in W tale che, se φ è una curva ammissibile che parte da un qualunque punto di V e termina in θ , allora φ termina in U .*

In [57] viene poi formulata una condizione sufficiente del second'ordine affinché un punto appartenga a θ_S^1 , basata su di una opportuna definizione di "Hessiana" della funzione a valori vettoriali u , già introdotta nello studio delle condizioni di ottimalità. Nel medesimo lavoro si evidenzia (senza però fornirne la dimostrazione) anche come, sotto opportune ipotesi di convessità e monotonia riguardanti la funzione u , $\theta_S^1 = \theta$. Segnaliamo infine che, successivamente, questa prima definizione di stabilità è stata estesa anche al caso generale di un ordine nello spazio immagine \mathbb{R}^m indotto da un arbitrario cono puntato, chiuso e convesso, con interno non vuoto (cfr. [42]), anche in questo caso viene poi individuata una condizione sufficiente affinché un punto critico sia stabile ed ancora una volta la dimostrazione di questo risultato è basata su elementi di teoria delle singolarità di mappe.

Nel lavoro [63] viene invece fornita una definizione di punto stabile parzialmente differente.

Definizione 67 *Un punto $x \in \theta$ si dice stabile (nel secondo senso) o $x \in \theta_S^2$, quando per ogni intorno U di x in W esiste un intorno V di x in U tale che ogni curva ammissibile $\varphi : [a, b) \rightarrow W$ con $\varphi(a) \in V$ soddisfi la relazione $\varphi([a, b)) \subset U$.*

In altre parole x è un punto stabile quando ogni curva ammissibile con punto di partenza vicino ad x , resta vicina ad x . Anche per questa definizione viene poi formulata una condizione del second'ordine sufficiente affinché un punto di θ sia stabile in questo secondo senso. Smale incorre inizialmente in un errore (cfr. [63], sec.3) affermando che ogni punto Pareto massimale per u in W , cioè ogni elemento di $S\text{Max}(u, W)$ secondo le nostre notazioni, risulta essere stabile nel secondo senso; questa affermazione viene però smentita in [59], sec.3, dove compare un controesempio che mette in evidenza come esistano punti Pareto-massimali che non sono stabili. L'esempio citato risulta però abbastanza complicato e, tutto sommato artificioso. Consideriamo pertanto i seguenti esempi originali in cui si fornisce un esempio di punto Pareto massimale che non è stabile in nessuno dei due sensi trattati.

Esempio 68 Sia \mathbb{R}^2 dotato dell'ordine paretiano. Consideriamo allora la funzione così definita

$$u : \mathbb{R}^2 \longrightarrow \mathbb{R}^2 \\ (x^1, x^2) \longmapsto (-(x^1)^2, -(x^1)^2),$$

allora

$$Du((x^1, x^2)) = \begin{bmatrix} -2x^1 & 0 \\ -2x^1 & 0 \end{bmatrix}.$$

E' inoltre agevole constatare che

$$u(\mathbb{R}^2) = \{(y^1, y^2) \in \mathbb{R}^2 : y^1 = y^2, y \leq 0\}$$

si osserva allora immediatamente che

$$\{(0, 0)\} = \text{Max}(u(\mathbb{R}^2), \mathbb{R}_+^2) = \text{IMax}(u(\mathbb{R}^2), \mathbb{R}_+^2)$$

e quindi

$$\{(x^1, x^2) \in \mathbb{R}^2 : x^1 = 0\} = \text{SMax}(u, \mathbb{R}^2) = \text{SIMax}(u, \mathbb{R}^2);$$

inoltre

$$\{(x^1, x^2) \in \mathbb{R}^2 : x^1 = 0\} = \theta.$$

Mostriamo ora che nessun punto di θ risulta essere stabile nei due sensi definiti precedentemente. Possiamo, senza perdere di generalità, considerare il solo punto $(0, 0)$; infatti in ogni altro punto di θ si possono ripetere i ragionamenti fatti in questo caso, a patto di effettuare un'opportuna traslazione che non ha nessuna influenza sul problema considerato.

Incominciamo con il mostrare che $(0, 0)$ non è stabile secondo la definizione 66. Sia $r > 0$ allora $U = B_r((0, 0))$ è un intorno di $(0, 0)$ in θ . Sia $\varepsilon > 0$, consideriamo allora la curva $\varphi_1 : [0, 1) \rightarrow \mathbb{R}^2$ così definita

$$\varphi_1(t) := (\varepsilon(1-t), -(r+1)t).$$

Poichè

$$\varphi_1'(t) := (-\varepsilon, -(r+1))$$

si ottiene che

$$\begin{aligned} Du(\varphi_1(t))\varphi_1'(t) &= \begin{bmatrix} -2\varepsilon(1-t) & 0 \\ -2\varepsilon(1-t) & 0 \end{bmatrix} \begin{bmatrix} -\varepsilon \\ -(r+1) \end{bmatrix} \\ &= \begin{bmatrix} 2\varepsilon^2(1-t) \\ 2\varepsilon^2(1-t) \end{bmatrix} \in \text{int}\mathbb{R}_+^2 \quad \forall t \in [0, 1). \end{aligned}$$

Dunque φ_1 è una curva ammissibile in base alla definizione 60, inoltre parte da $\varphi_1(0) = (\varepsilon, 0)$ e termina in $(0, -(r+1)) \in \theta \setminus U$. Si osservi ora che, a

patto di scegliere opportunamente ε , comunque si scelga un intorno V di $(0, 0)$ in \mathbb{R}^2 si ha che $(\varepsilon, 0) \in V$. In conclusione, dato U , non è possibile trovare un intorno V tale che ogni curva ammissibile che parte in V e termina in θ , termini anche in U .

Passiamo ora a provare che il punto $(0, 0)$ non è stabile neppure nel secondo senso definito in precedenza. Sia $\varepsilon > 0$, consideriamo allora la curva $\varphi_2 : [0, 1) \rightarrow \mathbb{R}^2$ così definita

$$\varphi_2(t) := (\varepsilon(1-t), \log(1-t)).$$

Poichè

$$\varphi_2'(t) := \left(-\varepsilon, -\frac{1}{1-t}\right)$$

si ottiene che

$$\begin{aligned} Du(\varphi_2(t))\varphi_2'(t) &= \begin{bmatrix} -2\varepsilon(1-t) & 0 \\ -2\varepsilon(1-t) & 0 \end{bmatrix} \begin{bmatrix} -\varepsilon \\ -\frac{1}{1-t} \end{bmatrix} \\ &= \begin{bmatrix} 2\varepsilon^2(1-t) \\ 2\varepsilon^2(1-t) \end{bmatrix} \in \text{int}\mathbb{R}_+^2 \quad \forall t \in [0, 1). \end{aligned}$$

Dunque φ_2 è una curva ammissibile in base alla definizione 60, inoltre parte da $\varphi_2(0) = (\varepsilon, 0)$ e soddisfa le seguenti relazioni

$$\lim_{t \rightarrow 1^-} \|\varphi_2(t)\| = +\infty,$$

$$\lim_{t \rightarrow 1^-} d(\varphi_2(t), \theta) = 0.$$

Analogamente al caso precedente, si vede che, a patto di scegliere opportunamente ε , comunque si fissi un intorno V di $(0, 0)$ in \mathbb{R}^2 , $(\varepsilon, 0) \in V$. In conclusione si constata che, scelto comunque V intorno di $(0, 0)$, esiste una curva ammissibile che presto o tardi uscirà da qualunque intorno U di $(0, 0)$.

Il precedente esempio, oltre ad essere molto più semplice e naturale di quello fornito in [59], mette anche in luce come entrambe le definizioni di stabilità per punti dell'insieme Pareto critiche considerate in precedenza non si adattino bene al caso vettoriale; risulta infatti quantomeno strano che un punto Pareto massimale *ideale* risulti instabile. La motivazione dell'inadeguatezza delle definizioni di stabilità nel caso di funzioni a valori vettoriali risiede principalmente nel fatto che queste sono troppo strettamente aderenti alle definizioni di punti di equilibrio stabili per sistemi dinamici (cfr. [30], cap.9). Non a caso infatti Smale sottolinea come, se si suppone la dimensione m dello spazio immagine essere pari a uno, allora i punti di θ stabili sono punti di massimo per $u : W \rightarrow \mathbb{R}$, ed inoltre se si considerano punti di θ non

degeneri vale anche il viceversa. Proprio l'ipotesi che un punto critico sia non degenero, ovvero che l'Hessiana in quel punto sia non singolare, non è adatta al caso vettoriale perchè il fatto che un punto critico sia non degenero implica che esso sia isolato (vd. [50], corollario 2.3 del lemma di Morse), ovvero che esista un intorno del punto in questione che non contenga altri punti critici. E questa situazione non si verifica quasi mai quando si considerano i punti dell'insieme θ per una funzione a valori vettoriali.

Queste riflessioni suggeriscono quindi l'interesse di studi futuri che portino ad una definizione stabilità per punti dell'insieme Pareto critico per una funzione $u : W \rightarrow \mathbb{R}^m$ ($m > 1$) che meglio tenga in conto le specificità del caso vettoriale. Balza inoltre agli occhi come per formulare un approccio di questo genere all'ottimizzazione vettoriale sia importante studiare il caso di punti critici non isolati.

A conclusione di questa sezione segnaliamo molto brevemente come l'impiego di strumenti matematici dell'analisi globale e della topologia differenziale sia stato particolarmente intenso e fruttuoso nell'ambito della teoria economica, ed in particolare nella teoria dell'equilibrio economico generale; a titolo d'esempio di quanto affermato si veda la monografia [44].

3.2 Il caso convesso: Aubin e Cellina

Nella monografia [3] di Aubin e Cellina, dedicata alle inclusioni differenziali, sono contenuti alcuni risultati concernenti un approccio dinamico al problema di determinare i punti Pareto minimali di una funzione $u = (u_1, \dots, u_m)$ definita in X , dove X è uno spazio di Hilbert ed ogni singola componente $u_i : X \rightarrow \mathbb{R} \cup \{+\infty\}$ è una funzione semicontinua inferiormente, convessa e propria (cioè tale che esista almeno un $x \in X$ tale che $u(x) \neq +\infty$). Un primo gruppo di risultati è legato all'applicazione ad opportune scalarizzazioni dei metodi sviluppati nell'ambito delle inclusioni differenziali con mappe monotone massimali ed in particolare delle inclusioni gradiente (cfr. 59); un secondo è basato invece sulla cosiddetta Viability Theory.

Incominciamo ad esporre i risultati che sono applicazioni delle inclusioni gradiente (cfr. [3], cap.3, sec.5); in questo caso supponiamo che le funzioni u_i siano limitate inferiormente e, senza perdita di generalità, possiamo assumere che

$$\inf_{x \in X} u_i(x) = 0 \quad \forall i = 1, \dots, m.$$

Come conseguenza di quanto visto nel paragrafo 1.3.1, si ottiene che \bar{x} è un punto Pareto minimale in Q sottoinsieme chiuso convesso di X se e solo se esistono $\lambda_1, \dots, \lambda_m$ tali che $\lambda_i \geq 0$ e $\sum_{i=1}^m \lambda_i = 1$ per cui valga la relazione

$$\sum_{i=1}^m \lambda_i u_i(\bar{x}) = \min_{x \in Q} \sum_{i=1}^m \lambda_i u_i(x).$$

Risulta perciò possibile associare ad ogni m -upla di parametri $\lambda = (\lambda_1, \dots, \lambda_m) \in \mathbb{R}_+^m$ e tali che la seguente inclusione gradiente

$$\begin{cases} x'(t) \in -\partial \left(\sum_{i=1}^m \lambda_i u_i + I_S \right) (x(t)) \\ x(0) = x_0 \end{cases} \quad (3.1)$$

dove I_Q indica la funzione indicatrice di Q ($I_Q(x) = 0$ per ogni $x \in Q$, $I_Q(x) = +\infty$ per ogni $x \in \mathbb{C}Q$) e ∂ il sottodifferenziale di una funzione convessa. Le traiettorie di questa inclusione convergono ai punti Pareto minimali quando $t \rightarrow +\infty$; è inoltre interessante osservare che data una soluzione $x(t)$ dell'inclusione differenziale considerata vale il fatto seguente:

$$t \mapsto \sum_{i=1}^m \lambda_i u_i(x(t)) \text{ è una funzione non crescente,}$$

mentre le soluzioni non soddisfano questa proprietà per le singole componenti della funzione obiettivo u , cioè non vale necessariamente la proprietà di monotonia seguente

$$t \mapsto u_i(x(t)) \text{ è una funzione non crescente.}$$

L'esigenza di far sì che la proprietà precedente valga, porta a sviluppare un discorso simile anche per una diversa scalarizzazione data dalla funzione

$$u_\lambda(x) := \max_{i=1, \dots, m} \frac{u_i(x)}{\lambda_i}$$

i cui punti di minimo, nel caso convesso, caratterizzano completamente i punti Pareto minimali della funzione u (cfr. [3], Lemma 1, pag. 169). Questa scalarizzazione permette di costruire l'inclusione differenziale seguente

$$\begin{cases} x'(t) \in -\text{co} \left(\bigcup_{i=1}^m \frac{\partial u_i(x(t))}{\lambda_i} \right) - N_S(x(t)) \\ x(0) = x_0 \end{cases} \quad (3.2)$$

dove $N_Q(x(t))$ indica il cono normale ad Q in $x(t)$ ($N_Q(x) := \partial I_Q(x)$, vd. [54]) e con $\text{co}(H)$ indichiamo l'involucro convesso dell'insieme $H \subset X$. Le soluzioni di quest'ultima inclusione sono tali che

$$t \mapsto u_i(x(t)) \text{ è una funzione non crescente}$$

ed inoltre convergono ai punti Pareto minimali di u . Questa inclusione però perde alcune importanti proprietà di struttura, possedute invece dall'inclusione 3.1, infatti il membro di destra di 3.2 non è più una mappa monotona massimale come invece era in 3.1.

Analoga all'inclusione differenziale 3.2 è l'inclusione differenziale connessa all'ottimizzazione vettoriale, studiata nell'ambito della teoria della viabilità (cfr. [3], pp. 271-274).

Sia $\Omega \subset \mathbb{R}^n$, dove Ω è un sottoinsieme aperto, consideriamo poi una funzione $u : \Omega \rightarrow \mathbb{R}^m$ tale che ogni sua componente $u_i : \Omega \rightarrow \mathbb{R}$ sia convessa e tale che $u_i \in C^2(\Omega, \mathbb{R})$ per ogni $i = 1, \dots, m$. Si dimostra allora il seguente risultato. Ricordiamo che con $T_Q(x)$ si indica il cono tangente ad S in suo punto x (per la definizione si veda ad esempio [54], cap.6)

Teorema 69 *Sia Q un sottoinsieme di Ω chiuso e convesso. Allora esistono soluzioni lente dell'inclusione differenziale*

$$\begin{cases} \forall t \geq 0, & x(t) \in Q; \quad x(0) = x_0 \text{ (fissato in } Q) \\ \text{per q.o. } t \geq 0 & x'(t) \in p_{T_Q(x(t))} \left(-\text{co} \left(\bigcup_{i=1}^m \nabla u_i(x(t)) \right) \right) \end{cases}$$

che soddisfano, per ogni $s \geq t$,

$$\max_{i=1, \dots, m} (u(x(s)) - u(x(t))) + \int_t^s \|x'(\tau)\|^2 d\tau \leq 0.$$

Inoltre, per $t \rightarrow \infty$, i punti d'accumulazione x_* di $x(t)$ sono punti debolmente Pareto minimali che soddisfano la relazione

$$\lim_{t \rightarrow \infty} u_i(x(t)) = u_i(x_*) \quad \forall i = 1, \dots, m.$$

(Per la definizione di $p_{T_Q(x(t))}$ vedi l'Appendice 1). Il teorema appena enunciato, in buona sostanza, fornisce un sistema dinamico le cui traiettorie soddisfano la proprietà di monotonia

$$t \longmapsto u_i(x(t)) \text{ è una funzione non crescente}$$

ed inoltre convergono ai punti debolmente Pareto minimali. Segnaliamo poi come, in [3] (pp .272-274), questo risultato sia riletto nell'ambito della teoria dei giochi.

In conclusione è interessante sottolineare come questo tipo di studi si discosti, almeno in parte, dall'analisi di Smale. Qui infatti si analizza solamente il caso convesso e ciò permette di esaminare direttamente l'insieme dei punti Pareto minimali, facendo a meno di introdurre la nozione di insieme dei punti critici. Si può inoltre osservare come i risultati concernenti le scalarizzazioni non siano particolarmente interessanti nella nostra ottica poichè essi inseriscono sì una componente dinamica nel problema ma sono essenzialmente lo studio di sistemi gradiente nel caso scalare; diverso invece è il discorso per il secondo approccio che davvero affronta il problema tenendo conto della sua natura vettoriale

3.3 L'approccio di Malivert

Una tappa importante nello studio di un approccio dinamico ai problemi di ottimizzazione vettoriale è segnata dall'articolo [40] di C. Malivert apparso nel 1982.

In esso vengono riprese alcune idee di Smale ed estendendole ad un ambiente più generale, si costruisce un metodo di discesa per l'ottimizzazione vettoriale. Si ha così il primo vero tentativo di generalizzare al caso vettoriale uno dei metodi classici di ricerca delle soluzioni nel caso scalare; naturalmente le modifiche introdotte sono considerevoli, basti pensare che il metodo di discesa scalare si basa sull'impiego del gradiente della funzione obiettivo, che nel caso vettoriale non è neppure definito. In questa sezione intendiamo esporre per sommi capi i risultati contenuti in [40]; per far questo dobbiamo premettere alcune notazioni.

Sia Y uno spazio di Banach ordinato mediante un cono P chiuso, convesso e con interno non vuoto; supponiamo poi che il cono P verifichi un'ulteriore ipotesi di natura tecnica consistente nel fatto che sia normale, cioè sia tale che esista una base di intorni V dell'origine tali che $V = (V - P) \cap (V + P)$. Definiamo ora una norma in Y ponendo

$$\|y\| := \inf \{t : t > 0, y \in [-e, e]\},$$

dove e è un elemento fissato in $\text{int}P$ e $[-e, e]$ denota l'intervallo d'ordine dato da $(-e + P) \cap (e - P)$. Definiamo poi per ogni $y \in Y$

$$\|y\|_+ := d(y, Y \setminus P) = d(y, Y \setminus \text{int}P)$$

dove d è la distanza indotta dalla norma definita sopra. Consideriamo infine un sottoinsieme chiuso Q di uno spazio di Hilbert X che soddisfi alcune ipotesi di regolarità riguardanti sostanzialmente la mappa punto insieme che ad ogni punto di Q associa il cono tangente a Q in quel punto (cfr. condizioni (H_1) e (H_2) in [40]). Possiamo quindi introdurre la funzione $f : Q \rightarrow Y$ che si richiede essere differenziabile con derivata localmente lipschitziana. Questa funzione è proprio la funzione obiettivo del problema di ottimizzazione vettoriale

$$\min f(x)$$

dove l'ordine in Y è indotto dal cono P . Si introduce allora l'insieme dei punti critici per la funzione f :

$$K_{(f,Q)} := \{x \in Q : f'(x)(T_x Q) \cap (-\text{int}P) = \emptyset\},$$

naturalmente si ha che $\text{Min}(f, Q) \subset K_{(f,Q)}$ (vd. [40]. prop. 3.1).

A questo punto si deve individuare, per ogni punto x di Q , un insieme $C(x)$ di "direzioni ammissibili", cioè di direzioni che giochino lo stesso ruolo giocato dalla direzione indicata da $-\nabla f$ nel caso scalare. Questo insieme è definito nel modo seguente

$$C(x) := \{v \in \text{int}T_x Q : \|-f'(x)v\|_+ > \|v\|^2, d(x, K_{(f,Q)}) > \|v\|\}$$

(per convenzione $d(x, K_{(f,Q)}) = +\infty$ se $K_{(f,Q)} = \emptyset$).

Studiando le proprietà della mappa punto insieme $C : Q \setminus K_{(f,Q)} \rightrightarrows X$ che associa ad ogni punto $x \in Q \setminus K$ l'insieme $C(x)$, si arriva a dimostrare, mediante una opportuna variante del teorema 37 (teorema di selezione di Michael), l'esistenza di una selezione $s : Q \setminus K_{(f,Q)} \rightarrow X$ della mappa punto insieme C localmente lipschitziana; si passa allora da esaminare il seguente problema di Cauchy

$$\begin{cases} x'(t) = s(x(t)) \\ x(0) = x_0 \end{cases} \quad x_0 \in Q, \text{ fissato} \quad (3.3)$$

le cui soluzioni sono indicate con $x(t, x_0)$; si denota inoltre con $[0, T(x_0))$ l'intervallo massimale di definizione della soluzione di 3.3, una volta fissata la condizione iniziale $x(0) = x_0$. In particolare si mostra che, sotto opportune ipotesi di limitatezza degli insiemi di sottolivello di f (cfr. [40], ipotesi (B)), la funzione continua $x : [0, \infty) \times Q \rightarrow Q$ definita, a partire dalle soluzioni di 3.3, nel modo seguente

$$x(t) := \begin{cases} x(t, x_0) & \text{se } t < T(x_0), x_0 \in Q \setminus K_{(f,Q)} \\ \lim_{t \rightarrow T(x_0)} x(t, x_0) & \text{se } t \geq T(x_0), x_0 \in Q \setminus K_{(f,Q)} \\ x_0 & \text{se } x_0 \in K_{(f,Q)} \end{cases}$$

converge, per $t \rightarrow \infty$, ai punti critici di f in Q ; inoltre gli elementi di $K_{(f,Q)}$ vengono caratterizzati come i punti limite di questa funzione per $t \rightarrow \infty$.

In conclusione ricordiamo che l'ultima parte di questo articolo è dedicata ad un argomento particolarmente innovativo: mediante l'utilizzo di una versione generalizzata delle categorie di Lusternik-Schnirelman (vd. [45], pag. 97 per una prima introduzione oppure si veda il classico lavoro [52]) vengono provati risultati riguardanti l'esistenza ed il numero dei punti critici di f .

3.4 Alcuni lavori recenti

In questa breve sezione, a conclusione del presente capitolo, ci limiteremo a fornire una sintetica panoramica su alcuni lavori molto recenti che hanno come oggetto approcci dinamici all'ottimizzazione vettoriale e teoria dei punti critici per funzioni a valori in spazi vettoriali parzialmente ordinati.

Il primo lavoro di cui intendiamo parlare è [24]; in esso viene sviluppato un metodo di massima discesa per l'ottimizzazione vettoriale nel caso in cui lo spazio immagine \mathbb{R}^m sia dotato dell'ordine paretiano; vengono esaminati sia il caso libero sia il caso vincolato ed in entrambi vengono assunte ipotesi di differenziabilità sulle funzioni in gioco. In particolare gli autori di questo articolo costruiscono una procedura per individuare le direzioni di massima discesa e basandosi su di essa formulano un algoritmo che dimostrano essere convergente a punti che soddisfano una condizione necessaria del prim'ordine per i punti Pareto minimali; è immediato osservare che questi punti altro non sono che i punti

dell'insieme Pareto critico definito da Smale nei suoi lavori. Si osservi che, pur non comparando esplicitamente i concetti di equazioni o inclusioni differenziali le cui soluzioni siano traiettorie continue che convergono ai punti critici, gli aspetti dinamici sono fondamentali anche in questo lavoro in quanto vengono individuate successioni di punti che convergono ai punti critici che in un'ultima analisi si possono vedere come "discretizzazioni" di traiettorie. I metodi sviluppati in questo lavoro sono poi estesi anche al caso in cui l'ordine nello spazio immagine \mathbb{R}^m sia quello associato ad un cono K generico che si suppone essere chiuso, convesso, puntato e con interno non vuoto ([27]). Sempre in questo lavoro si dimostra poi che con l'aggiunta di ipotesi di convessità si ottiene la convergenza dell'algoritmo studiato, non più solo ai punti critici, ma ai punti debolmente minimali. E' interessante sottolineare come questi lavori siano particolarmente incentrati sugli aspetti algoritmici del problema, a testimonianza di come l'approccio dinamico ai problemi di ottimizzazione vettoriale sia fruttuoso da questo punto di vista.

Su di un versante più strettamente teorico è da ricordare l'articolo [18] in cui viene formulata una teoria dei punti critici per funzioni a valori vettoriali. In questo lavoro vengono considerate ipotesi di regolarità sulle funzioni in gioco molto deboli; basta infatti che le funzioni considerate siano continue. Altra particolarità importante di questo lavoro è che il cono d'ordine nello spazio immagine non è detto abbia interno vuoto; questo fatto è particolarmente importante negli spazi infinito dimensionali dove spesso i coni d'ordine naturali hanno proprio interno vuoto (si ricordi l'esempio 9).

Spieghiamo ora in dettaglio cosa si intende, secondo questo articolo, per punti critici di una funzione $f : X \rightarrow Y$, dove (X, d) è uno spazio metrico ed Y uno spazio di Banach. A tal fine si consideri un cono $P \subset Y$ chiuso e convesso che sarà il cono d'ordine; si prenda poi un insieme $P_0 \subset P$ chiuso e convesso tale che non contenga lo zero di Y (quest'ultimo insieme è indispensabile per poter trattare coni con interno vuoto). Possiamo ora dare la seguente definizione.

Definizione 70 Per ogni $x \in X$, indichiamo con $|d_{P_0}f|(x)$ l'estremo superiore dei $\sigma \in [0, +\infty)$ tali per cui esista $\delta > 0$ ed una mappa continua $\mathcal{H} : B_\delta(x) \times [0, \delta] \rightarrow X$ tale che:

1. $d(\mathcal{H}(\xi, t), \xi) \leq t$,
2. $f(\mathcal{H}(\xi, t)) \in f(\xi) - \sigma t P_0 - P$,

per ogni $\xi \in B_\delta(x)$ e per ogni $t \in [0, \delta]$. Il numero reale esteso (cioè che può valere eventualmente $+\infty$) $|d_{P_0}f|(x)$ è chiamato pendenza debole di f in x , rispetto all'insieme P_0 .

La definizione precedente costituisce un adattamento al caso vettoriale di una recente definizione (si veda [12]) data per funzioni scalari e che ha aperto la strada allo studio della teoria dei punti critici per funzionali solamente continui.

Un *punto critico* per f (rispetto a P_0) è un punto $x \in X$ tale che $|d_{P_0}f|(x) = 0$.

Si osservi che quando X è uno spazio normato, $Y = \mathbb{R}^m$, $P = \mathbb{R}_+^m$, $P_0 = \{(1, \dots, 1)\}$, f è di classe C^1 si ottiene la seguente relazione

$$|d_{P_0}f|(x) \leq \min_{i=1, \dots, m} \|\nabla f_i(x)\|.$$

Quando $\text{int}P \neq \emptyset$, preso $P_0 = \{e\}$ con $e \in \text{int}P$, si dimostra che l'insieme dei punti critici di f non cambia se calcolato rispetto ad e oppure a $e + P$, inoltre l'insieme dei punti critici non dipende dalla scelta di e in $\text{int}P$.

Il lavoro citato procede poi provando un lemma di deformazione che permette di giungere ad un risultato che costituisce l'analogo, nel caso vettoriale, del noto teorema del passo di montagna (cfr. [45], pag. 92 oppure [12]). In questo articolo poi vengono provati anche alcuni risultati che si inseriscono nel filone, delineato in [40], che si occupa dell'applicazione delle categorie di Ljusternik-Schnirelman a funzioni a valori vettoriali.

A conclusione di questa breve sezione ricordiamo un ultimo lavoro ([1]) che affronta un problema di ottimizzazione vettoriale legando le soluzioni di quest'ultimo ai punti critici di un opportuno sistema nonlineare ed autonomo di equazioni differenziali.

Capitolo 4

Un nuovo approccio dinamico

In questo capitolo ci proponiamo di sviluppare un approccio, basato su considerazioni di carattere dinamico, completamente nuovo all'ottimizzazione vettoriale; in particolare ci proponiamo di arrivare a studiare una opportuna inclusione differenziale, le cui traiettorie convergono (in un senso da precisare) all'insieme dei punti critici di una funzione f definita in \mathbb{R}^n a valori in \mathbb{R}^m (ordinato in senso paretiano, cioè mediante il cono \mathbb{R}_+^m). Le motivazioni che stanno alla base di uno studio di questo tipo sono essenzialmente due:

- di natura applicativa, infatti in ogni applicazione sensata, specialmente in quelle economiche (si pensi ad esempio alla massimizzazione simultanea delle funzioni di utilità di m agenti oppure alla selezione del portafoglio), si può agire sulle variabili di partenza (\mathbb{R}^n) piuttosto che su quelle d'arrivo (\mathbb{R}^m); si capisce così l'importanza di ragionare nello spazio di partenza ed in particolare di individuare traiettorie lungo cui la funzione in esame cresca oppure decresca (rispetto all'ordine dato nello spazio d'arrivo); inoltre in questo contesto vengono studiati, oltre che le soluzioni del problema di ottimizzazione vettoriale, anche i punti d'equilibrio che spesso possono avere interpretazioni rilevanti all'interno dei modelli da cui proviene il problema di ottimizzazione vettoriale da esaminare;
- di natura teorica, in quanto lo studio di un'inclusione differenziale di questo tipo permette di approfondire meglio gli aspetti geometrico-topologici sottostanti al problema di ottimizzazione vettoriale ed inserisce un elemento dinamico in un problema che solitamente viene considerato solamente dal punto di vista statico; permette inoltre di collocare il problema di ottimizzazione vettoriale in un contesto più ampio costituito dalla teoria delle singolarità, suggerendo, nello stesso tempo, la possibilità di un approccio alla teoria dei punti critici per funzioni vettoriali simile a quello adottato dalla teoria dei punti critici moderni nel caso scalare (cfr. ad es. [45] oppure [14]) diverso da quello, ormai classico, che ha portato allo studio delle singolarità

di mappe ed esposto ad esempio in [26] ed in cui si inserisce anche la nota “teoria delle catastrofi”.

Sottolineiamo infine che l’approccio sviluppato in questo capitolo si differenzia da quelli delineati nel capitolo precedente essenzialmente per due motivi. Il primo consiste nel fatto che vengono esplicitamente indicate le traiettorie che portano ai punti critici e non ne viene solo garantita l’esistenza come invece accade in [40]; il secondo si basa invece sul fatto che viene esplicitata l’inclusione differenziale che determina le direzioni lungo cui la funzione obiettivo ”decesce” contrariamente agli articoli [24] e [27] dove le leggi che regolano la dinamica del problema non vengono formulate esplicitamente.

Ricordiamo ora le seguenti notazioni (in parte già introdotte in precedenza), che adotteremo in tutto questo capitolo.

Indichiamo con $f = (f_1, \dots, f_m) : \mathbb{R}^n \rightarrow \mathbb{R}^m$ una funzione tale che $f \in C^1(\mathbb{R}^n)$. Naturalmente con $f_i : \mathbb{R}^n \rightarrow \mathbb{R}$ ($i = 1, \dots, m$) indichiamo le componenti di f .

Denotiamo inoltre con $[f'(x)]$ la matrice jacobiana di f :

$$[f'(x)] = \begin{bmatrix} \nabla f_1(x) \\ \dots \\ \dots \\ \dots \\ \nabla f_m(x) \end{bmatrix}$$

dove $\nabla f_i(x)$ ($i = 1, \dots, m$) denota il gradiente della funzione reale f_i .

Si ricordi poi che, come già visto in 1.1, \mathbb{R}_+^m indica l’ortante positivo di \mathbb{R}^m dato dal cono $\{x = (x^1, \dots, x^m) \in \mathbb{R}^m : x^i \geq 0 \forall i = 1, \dots, m\}$ che induce in \mathbb{R}^m l’ordine paretiano (cfr. esempio 1 in 1.1.3).

4.1 Punti critici per una funzione vettoriale

Questa prima sezione è dedicata all’introduzione dell’insieme dei punti critici per una funzione f a valori vettoriali; vedremo poi le che relazioni intercorrono fra questo insieme e le soluzioni del problema di ottimizzazione libero riguardante la funzione f , una volta che si pensi lo spazio immagine ordinato mediante l’ortante positivo.

Definizione 71 *Il punto $x \in \mathbb{R}^n$ si dice critico per la funzione f quando*

$$[f'(x)](\mathbb{R}^n) \cap (-\text{int}(\mathbb{R}_+^m)) = \emptyset.$$

Denotiamo con K_f l’insieme dei punti critici della funzione f .

La definizione precedente riprende quella di Smale data in [57] e in [63] (a tale proposito si veda 3.1).

La seguente proposizione caratterizza i punti critici della funzione f .

Proposizione 72 *Il punto $x \in K_f$ se e solo se esistono $\lambda_1, \dots, \lambda_m$ numeri reali tali che $\lambda_i \geq 0 \forall i = 1, \dots, m$ e $\sum_{i=1}^m \lambda_i = 1$, per cui*

$$\sum_{i=1}^m \lambda_i \nabla f_i(x) = 0.$$

ovvero se i gradienti delle singole componenti di f risultano positivamente linearmente dipendenti.

Un'idea parziale e schematica della dimostrazione di questa proposizione era contenuta in [63] dove era richiesta anche un'ipotesi superflua, tuttavia la dimostrazione che si presenta in questa sede è almeno in parte originale in quanto rende rigorosa e completa quella prova iniziale (si veda la proposizione 63 e l'osservazione collegata). Come si vedrà dalla dimostrazione, in sostanza, il precedente teorema risulta essere un caso particolare di un teorema dell'alternativa.

Dimostrazione. Incominciamo supponendo per ipotesi che $x \in K_f$. Per la definizione 71 si ha che non esiste alcun vettore $h \in \mathbb{R}^n$ tale che

$$[f'(x)]h \in (\text{int}(\mathbb{R}_+^m)).$$

Questa affermazione equivale a dire che il sistema

$$\begin{cases} \langle \nabla f_1(x), h \rangle > 0 \\ \dots \\ \dots \\ \dots \\ \langle \nabla f_m(x), h \rangle > 0 \end{cases}$$

non ha soluzione. Utilizzando allora il teorema dell'alternativa di Gordan (cfr. [41]) si ha che esiste $\beta \in \mathbb{R}^m$ tale che

$$[f'(x)]^T \beta = 0 \quad \text{e} \quad \begin{cases} \beta_i \geq 0 \quad \forall i = 1, \dots, m \\ \beta \neq 0 \end{cases}$$

Ponendo ora

$$\lambda_i = \frac{\beta_i}{\sum_{s=1}^m \beta_s} \quad \forall i = 1, \dots, m$$

si ottiene la tesi.

Viceversa supponiamo ora, per assurdo, che $x \notin K_f$ ovvero che esista $v \in \mathbb{R}^n$ tale che

$$\langle \nabla f_i(x), v \rangle < 0 \quad \forall i = 1, \dots, m.$$

Da ciò segue immediatamente che

$$\left\langle \sum_{i=1}^m \lambda_i \nabla f_i(x), v \right\rangle < 0$$

comunque si scelgano i coefficienti λ_i purchè siano non tutti nulli e non negativi. Si giunge così ad un assurdo. ■

Osservazione 73 *E' immediato osservare che se x è un punto stazionario per una qualunque delle componenti della funzione f allora $x \in K_f$.*

Ricordiamo ora la seguente definizione, già enunciata in un contesto più generale nel capitolo 1 di questa tesi (cfr.1.2, in particolare l'esempio).

Definizione 74 *Un punto $x \in \mathbb{R}^n$ si dice punto Pareto minimale per la funzione f quando*

$$[f(\mathbb{R}^n) - f(x)] \cap (-\mathbb{R}_+^m) = \{0\}.$$

Un punto $x \in \mathbb{R}^n$ si dice punto Pareto massimale per la funzione f se

$$[f(\mathbb{R}^n) - f(x)] \cap (\mathbb{R}_+^m) = \{0\}.$$

Il seguente esempio mostra come una funzione possa avere simultaneamente sia punti Pareto minimali sia punti Pareto massimali.

Esempio 75 *Consideriamo la funzione $f : \mathbb{R}^2 \setminus \{(0,0)\} \rightarrow \mathbb{R}^2$ così definita*

$$f(x^1, x^2) = \left(\frac{x^1}{\sqrt{(x^1)^2 + (x^2)^2}}, \frac{x^2}{\sqrt{(x^1)^2 + (x^2)^2}} \right).$$

Si osserva immediatamente che tutti i punti di \mathbb{R}_+^2 sono punti Pareto massimali mentre tutti i punti di $-\mathbb{R}_+^2$ sono punti Pareto minimali.

La seguente proposizione evidenzia i legami tra la nozione di punto critico e quella di punto Pareto ottimali mettendo in luce come questi ultimi siano un sottoinsieme dei primi.

Proposizione 76 *Se $x \in \mathbb{R}^n$ è un punto Pareto ottimale per f allora $x \in K_f$.*

Dimostrazione. Sia x un punto Pareto minimo.

Dal momento che

$$[f'(x)]v = \lim_{t \rightarrow 0} \frac{1}{t} [f(x + tv) - f(x)] \quad \forall v \in \mathbb{R}^n \quad (t \in \mathbb{R})$$

si deduce immediatamente che $[f'(x)]v \notin -\text{int}(\mathbb{R}_+^m)$ per ogni $v \in \mathbb{R}^n$, cioè $x \in K_f$.

Sia ora x un punto Pareto massimo. Supponiamo per assurdo che $x \in K_f$, allora esiste $v \in \mathbb{R}^n$ tale che

$$[f'(x)]v \in -\text{int}(\mathbb{R}_+^m)$$

perciò si ha che

$$[f'(x)](-v) \in \text{int}(\mathbb{R}_+^m)$$

e questo va contro l'ipotesi che x sia un punto Pareto massimo. ■

La proposizione precedente non può essere invertita, infatti esistono punti critici che non sono nè punti Pareto minimali nè punti Pareto massimali come mostrano i seguenti esempi.

Esempio 77 Si consideri la funzione $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ così definita

$$f((x^1, x^2)) = ((x^1)^3, (x^1)^3).$$

Si osserva immediatamente che l'origine è un punto critico per f ma non è nè un punto Pareto minimo nè Pareto massimo.

Esempio 78 Si consideri la funzione $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ così definita

$$f((x^1, x^2)) = \left(x^2 - (x^1)^2, \frac{-x^2}{(x^1)^2 + 1} \right).$$

Sfruttando la caratterizzazione dei punti critici formulata nella proposizione 72, si prova che

$$K_f = \{x \in \mathbb{R}^2 : x^1 = 0\} \cup \{x \in \mathbb{R}^2 : x^2 = (x^1)^2 + 1\};$$

tuttavia i punti dell'insieme $\{x \in \mathbb{R}^2 : x^1 = 0\}$ risultano non essere Pareto ottimali.

4.2 Lo pseudogradiante generalizzato

In questa sezione ci proponiamo di definire e studiare uno strumento chiave per il nostro approccio ai problemi di ottimizzazione vettoriale; l'idea che sta alla base della sua definizione proviene dalla teoria dei punti critici scalare dove lo pseudogradiante gioca un ruolo fondamentale (vd. Appendice 2). Bisogna però sottolineare come l'impiego di questo strumento nell'ambito dei problemi di ottimizzazione vettoriale abbia richiesto profondi adeguamenti tanto che lo pseudogradiante generalizzato che viene qui impiegato si può dire essere solo un lontano parente di quello scalare.

Definiamo prima di tutto la seguente funzione $s_f : \mathbb{R}^n \rightarrow \mathbb{R}_+$

Definizione 79 $s_f(x) := \min \left\{ \left\| \sum_{i=1}^m \lambda_i \nabla f_i(x) \right\| : \lambda_i \in \mathbb{R}, \lambda_i \geq 0, \sum_{i=1}^m \lambda_i = 1 \right\}.$

Osservazione 80 *Si osserva immediatamente che la funzione $s_f(x)$ è ben definita in quanto fissato x si calcola il minimo di una funzione continua (rispetto alle variabili λ_i) su di un insieme compatto.*

Osservazione 81 *Si osservi che, in forza della proposizione 72 vale la seguente doppia implicazione:*

$$s_f(x) = 0 \iff x \in K_f$$

A questo punto possiamo definire esplicitamente l'oggetto di questa sezione.

Definizione 82 *Un vettore $v \in \mathbb{R}^n$ si dice pseudogradiante generalizzato di f in x se:*

- (1) $\|v\| \leq 2s_f(x),$
 - (2) $[f'(x)]v \in -[s_f(x)]^2 e - \mathbb{R}_+^m$
- dove $e = (1, \dots, 1) \in \text{int}(\mathbb{R}_+^m).$

Si può osservare che la precedente definizione, in realtà, non dipende dalla scelta del vettore e in $\text{int}(\mathbb{R}_+^m)$; noi abbiamo fissato e uguale al vettore di \mathbb{R}^m con tutte le componenti uguali ad 1 per esigenze di semplicità espositiva nelle considerazioni successive.

Definizione 83 *L'insieme*

$$V_f(x) := \{v \in \mathbb{R}^n : v \text{ è uno pseudogradiante generalizzato di } f \text{ in } x\}$$

prende il nome di insieme degli pseudogradianti generalizzati di f in x .

Incominciamo con il provare che in ogni punto di \mathbb{R}^n esiste almeno uno pseudogradiante generalizzato. Per far questo risulta utile richiamare la seguente notazione, che impiegheremo anche nel seguito di questo capitolo.

Notazione 84 *Dati un insieme $S \subset \mathbb{R}^n$ ed un punto $x \in \mathbb{R}^n$ con denotiamo, se esiste, l'elemento di S che realizza la minima distanza da x , ovvero tale che*

$$\|x - p_S(x)\| = \min_{y \in S} \|x - y\|;$$

in particolare con $m(S) := p_S(0)$ indichiamo l'elemento di minima norma dell'insieme S (coerentemente con quanto fatto nella definizione 42).

Proposizione 85 $V_f(x) \neq \emptyset$ per ogni punto $x \in \mathbb{R}^n$. Inoltre $x \in K_f$ se e solo se $V_f(x) = \{0\}$.

Dimostrazione. Consideriamo due casi distinti:

1. $x \notin K_f$; sia $v = m(H)$ dove H è l'insieme di \mathbb{R}^n così definito:

$$H = \left\{ \sum_{i=1}^m \lambda_i \nabla f_i(x) : \lambda_i \in \mathbb{R}, \lambda_i \geq 0, \sum_{i=1}^m \lambda_i = 1 \right\};$$

esso esiste ed è unico in base al teorema della migliore approssimazione (vd. Appendice 1) ed inoltre $\|v\| = s_f(x)$. Si osservi per inciso che $v \neq 0$ (dal momento che $x \notin K_f$). Allora sempre per il teorema della migliore approssimazione (in particolare per la caratterizzazione variazionale dell'elemento di minima norma) si ha che

$$\langle v, v - \nabla f_i(x) \rangle \leq 0 \quad \forall i = 1, \dots, m$$

da cui

$$-\langle v, \nabla f_i(x) \rangle \leq -[s_f(x)]^2 \quad \forall i = 1, \dots, m.$$

Ponendo ora $w = -v$, si ottiene

$$\langle w, \nabla f_i(x) \rangle \leq -[s_f(x)]^2 \quad \forall i = 1, \dots, m$$

perciò

$$[f'(x)]w \in -[s_f(x)]^2 e - \mathbb{R}_+^m.$$

Si conclude quindi che $w \in V_f(x)$.

2. $x \in K_f$; in virtù dell'osservazione 80 si ha che $s_f(x) = 0$, segue quindi immediatamente che $V_f(x) = \{0\}$. Sempre in virtù dell'osservazione 80, si prova anche il viceversa.

■

E'interessante osservare, che tranne nel caso in cui x sia un punto critico, esistono infiniti pseudogradienti generalizzati di f in x ; vale infatti la seguente osservazione.

Osservazione 86 *Sia $x \notin K_f$, allora si ha che $\text{int}(V_f(x)) \neq \emptyset$; infatti consideriamo $w \in V_f(x)$ costruito come nella proposizione 85, allora (per $\eta < 1$)*

$$(1 + \eta)w \in \text{int}(V_f(x)),$$

poichè

$$\begin{aligned} \|(1 + \eta)w\| &= (1 + \eta)s_f(x) < 2s_f(x), \\ [f'(x)](1 + \eta)w &\in -(1 + \eta)[s_f(x)]^2 e - \mathbb{R}_+^m \subseteq -[s_f(x)]^2 e - \text{int}(\mathbb{R}_+^m). \end{aligned}$$

Passiamo ora a studiare alcune proprietà riguardanti la struttura dell'insieme $V_f(x)$.

Proposizione 87 $V_f(x)$ è un insieme compatto e convesso per ogni $x \in \mathbb{R}^n$.

Dimostrazione. Incominciamo con il provare che $V_f(x)$ è convesso. Siano $v_1, v_2 \in V_f(x)$ e $\theta \in [0, 1]$.

Consideriamo la seguente catena di relazioni

$$\|\theta v_1 + (1 - \theta)v_2\| \leq \theta \|v_1\| + (1 - \theta) \|v_2\| \leq 2s_f(x);$$

abbiamo così provato che $\theta v_1 + (1 - \theta)v_2$ soddisfa la proprietà (1) della definizione 82.

Osserviamo ora che in base al punto (2) della definizione 82, si hanno le seguenti uguaglianze:

$$\begin{aligned} \theta [f'(x)] v_1 &= -\theta [s_f(x)]^2 e - \theta k_1 \\ (1 - \theta) [f'(x)] v_2 &= -(1 - \theta) [s_f(x)]^2 e - (1 - \theta) k_2 \end{aligned}$$

dove $k_1, k_2 \in \mathbb{R}_+^m$, quindi

$$[f'(x)] (\theta v_1 + (1 - \theta)v_2) = -[s_f(x)]^2 e - \theta k_1 - (1 - \theta) k_2.$$

Ricordando che $\theta k_1 + (1 - \theta) k_2 \in \mathbb{R}_+^m$, si ottiene la seguente relazione

$$[f'(x)] (\theta v_1 + (1 - \theta)v_2) \in -[s_f(x)]^2 e - \mathbb{R}_+^m,$$

la quale prova che $\theta v_1 + (1 - \theta)v_2$ soddisfa anche la proprietà (2) della definizione 82.

Dimostriamo ora che $V_f(x)$ è un insieme compatto. A tale proposito consideriamo l'insieme

$$W_f(x) = \{v \in \mathbb{R}^n : [f'(x)] v \in -[s_f(x)]^2 e - \mathbb{R}_+^m\};$$

si prova immediatamente che questo insieme è chiuso in quanto controimmagine di un chiuso mediante un'applicazione continua. Osserviamo ora che

$$V_f(x) = W_f(x) \cap \overline{B_{2s_f(x)}(0)}$$

è un insieme chiuso, quindi $V_f(x)$ è compatto dal momento che si può riguardare come un chiuso contenuto in un compatto. ■

Concludiamo questo paragrafo con alcuni esempi di calcolo esplicito degli pseudogradienti generalizzati.

Esempio 88 Consideriamo la funzione $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ così definita

$$f((x^1, x^2)) = ((x^1)^2, (x^1)^2);$$

allora si ha che

$$s_f((x^1, x^2)) = 2|x^1|.$$

Gli elementi di $V_f((x^1, x^2))$ sono quindi i vettori $v=(v,v)$ tali che soddisfino il punto 1. della definizione 82 ovvero

$$(v^1)^2 + (v^2)^2 \leq 16(x^1)^2$$

ed inoltre verifichino anche il punto 2. della stessa definizione, espresso dalla seguente disequazione

$$2x^1v^1 + 4(x^1)^2 \leq 0.$$

Si osserva allora che $V_f((x^1, x^2)) = \{(0,0)\}$ se e solo se $(x^1, x^2) \in \{(x^1, x^2) \in \mathbb{R}^2 : x^1 = 0\} = K_f$.

Esempio 89 Consideriamo la funzione $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ così definita

$$f((x^1, x^2)) = ((-x^1)^3, (x^1)^3);$$

allora si ha che

$$s_f((x^1, x^2)) = 0 \quad \forall (x^1, x^2) \in \mathbb{R}^2$$

e dunque

$$V_f((x^1, x^2)) = \{(0,0)\} \quad \forall (x^1, x^2) \in \mathbb{R}^2.$$

Si osservi infatti che in questo caso $K_f = \mathbb{R}^2$ (anzi ogni punto è sia Pareto minimale sia Pareto massimale).

4.3 L'inclusione differenziale e le sue traiettorie

Scopo di questo paragrafo è lo studio dell'inclusione differenziale

$$x'(t) \in V_f(x(t)) \tag{4.1}$$

e delle traiettorie determinate dalle sue soluzioni.

Una prima semplice, tuttavia molto significativa, osservazione rende chiaro il significato dell'inclusione differenziale precedente mostrando la stretta parentela che la lega ai sistemi di equazioni differenziali definiti mediante il gradiente di una funzione scalare (vedi cap.9, § 4 in [30]). Proveremo infatti che la funzione f , che altro non è che la funzione di cui si stanno cercando i punti Pareto ottimali, gode di una importante proprietà di monotonia lungo le traiettorie determinate dall'inclusione differenziale 4.1.

Proposizione 90 *La funzione f è strettamente decrescente lungo una qualsiasi traiettoria definita da 4.1. (f si intende monotona rispetto all'ordine indotto in \mathbb{R}^m dal cono \mathbb{R}_+^m - cfr. [37] cap.1, § 4).*

Dimostrazione. Si consideri la seguente uguaglianza (per $h > 0$):

$$f(x(t+h)) - f(x(t)) = \int_t^{t+h} f'(x(s)) x'(s) ds \quad \forall t$$

Ricordando che $f'(x(t)) x'(t) \in -[s_f(x(t))]^2 e - \mathbb{R}_+^m$ e che se $x(t) \notin K_f$, $-[s_f(x(t))]^2 e - \mathbb{R}_+^m \subset \text{int} \mathbb{R}_+^m$, si ottiene che

$$f(x(t+h)) - f(x(t)) \in -\text{int} \mathbb{R}_+^m.$$

■

Si osservi che nella dimostrazione precedente l'incremento di f deve essere espresso attraverso l'uso di un integrale poichè riguardo alla funzione $x(t)$ sappiamo solo che è assolutamente continua in quanto soluzione di 4.1 (cfr. definizione 46).

4.3.1 Esistenza delle soluzioni

Il primo problema che dobbiamo affrontare è quello inerente alla dimostrazione dell'esistenza di soluzioni del problema di Cauchy associato all'inclusione differenziale 4.1. Per far questo ci servirà, oltre alle proprietà di struttura dell'insieme $V_f(x)$ mostrate nel paragrafo precedente, provare che la mappa punto-insieme $V_f : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$, definita da

$$x \rightrightarrows V_f(x)$$

gode di alcune proprietà di regolarità. Alla dimostrazione vera e propria della validità delle proprietà di regolarità necessarie della mappa punto insieme V_f , dobbiamo però premettere un risultato riguardante la regolarità della funzione s_f .

Lemma 91 *La funzione s_f è continua per ogni $x \in \mathbb{R}^n$.*

Dimostrazione. Dal momento che $f \in C^1(\mathbb{R}^n)$, si ha che per ogni $\varepsilon > 0$, esiste $\delta > 0$ tale che se $\|x - y\| < \delta$ allora

$$\left| \frac{\partial f_i}{\partial x_j}(x) - \frac{\partial f_i}{\partial x_j}(y) \right| < \frac{\varepsilon}{\sqrt{n}} \quad \forall i = 1, \dots, m; \forall j = 1, \dots, n;$$

da questa disuguaglianza segue direttamente che

$$\|\nabla f_j(x) - \nabla f_j(y)\| < \varepsilon.$$

Dobbiamo ora stimare la quantità

$$|A(x) - A(y)| = \left| \min \left\{ \left\| \sum_{i=1}^m \lambda_i \nabla f_i(x) \right\| \right\} - \min \left\{ \left\| \sum_{i=1}^m \mu_i \nabla f_i(y) \right\| \right\} \right|$$

dove i minimi sono calcolati al variare di λ_i, μ_i nel semplice unitario

$$S = \left\{ \lambda \in \mathbb{R}^m : \lambda_i \geq 0, \sum_{i=1}^m \lambda_i = 1 \right\}.$$

Indichiamo ora con μ_i^* ($i = 1, \dots, m$) i valori tali che

$$\left\| \sum_{i=1}^m \mu_i^* \nabla f_i(y) \right\| = \min \left\{ \left\| \sum_{i=1}^m \mu_i \nabla f_i(y) \right\| \right\}.$$

Inoltre possiamo supporre, senza perdere di generalità, che $s_f(x) \geq s_f(y)$. Sotto queste condizioni si ottiene allora

$$\begin{aligned} |s_f(x) - s_f(y)| &\leq \left\| \sum_{i=1}^m \mu_i^* \nabla f_i(x) \right\| - \left\| \sum_{i=1}^m \mu_i^* \nabla f_i(y) \right\| \\ &\leq \left\| \sum_{i=1}^m \mu_i^* [\nabla f_i(x) - \nabla f_i(y)] \right\| \\ &\leq \sum_{i=1}^m \mu_i^* \|\nabla f_i(x) - \nabla f_i(y)\| < \varepsilon \sum_{i=1}^m \mu_i^* = \varepsilon. \end{aligned}$$

■

Osservazione 92 *In base all'osservazione 80 ed al lemma precedente, si vede immediatamente che K_f è un insieme chiuso in quanto controimmagine di un chiuso mediante una funzione continua.*

Passiamo ora a studiare la regolarità della mappa punto-insieme V_f .

Lemma 93 *La mappa punto-insieme V_f è continua in ogni punto $x \in \mathbb{R}^n$.*

Dimostrazione. La dimostrazione si divide in due parti distinte.

1. Incominciamo con il dimostrare che V_f è superiormente semicontinua.

Consideriamo le successioni $\{x_s\}$ e $\{v_s\}$ tali che $x_s \in B_{\frac{1}{s}}(x)$ e $v_s \in V_f(x_s)$ per $s \in \mathbb{N}$. Supponiamo ora, per assurdo, che esista un intorno U di $V_f(x)$ ($U \supset V_f(x)$) per cui si abbia $v_s \notin U$ (almeno definitivamente). Si osservi ora che la successione $\{v_s\} \subset \mathbb{R}^n$ è limitata; infatti poichè $x_s \rightarrow x$, esiste

$r_0 > 0$ tale che $x_s \in B_{r_0}(x) \quad \forall s \geq s_0$; poichè, per il lemma 91, s_f è continua, essa ammette un punto di massimo $x_0 \in \overline{B_{r_0}(x)}$; concludiamo quindi che

$$\|v_s\| \leq 2s_f(x_s) \leq 2s_f(x_0) \quad \forall s \geq s_0.$$

Possiamo allora costruire una sottosuccessione $\{v_{s_k}\}$ tale che $v_{s_k} \xrightarrow[k \rightarrow \infty]{} v$. Ricordando che $v_{s_k} \in V_f(x_{s_k}) \quad \forall k$, f è di classe C^1 , s_f è continua e \mathbb{R}_+^m è un chiuso si ottiene

$$\|v\| \leq 2s_f(x), \quad [f'(x)]v + [s_f(x)]^2 e \in -\mathbb{R}_+^m$$

ovvero che $v \in V_f(x)$, giungendo così ad un assurdo.

2. Passiamo ora a dimostrare la semicontinuità inferiore di $V_f(x)$, distinguendo due casi:

- sia $x \in K_f$; allora in base all'osservazione 81 $s_f(x) = 0$, consideriamo ora una successione $\{x_n\}$ tale che $x_n \rightarrow x$; per la continuità della funzione s_f si ha

$$s_f(x) \longrightarrow 0.$$

Costruiamo ora la successione v_n nel seguente modo

$$v_n = \begin{cases} w_n & \text{se } x_n \notin K_f \\ 0 & \text{se } x_n \in K_f \end{cases}$$

dove w_n è l'elemento di $V_f(x_n)$ costruito nella dimostrazione della proposizione 85; è immediato osservare che $v_n \in V(x_n)$ e che $\|v_n\| = s_f(x_n)$. Possiamo quindi concludere che $\|v_n\| \rightarrow 0$ e dunque $v_n \rightarrow 0$. Ricordando che $V_f(x) = \{0\}$ si vede che è stata così provata la semicontinuità inferiore di V_f in x .

- sia $x \notin K_f$; sia v un generico elemento di $V_f(x)$, consideriamo ora il seguente vettore

$$h_\alpha = \alpha(1 + \eta)w + (1 - \alpha)v$$

dove w è lo pseudogradiante generalizzato di f in x costruito nella proposizione 85, $\eta \in [0, \frac{1}{2}]$ ed $\alpha \in (0, 1)$. Osserviamo prima di tutto che fissato $\varepsilon > 0$ qualunque, si può determinare $0 < \tilde{\alpha} < 1$ tale che $h_\alpha \in B_\varepsilon(v)$ per ogni $\alpha \in (0, \tilde{\alpha})$.

Si osserva poi che in virtù del corollario 1 pag 35 [BS] $h_\alpha \in \text{int}V_f(x) \quad \forall \alpha \in (0, \tilde{\alpha})$; basta ricordare infatti che $V_f(x)$ è un insieme chiuso convesso (cfr. proposizione 87) ed inoltre per l'osservazione 86 $(1 + \eta)w \in \text{int}V_f(x)$.

Ricordando poi che la funzione f è di classe C^1 e che la funzione A è continua, si vede che esiste $\delta > 0$ tale che per ogni $y \in B_\delta(x)$ esiste

$\alpha \in (0, \tilde{\alpha})$ per cui $h_\alpha \in V_f(y)$; risulta allora verificata la seguente relazione

$$V_f(y) \cap B_\varepsilon(v) \neq \emptyset \quad \forall y \in B_\delta(x).$$

Abbiamo così provato la semicontinuità inferiore della mappa punto-insieme $V_f(x)$.

■

Possiamo a questo punto enunciare un teorema d'esistenza per il problema di Cauchy connesso all'inclusione differenziale (4.1). Per la definizione e le proprietà delle soluzioni lente di un'inclusione differenziale rimandiamo al capitolo sulle inclusioni differenziali, in particolare alla definizione 50 ed all'osservazione collegata.

Teorema 94 *Si consideri il problema di Cauchy dato dall'inclusione differenziale (4.1) e dalla condizione iniziale $x(0) = x_0$ con $x \in \mathbb{R}^n$ fissato. Allora esiste un intervallo (ω_-, ω_+) , con $\omega_- < 0 < \omega_+$, nel quale è definita una soluzione lenta $x(t)$ passante per x_0 . Inoltre o la soluzione lenta sopra citata è tale che $\|x(t)\| \rightarrow +\infty$ quando $t \rightarrow \omega_+$ oppure $\omega_+ = +\infty$.*

Dimostrazione. La dimostrazione si riduce a mostrare che valgono le ipotesi del teorema 49 del capitolo 2.

La mappa punto insieme V_f risulta infatti definita su tutto \mathbb{R}^n , continua per il lemma 93 ed a valori chiusi e convessi in virtù della proposizione 87. ■

Pur avendo dimostrato che tutte le soluzioni di 4.1 danno origine a traiettorie lungo le quali la funzione f gode di una proprietà di monotonia, noi abbiamo scelto di privilegiare le soluzioni lente che altro non sono che quelle che soddisfano la seguente equazione differenziale (almeno per quasi ogni t)

$$x'(t) = m(V_f(x)).$$

Questa scelta è dovuta anche al seguente fatto che rende più agevole il calcolo esplicito di queste soluzioni. E' interessante infatti osservare che l'elemento di norma minima in $V_f(x)$ gode di alcune proprietà; in particolare si può esprimere come combinazione convessa dei gradienti delle componenti della funzione f ed inoltre coincide proprio con la combinazione convessa di questi gradienti che ha norma minima. L'osservazione seguente permette quindi di calcolare l'elemento di norma minima dell'insieme $V_f(x)$ elemento in modo esplicito e più semplice rispetto all'esame diretto dell'intero insieme $V_f(x)$.

Osservazione 95 $m(V_f(x)) = -m(H)$ dove

$$H = \left\{ \sum_{i=1}^m \lambda_i \nabla f_i(x) : \lambda_i \in \mathbb{R}, \lambda_i \geq 0, \sum_{i=1}^m \lambda_i = 1 \right\}$$

(cfr. dimostrazione della proposizione 85).

Dimostrazione. Sia $v = m(H) = \sum_{i=1}^m \tilde{\lambda}_i \nabla f_i(x)$ con $\tilde{\lambda}_i \geq 0$, $\sum_{i=1}^m \tilde{\lambda}_i = 1$, si ricordi inoltre che $\|v\| = s_f(x)$ (cfr. dimostrazione della proposizione 85) e che $-m(H) \in V_f(x)$. Supponiamo ora, per assurdo, che esista $h \in V_f(x)$ tale che $\|h\| < s_f(x)$. Per definizione di pseudogradiente generalizzato si ha che

$$\langle h, \nabla f_i(x) \rangle \leq -[s_f(x)]^2 \quad \forall i = 1, \dots, m;$$

moltiplicando ambo i membri di ognuna delle precedenti disuguaglianze per il corrispondente coefficiente $\tilde{\lambda}_i$ ($i = 1, \dots, m$) e sommando membro a membro, si ottiene che

$$\langle h, v \rangle \leq -[s_f(x)]^2$$

che porta ad un assurdo (contro la disuguaglianza di Cauchy-Schwarz). Si conclude quindi che $\|m(V_f(x))\| = s_f(x)$; ricordando allora che l'elemento di norma minima di un insieme chiuso e convesso è unico, si giunge così alla tesi. ■

4.3.2 Proprietà asintotiche delle traiettorie

Terminiamo ora lo studio delle soluzioni delle inclusioni differenziali analizzando le proprietà asintotiche e mostrando che, sotto opportune ipotesi, esse "convergono" all'insieme dei punti critici. Il teorema 94 ci porta a considerare due diversi tipi di soluzioni lente del problema di Cauchy associato all'inclusione (4.1):

- traiettorie illimitate corrispondenti a soluzioni definite nell'intervallo limitato (ω_-, ω_+) ,
- traiettorie limitate corrispondenti a soluzioni definite su tutto $(\omega_-, +\infty)$.

Pr queste ultime traiettorie è particolarmente interessante studiarne il comportamento asintotico ovvero analizzare cosa accade della soluzione quando $t \rightarrow +\infty$. A tale proposito vale il seguente teorema.

Teorema 96 *Sia $x(t)$ una soluzione lenta di (4.1) definita almeno su $[0, +\infty)$. Allora ogni valore limite di $x(t)$ per $t \rightarrow +\infty$ è un punto critico per la funzione f .*

Dimostrazione. Dal teorema 94 segue immediatamente che la traiettoria $\{x(t)\}_{t \in [0, +\infty)}$ è limitata perciò, indicata con Λ la classe limite di $x(t)$ per $t \rightarrow +\infty$, si ha che anche Λ è un sottoinsieme non vuoto e limitato di \mathbb{R}^n .

Consideriamo allora $a \in \Lambda$, ciò significa che esiste una successione $\{t_j\}_{j \in \mathbb{N}}$ tale che $t_j \xrightarrow{j \rightarrow +\infty} +\infty$ ed inoltre

$$\lim_{j \rightarrow +\infty} x(t_j) = a.$$

Dobbiamo ora dimostrare che $a \in K_f$.

Incominciamo con l'osservare che $f(a) = \alpha$ è l'estremo inferiore (rispetto all'ordine in \mathbb{R}^m dato dal cono \mathbb{R}_+^m) dell'insieme $\{f(x(t)) : t \geq 0\}$; questo fatto segue dalla proposizione 90 e dalla regolarità di f .

Consideriamo ora l'orbita di a , cioè l'immagine in \mathbb{R}^n della soluzione $y(s)$ del seguente problema di Cauchy

$$\begin{cases} y'(s) = m(V_f(y(s))) \\ y(0) = a \end{cases}$$

si osserva immediatamente che ogni punto di quest'orbita appartiene a Λ , infatti si ha che

$$x(t_j + s) \longrightarrow y(s) \quad \forall s \geq 0.$$

Per quanto detto sopra si può affermare che $f(l) = \alpha$ per ogni $l \in \Lambda$, allora la funzione f è costante sull'orbita di a quindi

$$[f'(a)] m(V_f(a)) = 0$$

da cui $s_f(a) = 0$, che implica $a \in K_f$. ■

Dal teorema precedente deriva immediatamente il seguente corollario.

Corollario 97 *Sia $x(t)$ una soluzione lenta di (4.1) definita almeno su $[0, +\infty)$. Allora*

$$\lim_{t \rightarrow +\infty} d(x(t), K_f) = 0.$$

4.4 Linee di ricerca future

La presente trattazione per certi versi può ritenersi conclusa, tuttavia sotto molti punti di vista suggerisce possibili spunti per la prosecuzione di ricerche future. Di seguito elenchiamo i principali temi che vorremmo approfondire in futuro.

- Un primo problema aperto riguarda la definizione di una nozione di stabilità per i punti critici che sia adatta ai problemi di ottimizzazione vettoriale (a tale riguardo si vedano le considerazioni fatte in questa tesi a proposito delle definizioni di stabilità 66 e 67 date da Smale). Il problema più grosso che si incontra in questo studio è dovuto al fatto che nel caso vettoriale i punti critici di una funzione non sono quasi mai isolati; questo fatto non permette quindi di procedere come nel caso di sistemi dinamici definiti dal gradiente di una funzione scalare (cfr. [30]). Lo scopo di queste ricerche è, oltre allo studio della stabilità di per sè, il tentativo di caratterizzare i punti Pareto ottimali come punti critici stabili. Su questi argomenti abbiamo già ottenuto alcuni risultati, seppure molto parziali.
- Un secondo campo di ricerca è certamente l'applicazione dei risultati contenuti in questa tesi ad alcuni problemi concreti riguardanti, ad esempio, le applicazioni delineate nel primo capitolo della tesi.

- Un terzo indirizzo di ricerca costituisce l'elaborazione di un algoritmo per l'ottimizzazione vettoriale basato sulle traiettorie studiate. Naturalmente questo punto richiederà l'acquisizione di competenze in materia di analisi numerica e molto lavoro per utilizzare le idee presenti in questo lavoro al fine di creare un metodo efficiente dal punto di vista computazionale.
- Un quarto ed ultimo tema consiste nell'estendere i risultati contenuti in questa tesi a problemi di ottimizzazione vettoriale ambientati in spazi ordinati mediante coni differenti da quello dell'ordine paretiano o anche a spazi ad infinite dimensioni come quelli elencati in 1.1.3.

Appendice 1 - Il teorema della migliore approssimazione

Il teorema della migliore approssimazione costituisce un risultato classico della teoria degli spazi di Hilbert e garantisce che, dato un qualsiasi punto x , un sottoinsieme chiuso e convesso di uno spazio di Hilbert possiede sempre uno ed un solo elemento che rende minima la distanza di tale sottoinsieme dal punto x . Per i nostri scopi risulta però più utile la versione del medesimo teorema che fornisce anche una caratterizzazione variazionale dell'elemento di norma minima (vd. [3], pp.22-24). Per la dimostrazione di questo risultato seguiremo [3], dove si utilizza un teorema di natura generale, riguardante le funzioni convesse semicontinue inferiormente, di cui il teorema della migliore approssimazione risulta essere un corollario; ricordiamo comunque che la sola dimostrazione dell'esistenza e dell'unicità dell'elemento di norma minima si può ottenere direttamente (cfr. ad esempio [55] pag 306, teor. 2.1).

Il risultato concernente le funzioni convesse cui si accennava sopra è il seguente:

Teorema 98 *Siano X uno spazio di Hilbert ed $f : X \rightarrow \mathbb{R} \cup \{+\infty\}$ una funzione propria (cioè non identicamente uguale a $+\infty$), convessa e semicontinua inferiormente. Allora, per ogni $x \in X$, esiste un unico elemento $x_0 \in X$ tale che*

$$(*) \quad f(x_0) + \frac{1}{2} \|x_0 - x\|^2 = \inf_{y \in X} \left(f(y) + \frac{1}{2} \|y - x\|^2 \right).$$

Inoltre l'elemento x_0 è completamente caratterizzato dalla seguente disuguaglianza

$$(**) \quad \forall y \in X, \quad f(x_0) - f(y) + \langle x_0 - x, x_0 - y \rangle \leq 0.$$

Dimostrazione. Si consideri la funzione

$$g(x) := \inf_{y \in X} \left(f(y) + \frac{1}{2} \|y - x\|^2 \right);$$

si osserva immediatamente che $g(x) > 0$ per ogni $x \in X$.

1. Incominciamo con il provare l'esistenza di un elemento di X che sia soluzione dell'equazione

$$g(x) = \left(f(x_0) + \frac{1}{2} \|x_0 - x\|^2 \right).$$

Per far questo consideriamo una successione $\{y_n\} \subset X$ tale che

$$f(y_n) + \frac{1}{2} \|y_n - x\|^2 \leq g(x) + \frac{1}{n}.$$

Osserviamo ora che

$$\begin{aligned} \|y_n - y_m\|^2 &= 2 \|y_n - x\|^2 + 2 \|y_m - x\|^2 - 4 \left\| \frac{y_n + y_m}{2} - x \right\|^2 \\ &\leq 4 \left[\frac{1}{n} + \frac{1}{m} + 2g(x) - f(y_n) - f(y_m) \right] + 8 \left[f \left(\frac{y_n + y_m}{2} \right) - g(x) \right] \\ &\leq 4 \left[\frac{1}{n} + \frac{1}{m} + 2f \left(\frac{y_n + y_m}{2} \right) - f(y_n) - f(y_m) \right] \\ &\leq 4 \left(\frac{1}{n} + \frac{1}{m} \right), \end{aligned}$$

e questo mostra che la successione $\{y_n\}$ è una successione di Cauchy; possiamo perciò estrarne una sottosuccessione $\{y_{n_j}\}$ convergente ad un elemento $x_0 \in X$. Ricordando ora che f è semicontinua inferiormente, si ottiene infine che

$$f(x_0) + \frac{1}{2} \|x_0 - x\|^2 \leq \liminf_{j \rightarrow \infty} f(y_{n_j}) + \lim_{j \rightarrow \infty} \frac{1}{2} \|y_{n_j} - x\|^2 g(x).$$

2. Sia ora x_0 una soluzione del problema di minimizzazione dato dalla (*). Allora, prendendo $y = x_0 + \theta(x_0 - z)$, (con $\theta > 0$) si deduce che

$$\begin{aligned} &f(x_0) + \frac{1}{2} \|x_0 - x\|^2 \\ &\leq (1 - \theta)f(x_0) + \theta f(z) + \frac{1}{2} (\|x_0 - x\|^2 + 2\theta \langle x_0 - x, x_0 - z \rangle + \theta \|x_0 - z\|^2) \end{aligned}$$

da cui

$$f(x_0) - f(z) \leq \langle x_0 - x, x_0 - z \rangle + \frac{\theta}{2} \|x_0 - z\|^2.$$

Facendo ora tendere θ a 0, si giunge a mostrare la validità della (**).

3. Viceversa la disuguaglianza (**) implica che x_0 minimizza la funzione

$$y \mapsto \left(f(y) + \frac{1}{2} \|y - x\|^2 \right),$$

dal momento che

$$\frac{1}{2} \|x_0 - x\|^2 - \frac{1}{2} \|y - x\|^2 \leq \langle x_0 - x, x_0 - y \rangle.$$

4. Ci resta ora da provare l'unicità di x_0 ; supponiamo per assurdo che esista un'altra soluzione x_1 della disuguaglianza

$$(***) \quad \forall z \in X, \quad f(x_1) - f(z) + \langle x_1 - x, x_1 - z \rangle \leq 0.$$

Si prenda ora $y = x_1$ in (**), $z = x_0$ in (***) e si sommino membro a membro le due disuguaglianze; si ottiene allora che $\|x_0 - x_1\|^2 \leq 0$, cioè che $x_0 = x_1$.

■

Come corollario del teorema precedente si ottiene il teorema della migliore approssimazione.

Corollario 99 (Teorema della migliore approssimazione) *Sia $K \neq \emptyset$ un sottoinsieme chiuso e convesso di uno spazio di Hilbert X . Ad ogni $x \in X$ è possibile associare un unico elemento $p_K(x) \in K$ tale che*

$$\|x - p_K(x)\| = \min_{y \in K} \|x - y\|$$

Inoltre $p_K(x)$ è completamente caratterizzato dalla seguente disequazione variazionale:

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

Dimostrazione. Basta prendere $f = I_K$ nel teorema precedente, dove I_K indica la funzione indicatrice di K così definita

$$I_K(x) = \begin{cases} 0 & \text{se } x \in K \\ +\infty & \text{se } x \notin K \end{cases}$$

che risulta essere una funzione convessa dal momento che K è convesso e semi-continua inferiormente poichè K è chiuso (per le proprietà di i_K vd. [38]), inoltre I_K è propria in quanto $K \neq \emptyset$. ■

Osservazione 100 *Ricordando la notazione riguardante la mappa minimale introdotta in precedenza (definizione 42) possiamo indicare con $m(K)$ l'elemento $p_K(0)$, che altro non è che l'elemento di norma minima appartenente a K . Tale elemento risulta allora caratterizzato dalla seguente disequazione variazionale*

$$\forall y \in K, \quad \langle m(K), m(K) - y \rangle \leq 0.$$

Appendice 2 - Lo pseudogradiante (scalare)

In questa appendice intendiamo ricordare brevemente la definizione ed alcune proprietà dello pseudogradiante per funzioni a valori reali. Essa ha costituito lo spunto iniziale che ci ha portato alla definizione di pseudogradiante generalizzato (vd definizione 82) nel caso vettoriale che risulta essere tuttavia molto lontana da quella originaria per la presenza della funzione ausiliaria s_f , indispensabile tra l'altro per poter definire lo pseudogradiante generalizzato anche nei punti critici, fatto questo che non si verifica nel caso scalare.

Lo pseudogradiante nel caso scalare è uno strumento fondamentale della teoria dei punti critici moderni per funzionali a valori reali (in particolare della teoria di Morse estesa anche a spazi di Banach) perchè permette di dimostrare risultati di deformazione che sono di grande importanza nello sviluppo di questa teoria (cfr. [45], cap.6 e cap.8 oppure [14]). Da un punto di vista meramente intuitivo si può dire che i risultati di deformazione indagano il comportamento, dal punto di vista topologico, degli insiemi di sottolivello della funzione oggetto di studio. Nella teoria classica di Morse ([50]) si considera infatti la deformazione dello spazio lungo le linee di massima discesa associate alla funzione $f : X \rightarrow \mathbb{R}$, cioè lungo le traiettorie individuate dal seguente sistema di equazioni differenziali:

$$x'(t) = -\nabla f(x(t)).$$

Tuttavia questo non è sempre possibile perchè a volte le ipotesi in gioco non permettono di definire il gradiente di f a causa della mancanza di regolarità di f oppure a causa della struttura troppo debole dello spazio X . In questi casi si ricorre allora al campo vettoriale pseudo gradiente che gioca un ruolo analogo al campo definito da $-\nabla f$, cioè indica le direzioni di discesa per f ovvero quelle direzioni lungo cui la funzione f decresce.

La teoria dei punti critici moderna esula dai limiti di questo lavoro, perciò si rimanda alle monografie [45] e [14] per una sua esposizione dettagliata; in questa sede richiamiamo solo la definizione di campo vettoriale pseudogradiante ed un risultato riguardante la sua esistenza ed una prima sua proprietà di regolarità. Si ricorda che con f' si indica la derivata di Frechet di una funzione f pensata come elemento del duale di X e con \langle, \rangle il prodotto di dualità.

Definizione 101 Sia X uno spazio di Banach, $f \in C^1(X, \mathbb{R})$ e

$$Y = \{u \in X : f'(u) \neq 0\}.$$

Un campo vettoriale pseudogradiante per f su Y è una mappa continua e localmente lipschitziana $v : Y \rightarrow X$ tale che, per ogni $u \in Y$, si abbia

$$\|v(u)\| \leq 2 \|f'(u)\|, \quad (2)$$

$$\langle f'(u), v(u) \rangle \geq \|f'(u)\|^2. \quad (3)$$

Teorema 102 Valgano le ipotesi su X ed f della definizione precedente; allora esiste un campo vettoriale pseudogradiante per f su Y .

Dimostrazione. Sia $\tilde{u} \in Y$; poiché $f'(\tilde{u}) \neq 0$ si ha che esiste $w \in X$ tale che $\|w\| = 1$ e

$$\langle f'(\tilde{u}), w \rangle > \frac{2}{3} \|f'(\tilde{u})\|.$$

Poniamo $v := \frac{3}{2} \|f'(\tilde{u})\| w$, si ottiene allora che

$$\begin{aligned} \|v\| &= \frac{3}{2} \|f'(\tilde{u})\| < 2 \|f'(\tilde{u})\|, \\ \langle f'(\tilde{u}), v \rangle &= \frac{3}{2} \|f'(\tilde{u})\| \langle f'(\tilde{u}), w \rangle > \|f'(\tilde{u})\|^2. \end{aligned}$$

Dal momento che f' è continua esiste un intorno aperto $V_{\tilde{u}}$ di \tilde{u} in Y tale che, per ogni $u \in V_{\tilde{u}}$, valgano le seguenti relazioni:

$$\|v\| < 2 \|f'(u)\|, \quad (4)$$

$$\langle f'(u), v \rangle > \|f'(u)\|^2. \quad (5)$$

La famiglia d'insiemi $F = \{V_{\tilde{u}}\}_{\tilde{u} \in Y}$ costituisce una copertura aperta dello spazio metrico Y . Ricordando che ogni spazio metrico è paracompatto si può affermare che F possiede un raffinamento localmente finito $\{W_{\tilde{u}_i}\}_{i \in I}$, cioè una famiglia d'insiemi tale che per ogni $i \in I$ esista $\tilde{u} \in Y$ per cui $W_{\tilde{u}_i} = V_{\tilde{u}}$ ed inoltre per ogni $y \in Y$ esista un intorno U di y tale che $U \cap W_{\tilde{u}_i} \neq \emptyset$ solo per un numero finito di valori dell'indice i . A questo punto definiamo le seguenti funzioni

$$\rho_i(u) := d(u, \mathbb{C}W_{\tilde{u}_i}), \quad i \in I$$

e

$$v(u) := \sum_{i \in I} \frac{\rho_i(u)}{\sum_{j \in I} \rho_j(u)} v_i$$

dove gli elementi v_i sono costruiti a partire da \tilde{u}_i nello stesso modo in cui, all'inizio della dimostrazione, partendo da \tilde{u} si arrivava a v . In virtù del fatto

che $\{W_{\bar{u}_i}\}_{i \in I}$ è un ricoprimento localmente finito, si osserva immediatamente che tutte le sommatorie considerate nell'espressione precedente sono ben definite; si vede inoltre che v è una mappa continua e localmente lipschitziana. Risulta poi chiaro che $v(u)$ è dato dalla combinazione convessa di elementi che soddisfano le relazioni 4 e 5, infatti ogni funzione ρ_i si annulla fuori da $W_{\bar{u}_i}$. Si giunge così agevolmente a mostrare che le relazioni 2 e 3 sono soddisfatte per ogni $u \in Y$. ■

Nella dimostrazione precedente si utilizzano le nozioni di spazio paracompatto e di raffinamento di una copertura localmente finito per le quali si rimanda a [2], pagg. 67-69.

Osservazione 103 *Se supponiamo che X sia uno spazio di Hilbert e che f sia differenziabile, il gradiente ∇f di f definito dalla seguente uguaglianza*

$$(\nabla f(u), w) = \langle f'(u), w \rangle, \quad w \in X$$

(dove $(,)$ indica il prodotto scalare in X) è naturalmente un campo vettoriale pseudogradiente per f .

Bibliografia

- [1] ALI, F.M., Technique for Solving Multiobjective Nonlinear Programming Using Differential Equations Approach, *Journal of Information & Optimization Sciences*, Vol.18 (1997), pp. 351-357.
- [2] ALIPRANTIS, C.D., BORDER, K.C., "Infinite Dimensional Analysis:A Hitchhiker's Guide", *Studies in Economic Theory*, Vol. 4, Springer-Verlag, New York, 1994.
- [3] AUBIN, J.P., CELLINA A., "Differential Inclusions", *Grundlehren der mathematischen Wissenschaften*, Vol. 264, Springer-Verlag, Berlin Heidelberg, 1984.
- [4] AUBIN, J.P., FRANKOWSKA, H., "Set-Valued Analysis", *Systems & Control: Foundations and Applications*, Vol. 2, Birkhäuser; Boston, 1990.
- [5] ABRAHAM, R.H., MARSDEN, J.E., RATIU, T., "Manifolds, Tensor Analysis, and Applications" (second edition), *Applied Mathematical Sciences*, Vol. 75, Springer-Verlag, New York, 1988.
- [6] BATES, R., BOSIO, T., FONTANA, R., HATJIPULIDIS, A, HATZIS, K., LANDSHEERE, E., MASCOLO, J., PISTONE, G., PRONZATO, L., WYNN, H.P., (CE)²: Computer Experiments for Concurrent Engineering, presentato al convegno "ECCOMAS 2000", Barcellona, 11-14 settembre 2000.
- [7] BAZARAA, M.S., SHETTY, C.M., "Foundations of Optimization" - *Lecture Notes in Economics and Mathematical Systems*, Vol.122; Springer-Verlag, Berlin, 1976.
- [8] BERNARDI, F., Il problema della selezione del portafoglio: risoluzione con metodi di ottimizzazione multiobiettivo, preprint, Università degli Studi di Ferrara, Dipartimento di Matematica, 2001.
- [9] BLUME, L., SIMON C.P., "Mathematics for Economists", W.W. Norton & C., New York, 1994.

- [10] BOREL, E., The theory of play and of integral equations with skew symmetric kernels, *Comptes Rendus de l'Academie de Sciences de Paris*, Vol. **173** (1921), pp. 1304-1308.
- [11] BORTOT, P., MAGNANI, U., OLIVIERI, G., ROSSI, F.A., TORRIGIANI, M., "Matematica Finanziaria" (seconda edizione), Monduzzi Editore, Bologna, 1998.
- [12] CANINO, A., DEGIOVANNI M., Nonsmooth Critical Point Theory and Quasilinear Elliptic Equations, (pp. 1-50) in "Topological Methods in Differential Equations and Inclusions" (Montreal, PQ, 1994), A. Granas, M. Frigon, G. Sabidussi eds., *NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci.*, Vol. **472**, Kluwer Academic Publishers, Dordrecht, 1995.
- [13] CELLINA, A., A theorem on the approximation of compact multi-valued mappings, *Rendiconti dell'Accademia Nazionale dei Lincei*, Vol. **47** (1969), pp.429-433.
- [14] CHANG, K.C., "Infinite-Dimensional Morse Theory and Multiple Solution Problems", *Progress in Nonlinear Differential Equations and their Applications*, Vol. **6**, Birkhäuser; Boston, 1993.
- [15] CILIGOT-TRAVAIN, M., On Lagrange-Kuhn-Tucker Multipliers for Pareto Optimization Problem, *Numerical Functional Analysis and Optimization*, Vol. **15** (1994), pp.689-693.
- [16] CASTAGNOLI, E., PECCATI, L., "Introduzione alla selezione del portafoglio", Cooperativa di cultura L. Milani, Milano, 1991.
- [17] DEBREU, G., Valuation Equilibrium and Pareto Optimum, *Proceedings of National Academy of Sciences of U.S.A.*, Vol. **40** (1954), pp.588-592.
- [18] DEGIOVANNI, M., LUCCHETTI R., RIBARSKA N., Critical Point Theory for Vector Valued Functions, preprint n. **461/p**, Dipartimento di Matematica "Francesco Brioschi", Politecnico di Milano, giugno 2001.
- [19] DEIMLING, K., "Nonlinear Functional Analysis", Springer-Verlag, Berlin Heidelberg, 1985.
- [20] DEIMLING, K., "Multivalued Differential Equations", *de Gruyter Series in Nonlinear Analysis and Applications*, Vol. **1**, de Gruyter, Berlin, 1992.
- [21] EDGEWORTH, F.I., "Mathematical Physics", Kegan Paul, London, 1881.
- [22] EHRGOTT, M., "Multicriteria Optimization", *Lecture Notes in Mathematical Systems and Economics*, Vol. **491**, Springer-Verlag, Berlin, 2000.

- [23] ESCHENAUER, H., KOSKI, J., OSYCZKA, A., "Multicriteria Design Optimization", Springer-Verlag, Berlin, 1990.
- [24] FLIEGE, J., SVAITER, B.F., Steepest Descent Methods for Multicriteria Optimization, *Mathematical Methods of Operations Research*, Vol. 51 (2000), pp. 479-494.
- [25] GIANNESI, F., Some Remarks on Mimimum Principles, (pp. 75-103) in "Advances in Convex Analysis and Global Optimization", Hadjisavvas, P.M. Pardalos eds., *Nonconvex Optimization and its Applications*, Vol. 54, Kluwer Academic Publishers, Dordrecht, 2001.
- [26] GOLUBITSKY, M., GUILLEMIN, V., "Stable Mappings and Their Singularities", *Graduate Texts in Mathematics*, Vol. 14, Springer-Verlag, New York, 1973 (terza edizione riveduta e corretta 1986).
- [27] GRAÑA DRUMMOND, L.M., SVAITER, B.F., A Steepest Descent Method for Vector Optimization, preprint, luglio 2001.
- [28] GUERRAGGIO, A., MOLHO, E., ZAFFARONI, A., On the Notion of Proper Efficiency in Vector Optimization, *Journal of Optimization Theory and Applications*, Vol. 82 (1994), pp. 1-21.
- [29] HABETS, P., LALOY, M., ROUCHE, N., "Stability Theory by Liapunov's Direct Method", *Applied Mathematical Sciences*, Vol. 22, Springer-Verlag, New York, 1977.
- [30] HIRSCH, M.W., SMALE, S., "Differential Equations, Dynamical Systems, and Linear Algebra", *Pure and Applied Mathematics*, Vol. 60, Academic Press, San Diego, 1974.
- [31] JAHN, J., "Mathematical Vector Optimization in Partially Ordered Linear Spaces", *Methoden und Verfahren der mathematischen Physik*, Vol. 31, Verlag Peter Lang, Frankfurt am Main, 1986.
- [32] JAHN, J., Vector Optimization: Theory, Methods and Application to Design Problems in Engineering, (pp.129-150), in "Modern Methods of Optimization", W.Krabs, J.Zowe eds., *Lecture Notes in Economics and Mathematical Systems*, Vol. 378, Spriger-Verlag, Berlin, 1992.
- [33] JAMESON, G., "Ordered Linear Spaces", *Lecture Notes in Mathematics*, Vol.141, Springer- Verlag, Berlin, 1970.
- [34] KOOPMANS, T.C., Analysis of Production as an Efficient Combination of Activities, (pp.33-97) in "Activity Analysis of Production and Allocation", T.C. Koopmans ed., *Cowles Commission Monograph*, Vol. 13, John Wiley and Sons, New York, 1951.

- [35] KUHN, H.W., TUCKER A.W., Nonlinear Programming, (pp.481-492), in "Proceedings of the 2nd Berkeley Symposium on Mathematical Statistics and Probability", Berkeley, California, 1951.
- [36] LORENZI A., "Analisi Matematica" (3 volumi), CLUP, Milano, 1987.
- [37] LUC, D.T., "Theory of Vector Optimization", *Lecture Notes in Economics and Mathematical Systems*, Vol. **319**, Spriger-Verlag, Berlin, 1989.
- [38] LUCCHETTI, R.E., Ottimizzazione convessa, ipertopologie e problemi ben posti, *Studi Matematici*, quaderno n.53, Istituto di Metodi Quantitativi - Università "L.Bocconi", Milano, 1999.
- [39] MADERNA, C., SOARDI, P.M., "Lezioni di Analisi Matematica", CLUP, Milano, 1987.
- [40] MALIVERT, C., A Descent Method for Pareto Optimization, *Journal of Mathematical Analysis and Applications*, Vol. **88** (1982), pp. 610-631.
- [41] MANGASARIAN, O.L., "Nonlinear Programming", *Classics in Applied Mathematics*, Vol. **10**, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, 1994 (ristampa dell'edizione originale del 1969).
- [42] MARZOLLO, A., PASCOLETTI, A., SERAFINI, P., Differential Techniques for Cone Optimality and Stability, (pp.351-363), in "New Trends in Dynamic System Theory and Economics", M. Aoki and A. Marzollo eds., Academic Press, 1979.
- [43] MARZOLLO, A., PASCOLETTI, A., SERAFINI, P., Genericity and Singularities in Vector Optimization, (pp.201-213), in "Recent Developments in Variable Structure Systems, Economics and Biology", R.R. Mohler and A. Ruberti eds., *Lecture Notes in Economics and Mathematical Systems*, Vol. **162**, Spriger-Verlag, Berlin, 1978.
- [44] MAS-COLELL, A., "The theory of general economic equilibrium" (A differentiable approach), *Econometric Society monographs*, Vol. **9**, Cambridge University Press, New York, 1985.
- [45] MAWHIN, J., WILLEM, M., "Critical Point Theory and Hamiltonian Systems", *Applied Mathematical Sciences*, Vol. **74**, Springer-Verlag, New York, 1989.
- [46] MIETTINEN, K., "Nonlinear Multiobjective Optimization", *International Series in Operations Research and Management Science*, Vol. , Kluwer Academic Publishers, Dordrecht, 1999.

- [47] MIGLIERINA, E., Characterization of Solutions of Multiobjective Optimization Problems, *Rendiconti del Circolo Matematico di Palermo*, Vol. 50 (2001), pp. 153-164.
- [48] MIGLIERINA E., MOLHO E., Scalarization and its Stability in Vector Optimization, *Quaderni del Dipartimento di Economia*, n. 1/2001, Università dell'Insubria; accettato per la pubblicazione su *Journal of Optimization Theory and Applications* (ottobre 2001).
- [49] MILNOR, J., "Topology from the Differentiable Viewpoint", *Princeton Landmarks in Mathematics*, Princeton University Press, Princeton, 1997 (ristampa dell'edizione originale del 1965).
- [50] MILNOR, J., "Morse Theory", Princeton University Press, Princeton, 1963.
- [51] NAKAYAMA, H., SAWARAGI, Y. AND TANINO, T., "Theory of Multiobjective Optimization", *Pure and Applied Mathematics*, Vol. 176, Academic Press, New York and London, 1985.
- [52] PALAIS, R.S., Lusternik-Schnirelman Theory on Banach Manifolds, *Topology*, Vol. 5 (1966), pp. 115-132.
- [53] PARETO, V., "Manuale di Economia Politica", Società Editrice Libreria, Milano, 1906.
- [54] ROCKAFELLAR, R.T., WETS, R.J.-B., "Variational Analysis", *Grundlehren der mathematischen Wissenschaften*, Vol. 317, Springer-Verlag, Berlin Heidelberg, 1998.
- [55] ROUX, D., "Lezioni di Analisi Superiore", Masson, Milano, 1992.
- [56] SIMON, C.P., TITUS, C., Characterization of Optima in Smooth Pareto Economic Systems, *Journal of Mathematical Economics*, Vol. 2 (1975), pp.297-330.
- [57] SMALE, S., Global Analysis and Economics I: Pareto Optimum and a Generalization of Morse Theory, (pp.531-534), in "Dynamical Systems" (Proc.Sympos., Univ. of Bahia, Salvador, 1971), M. Peixoto ed., Academic Press, New York, 1973.
- [58] SMALE, S., Global Analysis and Economics IIA (Extension of a Theorem of Debreu), *Journal of Mathematical Economics*, Vol. 1 (1974), pp. 1-14.
- [59] SMALE, S., Global Analysis and Economics III (Pareto Optima and Price Equilibria), *Journal of Mathematical Economics*, Vol. 1 (1974), pp. 107-117.

- [60] SMALE, S., Global Analysis and Economics IV (Finiteness and stability of equilibria with general consumption sets and production), *Journal of Mathematical Economics*, Vol. 1 (1974), pp. 119-127.
- [61] SMALE, S., Global Analysis and Economics V (Pareto theory with constraints), *Journal of Mathematical Economics*, Vol. 1 (1974), pp. 213-221.
- [62] SMALE, S., Sufficient Condition for an Optimum, (pp.287-292), in "Dynamical Systems: Warwick 1974" (Proceedings of a Symposium on Applications of Topology and Dynamical Systems, University of Warwick, 1973-74), A. Manning ed., *Lecture Notes in Mathematics*, Vol. 468, Springer-Verlag, Berlin, 1975.
- [63] SMALE, S., Optimizing Several Functions, (pp.69-75), in "Manifolds - Tokyo 1973" (Proceedings of International Conference on Manifolds and Related Topics in Topology), University Tokyo Press, Tokyo, 1975.
- [64] SMALE, S., Global Analysis and Economics VI (Geometric Analysis of Pareto optima and price equilibria under classical hypotheses), *Journal of Mathematical Economics*, Vol. 3 (1976), pp. 1-14.
- [65] SMALE, S., Exchange Processes with Price Adjustment, *Journal of Mathematical Economics*, Vol. 3 (1976), pp. 211-226.
- [66] SMALE, S., Global Analysis and Economics, (pp.331-370), in *Handbook of Mathematical Economics*, Vol. 1, K.J. Arrow and M.D. Intriligator eds., North Holland, New York, 1981.
- [67] TAKAYAMA, A., "Mathematical Economics" (second edition), Cambridge University Press, Cambridge, 1985.
- [68] VON NEUMANN, J., On the Theory of Parlor Games, *Mathematische Annalen*, Vol. 100 (1928), pp.295-320.
- [69] WAN, Y.-H., On Local Pareto Optima, *Journal of Mathematical Economics*, Vol. 2 (1975), pp. 35-42.
- [70] WHITE, D.J., Epsilon-Dominating Solutions in Mean-Variance Portfolio Analysis, *European Journal of Operational Research*, Vol. 105 (1998), pp.457-466.
- [71] WIERZBICKI, A.P., The Use of References Objective in Multiobjective Optimization, (pp.468-486) in "Multiple Criteria Decision Making Theory and Applications", G. Fandel and T. Gal eds., *Lecture Notes in Economics and Mathematical Systems*, Vol. , Springer-Verlag, Berlin, 1980.

- [72] YOSHIKAWA, T., "Stability Theory by Liapunov's Second Method", The Mathematical Society of Japan, Tokyo, 1966.
- [73] YU, P.-L., "Multiple-Criteria Decision Making", *Mathematical Concepts and Methods in Science and Engineering*, Vol. 30, Plenum Press, New York, 1985.