Appunti delle Lezioni di Ottimizzazione
Programmazione Multiobiettivo

a cura di G. Liuzzi*

a.a. 1999-2000

1 Introduzione

Un problema matematico di ottimizzazione può essere definito come la minimizzazione o massimizzazione di una funzione a valori reali su un insieme specificato. L'importanza di questo modello matematico deriva ovviamente dal fatto che molti problemi reali vengono affrontati facendovi ricorso. Tuttavia quasi ogni problema reale di ottimizzazione è caratterizzato dalla presenza contemporanea di più obiettivi, cioè funzioni a valori reali da massimizzare e/o minimizzare, tipicamente in contrasto tra loro.
Nel seguito considereremo il seguente problema di ottimizzazione multiobiettivo:

$$\begin{array}{ll} \min & (f_1(x) f_2(x) \dots f_k(x))^\top \\ & x \in \mathcal{F} \end{array} \tag{MOP}$$

ove $k \geq 2$ e $f_i : \mathbb{R}^n \to \mathbb{R}$, per $i = 1, \dots, k$.
D'ora in avanti chiameremo $\mathbb{R}^k$ *spazio degli obiettivi* e $\mathbb{R}^n$ *spazio delle variabili di decisione*. Un vettore $x \in \mathbb{R}^n$ sarà pertanto un *vettore di decisioni* mentre $z \in \mathbb{R}^k$ un *vettore di obiettivi*. Indicheremo, inoltre, con $f(x)$ il vettore delle funzioni obiettivo $(f_1(x) f_2(x) \dots f_k(x))^\top$ e con $\mathcal{Z} = f(\mathcal{F})$ l'immagine della regione ammissibile $\mathcal{F}$ nello spazio degli obiettivi (vedi figura) e cioè

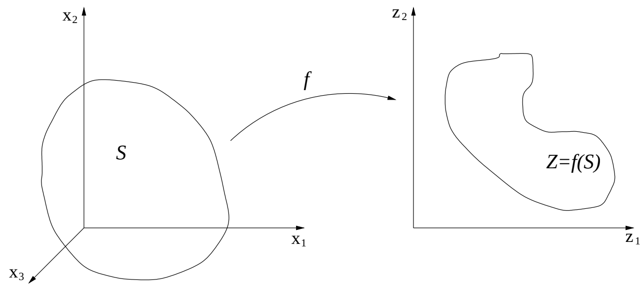
$$\mathcal{Z} = f(\mathcal{F}) = \left\{ z \in \mathbb{R}^k : \ \exists x \in \mathcal{F}, \ z = f(x) \right\}.$$

In particolare diremo che un vettore di obiettivi $z \in \mathbb{R}^k$ è ammissibile quando risulti $z \in \mathcal{Z}$.
Definiamo , inoltre, il vettore *ideale* degli obiettivi z^{id} come il vettore di componenti *Vettore ideale*

$$z_i^{id} = \begin{array}{l} \min \ f_i(x) \\ x \in \mathcal{F} \end{array}$$

Quello che vogliamo fare è minimizzare tutte le funzioni obiettivo simultaneamente. Se non ci fossero conflitti tra le funzioni obiettivo, una soluzione banale al problema sarebbe quella ottenibile risolvendo separatamente k problemi di ottimizzazione (uno per ogni funzione obiettivo) ottenendo quindi come soluzione proprio il vettore ideale z^{id}. Non sarebbe pertanto necessario applicare nessuna tecnica specifica di soluzione. Per evitare il sorgere di tale caso banale, supporremo che $z^{id} \notin \mathcal{Z}$. Questo significa assumere che le funzioni $f_1(x), f_2(x), \dots, f_k(x)$ siano, almeno in parte, in contrasto tra loro.

*`liuzzi@dis.uniroma1.it`, `http://www.dis.uniroma1.it/~liuzzi`



Con $\mathcal{B}(x,\delta) = \{y \in \mathbb{R}^n | \ \|x-y\| < \delta\}$ indicheremo la sfera aperta di centro $x \in \mathbb{R}^n$ e raggio $\delta > 0$. Dato un insieme A, indicheremo con ∂A la frontiera di A, con $\mathring{A}$ l'interno e con $\bar{A}$ la chiusura di A.
Dati due insiemi A e B definiamo l'insieme differenza $A \backslash B$ come quell'insieme costituito da tutti e soli gli elementi di A che non appartengono a B, ovvero

$$A \backslash B = \big\{x \in A : x \notin B\big\}.$$

Dato un vettore $b \in \mathbb{R}^n$ ed un insieme $A \subset \mathbb{R}^n$, definiamo *traslazione di A rispetto a b* l'insieme

$$b + A = \big\{x \in \mathbb{R}^n : x = b + a \text{ per ogni } a \in A\big\}.$$

In maniera del tutto analoga definiamo l'insieme

$$b - A = \big\{x \in \mathbb{R}^n : x = b - a \text{ per ogni } a \in A\big\}.$$

Se I è un insieme di indici e v un vettore, con v_I indicheremo il sottovettore costituito dalle componenti di v con indici in I. Sia infine $\mathbb{R}^k_+$ l'ortante positivo dello spazio degli obiettivi cioè

$$\mathbb{R}^k_+ = \big\{z \in \mathbb{R}^k : \ z \geq 0\big\}.$$

2 Ottimalità secondo Pareto

Prima di procedere, è necessario definire con chiarezza cosa si intende per soluzione ottima di un problema di programmazione multiobiettivo. La definizione che adottiamo e che è riportata nel seguito, è stata proposta per la prima volta da Edgeworth nel 1881 e successivamente ripresa da Vilfredo Pareto nel 1896 [2] che la approfondì ulteriormente.

Definizione 2.1 *Dati due vettori $z^1, z^2 \in \mathbb{R}^k$, diciamo che z^1 domina z^2 secondo Pareto ($z^1 \leq_P z^2$) quando risulta*

$$\begin{array}{ll} z^1_i \leq z^2_i & \text{per ogni } i = 1, 2, \ldots, k \text{ e} \\ z^1_j < z^2_j & \text{per almeno un indice } j \in \{1, \ldots, k\}. \end{array}$$

La relazione binaria $\leq_P$ è un ordinamento parziale (non riflessivo) nello spazio delle k-uple di numeri reali. Sfruttando la relazione $\leq_P$ possiamo dare la definizione di ottimalità secondo Pareto.

Definizione 2.2 *Un vettore di decisioni $x^\star \in \mathcal{F}$ è un ottimo di Pareto se non esiste un'altro vettore $x \in \mathcal{F}$ tale che:* *Ottimo di Pareto*

$$f(x) \leq_P f(x^\star).$$

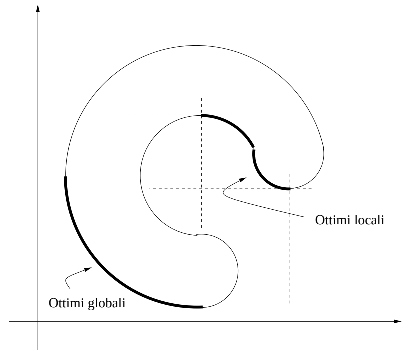


Figure 1: Ottimi locali e globali di Pareto.

Corrispondentemente diremo che un vettore di obiettivi $z^\star \in \mathcal{Z}$ è ottimo secondo Pareto quando non esiste un altro vettore $z \in \mathcal{Z}$ tale che

$$z \leq_P z^\star.$$

Quindi se ci troviamo in un punto ottimo secondo Pareto e vogliamo ulteriormente diminuire il valore di una o più funzioni obiettivo dobbiamo essere disposti ad accettare un conseguente aumento in alcune (o tutte) le rimanenti funzioni del problema. In tal senso possiamo affermare che, nello spazio degli obiettivi, gli ottimi di Pareto sono punti di *equilibrio* che si trovano sulla frontiera dell'insieme $\mathcal{Z}$.

Definizione 2.3 *Diciamo frontiera efficiente l'insieme degli ottimi di Pareto del problema (MOP)*

Frontiera efficiente

La definizione di ottimo secondo Pareto è ovviamente, una definizione di ottimo *globale* dato che si richiede la validità di una certa proprietà su *tutto* l'insieme ammissibile del problema (MOP). È evidentemente possibile, però, dare una definizione di ottimo **locale** secondo Pareto.

Definizione 2.4 *Un vettore di decisioni $x^\star \in \mathcal{F}$ è un ottimo locale di Pareto se esiste un numero $\delta > 0$ tale che $x^\star$ è ottimo secondo Pareto in $\mathcal{F} \cap \mathcal{B}(x^\star, \delta)$.*

Ottimo locale di Pareto

In figura 1 si riportano gli ottimi globali e locali di Pareto per un insieme $\mathcal{Z}$.
Ovviamente, ogni ottimo globale è anche un ottimo locale di Pareto. Il viceversa è vero solo se facciamo alcune ipotesi sulla struttura del problema (MOP). Se (MOP) è convesso cioè se

i- l'insieme ammissibile $\mathcal{F}$ è convesso e

ii- tutte le funzioni obiettivo $f_i(x)$ (con $i = 1, 2, \ldots, k$) sono convesse.

allora si può dimostrare che ogni ottimo locale di Pareto è anche un ottimo globale.

Teorema 2.5 *Sia (MOP) un problema di programmazione multiobiettivo convesso. Allora, ogni ottimo locale di Pareto è anche un ottimo globale.*

Equivalenza tra ottimi locali e globali di Pareto

Dim. Sia $x^\star \in \mathcal{F}$ un ottimo locale secondo Pareto del problema (MOP). Dalla definizione di ottimo locale segue che esiste un numero $\delta > 0$ tale che $\nexists\, x \in \mathcal{F} \cap \mathcal{B}(x^\star, \delta)$ t.c.

$$\begin{array}{ll} f(x) \leq f(x^\star) & e \\ f_j(x) < f_j(x^\star) & \text{per qualche } j \in \{1, 2, \ldots, k\}. \end{array}$$

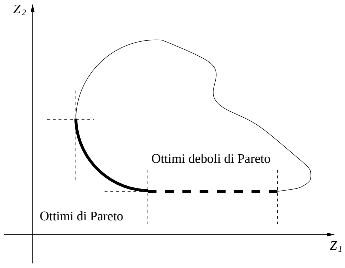


Figure 2: Gli ottimi di Pareto sono individuati dalla linea doppia a tratto continuo. La linea tratteggiata individua invece gli ottimi deboli che non sono ottimi di Pareto.

Ragioniamo per contraddizione e supponiamo che $x^\star$ non sia un ottimo globale. Se $x^\star$ non è ottimo globale allora esiste un vettore $x^\circ \in \mathcal{F}$ tale che

$$\begin{array}{ll} f(x^\circ) \leq f(x^\star) & e \\ f_j(x^\circ) < f_j(x^\star) & \text{per qualche } i \in \{1, 2, \dots, k\}. \end{array}$$

Poiché $\mathcal{F}$ è convesso, per ogni $\beta \in (0, 1)$, $\hat{x} = \beta x^\circ + (1 - \beta)x^\star$ è ammissibile. È inoltre possibile trovare un valore di β tale che $\hat{x} \in \mathcal{F} \cap \mathcal{B}(x^\star, \delta)$.
Dalle ipotesi di convessità sulle funzioni obiettivo segue che

$$f(\hat{x}) \leq \beta f(x^\circ) + (1 - \beta)f(x^\star) \leq f(x^\star).$$

Siccome $x^\star$ è un ottimo locale di Pareto, si deve necessariamente avere $f(\hat{x}) = f(x^\star)$, dato che per nessun indice $i \in \{1, 2, \dots, k\}$ può essere $f_i(\hat{x}) < f_i(x^\star)$. Quindi possiamo scrivere

$$f(\hat{x}) = f(x^\star) \leq \beta f(x^\circ) + (1 - \beta)f(x^\star),$$

ovvero $f(x^\star) \leq f(x^\circ)$. Ricordando, però, l'ipotesi fatta si dovrebbe contemporaneamente avere $f_j(x^\star) > f_j(x^\circ)$ per qualche indice j, il che è ovviamente una contraddizione. □

La definizione di ottimo secondo Pareto può essere leggermente indebolita ottenendo così la definizione di ottimo *debole* secondo Pareto.

Definizione 2.6 *Un vettore $x^\star \in \mathcal{F}$ è un ottimo di Pareto debole per il problema (MOP) se non esiste un punto $x \in \mathcal{F}$ tale che* *Ottimalità debole secondo Pareto*

$$f(x) < f(x^\star).$$

Ovviamente, l'insieme degli ottimi secondo Pareto è contenuto nell'insieme degli ottimi deboli di Pareto.
Anche qui, come gia si era fatto per gli ottimi di Pareto, è possibile dare una definizione di ottimo locale debole.

Definizione 2.7 *Un vettore di decisioni $x^\star \in \mathcal{F}$ è un ottimo locale debole (secondo Pareto) se esiste un numero $\delta > 0$ tale che $x^\star$ è ottimo debole di Pareto in $\mathcal{F} \cap \mathcal{B}(x^\star, \delta)$.* *Ottimo locale debole di Pareto*

In figura 2 sono riportati, per maggiore chiarezza, ottimi e ottimi deboli secondo Pareto.
Anche per gli ottimi deboli secondo Pareto vale una proprietà analoga a quelle espressa dal teorema 2.5 e cioè se il problema (MOP) è convesso ogni ottimo locale (debole) è anche ottimo globale (debole) di Pareto.

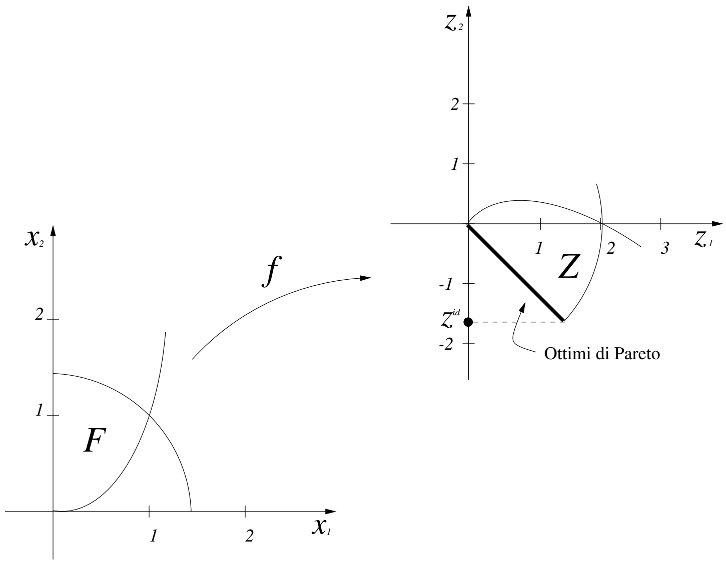


Figure 3: Esempio

2.1 Esercizio

Sia dato il seguente problema di ottimizzazione vettoriale:

$$\begin{aligned} &\min\ (x_1 + x_2,\ x_1 - x_2)^\top \\ &\begin{cases} x_1^2 + x_2^2 \leq 2 \\ x_1^2 - x_2 \leq 0 \\ x_1 \geq 0 \end{cases} \end{aligned} \tag{1}$$

In questo caso, in cui il vettore delle decisioni ha dimensione due, è semplicissimo tracciare la regione ammissibile nello spazio delle variabili di decisione. Inoltre, poiché il vettore degli obiettivi ha dimensione due e le funzioni obiettivo sono lineari, possiamo determinare la rappresentazione grafica di $\mathcal{Z} = f(\mathcal{F})$. Vediamo nel dettaglio come fare.
Le relazioni che dobbiamo prendere in considerazione sono le seguenti due trasformazioni lineari:

$$\begin{cases} z_1 = f_1(x_1, x_2) = x_1 + x_2 \\ z_2 = f_2(x_1, x_2) = x_1 - x_2 \end{cases} \tag{2}$$

La regione ammissibile degli obiettivi è ottenibile da quella delle decisioni mediante rotazione e scalatura, come messo in evidenza dalla figura 3
Sempre per via grafica, è facile risolvere i sottoproblemi ad un solo obiettivo associati al problema (1) che sono:

$$\begin{aligned} &z_1^{id} = \min\ x_1 + x_2 \\ &\begin{cases} x_1^2 + x_2^2 \leq 2 \\ x_1^2 - x_2 \leq 0 \\ x_1 \geq 0 \end{cases} \end{aligned} \qquad \text{e} \qquad \begin{aligned} &z_2^{id} = \min\ x_1 - x_2 \\ &\begin{cases} x_1^2 + x_2^2 \leq 2 \\ x_1^2 - x_2 \leq 0 \\ x_1 \geq 0 \end{cases} \end{aligned}$$

$$x^{1\star} = (0, 0)^\top \qquad\qquad x^{2\star} = (0, \sqrt{2})^\top$$

ricavando, in tal modo, il vettore ideale degli obiettivi $z^{id} = (0, -\sqrt{2})^\top$. Notiamo subito che non essendo z^{id} contenuto in $\mathcal{Z} = f(\mathcal{F})$, il problema vettoriale non è risolvibile semplicemente minimizzando separatamente le due funzioni obiettivo.
È altresì facile individuare, in figura 3, la frontiera efficiente secondo Pareto dell'insieme $\mathcal{Z}$. Come ci aspettavamo, essendo il problema (1) convesso, tutti gli ottimi locali di Pareto sono anche globali.

3 Punti Efficienti e Punti Dominati

È bene ribadire il fatto che nell' ambito della programmazione vettoriale esistono vari modi di definire il concetto di ottimalità. Questo dipende dal fatto che nello spazio delle k-uple di numeri reali non esiste nessun ordinamento completo. D'altra parte, esistono molteplici modi di definire un ordinamento parziale tra le k-uple. Ciascuno di questi ordinamenti parziali induce una differente definizione di ottimalità.
In sezione 2 abbiamo visto come, sfruttando la relazione $\leq_P$ sia possibile definire un ordinamento parziale su $\mathbb{R}^k$ e conseguentemente il concetto di ottimo secondo Pareto. Tuttavia, è possibile dotare $\mathbb{R}^k$ di altri ordinamenti parziali e quindi dare altre definizioni di ottimalità per il problema (MOP). In particolare vedremo come, sfruttando il concetto di *cono*, sia possibile generalizzare le definizioni di ottimo e ottimo debole secondo Pareto.

Definizione 3.1 *Un vettore $y \in \mathbb{R}^n$ è combinazione conica di m vettori $\{x^1, x^2, \ldots, x^m\}$ di $\mathbb{R}^n$ quando è possibile trovare m numeri reali $\lambda_1, \lambda_2, \ldots, \lambda_m$ tali che:*

$$\sum_{i=1}^{m} \lambda_i x^i = y$$
$$\lambda_i \geq 0 \quad \forall\ i = 1, 2, \ldots, m$$

Definizione 3.2 *Un insieme $D \in \mathbb{R}^k$ è un cono se la combinazione conica dei vettori di un qualsiasi sottoinsieme finito di D appartiene a D.*

Sia, quindi, D un cono nello spazio degli obiettivi $\mathbb{R}^k$. D induce un ordinamento parziale nello spazio delle k-uple di numeri reali.

Definizione 3.3 *Presi due vettori z^1 e z^2 in $\mathbb{R}^k$, diciamo che z^1* domina *z^2 ($z^1 \leq_D z^2$) se $z^2 - z^1 \in D \backslash \{0\}$.*

Dalla definizione precedente segue che $z^1 \leq_D z^2$ se e solo se $z^2 \in z^1 + (D\backslash\{0\})$ ovvero quando $z^1 \in z^2 - (D\backslash\{0\})$. Ovviamente, essendo l'ordinamento così definito, solamente parziale, esistono certamente coppie di vettori (z^1, z^2) che non sono in relazione tra loro (tali che cioè $z^2 - z^1 \notin D\backslash\{0\}$ e $z^1 - z^2 \notin D\backslash\{0\}$).
A questo punto possiamo dare la definizione di punto efficiente rispetto ad un cono D.

Definizione 3.4 *Un vettore di obiettivi $z^\star \in \mathcal{Z}$ è efficiente rispetto ad un cono D se non esiste alcun vettore $z \in \mathcal{Z}$ tale che $z \leq_D z^\star$ ovvero se e solo se $(z^\star - D\backslash\{0\}) \cap \mathcal{Z} = \phi$.* *Punto efficiente rispetto ad un cono*

Corrispondentemente diciamo che un vettore di decisioni $x^\star \in \mathcal{F}$ è efficiente rispetto ad un cono $D \subset \mathbb{R}^k$ se e solo se non esiste alcun vettore $x \in \mathcal{F}$ tale che $f(x) \leq_D f(x^\star)$. In figura 4 sono riportati i punti efficienti rispetto ad un cono D.
Notiamo che ponendo $D = \mathbb{R}^k_+$ (ortante positivo dello spazio degli obiettivi) si ottiene esattamente la definizione di ottimalità secondo Pareto vista in precedenza.

Definizione 3.5 *Un vettore di decisioni $x^\star \in \mathcal{F}$ è debolmente efficiente rispetto ad un cono $D \subset \mathbb{R}^k$ se non esiste alcun vettore $x \in \mathcal{F}$ tale che risulti $f(x^\star) - f(x) \in \overset{\circ}{D}$.* *Efficienza debole*

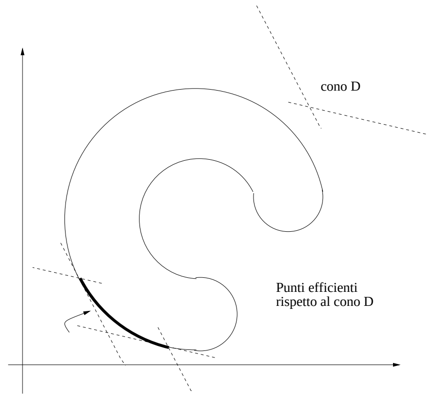


Figure 4: Frontiera efficiente

In maniera analoga avremmo potuto definire $x^\star \in \mathcal{F}$ debolmente efficiente rispetto a D quando risulti $(f(x^\star) - \overset{\circ}{D}) \cap \mathcal{Z} = \phi$.
Scegliendo $D = \mathbb{R}^k_+$, ritroviamo nuovamente la definizione di ottimo debole secondo Pareto.
Anche per i punti efficienti valgono considerazioni analoghe a quelle fatte per gli ottimi di Pareto. Quindi, se le funzioni obiettivo $f_i(x)$ sono continue, essi si trovano sicuramente sulla frontiera della regione amissibile $\mathcal{F}$.

4 Condizioni di Ottimalità

Nelle sezioni precedenti abbiamo dato delle definizioni fondamentali della programmazione multiobiettivo. In particolare, dato che lo spazio delle k-uple di numeri reali è solo parzialmente ordinato, abbiamo dovuto definire cosa si intende per minimo di un vettore di funzioni.
Abbiamo altresì visto che le definizioni di ottimo e di ottimo debole secondo Pareto non sono le uniche che è possibile dare e che, anzi, queste sono un caso particolare di una definizione più generale, quella di punto efficiente rispetto ad un cono.
Quello che dobbiamo fare ora, è dare una caratterizzazione analitica dei punti di ottimo secondo Pareto. Come vedremo, tutte le condizioni di ottimo per la programmazione multiobiettivo, comprendono come caso particolare quelle per la programmazione nonlineare (con una sola funzione obiettivo). Per ulteriori approfondimenti sull'argomento di questa sezione si rimanda al testo [3] citato in bibliografia.
Nel seguito consideriamo un problema in cui $\mathcal{F}$ è definito da vincoli di disuguaglianza; cioè:

$$\begin{aligned} \min \quad & f(x) \\ & g(x) \le 0 \end{aligned} \tag{P}$$

ove $f : \mathbb{R}^n \to \mathbb{R}^k$ $(k \ge 2)$ e $g : \mathbb{R}^n \to \mathbb{R}^m$ sono funzioni continuamente differenziabili ed $\mathcal{F}$ assume la seguente struttura:

$$\mathcal{F} = \{x \in \mathbb{R}^n : \; g(x) \le 0\}.$$

Indichiamo con

$$I_0(x) = \{i : g_i(x) = 0\}$$

l'insieme degli indici dei vincoli attivi nel punto x. Sia, inoltre, $L : \mathbb{R}^{n \times k \times m} \to \mathbb{R}$ così definita

$$L(x, \lambda, \mu) = \lambda^\top f(x) + \mu^\top g(x),$$

la funzione Lagrangiana associata al problema (P).
Per completezza richiamiamo brevemente il teorema dell'alternativa di Gordan che utilizzeremo nel dare le condizioni necessarie di ottimo secondo Pareto.

Teorema 4.1 (Gordan) *Uno ed uno solo dei seguenti due sistemi ammette soluzione:*

$$Bz < 0, \qquad \begin{cases} B^\top y = 0 \\ y \geq 0, \quad y \neq 0 \end{cases}$$

dove $B \in \mathbb{R}^{s \times n}$, $z \in \mathbb{R}^n$ e $y \in \mathbb{R}^s$. □

4.1 Condizioni di Fritz-John

Definizione 4.2 (direzione ammissibile) *Dato un punto $\bar{x} \in \mathcal{F}$, una direzione ammissibile in $\bar{x}$ è un vettore $d \in \mathbb{R}^n$ $(d \neq 0)$ per cui esiste un $\delta > 0$, tale che*

$$\bar{x} + \lambda d \in \mathcal{F} \quad \forall \lambda \in (0, \delta).$$

Quindi se dal punto $\bar{x}$ ci spostiamo di poco lungo la direzione individuata dal vettore d andiamo su punti che sono ancora ammissibili. Indichiamo con

$$C(\bar{x}) = \Big\{ d \in \mathbb{R}^n |\ d \neq 0,\ \bar{x} + \lambda d \in \mathcal{F}\ \forall \lambda \in (0, \delta) \text{ per qualche } \delta > 0 \Big\}$$

l'insieme di tutte le *direzioni ammissibili* in $\bar{x}$.

Definizione 4.3 (direzione efficiente) *Dato un punto $\bar{x} \in \mathcal{F}$, una direzione efficiente in $\bar{x}$ è un vettore $d \in \mathbb{R}^n$ $(d \neq 0)$ per cui esiste un $\delta > 0$, tale che*

$$f(\bar{x} + \lambda d) \leq_P f(\bar{x}) \quad \forall \lambda \in (0, \delta).$$

Pertanto, spostandoci da $\bar{x}$ lungo la direzione individuata dal vettore d e per spostamenti sufficientemente piccoli, siamo certi di trovare punti che migliorano il valore di almeno una funzione obiettivo senza, al tempo stesso, peggiorare il valore delle rimanenti. Indichiamo con

$$F(\bar{x}) = \Big\{ d \in \mathbb{R}^n |\ d \neq 0, f(\bar{x} + \lambda d) \leq_P f(\bar{x}) \quad \forall\ \lambda \in (0, \delta) \text{ e qualche } \delta > 0 \Big\}$$

l'insieme di tutte le *direzioni efficienti* in $\bar{x}$.
Evidentemente, se $\bar{x}$ è un ottimo (locale o globale) di Pareto risulterà $C(\bar{x}) \cap F(\bar{x}) = \phi$ cioè nessuna direzione ammissibile porta a punti tali che $f(x) \leq_P f(\bar{x})$.
Fin qui abbiamo dato una caratterizzazione geometrica dei punti di ottimo secondo Pareto. Quello che adesso vogliamo fare è caratterizzare analiticamente gli ottimi di Pareto. Per fare questo definiamo i seguenti altri due insiemi:

$$\begin{aligned} F_0(\bar{x}) &= \big\{ d \in \mathbb{R}^n |\ \nabla f_i(\bar{x})^\top d < 0,\ \forall\ i = 1, 2, \ldots, k \big\}, \\ G_0(\bar{x}) &= \big\{ d \in \mathbb{R}^n |\ \nabla g_j(\bar{x})^\top d < 0,\ \forall\ j \in I_0(\bar{x}) \big\}. \end{aligned}$$

È facile dimostrare che, per ogni $x \in \mathcal{F}$, risulta $G_0(x) \subseteq C(x)$ e $F_0(x) \subseteq F(x)$ (cfr. [3]) e quindi che:

> *condizione necessaria affinché il punto $\bar{x} \in \mathcal{F}$ sia ottimo di Pareto (locale o globale) è che risulti: $G_0(\bar{x}) \cap F_0(\bar{x}) = \phi$.*

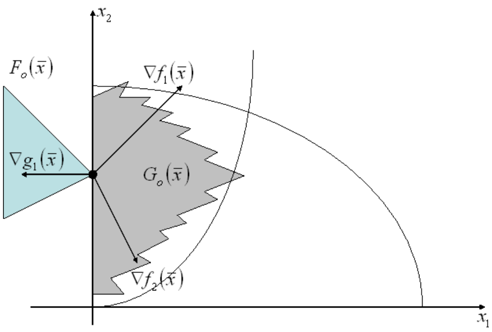
Con queste premesse, possiamo facilmente dimostrare il teorema di Fritz-John per la programmazione multiobiettivo (cfr. [1], [3]).

Teorema 4.4 *Condizione necessaria affinché $\bar{x} \in \mathcal{F}$ sia ottimo secondo Pareto è che esistano dei vettori $\lambda \in \mathbb{R}^k$ e $\mu \in \mathbb{R}^m$ tali che sia soddisfatto il seguente sistema:* *CN di Fritz-John*

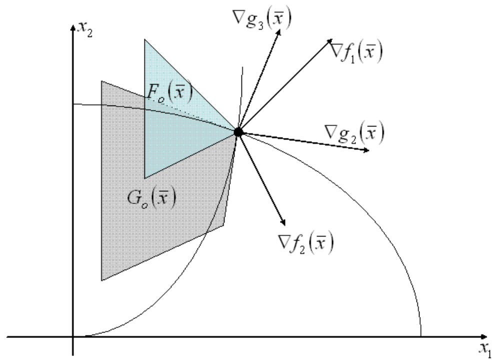
$$\sum_{i=1}^{k} \lambda_i \nabla f_i(\bar{x}) + \sum_{j=1}^{m} \mu_j \nabla g_j(\bar{x}) = 0 \tag{3a}$$

$$\mu^\top g(\bar{x}) = 0, \tag{3b}$$

$$(\lambda, \mu) \geq 0, \quad (\lambda, \mu) \neq (0, 0) \tag{3c}$$



Esempio di punto che soddisfa le condizioni necessarie di ottimalità.



Esempio di punto che NON soddisfa le condizioni necessarie di ottimalità.

Dim. Essendo $\bar{x}$ un ottimo di Pareto del problema (P) risulta necessariamente $G_0(\bar{x}) \cap F_0(\bar{x}) = \phi$ cioè è inammissibile il seguente sistema:

$$\begin{aligned} &\nabla f_i(\bar{x})^\top d < 0 \quad \forall i = 1, 2, \ldots, k \\ &\nabla g_j(\bar{x})^\top d < 0 \quad \forall j \in I_0(\bar{x}). \end{aligned}$$

Ponendo $B = \begin{pmatrix} \nabla f_i(\bar{x})^\top \\ \nabla g_j(\bar{x})^\top \end{pmatrix}_{\substack{i = 1, \ldots, k \\ j \in I_0(\bar{x})}}$ e sfruttando il teorema dell'alternativa di Gordan, sappiamo che esistono dei numeri $\lambda_i \geq 0$ $(i = 1, 2, \ldots, k)$ e $\eta_j \geq 0$ $(j \in I_0(\bar{x}))$ non tutti nulli e tali che

$$\sum_{i=1}^{k} \lambda_i \nabla f_i(\bar{x}) + \sum_{j \in I_0(\bar{x})} \eta_j \nabla g_j(\bar{x}) = 0. \tag{4}$$

Sia ora $\mu \in \mathbb{R}^m$ un vettore così fatto:

$$\mu_i = \begin{cases} \eta_i & i \in I_0(\bar{x}) \\ 0 & i \notin I_0(\bar{x}) \end{cases}$$

Ovviamente, per come è stato definito il vettore μ, risulta $\mu^\top g(\bar{x}) = 0$. Quindi abbiamo dimostrato che esistono dei vettori λ e μ tali che è soddisfatto il sistema (3). □

Corollario 4.5 *Le condizioni del teorema 4.4 sono necessarie anche per l'ottimalità debole (secondo Pareto) di un punto $\bar{x}$.* □

Definizione 4.6 (tripla di FJ) *Una tripla $(x, \lambda, \mu) \in \mathbb{R}^{n \times k \times m}$ è una tripla di Fritz-John* *quando soddisfa il sistema (3) cioè:*

$$
\begin{aligned}
&\nabla_x L(x, \lambda, \mu) = 0\\
&g(x) \le 0\\
&\mu^\top g(x) = 0\\
&(\lambda, \mu) \ge 0 \quad (\lambda, \mu) \ne (0, 0).
\end{aligned}
$$

Definizione 4.7 (punto di FJ) *Un vettore di decisioni $x \in \mathbb{R}^n$ è un punto di FJ se esistono dei vettori $\lambda \in \mathbb{R}^k$ e $\mu \in \mathbb{R}^m$ tali che (x, λ, μ) è una tripla di FJ.*

4.2 Condizioni di Karush-Kuhn-Tucker

In base alle definizioni precedenti e ricordando che il teorema di Gordan dà condizioni necessarie e sufficienti di ammissibilità (o inammissibilità), possiamo dire che sono punti di FJ tutti e soli i vettori $x \in \mathcal{F}$ tali che $F_0(x) \cap G_0(x) = \phi$. In particolare, questa condizione è vera in ogni $\bar{x}$ ammissibile tale che $G_0(\bar{x}) = \phi$, indipendentemente dai valori delle funzioni obiettivo.
Motivati da questa considerazione introduciamo in questo paragrafo le condizioni di KKT con le quali in pratica forziamo i gradienti di alcune funzioni obiettivo ad avere un peso non nullo nell'espressione (3a). A questo proposito vale la seguente proprietà:

Proposizione 4.8 *Sia $\bar{x}$ un punto di FJ per il problema (P) tale che $G_0(\bar{x}) \ne \phi$. Si dimostra che i moltiplicatori di FJ associati alle funzioni obiettivo non possono essere tutti nulli ($\lambda \ne 0$).*

Dim. Ragioniamo per contraddizione e supponiamo che risulti $\lambda = 0$. Sotto tali ipotesi, il teorema 4.4 ci assicura che esiste un vettore $\mu \in \mathbb{R}^m$ tale che

$$
\sum_{j=1}^{m} \mu_j \nabla g_j(\bar{x}) = 0, \tag{5a}
$$

$$
\mu^\top g(\bar{x}) = 0, \tag{5b}
$$

$$
\mu \ge 0, \quad \mu \ne 0. \tag{5c}
$$

Dalla (5b) segue che $\mu_j = 0$ per ogni indice j tale che $g_j(\bar{x}) < 0$ (cioè per ogni $j \notin I_0(\bar{x})$). Possiamo dunque riscrivere il sistema (5) come

$$
\begin{aligned}
&\sum_{j \in I_0(\bar{x})} \mu_j \nabla g_j(\bar{x}) = 0,\\
&\mu_{I_0} \ge 0, \quad \mu_{I_0} \ne 0.
\end{aligned}
$$

Ponendo $B = \left(\nabla g_j(\bar{x})^\top\right)_{j \in I_0(\bar{x})}$ e sfruttando il teorema dell'alternativa di Gordan otteniamo $G_0(\bar{x}) = \phi$, il che è ovviamente assurdo. □

La condizione $G_0(\bar{x}) \ne \phi$ è una condizione sul comportamento dei vincoli del problema (P) nell'intorno del punto di ottimo $\bar{x}$.

Proposizione 4.9 *Condizione sufficiente affinché nel punto $\bar{x} \in \mathcal{F}$ risulti $G_0(\bar{x}) \ne \phi$ è che siano linearmente indipendenti i gradienti dei vincoli attivi.* *LICQ*

Dim. Vedi appendice. □

Definizione 4.10 *Un punto ammissibile $x \in \mathcal{F}$ è un punto di regolarità per i vincoli del problema (P) se in x risulta $G_0(x) \ne \phi$.*

La proposizione 4.9 ci permette di certificare la regolarità di un punto ammissibile e dimostrare (cfr. [1], [3]) il seguente

Teorema 4.11 *Sia $\bar{x}$ un punto ammissibile per il problema (P) e siano linearmente indipendenti i gradienti dei vincoli attivi in $\bar{x}$. Allora, condizione necessaria affinché $\bar{x}$ sia un ottimo di Pareto (locale o globale) è che sia ammissibile il seguente sistema:* *CN di KKT*

$$\sum_{i=1}^{k} \lambda_i \nabla f_i(\bar{x}) + \sum_{j=1}^{m} \mu_j \nabla g_j(\bar{x}) = 0, \tag{6a}$$

$$\mu^\top g(\bar{x}) = 0, \tag{6b}$$

$$(\lambda, \mu) \geq 0, \quad \lambda \neq 0. \tag{6c}$$

Dim. La dimostrazione è lasciata per esercizio al lettore (vedi appendice). □

Corollario 4.12 *Le condizioni del teorema 4.11 sono necessarie anche per l'ottimalità debole (secondo Pareto) di un punto $\bar{x}$.* □

Le condizioni di KKT (come quelle di FJ) sono condizioni solo necessarie di ottimo il che vuol dire che potrebbero essere verificate anche in punti non ottimi secondo Pareto. È tuttavia possibile dare condizioni sufficienti di ottimalità, senza ricorrere all'uso delle derivate seconde, a patto però di fare alcune ipotesi aggiuntive sulla struttura del problema (P).

Teorema 4.13 *Siano le $f_i(x)$ (per ogni $i = 1, 2, \dots, k$) e $g_j(x)$ (per ogni $j = 1, 2, \dots, m$) convesse. Condizione sufficiente affinché un punto $\bar{x} \in \mathcal{F}$ sia ottimo secondo Pareto è che esistano dei vettori di moltiplicatori $\lambda \in \mathbb{R}^k$ e $\mu \in \mathbb{R}^m$ tali che* *CS di ottimo secondo Pareto*

$$\nabla_x L(\bar{x}, \lambda, \mu) = 0, \tag{7a}$$

$$\mu^\top g(\bar{x}) = 0, \tag{7b}$$

$$\lambda > 0, \ \mu \geq 0. \tag{7c}$$

Dim. Siano λ e μ dei vettori tali che sia soddisfatto il sistema (7). Definiamo, sfruttando i moltiplicatori λ_i, la funzione ausiliaria

$$F(x) = \sum_{i=1}^{k} \lambda_i f_i(x).$$

Ovviamente, la funzione $F(x)$ è convessa (combinazione conica di funzione convesse). Sempre dal sistema (7) otteniamo

$$\nabla F(\bar{x}) + \sum_{j=1}^{m} \mu_j \nabla g_j(\bar{x}) = 0,$$

$$\mu^\top g(\bar{x}) = 0, \ \mu \geq 0.$$

In pratica, sono verificate le condizioni sufficienti affinché la funzione $F(x)$ abbia in $\bar{x}$ un minimo globale e quindi risulta $F(\bar{x}) \leq F(x)$ per ogni $x \in \mathcal{F}$.
Supponiamo ora che il punto $\bar{x}$ non sia un minimo di Pareto per il problema (P). Se $\bar{x}$ non è ottimo allora, dalla definizione, deve esistere un punto $x \in \mathcal{F}$ tale che

$$f_i(x) \leq f_i(\bar{x}) \ \ \forall i = 1, 2, \dots, k$$

$$f_j(x) < f_j(\bar{x}) \ \text{ per qualche } j \in \{1, 2, \dots, k\}.$$

Siccome ogni λ_i è positivo per ipotesi, questo vuol dire che $\sum_{i=1}^{k} \lambda_i f_i(x) < \sum_{i=1}^{k} \lambda_i f_i(\bar{x})$ il che contraddice il fatto che $\bar{x}$ è un minimo globale della $F(x)$. □

Si noti che nelle condizioni sufficienti di KKT appena viste si richiede che tutti i moltiplicatori delle funzioni obiettivo siano strettamente positivi mentre nelle condizioni necessarie almeno un λ_i è strettamente positivo.
Per quanto riguarda i punti di ottimo debole secondo Pareto, è possibile stabilire, sempre sotto le ipotesi di convessità, un risultato ancora più forte. Per tali punti si possono dare condizioni necessarie e sufficienti di ottimo.

Teorema 4.14 *Siano le $f_i(x)$ (per ogni $i = 1, 2, \ldots, k$) e $g_j(x)$ (per ogni $j = 1, 2, \ldots, m$) convesse. Condizione necessaria e sufficiente affinché un punto $\bar{x} \in \mathcal{F}$ sia ottimo debole secondo Pareto è che esistano dei moltiplicatori $\lambda \in \mathbb{R}^k$ e $\mu \in \mathbb{R}^m$ tali che* *CNS di KKT*

$$\nabla_x L(\bar{x}, \lambda, \mu) = 0,$$
$$\mu^\top g(\bar{x}) = 0,$$
$$(\lambda, \mu) \geq 0, \ \ \lambda \neq 0.$$

Dim. La dimostrazione è molto simile a quella del teorema 4.13, ed è lasciata per esercizio al lettore (vedi appendice).

5 Metodi di Soluzione

Generare le soluzioni ottime secondo Pareto costituisce una parte essenziale della programmazione vettoriale ed anzi, matematicamente parlando, nella maggior parte dei casi, il problema (P) si considera risolto una volta che sia stato individuato l'insieme degli ottimi di Pareto. Tuttavia, non sempre ci si può accontentare semplicemente di aver trovato l'insieme degli ottimi secondo Pareto. Alcune volte è infatti necessario ordinare tutte le soluzioni trovate e quindi selezionare la migliore rispetto a tale ordinamento. Per questo motivo abbiamo bisogno di un *decisore* cioè di qualcuno che ci dica, in base alle sue preferenze, come ordinare l'insieme degli ottimi di Pareto del problema (P).
In base al ruolo svolto dal decisore nella strategia di soluzione del problema, i metodi risolutivi della programmazione multiobiettivo vengono spesso suddivisi in quattro grandi categorie.

- *Metodi senza preferenze* nei quali il decisore non ha nessun ruolo e si considera soddisfacente l'aver trovato un qualunque ottimo di Pareto.
- *Metodi a posteriori* nei quali si genera l'insieme di tutti gli ottimi di Pareto e poi lo si presenta al decisore che sceglie la soluzione per lui migliore.
- *Metodi a priori* nei quali il decisore specifica le sue preferenze prima che abbia inizio il processo risolutivo. In base alle informazioni avute dal decisore viene direttamente trovata la soluzione ottima migliore, senza dover dunque generare tutti gli ottimi di Pareto.
- *Metodi interattivi* nei quali il decisore specifica le sue preferenze mano a mano che l'algoritmo procede, guidando in tal modo il processo risolutivo verso la soluzione per lui più soddisfacente.

Al di là di questa distinzione, tutti i metodi di soluzione per la programmazione multiobiettivo si basano sulla medesima idea di fondo, ovvero quella di trasformare il problema originario in uno con *una sola* funzione obiettivo. La tecnica mediante la quale si ottiene il problema mono obiettivo a partire dal problema (P) è nota come *scalarizzazione.*

5.1 Metodi Senza Preferenze

Nei metodi senza preferenze ci si accontenta di generare una soluzione ottima di Pareto, qualunque essa sia, senza tenere in considerazione le indicazioni del decisore.
Il metodo che presentiamo è noto come metodo *GOAL* (cfr. [1, sez. 2.1]). Quello che si fa è cercare la soluzione che minimizza, nello spazio degli obiettivi, la distanza tra la regione ammissibile ($\mathcal{Z}$) e un qualunque punto di riferimento $z^{ref} \notin \mathcal{Z} = f(\mathcal{F})$. Il vettore di riferimento sarà costituito dai valori auspicabili per le singole funzioni obiettivo. In particolare, una possibile scelta di z^{ref} è $z^{ref} = z^{id}$. Il problema che otteniamo è perciò il seguente:

$$\begin{aligned} \min \quad & \|f(x) - z^{id}\|_p \\ & g(x) \leq 0 \end{aligned} \tag{P$_p$}$$

ove $\|\cdot\|_p$ indica la norma p di un vettore (con $1 \leq p \leq \infty$). In particolare, se $p = \infty$, il problema (P$_p$) è noto come problema di *Tchebycheff*. Supponiamo di conoscere il vettore ideale *globale* degli obiettivi. Sotto tali ipotesi, il problema (P$_p$) ammette sempre soluzione. Valgono le seguenti proprietà.

Proposizione 5.1 *Ogni soluzione globale del problema (P$_p$) (con $1 \leq p < \infty$) è un ottimo globale di Pareto per il problema (P).*

Dim. Sia $x^\star$ una soluzione globale di (P$_p$) e supponiamo, per assurdo, che non sia un ottimo globale di Pareto. Deve allora esistere un punto ammissibile $x \in \mathcal{F}$ tale che

$$\begin{aligned} & f_i(x) \leq f_i(x^\star) \quad \forall i = 1, 2, \ldots, k \\ & f_j(x) < f_j(x^\star) \quad \text{per qualche } j \in \{1, 2, \ldots, k\}. \end{aligned}$$

Dalle precedenti disuguaglianze otteniamo che

$$\begin{aligned} & f_i(x) - z^{id} \leq f_i(x^\star) - z^{id} \quad \forall i = 1, 2, \ldots, k \\ & f_j(x) - z^{id} < f_j(x^\star) - z^{id} \quad \text{per qualche } j \in \{1, 2, \ldots, k\}. \end{aligned}$$

e quindi, sommando ed elevando alla potenza p-esima

$$\sum_{i=1}^{k} \left(f_i(x) - z^{id} \right)^p < \sum_{i=1}^{k} \left(f_i(x^\star) - z^{id} \right)^p$$

il che è ovviamente assurdo, dovendo essere $x^\star$ minimo globale del problema (P$_p$). □

Estendiamo ora il risultato precedente al caso dei minimi locali (cfr. [1]).

Proposizione 5.2 *Ogni ottimo locale del problema (P$_p$) (con $1 \leq p < \infty$) è un ottimo locale di Pareto per il problema (P).* □

Nel caso in cui $p = \infty$ vale invece la seguente proprietà (cfr. [1]).

Proposizione 5.3 *Ogni ottimo locale (globale) del problema di Tchebycheff (P$_\infty$) è un ottimo locale (globale) debole di Pareto del problema (P).* □

Tuttavia, la seguente proposizione, ci assicura l'esistenza di almeno una soluzione di (P$_\infty$) ottima secondo Pareto per il problema (P).

Proposizione 5.4 *Il problema di Tchebycheff (P$_\infty$) ha almeno una soluzione che è ottima secondo Pareto.* □

Le scelte di $p = 1$ e $p = \infty$ sono particolarmente vantaggiose nel caso in cui il problema multiobiettivo originario è lineare ($f_i(x), g_j(x)$ lineari per ogni i e j). Mediante semplici manipolazioni sul problema (P_p) è infatti possibile ottenere ancora un problema lineare e quindi adottare le ben note tecniche della PL per la sua soluzione. Supponiamo che (P) sia lineare ovvero

$$\begin{array}{ll} \min & (c_1^\top x, c_2^\top x, \dots, c_k^\top x) \\ & Ax \le b \end{array}$$

• *Norma* $p = 1$.
Il problema scalarizzato

$$\begin{array}{ll} \min & \displaystyle\sum_{i=1}^{k} |c_i^\top x - z_i^{id}| \\ & Ax \le b \end{array}$$

può essere facilmente trasformato in un problema di PL con l'aggiunta di k variabili ausiliarie, α_i per $i = 1, 2, \dots, k$, ottenendo:

$$\begin{array}{ll} \min & \displaystyle\sum_{i=1}^{k} \alpha_i \\ & \begin{cases} |c_i^\top x - z_i^{id}| \le \alpha_i & i = 1, 2, \dots, k \\ Ax \le b \end{cases} \end{array}$$

• *Norma* $p = \infty$.
In questo caso, il problema scalarizzato

$$\begin{array}{ll} \min & \displaystyle\max_{i=1,\dots,k} \{|c_i^\top x - z_i^{id}|\} \\ & Ax \le b \end{array}$$

può essere facilmente trasformato in un problema di PL con l'aggiunta di una sola variabile ausiliaria, α, ottenendo:

$$\begin{array}{ll} \min & \alpha \\ & \begin{cases} |c_i^\top x - z_i^{id}| \le \alpha & i = 1, 2, \dots, k \\ Ax \le b \end{cases} \end{array}$$

5.1.1 Esempio

Si consideri il seguente problema di programmazione multiobiettivo:

$$\begin{array}{ll} \min & (x_1, x_2) \\ & \begin{cases} x_1^2 + x_2^2 \ge 1 \\ 0 \le x_1 \le 2 \\ 0 \le x_2 \le 2 \end{cases} \end{array} \tag{P_{es}}$$

In questo esempio, vedi figura 5, le regioni ammissibili nello spazio delle variabili di decisione ed in quello degli obiettivi coincidono essendo

$$\begin{cases} z_1 = x_1 \\ z_2 = x_2 \end{cases}$$

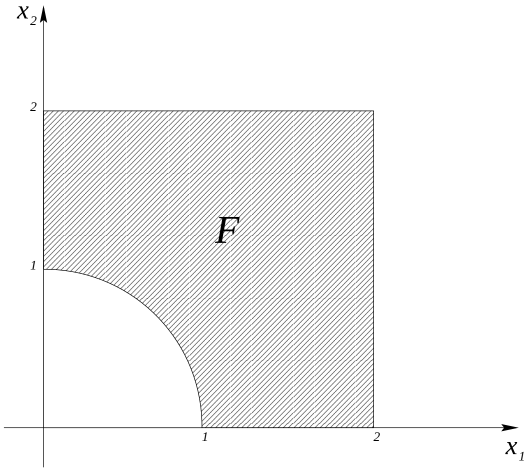


Figure 5: Regione ammissibile

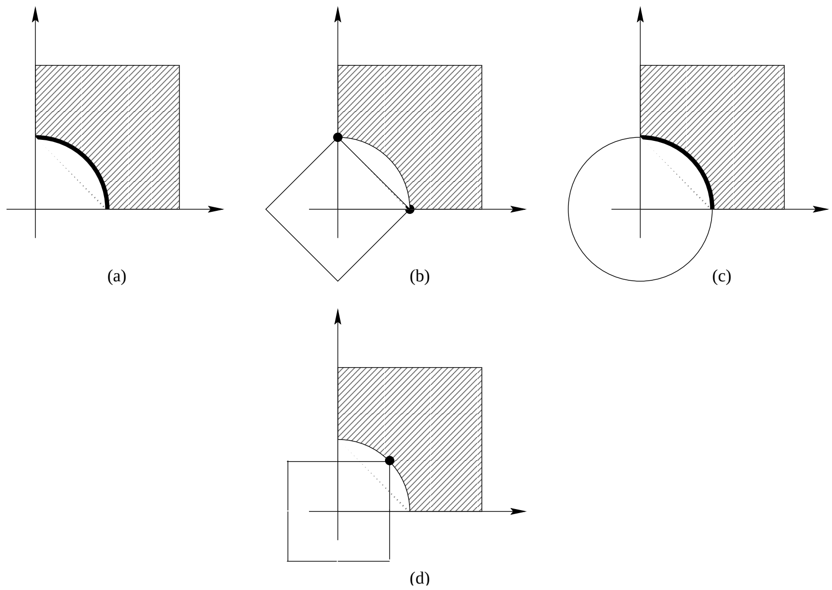


Figure 6: Differenti soluzioni in corrispondenza a differenti tipi di norme.

Possiamo inoltre facilmente calcolare il vettore ideale degli obiettivi ottenendo così: $z^{id} = (0, 0)^\top$. In figura 6a sono riportati gli ottimi di Pareto del problema (P_{es}).
A questo punto applichiamo il metodo GOAL con $p = 1, 2, \infty$.

- $\underline{p = 1}$. Il problema scalarizzato con questo tipo di norma ha due soluzioni (vedi figura 6b).
- $\underline{p = 2}$. Il problema che otteniamo usando la norma euclidea ha infinite soluzioni (come si vede in figura 6c) e cioè tutte le soluzioni del problema (P_{es}).
- $\underline{p = \infty}$. Il problema di Tchebycheff ha una sola soluzione (vedi figura 6d)

5.2 Metodi a Posteriori

I metodi appartenenti a questa classe sono anche noti come metodi per generare l'insieme delle soluzioni di Pareto. Infatti, siccome le preferenze del decisore vengono considerate solo al termine del processo risolutivo, quello che si fa è generare tutti i punti ottimi secondo Pareto. Una volta che l'insieme delle soluzioni di Pareto è stato generato, esso viene presentato al decisore che seleziona il o i vettori per lui migliori.
L'inconveniente principale di questa strategia sta nel fatto che il processo di generazione degli ottimi di Pareto è, molto spesso, computazionalmente oneroso. Inoltre, potrebbe non essere semplice, per il decisore, scegliere una soluzione tra gli ottimi che gli vengono presentati, in special modo se questi sono numerosi. Per questo motivo, è molto importante il modo con il quale le soluzioni vengono presentate al decisore.

5.2.1 Metodo dei Pesi (cfr. [1, sez. 3.1])

Consideriamo il seguente problema

$$\begin{aligned} \min \quad & \sum_{i=1}^{k} w_i f_i(x) \\ & g(x) \leq 0 \end{aligned} \tag{P_w}$$

ove $w \in \mathbb{R}^k_+$ e i coefficienti w_i si intendono normalizzati cioè tali che

$$\sum_{i=1}^{k} w_i = 1.$$

Esiste una relazione tra le soluzioni di (P_w) e i punti di Pareto di (P). La seguente proposizione mette in evidenza proprio questa relazione (cfr. [1]).

Proposizione 5.5 *Ogni soluzione locale (globale) del problema (P_w) è un ottimo debole locale (globale) di Pareto per il problema (P).* □

Nel caso in cui il problema (P_w) ammette una unica soluzione allora si può stabilire un risultato un po' più forte del precedente (cfr. [1]).

Proposizione 5.6 *Se il problema (P_w), fissato il vettore dei pesi $w \geq 0$, ammette una unica soluzione allora essa è un ottimo di Pareto per il problema (P).*

Dim. Sia $x^\star \in \mathcal{F}$ l'unica soluzione di (P_w) e supponiamo, per assurdo, che $x^\star$ non sia un ottimo di Pareto di (P). In tal caso deve necessariamente esistere un altro punto $x \in \mathcal{F}$ tale che

$$\begin{aligned} & f_i(x) \le f_i(x^\star) \quad \forall i = 1, 2, \ldots, k \\ & f_j(x) < f_j(x^\star) \quad \text{per qualche } j \in \{1, 2, \ldots, k\}. \end{aligned} \tag{8}$$

Ricordando che i pesi w_i sono nonnegativi abbiamo, moltiplicando ciascuna delle (8) per w_i e sommando tutto quanto, che

$$\sum_{i=1}^{k} w_i f_i(x) \le \sum_{i=1}^{k} w_i f_i(x^\star). \tag{9}$$

Siccome, però, $x^\star$ è l'unico punto di minimo di (P_w) risulta

$$\sum_{i=1}^{k} w_i f_i(x^\star) < \sum_{i=1}^{k} w_i f_i(\hat{x}) \quad \forall \hat{x} \in \mathcal{F}$$

che è ovviamente in contrasto con la (9). □

Allo stesso modo, se i pesi w_i sono tutti strettamente positivi è possibile dimostrare la seguente

Proposizione 5.7 *Se $w_i > 0$ per ogni indice i, ogni soluzione locale (globale) del problema (P_w) è un ottimo locale (globale) di Pareto per il problema (P).* □

Nell'ipotesi in cui il problema multiobiettivo (P) è convesso, è possibile stabilire la seguente proprietà di esistenza (cfr. [1])

Proposizione 5.8 *Sia $x^\star$ un ottimo di Pareto per il problema (P). Se (P) è convesso allora esistono dei pesi $w \in \mathbb{R}^k_+$ con*

$$\sum_{i=1}^{k} w_i = 1$$

e tali che $x^\star$ è soluzione anche del problema (P_w). □

5.2.2 Metodo degli ε-vincoli (cfr. [1, sez. 3.2])

Si seleziona una funzione obiettivo $f_l(x)$ tra gli obiettivi di (P) e poi si trasformano tutte le altre funzioni $f_i(x)$ (con $i = 1, 2, \ldots, k$ $i \neq l$) in vincoli, imponendo degli *upper bound* sui loro valori. Il problema che otteniamo è allora il seguente:

$$\begin{aligned} \min \quad & f_l(x) \\ & f_i(x) \le \varepsilon_i \quad \forall i = 1, 2, \ldots, k \; i \neq l \\ & g(x) \le 0 \end{aligned} \tag{P_ε}$$

ove $l \in \{1, 2, \ldots, k\}$.

Proposizione 5.9 (cfr. [1]) *Ogni soluzione di (P_ε) è un ottimo debole secondo Pareto per il problema (P).* □

La prossima proposizione fornisce condizioni necessarie e sufficienti di ottimalità (secondo Pareto) per il problema (P) delle soluzioni di (P_ε).

Proposizione 5.10 *Un vettore* $x^\star \in \mathcal{F}$ *è ottimo secondo Pareto di (P) se e solo se è soluzione di* (P$_\varepsilon$) *per ogni scelta di* $l \in \{1, 2, \dots, k\}$ *ed essendo* $\varepsilon_i = f_i(x^\star)$ *con* $i \neq l$*.*

Dim. (Necessità) Sia $x^\star$ un otimo di Pareto di (P) e supponiamo per assurdo che esso non sia soluzione del problema (P$_\varepsilon$) per qualche $l \in \{1, 2, \dots, k\}$ con $\varepsilon_i = f_i(x^\star)$ (per ogni $i \neq l$). Se questo è vero, allora deve esistere un punto $x \in \mathcal{F}$ tale che:

$$
\begin{aligned}
&f_l(x) < f_l(x^\star) \quad \text{e} \\
&f_i(x) \leq \varepsilon_i = f_i(x^\star) \quad \forall i \neq l.
\end{aligned}
$$

Pertanto abbiamo ottenuto che $x^\star$ non è un ottimo di Pareto per (P), il che è assurdo.
(Sufficienza) Sia ora $x^\star \in \mathcal{F}$ una soluzione di (P$_\varepsilon$) per ogni $l \in \{1, 2, \dots, k\}$. Da questo fatto segue che allora non può esistere alcun vettore $x \in \mathcal{F}$ tale che

$$
\begin{aligned}
&f_l(x) < f_l(x^\star) \quad \text{e} \\
&f_i(x) \leq f_i(x^\star) = \varepsilon_i \quad \forall i \neq l.
\end{aligned}
$$

Il punto $x^\star$ è quindi, in base alle precedenti relazioni, Pareto ottimo. □

Proposizione 5.11 (cfr. [1]) *Se il punto* $x^\star \in \mathcal{F}$ *è l'unica soluzione del problema* (P$_\varepsilon$) *per qualche* $l \in \{1, 2, \dots, k\}$ *e con* $\varepsilon_j = f_j(x^\star)$ *per ogni* $j \neq l$ *allora esso è Pareto ottimo per il problema (P).* □

5.3 Metodi a Priori

Nei metodi a priori il decisore ci fornisce le informazioni di cui abbiamo bisogno prima che il processo risolutivo abbia inizio. Una volta ottenute queste informazioni l'algoritmo si ferma solo quando è stato individuato un ottimo di Pareto. Di seguito presentiamo due metodi a priori e precisamente il metodo della *Value Function* e il metodo *lessicografico.*

5.3.1 Metodo della "value function" (cfr. [1, sez. 4.1])

Il decisore specifica l'espressione analitica di una funzione di utilità degli obiettivi $U(z)$. il problema da risolvere è pertanto il seguente

$$
\begin{aligned}
\min \quad & U(f(x)) \\
& g(x) \leq 0
\end{aligned}
\tag{P$_U$}
$$

Notiamo che se la $U(z)$ fosse una funzione lineare otterremmo nuovamente il problema (P$_w$) del metodo dei pesi. Mentre però il metodo dei pesi è un metodo a posteriori, in cui si vogliono generare tutte le soluzioni di Pareto (cambiando di volta in volta i pesi w_i), il metodo della "value function" è un metodo a priori, in cui cioè il decisore ci comunica le sue preferenze imponendo attraverso i pesi w_i le utilità relative delle funzioni obiettivo. La funzione $U(z)$, nel caso generale, potrebbe non essere lineare.

5.3.2 Metodo dell'ordinamento lessicografico (cfr. [1, sez 4.2])

Il decisore ordina le funzioni obiettivo in base alla loro importanza relativa. Una volta che le funzioni obiettivo siano state ordinate, il processo risolutivo inizia con la minimizzazione della prima funzione obiettivo sull'insieme ammissibile originario $\mathcal{F}$, cioè si risolve il problema

$$
\begin{aligned}
\min \quad & f_1(x) \\
& g(x) \leq 0
\end{aligned}
\tag{P$_1$}
$$

Se il problema (P_1) ha un'unica soluzione allora questa è anche soluzione di (P) e l'algoritmo termina. Altrimenti si minimizza la seconda funzione obiettivo nell'ordinamento lessicografico. Questa volta però oltre ai vincoli originari, si aggiunge un ulteriore vincolo il cui scopo è quello di garantire che all'ottimo non peggiori il valore della prima funzione obiettivo. Il problema è quindi il seguente

$$\begin{aligned} \min \quad & f_2(x) \\ & f_1(x) \le f_1(x^{1\star}) \\ & g(x) \le 0 \end{aligned} \tag{P$_2$}$$

ove $x^{1\star}$ è la soluzione di (P_1). Se (P_2) ha un'unica soluzione ci si ferma, altrimenti si procede come prima, selezionando la successiva funzione obiettivo nell'ordinamento assegnato dal decisore. Al generico passo $h \le k$ avremo il seguente problema

$$\begin{aligned} \min \quad & f_h(x) \\ & f_i(x) \le f_i(x^{h-1\star}) \quad i = 1, 2, \dots, h-1 \\ & g(x) \le 0 \end{aligned} \tag{P$_h$}$$

ove $x^{i\star}$ è una soluzione del problema (P_i).

Proposizione 5.12 *Ogni soluzione ottenuta con il metodo lessicografico è ottima secondo Pareto per il problema (P).*

Dim. Supponiamo per assurdo che la soluzione $x^\star$ ottenuta con il metodo lessicografico non sia un ottimo di Pareto e che esista, quindi, un vettore $x \in \mathcal{F}$ tale che

$$f(x) \le f(x^\star) \tag{10}$$

$$f_j(x) < f_j(x^\star) \quad \text{per qualche } j \in \{1, 2, \dots, k\}. \tag{11}$$

Nel determinare la soluzione $x^\star$ ci sono due sole alternative possibili:

- il punto $x^\star$ viene determinato in una minimizzazione intermedia, cioè $x^\star$ è l'unico punto di minimo di (P_i) per qualche $i \in \{1, 2, \dots, k\}$;
- tutte le minimizzazioni intermedie vengono effettuate senza mai trovare una unica soluzione.

Nella seconda ipotesi, abbiamo che per ogni $i = 1, 2, \dots, k$

$$f_i(x^\star) \le f_i(x) \quad \forall x \in \mathcal{F}$$

che insieme con la (10) ci da

$$f_i(x^\star) = f_i(x) \quad \forall x \in \mathcal{F}$$

che è in contrasto con la (11).
Nella prima ipotesi, invece, abbiamo che $x^\star$ è l'unico punto di minimo di un problema intermedio, sia esso (P_i). Quindi abbiamo che

$$f_i(x^\star) < f_i(x) \quad \forall x \ne x^\star$$

che, ancora una volta, è in contrasto con la (10). □

5.4 Metodi Interattivi

Lo schema generale di un metodo interattivo è il seguente

(1) Trova una soluzione ammissibile iniziale,

(2) Presenta la soluzione trova al decisore,

(3) Se la soluzione trovata va bene allora STOP.

    Altrimenti, sulla base delle indicazioni ottenute, trova una nuova soluzione e torna al punto (2).

Il metodo che presentiamo di seguito è noto con il nome di *STEP method*.

5.4.1 STEP method (cfr. [1, sez. 5.5])

Supponiamo che in un punto ottimo secondo Pareto il decisore sappia indicare quali funzioni obiettivo hanno un valore accettabile e quali no. Supponiamo inoltre che tutte la funzioni obiettivo siano limitate sulla regione ammissibile.
Ad ogni iterazione del metodo un ottimo di Pareto viene presentato al decisore che sulla base delle sue conoscenze e dei valori delle funzioni obiettivo nel punto corrente, specifica le funzioni obiettivo per le quali è accettabile un aumento del valore al fine di poter ulteriormente ridurre i valori delle restanti funzioni.
In sostanza, l'insieme degli indici $J = \{1, 2, \dots, k\}$ viene partizionato, ad ogni iterazione, in due sottoinsiemi

1. $J_<$, insieme degli indici delle f.ob. i cui valori sono insoddisfacienti per il decisore e
2. $J_> = J \backslash J_<$.

Se $J_< = \phi$ allora l'algoritmo si ferma avendo trovato l'ottimo di Pareto che meglio soddisfa il decisore. Altrimenti si chiede al decisore di specificare dei bounds ε_i sulle funzioni obiettivo con indici in $J_>$ e quindi si risolve il problema

$$
\begin{array}{lll}
\min & \left(\displaystyle\sum_{i \in J_<} |f_i(x) - z_i^{id}|^p \right)^{\frac{1}{p}} & \\
& f_i(x) \leq \varepsilon_i & i \in J_> \\
& f_i(x) \leq f_i(x^\star) & i \in J_< \\
& g(x) \leq 0 &
\end{array}
$$

ove $1 \leq p < \infty$ e $x^\star$ è l'ottimo di Pareto trovato nella iterazione precedente.
Notiamo che l'algoritmo si deve fermare, al fine di evitare problemi di ciclaggio, anche quando è vuoto l'insieme $J_>$ cioè quando tutte le funzioni obiettivo hanno valori insoddisfacienti.
L'algoritmo è quindi il seguente:

Algoritmo STEP

(1) Con il metodo *GOAL* (o un qualunque altro metodo) si trova un primo ottimo di Pareto $x^1 \in \mathcal{F}$ e si pone $h = 1$.

(2) Si chiede al decisore di partizionare l'insieme J in $J^h_>$ e $J^h_<$; Se $J^h_> = \phi$ oppure $J^h_< = \phi$ allora vai al punto (4).

(3) Risolvi il problema

$$\begin{array}{lll} \min & \left(\displaystyle\sum_{i \in J^h_<} |f_i(x) - z_i^{id}|^p \right)^{\frac{1}{p}} & \\ & f_i(x) \le \varepsilon_i^h & i \in J^h_> \\ & f_i(x) \le f_i(x^h) & i \in J^h_< \\ & g(x) \le 0 & \end{array} \tag{12}$$

con $1 \le p < \infty$. Sia x^{h+1} la soluzione di (12). Si pone $h = h + 1$ e si torna al punto (2).

(4) Si pone $x^\star = x^h$ e STOP.

6 Appendice

6.1 Dimostrazione della proposizione 4.9

La dimostrazione è per assurdo. Supponiamo, cioè, che nel punto $\bar{x}$ risulti $G_0(\bar{x}) = \phi$. Questo vuol dire (semplicemente richiamando la definizione dell'insieme $G_0(x)$) che il sistema

$$\nabla g_j(\bar{x})^\top d < 0 \quad \forall\ j \in I_0(\bar{x})$$

è inammissibile. Per il teorema di Gordan esistono dei numeri μ_j non tutti nulli e tali che

$$\sum_{j \in I_0(\bar{x})} \mu_j \nabla g_j(\bar{x}) = 0$$
$$\mu_j \ge 0 \quad \forall\ j \in I_0(\bar{x}).$$

Questo fatto è ovviamente assurdo dato che per ipotesi i vettori $\nabla g_j(\bar{x})$ (con $j \in I_0(\bar{x})$) sono linearmente indipendenti.

6.2 Dimostrazione del teorema 4.11

Per ipotesi abbiamo che $\bar{x}$ è un ottimo di Pareto e che $G_0(\bar{x}) \neq \phi$ (cioè $\bar{x}$ è un punto di regolarità per i vincoli del problema (P)).
La tesi segue banalmente sfruttando le condizioni necessarie di FJ e la proposizione 4.8.

6.3 Dimostrazione del teorema 4.14

La dimostrazione della necessità è banale se consideriamo il fatto che le condizioni necessarie di KKT valgono anche per punti di ottimo debole secondo Pareto.
La dimostrazione della sufficienza è analoga a quella del teorema 4.13. Definiamo, infatti, sfruttando i moltiplicatori λ_i la funzione $F(x) = \sum_{i=1}^k \lambda_i f_i(x)$ che, in quanto combinazione

conica di funzioni convesse è a sua volta convessa. Ragionando come nel teorema 4.13 abbiamo che valgono per la $F(x)$ le condizioni sufficienti di ottimalità. Il punto $\bar{x}$ è un ottimo globale per la $F(x)$ e quindi risulta

$$F(\bar{x}) \leq F(x) \quad \forall x \in \mathcal{F}.$$

Supponiamo, per assurdo, che $\bar{x}$ non sia un ottimo debole di Pareto per il problema (P). In questo caso dovrà esistere un altro punto ammissibile $x \in \mathcal{F}$ tale che

$$f_i(x) < f_i(\bar{x}) \quad \forall i = 1, 2, \ldots, k.$$

A questo punto, esistendo almeno un indice j tale che $\lambda_j > 0$ per ipotesi, dalle precedenti relazioni otteniamo

$$\sum_{i=1}^{k} \lambda_i f_i(x) < \sum_{i=1}^{k} \lambda_i f_i(\bar{x})$$

che è in contrasto con il fatto che la $F(x)$ ha in $\bar{x}$ un punto di ottimo globale.

Contents

References

[1] K. Miettinen, Nonlinear Multiobjective Optimization, Kluwer Academic Publishers, Boston, 1999.

[2] V. Pareto, *"Cours d'economie Politique"*, Rouge, Lausanne, Switzerland, 1896.

[3] M.S. Bazaraa, H.D. Sherali, C.M. Shetty, Nonlinear Programming: Theory and Algorithms , Wiley, New York, 1979.