

Capitolo 4

Analisi convessa

L'analisi convessa sta alla base di molta parte della teoria dell'ottimizzazione e trova applicazione sia nello studio dei problemi continui che di quelli discreti. Un risultato così apparentemente 'innocuo' come il teorema di separazione fra insiemi convessi costituisce il fondamento della teoria della dualità e quindi di gran parte dell'ottimizzazione. Particolari insiemi convessi come i poliedri rivestono un ruolo centrale nella programmazione lineare e nell'ottimizzazione combinatoria. L'estensione del concetto di gradiente a quello di subgradiente è spesso indispensabile per progettare algoritmi di risoluzione.

Per questi motivi si è scelto di fornire una trattazione dell'analisi convessa più ampia di quanto non avvenga normalmente in testi di ottimizzazione includendo comunque solo i risultati pertinenti all'ottimizzazione. Per quanto possibile si è cercato di presentare il materiale in modo autonomo. Tutte le dimostrazioni sono svolte e fanno riferimento unicamente a risultati già esposti. Per uno studio più approfondito si consiglia Rockafellar [1970], testo fondamentale sull'analisi convessa, in cui c'è praticamente tutto quello che può servire in spazi a dimensione finita.

4.1. Definizioni fondamentali

4.1 DEFINIZIONE. *Un insieme $K \subset \mathbb{R}^n$ è convesso se $\forall x_1, x_2 \in K$ e $\forall \alpha \in [0, 1]$, vale $\alpha x_1 + (1 - \alpha)x_2 \in K$.* ■

Dati due punti x_1 e x_2 il punto $\alpha x_1 + (1 - \alpha)x_2$ con $\alpha \in [0, 1]$ appartiene al segmento congiungente i due punti. Questo fatto è facilmente dimostrabile usando opportune similitudini per il triangolo $(0, x_1, x_2)$. Quindi un insieme è convesso solo quando il segmento congiungente una coppia arbitraria di punti dell'insieme è tutto contenuto nell'insieme. Per convenzione l'insieme vuoto è convesso. Anche un insieme costituito da un singolo punto è convesso, e, si noti, è l'unico insieme finito e convesso. Si vedano in figura 4.1 un insieme convesso (a destra) e uno non convesso (a sinistra).



FIGURA 4.1

Alcuni insiemi, dotati di particolare struttura intervengono di frequente in problemi di ottimizzazione. Abbiamo quindi:

4.2 DEFINIZIONE. Un insieme $C \subset \mathbb{R}^n$ è un cono se $\forall x \in C$ e $\forall \alpha \geq 0$ vale $\alpha x \in C$. ■

Quindi un cono è costituito da un fascio di semirette uscenti dall'origine e, ad esclusione del cono costituito dall'origine stessa, ogni cono è un insieme illimitato.

4.3 ESERCIZIO. Verificare che un cono è anche convesso se $\forall x_1, x_2 \in C$ vale $x_1 + x_2 \in C$. ■

4.4 ESEMPIO. Dato un insieme convesso K ed un punto $x \in K$, ci si può chiedere quali 'movimenti' siano concessi a partire da x senza uscire da K . Questo porta alla definizione di un cono particolare, detto *cono ammissibile*, frequentemente utilizzato (figura 4.2):

$$D(K, x) := \{h \in \mathbb{R}^n : \exists \alpha > 0 : x + \alpha h \in K\}$$

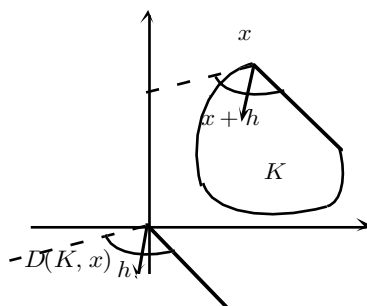


FIGURA 4.2

4.5 DEFINIZIONE. Un insieme $S \subset \mathbb{R}^n$ è un sottospazio se $\forall x_1, x_2 \in S$ e $\forall \alpha_1, \alpha_2 \in \mathbb{R}$ vale $\alpha_1 x_1 + \alpha_2 x_2 \in S$. ■

Si noti che ogni sottospazio deve contenere l'origine (basta porre $\alpha_1 = \alpha_2 = 0$) e che un sottospazio è anche un cono convesso. Spesso si ha a che fare con insiemi che 'assomigliano' a sottospazi pur non contenendo l'origine. Questi insiemi vengono definiti come:

4.6 DEFINIZIONE. Un insieme $M \subset \mathbb{R}^n$ è affine se $\forall x_1, x_2 \in M$ e $\forall \alpha \in \mathbb{R}$, vale $\alpha x_1 + (1 - \alpha)x_2 \in M$. ■

Per gli insiemi affini vale la seguente importante caratterizzazione:

4.7 TEOREMA. Un insieme $M \subset \mathbb{R}^n$ è affine se e solo se esiste un sottospazio S tale che $M = \{\bar{x}\} + S$ con \bar{x} un qualsiasi elemento fissato di M . Il sottospazio S è unico e prende il nome di sottospazio generatore di M .

DIMOSTRAZIONE. (\Leftarrow) Se esiste S per cui $M = \{\bar{x}\} + S$ con $\bar{x} \in M$, si ha che $(x_1 - \bar{x})$ e $(x_2 - \bar{x})$ appartengono ad S . Quindi $\forall \alpha$ vale $\alpha(x_1 - \bar{x}) + (1 - \alpha)(x_2 - \bar{x}) + \bar{x} \in M$ e cioè $\alpha x_1 + (1 - \alpha)x_2 \in M$.

(\Rightarrow) Sia M affine e $\bar{x} \in M$ arbitrario e sia $S := \{x - \bar{x} : x \in M\}$. Vogliamo dimostrare che S è un sottospazio, cioè che dati due elementi generici $y_1 = x_1 - \bar{x}$ e $y_2 = x_2 - \bar{x}$ di S si ha $\alpha y_1 + \beta y_2 \in S, \forall \alpha, \beta$. Allora

$$\alpha(x_1 - \bar{x}) + \beta(x_2 - \bar{x}) = \left(2 \left(\frac{1}{2}(\alpha x_1 + (1 - \alpha)\bar{x}) + \frac{1}{2}(\beta x_2 + (1 - \beta)\bar{x}) \right) - \bar{x} \right) - \bar{x}$$

e applicando ricorsivamente la definizione di insieme affine ai termini in parentesi discende la tesi.

Per l'unicità di S siano x_1 e x_2 due elementi diversi di M e siano $S_1 := \{x - x_1 : x \in M\}$ e $S_2 := \{x - x_2 : x \in M\}$. Vogliamo dimostrare che $S_1 = S_2$. Sia $\bar{y} \in S_1$. Quindi esiste $\bar{x} \in M$ tale che $\bar{y} = \bar{x} - x_1$ ovvero $\bar{y} = \bar{x} - x_1 + x_2 - x_2$. Ma $\bar{x} - x_1 + x_2 = 2(1/2\bar{x} + 1/2x_2) - x_1 \in M$ per definizione di insieme affine e quindi $\bar{y} \in S_2$, cioè $S_1 \subset S_2$. Analogamente si dimostra l'inclusione opposta. ■

La seguente proprietà è di immediata dimostrazione.

4.8 TEOREMA. *Insiemi convessi, coni, insiemi affini e sottospazi sono chiusi rispetto all'intersezione.* ■

Quindi data una famiglia arbitraria di insiemi convessi (coni, insiemi affini, sottospazi) che contengono un insieme dato A , ha senso parlare del 'più piccolo' insieme convesso (rispettivamente cono, insieme affine, sottospazio) che contiene A , proprio perché tale insieme è semplicemente dato dall'intersezione di tutti i membri della famiglia. In base a queste considerazioni possiamo pertanto definire:

4.9 DEFINIZIONE. *Si definisce inviluppo convesso (radiale, conico, affine, lineare) di A e si indica con $\text{conv}(A)$ (rispettivamente $\text{ray}(A)$, $\text{cone}(A)$, $\text{aff}(A)$, $\text{lin}(A)$) il più piccolo insieme convesso (cono, cono convesso, insieme affine, sottospazio) contenente A .* ■

Si veda in figura 4.3 un insieme A costituito da due punti. L'insieme $\text{ray}(A)$ è costituito da due semirette (figura 4.4), mentre $\text{cone}(A)$ contiene anche i punti all'interno fra le due semirette (figura 4.5). L'inviluppo convesso $\text{conv}(A)$ è il segmento che congiunge i due punti (figura 4.6), mentre $\text{aff}(A)$ è la retta che contiene i due punti (figura 4.7). Infine l'inviluppo lineare è dato da tutto \mathbb{R}^2 (figura 4.8).

4.10 ESERCIZIO. Siano dati $x_1, \dots, x_m \in \mathbb{R}^n$. Si dimostri che

$$\text{aff}\{x_1, \dots, x_m\} = x_1 + \text{lin}\{x_2 - x_1, \dots, x_m - x_1\}$$

Le definizioni date corrispondono a considerare gli inviluppi come costruiti 'esternamente' agli insiemi dati. È possibile tuttavia anche una caratterizzazione dall'interno. Si considerino dapprima le seguenti definizioni:

4.11 DEFINIZIONE. *Dati m elementi $x_1, x_2, \dots, x_m \in \mathbb{R}^n$, si dice loro a) combinazione convessa, un qualsiasi x esprimibile come*

$$x = \sum_{i=1}^m \alpha_i x_i \quad \text{con } \alpha_i \geq 0 \quad \sum_{i=1}^m \alpha_i = 1$$

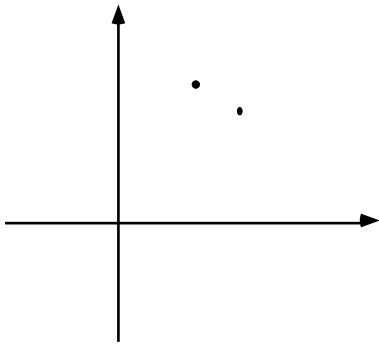


FIGURA 4.3

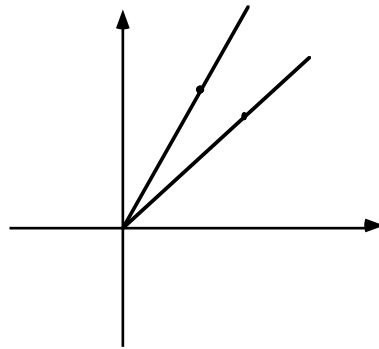


FIGURA 4.4

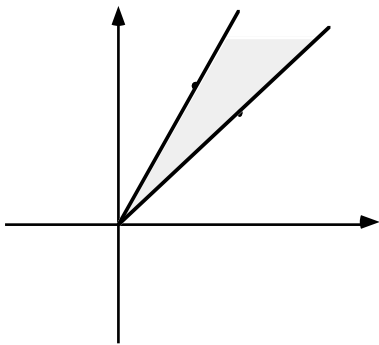


FIGURA 4.5

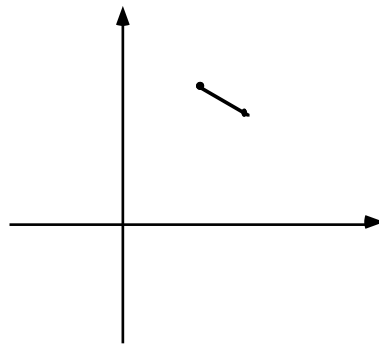


FIGURA 4.6

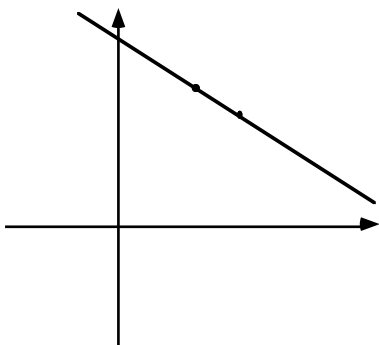


FIGURA 4.7

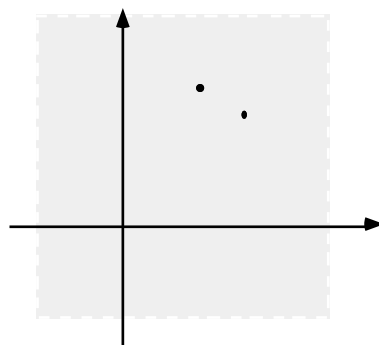


FIGURA 4.8

b) combinazione conica, un qualsiasi x esprimibile come

$$x = \sum_{i=1}^m \alpha_i x_i \quad \text{con } \alpha_i \geq 0$$

c) combinazione affine, un qualsiasi x esprimibile come

$$x = \sum_{i=1}^m \alpha_i x_i \quad \text{con} \quad \sum_{i=1}^m \alpha_i = 1$$

d) combinazione lineare, un qualsiasi x esprimibile come

$$x = \sum_{i=1}^m \alpha_i x_i$$

4.12 TEOREMA. L'involuppo convesso (conico, affine, lineare) di un insieme $A \subset \mathbb{R}^n$ è l'insieme di tutte le combinazioni convesse (coniche, affini, lineari) di arbitrari sottoinsiemi finiti di A .

DIMOSTRAZIONE. Il teorema viene dimostrato solo per il caso convesso in quanto le dimostrazioni per gli altri casi sono analoghe.

Bisogna dapprima dimostrare che un insieme K è convesso se e solo se contiene tutte le combinazioni convesse di suoi elementi. La sufficienza della condizione è vera, per definizione di convessità, limitandosi a considerare le combinazioni convesse di due elementi. La necessità verrà dimostrata per induzione completa. Il caso $m = 2$ deriva dalla definizione. Si supponga che una qualsiasi combinazione convessa di x_1, \dots, x_m appartenga a K . Sia $x_{m+1} \in K$ e $\sum_{k=1}^{m+1} \alpha_k x_k$ una qualsiasi combinazione convessa. Senza perdita di generalità possiamo supporre $0 < \alpha_{m+1} < 1$, altrimenti ricadiamo nel caso m . Quindi

$$\sum_{k=1}^{m+1} \alpha_k x_k = \alpha_{m+1} x_{m+1} + (1 - \alpha_{m+1}) \sum_{k=1}^m \frac{\alpha_k}{1 - \alpha_{m+1}} x_k$$

ed allora la combinazione convessa di x_1, \dots, x_{m+1} appartiene a K . Siccome la proprietà è vera per $m = 2$ discende l'asserto.

Per dimostrare il teorema si indichi con Q l'insieme delle combinazioni convesse e si noti che $\text{conv}(A) \supset A$ e siccome è convesso contiene, in base a quanto appena dimostrato, tutte le combinazioni convesse di elementi di A , cioè $\text{conv}(A) \supset Q$.

D'altra parte l'insieme Q delle combinazioni convesse è convesso. Infatti se $x \in Q$ e $y \in Q$ esistono $x_1, \dots, x_k \in A$, $\alpha_1, \dots, \alpha_k \geq 0$ con $\sum \alpha_i = 1$ e $y_1, \dots, y_h \in A$, $\beta_1, \dots, \beta_h \geq 0$ con $\sum \beta_j = 1$, tali che $x = \sum \alpha_i x_i$ e $y = \sum \beta_j y_j$. Ogni elemento del tipo $\gamma x + (1 - \gamma)y$ con $0 \leq \gamma \leq 1$ è, come si vede facilmente, una combinazione convessa di $\{x_1, \dots, x_k, y_1, \dots, y_h\}$. Allora Q è convesso e contiene A . Per definizione $Q \supset \text{conv}(A)$. Quindi $Q = \text{conv}(A)$. ■

Uno dei concetti più importanti è quello dell'indipendenza lineare di elementi di \mathbb{R}^n .

4.13 DEFINIZIONE. x_1, \dots, x_m si dicono linearmente indipendenti se $\sum_{i=1}^m \alpha_i x_i = 0$ implica $\alpha_i = 0 \quad \forall i$. Viceversa si dicono linearmente dipendenti se esistono α_i non tutti nulli tali che $\sum_{i=1}^m \alpha_i x_i = 0$. ■

Si noti che se l'origine fa parte di un qualsiasi insieme di vettori, l'insieme è linearmente dipendente. Il concetto di dimensione viene costruito a partire da quello di indipendenza lineare.

4.14 DEFINIZIONE. Un insieme linearmente indipendente $\{x_1, \dots, x_m\} \subset S$ si dice base di S se $\text{lin}\{x_1, \dots, x_m\} = S$. ■

4.15 TEOREMA. Se $\{x_1, \dots, x_m\}$ è una base di S , ogni insieme $\{y_1, \dots, y_k\} \subset S$ con $k > m$ è linearmente dipendente.

DIMOSTRAZIONE. Possiamo supporre $y_i \neq 0, \forall i$, (altrimenti sono automaticamente linearmente dipendenti). Allora $y_1 = \sum \alpha_i x_i$ con almeno un α_i diverso da zero (ad esempio $i = 1$). Quindi $x_1 = y_1/\alpha_1 + \sum_{i=2}^m (-\alpha_i/\alpha_1)x_i$ ed allora $y_2 = \sum \beta_i x_i = \beta'_1 y_1 + \sum_{i=2}^m \beta'_i x_i$. Analogamente almeno un $\beta'_i \neq 0$ per $i = 2, \dots, m$. Continuando si arriva a $y_{m+1} = \sum_{i=1}^m \mu_i y_i$. ■

4.16 COROLLARIO. Tutte le basi di un sottospazio S hanno la medesima cardinalità. ■

Ha senso allora la seguente definizione.

4.17 DEFINIZIONE. La dimensione di un sottospazio è data dalla cardinalità di una qualsiasi base. ■

È facile ora estendere il concetto di dimensione ad insiemi affini e poi ad insiemi convessi.

4.18 DEFINIZIONE. La dimensione di un insieme affine è la dimensione del suo sottospazio generatore. ■

4.19 DEFINIZIONE. La dimensione di un insieme convesso A è data dalla dimensione di $\text{aff}(A)$. ■

Volendo estendere questa definizione ad insiemi non convessi bisogna adottare una certa cautela perché si possono generare discordanze con altri tipi di definizione. Ad esempio l'insieme $A := \{x \in \mathbb{R}^2 : x_1 = t, x_2 = t^2, 0 < t < 1\}$ viene di solito considerato unidimensionale perché esiste un diffeomorfismo fra A e l'intervallo $(0, 1)$, mentre sarebbe bidimensionale secondo la definizione 4.19 estesa ad insiemi non convessi. Insiemi convessi della stessa dimensione dello spazio ambiente rivestono un ruolo speciale.

4.20 DEFINIZIONE. Un corpo convesso in \mathbb{R}^n è un insieme convesso di dimensione n . ■

Fra gli insiemi affini in \mathbb{R}^n , alcuni di questi hanno particolare rilevanza tanto da essere contraddistinti da un nome particolare.

4.21 DEFINIZIONE. Si dice piano un qualsiasi insieme affine di \mathbb{R}^n di dimensione $n - 1$. Si dice retta un qualsiasi insieme affine di dimensione 1. ■

Si noti che, se $n = 2$, un piano è anche una retta. Spesso si indica come iperpiano ciò che viene qui definito semplicemente piano. Questa distinzione ha origine pensando alla geometria dello spazio fisico, in cui un piano ha dimensione 2. Da un punto di vista matematico non ha molto senso distinguere fra \mathbb{R}^3 e \mathbb{R}^n con $n > 3$, per cui tanto vale usare direttamente il più semplice termine 'piano'.

Indichiamo sinteticamente con $a x$ il prodotto scalare $\sum_{i=1}^n a_i x_i$. Per non appesantire la notazione useremo quasi sempre questo modo di indicare il prodotto scalare anziché altri usati in letteratura come $a^T x, a' x, \langle a, x \rangle$. Il contesto renderà la notazione non ambigua.

4.22 TEOREMA. Un insieme $H \subset \mathbb{R}^n$ è un piano se e solo se esistono $a \in \mathbb{R}^n$, $b \in \mathbb{R}$, tali che $H = \{x \in \mathbb{R}^n : ax = b\}$.

DIMOSTRAZIONE. Se H è un piano, è generato da un sottospazio S di dimensione $n - 1$. Allora S^\perp (il complemento ortogonale di S) ha dimensione 1 e, per ogni $a \in S^\perp$ e $x \in S$ si ha $ax = 0$. Quindi se \bar{x} è un generico punto di H si ha $a(x - \bar{x}) = 0, \forall x \in H$ ovvero $ax = a\bar{x} =: b, \forall x \in H$.

Viceversa sia $H := \{x : ax = b\}$. Sia $\bar{x} \in H$. Per ogni $x \in H$, $a(\bar{x} - x) = 0$. Quindi $S := \{y : y = \bar{x} - x, x \in H\}$ è un sottospazio e allora $S + \bar{x}$ è affine. La dimensione di S è chiaramente $n - 1$ e quindi H è un piano. ■

Gli insiemi $\{x : ax \geq b\}$ e $\{x : ax \leq b\}$ prendono il nome di *semispazio chiuso* e gli insiemi $\{x : ax > b\}$ e $\{x : ax < b\}$ quello di *semispazio aperto*.

È importante notare che un piano ripartisce l'intero spazio in tre insiemi disgiunti costituiti da due semispazi aperti ed il piano stesso, e che il piano separa i due semispazi aperti, nel senso che ogni segmento congiungente un punto in un semispazio con un punto nell'altro semispazio deve attraversare il piano.

Anche un insieme affine M di dimensione m contenuto in un insieme affine M' di dimensione $m + 1$ ripartisce M' in tre insiemi disgiunti, di cui uno è M stesso e gli altri due sono due insiemi che per analogia chiameremo *semisottospazi*.

Più in generale si ha la seguente caratterizzazione di insiemi affini la cui dimostrazione viene lasciata come esercizio.

4.23 TEOREMA. Un insieme $M \subset \mathbb{R}^n$ è un insieme affine di dimensione m se e solo se esiste una matrice $(n - m) \times n$ A di rango $n - m$, ed un vettore $b \in \mathbb{R}^{n-m}$, tali che $M = \{x \in \mathbb{R}^n : Ax = b\}$. ■

Analogamente al concetto di indipendenza lineare si può introdurre quello di *indipendenza affine*:

4.24 DEFINIZIONE. Un insieme $\{x_1, \dots, x_{m+1}\}$ si dice *indipendente in modo affine* se

$$\dim \text{aff}\{x_1, \dots, x_{m+1}\} = m$$

■

Le seguenti proprietà sono di facile dimostrazione e vengono lasciate come esercizio.

4.25 ESERCIZIO. Verificare che due punti distinti sono sempre indipendenti in modo affine, mentre tre punti sono indipendenti in modo affine se e solo se non sono allineati. ■

4.26 ESERCIZIO. Dimostrare che $\{x_1, \dots, x_{m+1}\}$ sono indipendenti in modo affine se e solo se $x_1 - x_{m+1}, \dots, x_m - x_{m+1}$ sono linearmente indipendenti. ■

4.27 ESERCIZIO. Dimostrare che $\{x_1, \dots, x_{m+1}\}$ sono indipendenti in modo affine se e solo se nessuno di essi può essere espresso come combinazione affine degli altri. ■

4.28 ESERCIZIO. Un insieme $\{x_1, \dots, x_{m+1}\}$ è dipendente in modo affine se e solo se esistono $\alpha_1, \dots, \alpha_{m+1}$ non tutti nulli e tali che $\sum \alpha_i = 0$, per cui si abbia $\sum \alpha_i x_i = 0$.

4.29 ESERCIZIO. Verificare che i punti di un insieme affine si possono rappresentare come unica combinazione affine di un insieme indipendente in modo affine. ■

Da queste osservazioni si deduce che per ogni insieme A , con $m = \dim(A)$, esistono $m + 1$ punti di A indipendenti in modo affine. Questo fatto ha una conseguenza molto importante. Come si ricorderà, ogni punto di $\text{conv}(A)$ e in particolare di $\text{conv}(A) \setminus A$ si può rappresentare come combinazione convessa di elementi di A . Tuttavia sorge la domanda: quanti elementi in generale? Ci sono delle limitazioni a questo numero? La risposta, fortunatamente, è che bastano tanti punti di A quant'è la dimensione di A più uno. Naturalmente non si dice con questo che questi punti debbano essere sempre gli stessi per tutte le combinazioni convesse.

4.30 TEOREMA. (Caratheodory) Sia $A \subset \mathbb{R}^n$ e $\dim \text{conv}(A) = m$. Ogni punto di $\text{conv}(A)$ è rappresentabile come combinazione convessa di al più $(m + 1)$ punti di A .

DIMOSTRAZIONE. Sia $x \in \text{conv}(A)$, $x = \sum_{k=1}^h \alpha_k x_k$ con $x_1, \dots, x_h \in A$, $\sum_{k=1}^h \alpha_k = 1$, $\alpha_k > 0$, e $h > m + 1$. Quindi x_1, \dots, x_h sono dipendenti in modo affine. Esistono quindi β_1, \dots, β_h non tutti nulli tali che $\sum \beta_i = 0$ e $\sum \beta_i x_i = 0$. Allora $\forall \mu$ si ha

$$x = \sum_{k=1}^h (\alpha_k + \mu \beta_k) x_k \quad \text{e} \quad \sum_{k=1}^h \alpha_k + \mu \beta_k = 1$$

Siccome $\alpha_k > 0, \forall k$, è sempre possibile scegliere μ tale che $\alpha_k + \mu \beta_k \geq 0 \quad \forall k$ e $\alpha_k + \mu \beta_k = 0$ per almeno un k . Quindi si è trovata una combinazione convessa di $h - 1$ elementi. Il ragionamento quindi prosegue fino a $m + 1$ elementi. ■

4.2. Operazioni algebriche

Per arbitrari insiemi di \mathbb{R}^n si possono definire alcune operazioni algebriche. Ad esempio

$$A + B := \{x \in \mathbb{R}^n : x = a + b \text{ con } a \in A \text{ e } b \in B\},$$

$$\alpha A := \{x \in \mathbb{R}^n : x = \alpha a \text{ con } a \in A\}.$$

Si vedano in figura 4.9 due insiemi e la loro somma algebrica.

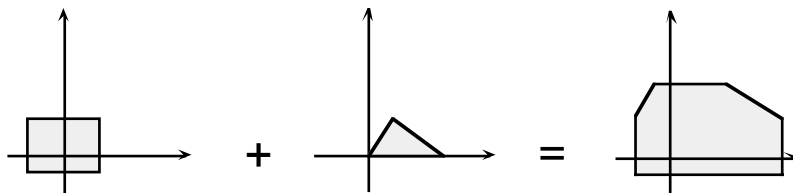


FIGURA 4.9

È facile dimostrare che se A e B sono convessi allora $A + B$ e αA sono convessi. Inoltre si ha che $(1 - \alpha) K + \alpha K = K$, per ogni $0 < \alpha < 1$, se e solo se K è convesso (che relazione c'è fra $\bigcup_{0 \leq \alpha \leq 1} ((1 - \alpha) K + \alpha K)$ e $\text{conv}(K)$? Si prenda ad esempio K l'insieme dei vertici di un triangolo), che $\alpha K = K$, per $\alpha > 0$, se e solo se K è un cono e, infine, che, per un cono K , $K + K = K$ se e solo se K è convesso.

Inoltre, per insiemi A e B qualsiasi valgono $\alpha A + \alpha B = \alpha(A + B)$ e $(\alpha + \beta) A \subset \alpha A + \beta A$, potendo l'inclusione essere stretta (si prenda ad esempio $\alpha = 1$ e $\beta = -1$ oppure si veda in figura 4.10 la differenza fra $2A$ e $A + A$ per un particolare A). L'inclusione diventa uguaglianza per insiemi convessi:

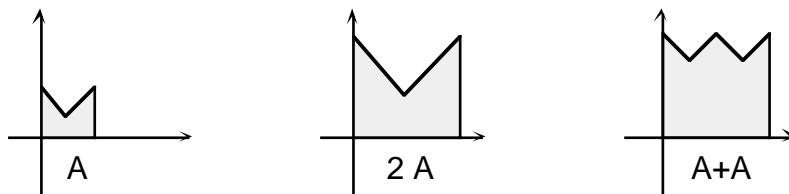


FIGURA 4.10

4.31 TEOREMA. $(\alpha + \beta)A = \alpha A + \beta A$ se e solo se A è convesso e $\alpha, \beta \geq 0$. ■

DIMOSTRAZIONE. Sia A convesso. Si è già detto che vale l'inclusione $(\alpha + \beta)A \subset \alpha A + \beta A$. Pertanto dimostriamo l'inclusione $(\alpha + \beta)A \supset \alpha A + \beta A$. Sia $z = \alpha x + \beta y$ con $x, y \in A$. Dobbiamo far vedere che esiste un elemento $w \in A$ tale che $z = (\alpha + \beta)w$. Da

$$\alpha x + \beta y = (\alpha + \beta) \left(\frac{\alpha}{\alpha + \beta} x + \frac{\beta}{\alpha + \beta} y \right) =: (\alpha + \beta)w$$

e da $\alpha, \beta \geq 0$ si vede che w è combinazione convessa di x e y e, per convessità di A , appartiene ad A .

Viceversa sia $(\alpha + \beta)A = \alpha A + \beta A$ per ogni $\alpha, \beta \geq 0$. Prendendo in particolare un valore $0 \leq \alpha \leq 1$ e $\beta := 1 - \alpha$ si ha $A = \alpha A + (1 - \alpha)A$, cioè la definizione di convessità. ■

4.3. Proprietà topologiche

In \mathbb{R}^n pensiamo definita l'usuale topologia metrica (ad esempio tramite una qualsiasi norma). Di conseguenza possiamo identificare alcuni insiemi come aperti o come chiusi e possiamo anche parlare di frontiera, di interno e di chiusura di un insieme arbitrario. Tuttavia nel caso di insiemi convessi si rivela più importante il concetto di frontiera relativa e di interno relativo.

Più esattamente sia $K \subset \mathbb{R}^n$ convesso; $\text{aff}(K)$ come sottoinsieme di \mathbb{R}^n ha la sua topologia relativa; pertanto si ha:

4.32 DEFINIZIONE. L'interno relativo $\overset{\circ}{K}$ di K è l'interno di K nella topologia relativa di $\text{aff}(K)$. ■

Ovviamente la chiusura \overline{K} di K è la medesima sia nella topologia di \mathbb{R}^n che in quella relativa di $\text{aff}(K)$. Per quel che riguarda la frontiera relativa si ha:

4.33 DEFINIZIONE. La frontiera relativa ∂K di K è data da $\overline{K} \setminus \overset{\circ}{K}$. ■

Si noti che $\text{aff}(K)$ è relativamente sia chiuso che aperto. È facile dimostrare che se K è convesso anche \overline{K} e $\overset{\circ}{K}$ sono convessi. La frontiera relativa ∂K è invece convessa in due soli casi, quando K sia un insieme affine oppure un semisottospazio. La dimostrazione di questo fatto non è elementare e verrà fatta in seguito utilizzando il teorema di separazione.

Gli interni relativi sono pertanto punti interni 'intrinseci' dell'insieme convesso senza alcun riferimento allo spazio in cui l'insieme è immerso. A meno che un insieme convesso non sia vuoto, il suo interno relativo non è mai vuoto, come viene ora dimostrato.

4.34 TEOREMA. $\dim K = \dim \overset{\circ}{K}$

DIMOSTRAZIONE. Ovviamente $\dim K \geq \dim \overset{\circ}{K}$. Sia $\dim K = n$. Quindi K contiene x_1, \dots, x_{n+1} indipendenti in modo affine. Allora $\text{conv}\{x_1, \dots, x_{n+1}\} \subset K$. Definiamo $y := \sum x_i / (n+1)$. Quindi $(x_1 - y), \dots, (x_n - y)$ sono linearmente indipendenti e gli insiemi $B_\varepsilon := \{u \in \mathbb{R}^n : u = \sum_{i=1}^n \alpha_i (x_i - y), |\alpha_i| < \varepsilon\}$ sono intorno in $\text{aff}(K)$. Si indichi $\beta := \sum_{i=1}^n \alpha_i$. Ora $\forall u \in B_\varepsilon$ si ha che

$$u + y = \sum_{i=1}^n \left(\frac{1-\beta}{n+1} + \alpha_i \right) x_i + \frac{1-\beta}{n+1} x_{n+1}$$

Quindi per ε sufficientemente piccolo si ha $u + y \in \text{conv}\{x_1, \dots, x_{n+1}\} \subset K$, e allora $y \in \overset{\circ}{K}$ ed anche $(1-\varepsilon)y + \varepsilon x_k \in \overset{\circ}{K}$ per ε sufficientemente piccolo. È immediato notare che questi $n+1$ punti sono indipendenti in modo affine e perciò $\dim \overset{\circ}{K} \geq n = \dim K$. ■

4.35 COROLLARIO. K non è vuoto se e solo se $\overset{\circ}{K}$ non è vuoto. ■

Si noti che se K è un punto è allo stesso tempo chiuso ed aperto. Un'importante proprietà dei punti interni è data dal seguente teorema:

4.36 TEOREMA. Sia K convesso, $x \in \overset{\circ}{K}$, $y \in \partial K$. Allora per ogni $0 < \alpha \leq 1$ si ha $\alpha x + (1-\alpha)y \in \overset{\circ}{K}$.

DIMOSTRAZIONE. Si tratta di dimostrare che $\alpha x + (1-\alpha)y + \varepsilon B \subset K$ per un certo $\varepsilon > 0$, dove B è la sfera unitaria nel sottospazio generatore di $\text{aff}(K)$. Siccome $y \in \partial K$, $(y + \varepsilon B) \cap K \neq \emptyset$, $\forall \varepsilon > 0$, e cioè $y \in K + \varepsilon B$. Quindi per ogni $\varepsilon > 0$

$$\alpha x + (1-\alpha)y + \varepsilon B \subset \alpha x + (1-\alpha)(K + \varepsilon B) + \varepsilon B = \alpha \left(x + \varepsilon \frac{2-\alpha}{\alpha} B \right) + (1-\alpha)K$$

Per ogni $\alpha > 0$ esiste $\varepsilon > 0$ tale che $x + \varepsilon((2-\alpha)/\alpha)B \subset K$. Quindi $\alpha x + (1-\alpha)y + \varepsilon B \subset \alpha K + (1-\alpha)K = K$. ■

4.37 ESERCIZIO. Dimostrare che, se $x \in \overset{\circ}{K}$, allora il cono ammissibile $D(K, x)$ è uguale al sottospazio generatore di $\text{aff} K$. ■

4.4. Separabilità

4.38 DEFINIZIONE. Due insiemi $A, B \subset \mathbb{R}^n$ sono separabili se esiste un piano i cui due semispazi chiusi contengano rispettivamente A e B . Inoltre A e B sono separabili strettamente se esistono due semispazi chiusi senza punti in comune che contengano rispettivamente A e B . ■

Ovviamente la separabilità di due insiemi è collegata in qualche modo con la non esistenza di punti in comune fra i due insiemi. Tuttavia è facile dare esempi in cui insiemi disgiunti non sono separabili. Come si intuisce, la convessità gioca un ruolo determinante. Vedremo infatti fra poco che insiemi convessi e disgiunti sono sempre separabili.

Siccome un piano si può rappresentare con l'equazione $ax = b$, possiamo dire che due insiemi A e B sono separabili se esistono $a \in \mathbb{R}^n$ e $b \in \mathbb{R}$ tali che $ax \geq b, \forall x \in A$ e $ax \leq b, \forall x \in B$. Inoltre sono strettamente separabili se $\inf_{x \in A} ax > \sup_{x \in B} ax$ (i due semispazi devono necessariamente essere generati da due piani paralleli altrimenti la loro intersezione non è vuota). Cominciamo con alcuni lemmi.

4.39 LEMMA. Sia K convesso, $0 \notin \overline{K}$, allora 0 e K sono separabili strettamente.

DIMOSTRAZIONE. In base al teorema di Weierstrass esiste $x_0 \in \overline{K}$ tale che $\|x_0\| \leq \|x\|$, $\forall x \in K$. In base all'ipotesi $\|x_0\| > 0$. Sia $x_1 \in \overset{\circ}{K}$, allora $\alpha x_1 + (1 - \alpha)x_0 \in K$, con $0 < \alpha \leq 1$, per il teorema 4.36, e quindi $\|\alpha x_1 + (1 - \alpha)x_0\|^2 = \|x_0 + \alpha(x_1 - x_0)\|^2 \geq \|x_0\|^2$ per la minimalità di $\|x_0\|$, ovvero

$$(x_0 + \alpha(x_1 - x_0))(x_0 + \alpha(x_1 - x_0)) \geq x_0 x_0$$

$$\|x_0\|^2 + \alpha^2 \|x_1 - x_0\|^2 + 2\alpha x_0(x_1 - x_0) \geq \|x_0\|^2$$

da cui, eliminando $\|x_0\|^2$ e dividendo per $\alpha > 0$ si ottiene

$$\|x_0\|^2 - x_0 x_1 \leq \frac{\alpha}{2} \|x_1 - x_0\|^2 \quad \forall 0 < \alpha < 1$$

e siccome α è arbitrario si ha $\|x_0\|^2 - x_0 x_1 \leq 0$ cioè $x_0 x_1 \geq \|x_0\|^2 > 0$, $\forall x_1 \in \overset{\circ}{K}$. Estendendo per continuità a tutto K si ha quindi $x_0 x > 0$, $\forall x \in K$. Siccome $x_0 \neq 0$, l'equazione $x_0 x = \|x_0\|^2/2$ fornisce direttamente un piano separatore. ■

Il teorema è illustrato in figura 4.11. In figura 4.12 si vede che l'origine e K non possono essere separati a causa della mancanza di convessità.

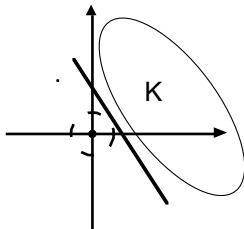


FIGURA 4.11

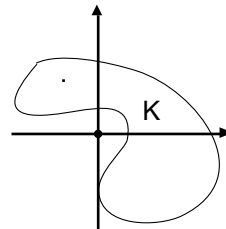


FIGURA 4.12

4.40 LEMMA. Sia K convesso, $0 \notin K$, allora 0 e K sono separabili.

DIMOSTRAZIONE. Se $0 \notin \overline{K}$ basta applicare il lemma precedente. Se $0 \in \overline{K}$, allora, da $0 \notin K$, abbiamo $0 \in \partial K$. Quindi esiste una successione $\{y_k\} \rightarrow 0$ tale che $y_k \notin \overline{K}$. In base al lemma precedente ogni y_k è separabile strettamente da K (cioè 0 è separabile da $K - y_k$) tramite un opportuno piano; sia questo $a_k x = b_k$ e senza perdita di generalità possiamo supporre $|a_k| + |b_k| = 1$. Essendo la successione (a_k, b_k) limitata, esiste una sottosuccessione convergente ad un certo (\bar{a}, \bar{b}) tale che $|\bar{a}| + |\bar{b}| = 1$. Allora si ha:

$$\begin{aligned} a_k y &> b_k \quad \forall y \in K \\ a_k y_k &< b_k \end{aligned}$$

Passando al limite sulla sottosuccessione si ha:

$$\begin{aligned} \bar{a} y &\geq \bar{b} \quad \forall y \in K \\ \bar{a} 0 &\leq \bar{b} \end{aligned}$$

Inoltre, siccome $0 \in \overline{K}$ vale $\bar{a} 0 \geq \bar{b}$ e quindi $\bar{b} = 0$ e $\bar{a} y \geq 0$, $\forall y \in K$. ■

Come si vede la separazione non può essere stretta se $0 \in \partial K$.

4.41 TEOREMA. *Siano A e B convessi. Se $A \cap B = \emptyset$ allora A e B sono separabili. Se inoltre \bar{A} e \bar{B} non contengono punti arbitrariamente vicini allora A e B sono strettamente separabili.*

DIMOSTRAZIONE. Si ponga $K = A - B$. Allora la condizione $A \cap B = \emptyset$ è equivalente a $0 \notin K$, mentre la condizione su \bar{A} e \bar{B} è equivalente a $0 \notin \bar{K}$. Nel primo caso quindi esiste un piano separatore tale che $ax \geq 0, \forall x \in K$, cioè $x = z - y$ con $z \in A$ e $y \in B$ e quindi si ha $az \geq ay, \forall z \in A, \forall y \in B$. Nel secondo caso si ha $\inf_{x \in K} ax \geq b > 0$ cioè:

$$\begin{aligned} \inf_{z \in A, y \in B} (az - ay) &> 0 \\ \inf_{z \in A} az - \sup_{y \in B} ay &> 0 \end{aligned}$$

■

Come esempio di insiemi chiusi e disgiunti ma non strettamente separabili si consideri $A := \{(x_1, x_2) : x_1 x_2 \geq 0, x_1 \geq 0\}$ e $B := \{(x_1, x_2) : x_1 x_2 \leq 0, x_1 \leq 0\}$, per i quali l'unico piano separatore è $x_1 = 0$ e $\inf_{z \in A} z_1 = \sup_{y \in B} y_1 = 0$.

4.42 COROLLARIO. *Siano A e B convessi. Se $\overset{\circ}{A} \cap \overset{\circ}{B} = \emptyset$ allora A e B sono separabili.*

(la dimostrazione è facile e viene lasciata come esercizio). ■

Facciamo notare che la definizione data di separazione può non essere soddisfacente nel caso gli insiemi da separare appartengano ad un insieme affine proprio. In questo caso gli insiemi potrebbero essere separabili pur avendo punti in comune, in quanto il piano contenente l'insieme affine risulta essere separatore. Bisognerebbe allora escludere questo caso dalla definizione di separazione. Si parla allora di *separazione propria* se il piano separatore non contiene gli insiemi da separare. Per attuare questa condizione basta aggiungere alla condizione $\inf_{x \in A} ax \geq \sup_{x \in B} ax$ la condizione $\sup_{x \in A} ax > \inf_{x \in B} ax$. In ogni caso le ipotesi del teorema 4.41 garantiscono la separazione propria. Ovviamente non vi sono problemi di questo genere per la separazione stretta.

I piani che separano punti di ∂K da K hanno una speciale importanza e prendono il nome di piani di supporto.

4.43 DEFINIZIONE. *Si definisce piano di supporto all'insieme K ogni piano di equazione $ax = \inf_{y \in K} ay > -\infty$. Si definisce semispazio di supporto all'insieme K ogni insieme del tipo $\{x : ax \geq \inf_{y \in K} ay > -\infty\}$.* ■

Si noti che non è necessaria la convessità di K per definire i piani di supporto a K . In presenza di convessità la frontiera e i piani di supporto sono legati fra loro dal seguente risultato, che discende immediatamente dai teoremi di separazione.

4.44 TEOREMA. *Sia K convesso. Per ogni $x \in \partial K$ passa almeno un piano di supporto. Ogni piano di supporto per x separa (non strettamente) x da K .* ■

4.45 ESERCIZIO. (opposto dell'esercizio 4.37) Dimostrare che se $\bar{x} \in \partial K$ il cono ammissibile $D(K, \bar{x})$ è contenuto in un semisottospazio. ■

4.46 ESERCIZIO. Dimostrare che, dato un insieme convesso K e un punto $\bar{x} \in \partial K$, l'insieme

$$\{a \in \mathbb{R}^n : a x = a \bar{x} \text{ è un piano di supporto a } K\}$$

è un cono convesso e chiuso. Che relazione c'è fra questo cono e $D(K, \bar{x})$? ■

Nella letteratura un piano di supporto viene di solito definito tramite il 'sup' anziché l'inf'. Si tratta di una differenza solo formale. In questa sede è più conveniente la definizione con l'inf' perché vengono considerati soprattutto problemi di minimizzazione.

Precedentemente si era visto che un insieme convesso K si poteva rappresentare 'esternamente' come intersezione di insiemi convessi contenenti K , e 'internamente' come unione di combinazioni convesse di punti di K . Queste rappresentazioni sono tuttavia ridondanti. Infatti è sufficiente considerare i semispazi di supporto per la rappresentazione esterna e, tranne alcuni casi particolari, è sufficiente considerare la frontiera per la rappresentazione interna. La dimostrazione di questo fatto sfrutterà il teorema di separazione.

4.47 TEOREMA. Sia K convesso e chiuso, allora K è eguale all'intersezione di tutti i suoi semispazi di supporto.

DIMOSTRAZIONE. È ovvio che l'intersezione contiene K . Sia pertanto $x \notin K$. Siccome K è chiuso, allora x e K sono strettamente separabili da un piano. A maggior ragione quindi x non appartiene al semispazio che contiene K , individuato dal piano di supporto parallelo al piano separatore. ■

4.48 LEMMA. ∂K è convessa se e solo se \bar{K} è vuoto, oppure è un insieme affine, oppure è un semisottospazio.

DIMOSTRAZIONE. (\Rightarrow) Se $\partial K = \emptyset$ il teorema è banale perché un insieme affine è l'unico insieme convesso senza frontiere. Sia $\partial K \neq \emptyset$ e convesso. Senza perdita di generalità possiamo supporre $\dim K = n$. Siano $x \in \overset{\circ}{K}$ e $y \notin \bar{K}$. Quindi esistono due piani P_x e P_y che separano rispettivamente x da ∂K e y da \bar{K} . Sia P_x^+ il semispazio contenente x e P_y^- il semispazio contenente y . Quindi $P_x^+ \subset \overset{\circ}{K}$ (perché P_x^+ è separato dalla frontiera) e $P_y^- \subset R^n \setminus \bar{K}$. Allora $P_x^+ \cap P_y^- = \emptyset$. Ciò può avvenire solo se P_x e P_y sono paralleli. Questo implica anche che P_z e P_x sono paralleli $\forall z \in \overset{\circ}{K}$. Allora $H_1 = \cup_{x \in \overset{\circ}{K}} P_x^+$ e $H_2 = \cup_{y \notin \bar{K}} P_y^-$ sono semispazi. Ovviamente $H_1 \supset \overset{\circ}{K}$ e $H_2 \supset R^n \setminus \bar{K}$. Quindi $H_1 = \overset{\circ}{K}$ e $H_2 = R^n \setminus \bar{K}$ e quindi \bar{K} è un semispazio chiuso.

(\Leftarrow) banale. ■

4.49 COROLLARIO. Sia K convesso. Se, per un certo $\bar{x} \in K$ e un certo $d \in \mathbb{R}^n$ avviene $\bar{x} + \alpha d \in K$, $\forall \alpha \geq 0$, allora $x + \beta d \in \bar{K}$ per ogni $x \in K$ e ogni $\beta \geq 0$.

DIMOSTRAZIONE. Da $x \in K$, $\bar{x} + \alpha d \in K$ e la convessità di K si ha

$$y(\alpha) := \frac{\alpha}{\alpha + \beta} x + \frac{\beta}{\alpha + \beta} (\bar{x} + \alpha d) \in K$$

e $y(\alpha) \rightarrow x + \beta d$ se $\alpha \rightarrow \infty$. ■

4.50 TEOREMA. Se K è chiuso, convesso e non vuoto e inoltre non è un insieme affine né un semispazio affine, allora $K = \text{conv } \partial K$.

DIMOSTRAZIONE. In base alle ipotesi e al lemma precedente ∂K non è convesso, quindi esistono due punti $x_1, x_2 \in \partial K$ tali che $y = \alpha x_1 + (1 - \alpha)x_2 \in \overset{\circ}{K}$ per un certo $\alpha \in (0, 1)$. In base al teorema 4.36 si ha $\{y : y = \alpha x_1 + (1 - \alpha)x_2, \forall \alpha \in (0, 1)\} \subset \overset{\circ}{K}$. Il segmento $\text{conv}(x_1, x_2) = \text{aff}(x_1, x_2) \cap K$ è ovviamente limitato, ma ciò avviene anche per $H \cap K$ dove H è una qualsiasi retta parallela ad $\text{aff}(x_1, x_2)$, per il corollario precedente. Quindi è sufficiente, per ogni punto y di $\overset{\circ}{K}$, considerare la retta passante per y e parallela ad $\text{aff}(x_1, x_2)$ e le due intersezioni y_1 e y_2 della retta con ∂K . Allora y è combinazione convessa di y_1 e y_2 . ■

Le due rappresentazioni ottenute sono ancora ridondanti. Tuttavia prima di operare un ulteriore raffinamento è necessario introdurre alcuni concetti nuovi.

4.5. Facce

4.51 DEFINIZIONE. Una faccia F di un insieme convesso K è un qualsiasi sottoinsieme convesso tale che:

$$x \in F, z, y \in K, 0 < \alpha < 1, x = \alpha z + (1 - \alpha)y \Rightarrow z, y \in F. \quad \blacksquare$$

In base alla definizione anche K e l'insieme vuoto sono facce. Queste facce vengono dette *improprie*. In tutti gli altri casi le facce vengono dette *proprie* e per queste si ha:

4.52 TEOREMA. Sia F una faccia propria di K . Allora $F \subset \partial K$, $\text{aff}(F) \cap \overset{\circ}{K} = \emptyset$ e $\dim F < \dim K$.

DIMOSTRAZIONE. Sia $z \in K \cap \partial K$ e $y \in \overset{\circ}{K}$. Siccome $y \in \overset{\circ}{K}$ esiste $\varepsilon > 0$ tale che $y_\varepsilon := y + \varepsilon(y - z) \in \overset{\circ}{K}$. Ora $y = \varepsilon/(1 + \varepsilon)z + 1/(1 + \varepsilon)y_\varepsilon$ e, se $y \in F$, per definizione di faccia si deve avere anche $z, y_\varepsilon \in F$. Essendo z, y arbitrari in K si ha $K \subset F$ cioè $K = F$. Avendo escluso quest'ipotesi non può quindi essere che $y \in \overset{\circ}{K}$ e $y \in F$. Quindi $F \subset \partial K$.

Siccome F è convesso e $F \cap \overset{\circ}{K} = \emptyset$ esiste un piano separatore fra F e $\overset{\circ}{K}$. Si noti che questo piano separatore contiene F in quanto $F \subset \partial K$. Allora $\text{aff}(F)$ è contenuto nel piano separatore e da questo discende il resto della tesi. ■

Vale inoltre il teorema inverso.

4.53 TEOREMA. Sia K chiuso e convesso, $\dim K > 0$, H piano tale che $\emptyset \neq H \cap K \subset \partial K$. Allora $H \cap K$ è una faccia propria di K .

DIMOSTRAZIONE. Sia $x \in H \cap K$ e $z, y \in K$ con $x = \alpha z + (1 - \alpha)y$ e $0 < \alpha < 1$. Si noti che $\overset{\circ}{K} \subset H^+$ dove H^+ è uno dei due semispazi chiusi individuati da H . Infatti se $u \in \overset{\circ}{K} \cap H^-$ e $v \in \overset{\circ}{K} \cap H^+$ allora esiste un punto $w \in \text{conv}(u, v)$ tale che $w \in H$ e $w \in \overset{\circ}{K}$ (perché $\overset{\circ}{K}$ è convesso), ma $H \cap \overset{\circ}{K} = \emptyset$ dato che $H \cap K \subset \partial K$. Sempre in base al teorema 4.36 si ha $z \notin \overset{\circ}{H}^+, y \notin \overset{\circ}{H}^+$ altrimenti $x \in \overset{\circ}{H}^+$ (mentre $x \in H$). Siccome $z, y \in K$ si ha che $z \notin \overset{\circ}{H}^-$ e $y \notin \overset{\circ}{H}^-$ e quindi $z \in H$ e $y \in H$. ■

4.54 COROLLARIO. Sia K chiuso e convesso, $\dim K > 0$. Se $\partial K \neq \emptyset$ esiste almeno una faccia propria di K .

DIMOSTRAZIONE. Sia $\bar{x} \in \partial K$. Il piano che separa \bar{x} da $\overset{\circ}{K}$ soddisfa alle ipotesi del precedente teorema. ■

Inoltre in base al successivo teorema vedremo che l'insieme delle facce di un insieme convesso forma un reticolo ordinato secondo la dimensione.

4.55 TEOREMA. Se F è una faccia di K e G è una faccia di F , allora G è una faccia di K .

DIMOSTRAZIONE. Sia $x \in G$ e $z, y \in K$ tali che $x = \alpha z + (1 - \alpha)y$ e $0 < \alpha < 1$. Siccome G è una faccia di F , si ha $x \in F$. Allora $z, y \in F$ perché F è una faccia di K , e siccome G è una faccia di F , vale anche $z, y \in G$. ■

Le facce di dimensione 1 prendono il nome di *spigoli*, quelle di dimensione 0 il nome di *punti estremi*. Si noti in particolare che un punto estremo non può essere espresso come combinazione convessa stretta di altri punti dell'insieme.

I punti estremi di un insieme convesso hanno un'importanza fondamentale perché, come vedremo fra poco, sono in grado di generare tramite combinazioni convesse tutto l'insieme se certe condizioni, molto blande, sono verificate. Innanzitutto bisogna chiedersi sotto quali condizioni esistano punti estremi.

4.56 TEOREMA. Un insieme chiuso e convesso ha punti estremi se e solo se non contiene rette.

DIMOSTRAZIONE. (\Rightarrow) Sia R una retta contenuta nell'insieme chiuso e convesso K . Per definizione di convessità $\alpha K + (1 - \alpha)K = K$, $\forall 0 \leq \alpha \leq 1$. Quindi $\alpha R + (1 - \alpha)x \subset K$, $\forall x \in K$, $\forall \alpha \in [0, 1]$. Ma $\alpha R = R$ se $\alpha \neq 0$. Quindi $R + (1 - \alpha)x \subset K$, $\forall x \in K$, $\forall \alpha \in (0, 1]$. Siccome K è chiuso anche $R + x \subset K$, quindi x non può essere punto estremo dato che può essere espresso come combinazione convessa di altri punti.

(\Leftarrow) Se K non ha punti estremi esiste almeno una faccia F (eventualmente K stesso) tale che $\dim F > 0$. Dal corollario 4.54 si deduce $\partial F = \emptyset$, ovvero $F = \overset{\circ}{F}$. Siccome F è chiusa, non può essere che un insieme affine di dimensione maggiore di zero. ■

In base alle considerazioni sin qui svolte, per ogni punto estremo passa almeno un piano di supporto. Questo punto è tuttavia l'unico punto di K ad appartenere al piano di supporto?

Ovviamente se l'intersezione di un piano di supporto con K si riduce ad un punto, allora questo è anche estremo. Il contrario non è però vero in generale. Punti estremi che godano anche di questa proprietà prendono il nome di punti esposti. Più esattamente si ha:

4.57 DEFINIZIONE. Sia dato un insieme convesso K . Un punto $x \in \partial K$ per il quale esiste un piano di supporto H tale che $K \cap H = x$, prende il nome di punto esposto. ■

Quindi, indicando con $\text{ext}(K)$ l'insieme dei punti estremi di K e con $\text{exp}(K)$ quello dei punti esposti, vale sempre $\text{exp}(K) \subset \text{ext}(K)$. L'inclusione può anche essere stretta come il seguente esempio dimostra:

$$K = \{(x, y) \in \mathbb{R}^2 : y \geq x^2 \text{ se } -1 \leq x \leq 1; y \geq 2|x| - 1 \text{ altrimenti}\}$$

I punti $(1, 1)$ e $(-1, 1)$ sono estremi ma non esposti. Tuttavia l'inclusione è di tipo particolare. Infatti i punti esposti risultano essere densi in quelli estremi (la dimostrazione è omessa).

4.58 TEOREMA. Sia K convesso e chiuso. Allora $\overline{\text{exp}}(K) = \text{ext}(K)$. ■

Se un insieme convesso è illimitato può succedere che alcune facce abbiano dimensione unitaria e siano in particolare delle semirette emananti da un punto estremo. È come se si trattasse di uno spigolo con un punto estremo all'infinito. È chiaro che combinazioni convesse di punti estremi non possono generare tale semiretta a meno che non si tenga conto in qualche modo del punto all'infinito. A tale scopo si introducono i concetti di direzione e direzione estrema.

4.59 DEFINIZIONE. Una direzione di un insieme convesso $K \subset R^n$ è un elemento $d \in R^n$ tale che $x + \alpha d \in \overline{K}$, $\forall \alpha \geq 0$ per $x \in K$. ■

Si ricorda che, dal lemma 4.49, se $\bar{x} + \alpha d \in \overline{K}$, $\forall \alpha \geq 0$, è soddisfatta per un certo $\bar{x} \in K$ allora deve essere soddisfatta per ogni $x \in K$.

4.60 ESERCIZIO. Dimostrare che le direzioni di un insieme convesso formano un cono convesso. ■

4.61 DEFINIZIONE. Una direzione estrema di un insieme chiuso e convesso $K \subset R^n$ è un elemento $d \in R^n$ tale che esiste uno spigolo S di K per cui $x + \alpha d \in S$, $\forall \alpha \geq 0$, $\forall x \in S$. ■

Anche per le direzioni si può distinguere fra direzioni estreme ed esposte. Analogamente si può definire:

4.62 DEFINIZIONE. Una direzione esposta di un insieme convesso K è una direzione di K tale che il corrispondente spigolo illimitato è eguale all'intersezione di K e un piano di supporto a K . ■

E, in modo simile, vale il teorema:

4.63 TEOREMA. Le direzioni esposte sono dense in quelle estreme. ■

L'insieme delle direzioni estreme di K viene indicato con $\text{dext}(K)$ e quello delle direzioni esposte con $\text{dexp}(K)$.

La relazione che passa fra punti di frontiera e punti esposti può essere traslata ai piani di supporto e a particolari piani di supporto, semplicemente invertendo i termini nella definizione.

4.64 DEFINIZIONE. Sia dato un insieme convesso K . Un piano di supporto H a K per il quale esiste un punto $x \in K$ tale che H è l'unico piano di supporto contenente x , prende il nome di piano tangente. ■

Il semispazio di supporto di un piano tangente prende il nome di *semispazio tangente*. Possiamo a questo punto dimostrare il seguente teorema fondamentale:

4.65 TEOREMA. Se K non contiene rette ed è chiuso e convesso allora

$$K = \text{conv}(\text{ext}(K)) + \text{cone}(\text{dext}(K))$$

DIMOSTRAZIONE. Si era già dimostrato precedentemente che $K = \text{conv}(\partial K)$. Basta tener presente che ∂K è l'unione di facce che, essendo convesse, chiuse e senza rette, sono a loro volta rappresentabili come combinazioni convesse di loro facce. Continuando si arriva a facce di dimensione unitaria che sono quindi, o combinazione convessa di punti estremi oppure direzioni estreme. Siccome la catena di facce e sottofacce è finita, perché la dimensionalità delle facce è almeno di uno superiore a quello delle sottofacce, si vede che ogni punto di K è rappresentabile come combinazione convessa finita di punti estremi di K più una combinazione conica di direzioni estreme. ■

Il teorema 4.65 è illustrato schematicamente in figura 4.13. Si noti che, se K contiene rette, o più in generale un sottospazio S , allora si ha $K \cap S^\perp + S = K$, e se B è una base di S , possiamo scrivere

$$K = \text{conv}(\text{ext}(K \cap S^\perp)) + \text{cone}(\text{dext}(K \cap S^\perp)) + \text{lin}(B)$$

Il seguente teorema (di facile dimostrazione) lega la rappresentazione interna dell'insieme ai suoi punti esposti:

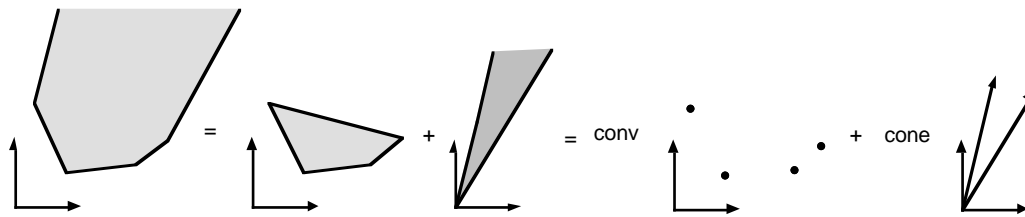


FIGURA 4.13

4.66 TEOREMA. *Sia K chiuso, convesso e non contenente rette, allora*

$$K = \overline{\text{conv}}(\text{exp}(K) + \overline{\text{cone}}(\text{dexp}(K)))$$

■

Questi ultimi teoremi costituiscono il massimo raffinamento possibile della rappresentazione interna di un insieme convesso. Per ciò che riguarda la rappresentazione esterna si era visto che un insieme convesso poteva essere visto come l'intersezione di tutti i semispazi di supporto. Siccome dualmente piani di supporto corrispondono a punti generici della frontiera, viene da chiedersi se bastano alcuni piani di supporto particolari a rappresentare esternamente l'insieme. Ciò è effettivamente possibile. Infatti, come si dimostrerà fra poco, sono sufficienti i piani tangenti. La dimostrazione di questa proprietà richiede tuttavia alcune nozioni preliminari.

4.6. Polarità

4.67 DEFINIZIONE. *Sia $K \subset \mathbb{R}^n$ un insieme chiuso e convesso contenente l'origine. Si definisce suo insieme polare l'insieme:*

$$K^* = \{w \in \mathbb{R}^n : wx \leq 1, \forall x \in K\}$$

■

L'insieme polare gode delle seguenti proprietà di facile dimostrazione:

- 1) K^* è chiuso e convesso;
- 2) $K^{**} = \overline{\text{conv}}(K)$ ($= K$ se K è chiuso e convesso);
- 3) K^* è limitato se e solo se $0 \in \overset{\circ}{K}$;
- 4) $0 \in \overset{\circ}{K}^*$ se e solo se K è limitato;
- 5) $0 \in K^*$;
- 6) $\dim K = n$ se e solo se K^* non contiene rette;
- 7) K non contiene rette se e solo se $\dim K^* = n$;
- 8) Sia $K = \{x : wx \leq 1\}$; allora $K^* = \text{conv}(0, w)$;

Inoltre si ha:

4.68 TEOREMA. Siano A e B convessi, $0 \in A \cap B$. Allora

$$(A \cap B)^* = \text{conv}(A^* \cup B^*)$$

DIMOSTRAZIONE. Sia $w \in \text{conv}(A^* \cup B^*)$. Allora dal teorema di Caratheodory esistono al più $m + 1$ punti $w_k \in A^* \cup B^*$ tali che $w \in \text{conv}\{w_k\}$. Siccome per ogni k vale $w_k x \leq 1$, $\forall x \in A$, oppure $w_k x \leq 1$, $\forall x \in B$, si ha che per ogni k vale $w_k x \leq 1$, $\forall x \in A \cap B$, e quindi si deduce facilmente $w \in (A \cap B)^*$.

Sia $w \notin \text{conv}(A^* \cup B^*)$. Siccome $\text{conv}(A^* \cup B^*)$ è chiuso, esiste un piano separatore, cioè esiste z tale che

$$wz > 1 \text{ e } w'z \leq 1 \quad \forall w' \in \text{conv}(A^* \cup B^*)$$

Questo implica $w'z \leq 1 \quad \forall w' \in A^*$ e $w'z \leq 1 \quad \forall w' \in B^*$, cioè $z \in \overline{A} \cap \overline{B} = \overline{(A \cap B)}$. Ma siccome $wz > 1$, $w \notin ((A \cap B))^* = (A \cap B)^*$. ■

Si veda in figura 4.14 un insieme ottenuto come intersezione dei due semispazi $A := \{x : ax \leq 1\}$ e $B := \{x : bx \leq 1\}$. Il suo polare è rappresentato in figura 4.15 come combinazione convessa di $A^* = \text{conv}(0, a)$ e $B^* = \text{conv}(0, b)$.

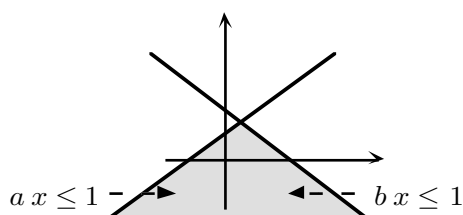


FIGURA 4.14

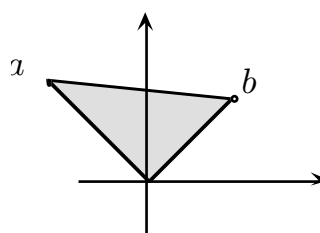


FIGURA 4.15

Non è difficile estendere il teorema ad arbitrarie famiglie di insiemi convessi.

4.69 COROLLARIO. Siano A_α insiemi convessi tali che $0 \in \bigcap_\alpha A_\alpha$. Allora

$$\left(\bigcap_\alpha A_\alpha\right)^* = \overline{\text{conv}}\left(\bigcup_\alpha A_\alpha^*\right)$$

DIMOSTRAZIONE. Come nel teorema $\text{conv}(\bigcup_\alpha A_\alpha^*) \subset (\bigcap_\alpha A_\alpha)^*$. Siccome $(\bigcap_\alpha A_\alpha)^*$ è chiuso, si ha anche $\overline{\text{conv}}(\bigcup_\alpha A_\alpha^*) \subset (\bigcap_\alpha A_\alpha)^*$. Per il resto basta applicare il teorema di separazione a $\overline{\text{conv}}(\bigcup_\alpha A_\alpha^*)$. ■

4.70 TEOREMA. Sia $\dim K = n$ e $0 \in \overset{\circ}{K}$, allora i punti esposti di K^* , esclusa eventualmente l'origine, sono tutti e soli i piani tangenti a K .

DIMOSTRAZIONE. Siccome $0 \in \overset{\circ}{K}$ e $\dim K = n$, ogni piano di supporto si può rappresentare come $wx = -1$. Sia $W := \{w : wx = -1 \text{ è un piano di supporto}\}$. Sia \bar{w} corrispondente ad un piano tangente. Per definizione esiste \bar{x} tale che $\bar{w}\bar{x} = -1$ e $w\bar{x} > -1$, $\forall w \in W$, $w \neq \bar{w}$. Inoltre, dalla definizione di K^* si ha che $wx = -1$ per qualche $-w \in K^*$ e $x \in K$ implica $w \in W$. Quindi non esiste $-w \in K^*$, $w \neq \bar{w}$, tale che $w\bar{x} = -1$ e quindi $-\bar{w}$ è un punto esposto di K^* .

Viceversa se $w \neq 0$ è un punto esposto di K^* esiste un piano di supporto a K^* contenente solo w come elemento di K^* . Siccome $0 \in K^*$ e $w \neq 0$, il piano di supporto non contiene l'origine e può essere rappresentato da un'equazione del tipo $w(-z) = 1$. Quindi $w'(-z) < 1$, $\forall w' \in K^*$. Questo implica $-z \in K$ con w per definizione piano tangente a K in $-z$. ■

L'importanza del precedente teorema risiede nel fatto che affermazioni riguardanti piani di supporto si trasformano in affermazioni riguardanti punti estremi. Infatti si ha:

4.71 TEOREMA. Sia K un corpo convesso chiuso. Allora K è dato dall'intersezione dei suoi semispazi tangenti.

DIMOSTRAZIONE. Si trasli K in modo che $0 \in \overset{\circ}{K}$. Sia $W_\alpha := \{x : w_\alpha x \geq -1\}$ un semispazio di supporto. Dal teorema 4.47 si ha $K = \bigcap_\alpha W_\alpha$ e dal corollario 4.69 $K^* = \overline{\text{conv}}(0, \bigcup -w_\alpha)$. Quindi (indicando $[w, 0] := \text{conv}(w, 0)$):

$$K = \bigcap_{w \in K^*} [w, 0]^* = \bigcap_{w \in \text{conv}(\text{ext } K^*)} [w, 0]^* = \bigcap_{w \in \overline{\text{conv}}(\text{exp } K^*)} [w, 0]^* =$$

(la seconda uguaglianza dal teorema 4.65 in quanto K^* è limitato e la terza dal teorema 4.66)

$$= \bigcap_{w \in \text{conv}(\text{exp } K^*)} [w, 0]^* = \bigcap_{w \in \text{exp } K^*} [w, 0]^* = \bigcap_{w \in \text{exp } K^* \setminus 0} [w, 0]^*$$

(la prima uguaglianza da $x : wx \leq 1, \forall w \in Q, \implies wx \leq 1, \forall w \in \overline{Q}$, la seconda da $x : wx \leq 1, \forall w \in Q \iff x : wx \leq 1, \forall w \in \text{conv } Q$ e la terza da $[0, 0]^* = \mathbb{R}^n$) Quindi in base al precedente teorema vale la tesi. ■

Il teorema 4.66 è duale rispetto al teorema 4.71. Si notino in particolare le ipotesi di non esistenza di rette appartenenti a K e di $\dim K = n$. Anche queste ipotesi sono duali fra di loro come si vede dalla relazione intercorrente fra K e K^* .

L'ipotesi $\dim K = n$ non può essere eliminata. Si può tuttavia estendere il precedente risultato nel seguente modo: Sia S il sottospazio corrispondente ad $\text{aff } K$. Allora il teorema precedente vale in $\text{aff } K$. Se \hat{H} è un semispazio di $\text{aff } K$ allora $\hat{H} + S^\perp$ è un semispazio di \mathbb{R}^n . Inoltre se $B \subset S^\perp$ è una base in S^\perp le disequazioni $bx \geq bx_0$ e $bx \leq bx_0, \forall b \in B$ e per un generico $x_0 \in \text{aff } K$, definiscono esattamente $\text{aff } K$. E allora K è definito dall'intersezione dei sottospazi 'quasi tangenti' $\hat{H} + S^\perp$ più i $2(n - \dim S)$ semispazi definenti $\text{aff } K$.

Nel caso C sia un cono la definizione di polarità viene modificata in:

4.72 DEFINIZIONE. Sia $C \subset \mathbb{R}^n$ un cono. Si definisce suo cono polare l'insieme:

$$C^* := \{y : yx \leq 0, \forall x \in C\}$$

Un esempio di coni polari è già stato introdotto nell'esercizio 4.46. Il cono (opposto di quello definito nell'esercizio)

$$\{a \in \mathbb{R}^n : (-a)x = (-a)\bar{x} \text{ è un piano di supporto a } K\}$$

è proprio $D^*(K, \bar{x})$. Si veda la figura 4.16.

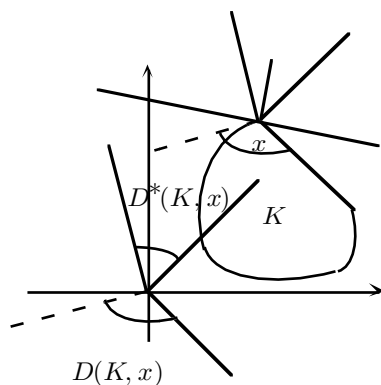


FIGURA 4.16

Di fatto il cono polare è l'insieme polare dato che non vi può essere nessun elemento $y \in C^*$ tale che $0 < yx \leq 1$ per qualche elemento $x \in C$. Ovviamente C^* è un cono e vale $C^{**} = C$ se C è chiuso e convesso. Se $C := \{x : wx \leq 0\}$ allora $C^* = \text{cone}\{w\}$. Anche se il seguente risultato è una conseguenza del teorema 4.68 ne diamo comunque una dimostrazione indipendente.

4.73 TEOREMA. Sia $C := \{x : a^i x \leq 0, i = 1, \dots, m\}$. Allora $C^* = \text{cone}\{a^i\}_{i=1}^m =: D$.

DIMOSTRAZIONE. Dimostriamo che $D \subset C^*$. Sia $y \in D$. Quindi $y = \sum_i \alpha_i a^i$, con $\alpha_i \geq 0$, e ovviamente $yx \leq 0$ per ogni $x \in C$. Quindi $y \in C^*$.

Per dimostrare che $C^* \subset D$ bisogna invocare il teorema di separazione. Sia pertanto \bar{y} tale che $\bar{y} \in C^*$ e $\bar{y} \notin D$. Per la chiusura di D esiste un piano che separa strettamente \bar{y} da D , cioè esiste \bar{x} tale che $y\bar{x} \leq \beta < \bar{y}\bar{x}$ per ogni $y \in D$. Si noti che $y\bar{x} \leq 0, \forall y \in D$, altrimenti la disuguaglianza $y\bar{x} \leq \beta$ verrebbe violata (D è un cono) e quindi $\bar{x} \in D^*$. Siccome poi $0 \in D$ si ha $\beta = 0$. Allora $\bar{y}\bar{x} > 0$ con $\bar{y} \in C^*$. Quindi $\bar{x} \notin C$. Dimostriamo ora che $D^* \subset C$, con il che raggiungeremo la contraddizione. Si ha per definizione $D^* = \{x : \sum_i \alpha_i a^i x \leq 0, \forall \alpha_i \geq 0\}$, quindi $x \in D^*$ implica $a^i x \leq 0, \forall i$, e cioè $x \in C$. ■

Il seguente risultato è un'importante conseguenza del precedente teorema e costituisce una parte del celebre lemma di Farkas (Farkas [1895]).

4.74 LEMMA. Sia $a^0 x \leq 0$ per ogni x che soddisfa $a^i x \leq 0, i = 1, \dots, m$. Allora esistono coefficienti $\alpha_i \geq 0$ tali che $a^0 = \sum_i \alpha_i a^i$.

DIMOSTRAZIONE. Sia $C := \{x : a^i x \leq 0, \forall i\}$ e $D := \text{cone}\{a^i\}_{i=1}^m$. L'ipotesi quindi afferma che $a^0 x \leq 0$ per ogni $x \in C$ che è quanto dire $a^0 \in C^*$ e, dal precedente teorema $a^0 \in D$ da cui la tesi. ■

La seconda parte del lemma di Farkas corrisponde all'implicazione inversa:

4.75 LEMMA. Esista \hat{x} tale che $a^0 \hat{x} > 0$ e $a^i \hat{x} \leq 0, i = 1, \dots, m$. Allora non esistono coefficienti $\alpha_i \geq 0$ tali che $a^0 = \sum_i \alpha_i a^i$.

DIMOSTRAZIONE. Il vettore \hat{x} corrisponde ad un piano di equazione $a\hat{x} = 0$ che separa strettamente il punto a^0 dai punti a^i . Inoltre il piano separatore contiene l'origine. La separazione si estende quindi agli involuppi conici di a^0 e a^i rispettivamente, da cui la tesi. ■

Riunendo assieme i due risultati si ha pertanto

4.76 TEOREMA. (lemma di Farkas) Siano $a^0, a^1, \dots, a^m \in \mathbb{R}^n$. Si considerino le due seguenti affermazioni:

- $\{x \in \mathbb{R}^n : a^0 x > 0, a^i x \leq 0, i = 1, \dots, m\} = \emptyset$
- $\{\alpha \in \mathbb{R}^m : a^0 = \sum_i \alpha_i a^i, \alpha_i \geq 0\} = \emptyset$

Una qualsiasi delle due affermazioni è vera se e solo se l'altra è falsa. ■

Fra le molti varianti del lemma di Farkas la seguente verrà sfruttata nella prossima sezione:

4.77 COROLLARIO. Si considerino le due seguenti affermazioni:

- $\{x : x \geq 0, a^0 x < 0, a^i x = 0, i = 1, \dots, m\} = \emptyset$
- $\{\alpha : a^0 \geq \sum_i \alpha_i a^i\} = \emptyset$

Una qualsiasi delle due affermazioni è vera se e solo se l'altra è falsa. ■

DIMOSTRAZIONE. Si scriva il primo insieme come

$$\{x : (-a^0)x > 0, (-a^i)x \leq 0, a^i x \leq 0, (-e^i)x \leq 0, i = 1, \dots, m\}$$

dove e^i è la riga i -ma della matrice identica. A questo punto possiamo applicare il lemma di Farkas ottenendo l'altro insieme

$$\left\{ (\alpha^+, \alpha^-, \beta) : -a^0 = \sum_i \alpha_i^- (-a^i) + \sum_i \alpha_i^+ a^i - \beta, \alpha^+ \geq 0, \alpha^- \geq 0, \beta \geq 0 \right\}$$

da cui, ponendo $\alpha := \alpha^+ - \alpha^-$ e notando che α non ha vincoli di segno, e trasformando la presenza di $\beta \geq 0$ in disequaglianza si ottiene la tesi. ■

4.7. Poliedri

Quando la rappresentazione di un insieme convesso è affidata ad un numero finito di elementi, allora l'insieme convesso gode di particolari proprietà.

4.78 DEFINIZIONE. *L'intersezione di un numero finito di semispazi chiusi prende il nome di poliedro.* ■

Quindi un poliedro è un insieme convesso con rappresentazione esterna finita. Si noti che non necessariamente un poliedro è limitato. Spesso si usa il termine *polìtopo* per designare un poliedro limitato. Sorge la domanda se una rappresentazione esterna finita sia condizione necessaria e sufficiente per una rappresentazione interna finita, cioè come combinazione convessa di un numero finito di punti e conica di un numero finito di direzioni. La risposta è affermativa, come si può intuire, tuttavia la dimostrazione non è immediata. Dimostriamo ora che un poliedro ha una rappresentazione interna finita. Sia $P := \{x \in R^n : Ax \leq b\}$ un poliedro dove A è una matrice $m \times n$.

4.79 TEOREMA. *ext P e dext P sono insiemi finiti.*

DIMOSTRAZIONE. Sia $\bar{x} \in P$ e siano A e b opportunamente ripartiti in modo da avere:

$$A\bar{x} = \begin{pmatrix} A_1 \\ A_2 \end{pmatrix} \bar{x} \quad \text{con} \quad \begin{array}{l} A_1 \bar{x} = b_1 \\ A_2 \bar{x} < b_2 \end{array}$$

Dimostriamo dapprima che \bar{x} è un punto estremo se e solo se $\text{rank } A_1 = n$. Infatti se $\text{rank } A_1 < n$ allora esiste $y \neq 0$ tale che $A_1 y = 0$ e quindi per $\varepsilon > 0$ sufficientemente piccolo si ha $\bar{x} \pm \varepsilon y \in P$ e \bar{x} non può essere punto estremo. Viceversa se \bar{x} non è punto estremo esistono $z, y \in P$, $z \neq y \neq \bar{x}$ tali che $\bar{x} = \alpha z + (1 - \alpha)y$ con $0 \leq \alpha \leq 1$. Quindi

$$b_1 = A_1 \bar{x} = \alpha A_1 z + (1 - \alpha) A_1 y \leq \alpha b_1 + (1 - \alpha) b_1 = b_1 \quad \Rightarrow \quad A_1 z = A_1 y = b_1,$$

cosa impossibile se $\text{rank } A_1 = n$.

Inoltre le direzioni di P , cioè quei punti d tali che $x + \alpha d \in P$, $\forall x \in P$, $\forall \alpha \geq 0$, sono dati dalle soluzioni di $A d \leq 0$ come si vede facilmente. In modo analogo si può dimostrare che se \bar{d} è tale che $A_1 \bar{d} = 0$ e $A_2 \bar{d} < 0$ allora \bar{d} è una direzione estrema se e soltanto se $\text{rank } A_1 = n - 1$.

Si tenga ora presente che due punti estremi o due direzioni estreme sono distinti solo se le ripartizioni di A sopra indicate sono distinte. Pertanto, siccome il numero di ripartizioni distinte di A è finito, discende che il numero di punti estremi e di direzioni estreme è finito. ■

I punti estremi di un poliedro prendono il nome di *vertici* e, dato che sono in numero finito, coincidono con i punti esposti.

4.80 ESERCIZIO. Siano A e B due poliedri e sia $C = A + B$. Si dimostri:

- che C è un poliedro;
- che ogni vertice di C è ottenuto soltanto come somma di un vertice di A e di un vertice di B ;
- che vertici $\hat{x} \in A$ e $\hat{y} \in B$ generano un vertice $\hat{z} = \hat{x} + \hat{y} \in C$ se e solo se esiste una funzione lineare, per la quale \hat{x} , \hat{y} e \hat{z} sono minimi stretti nei rispettivi poliedri. ■

Si supponga ora di avere un numero finito di punti e direzioni. La loro combinazione convessa dà luogo ad un poliedro, cioè un insieme che sia l'intersezione di un numero finito di semispazi? Il seguente teorema dà appunto una risposta affermativa.

4.81 TEOREMA. $K = \text{conv}(x_1, \dots, x_h) + \text{cone}(d_1, \dots, d_k)$ è un poliedro.

DIMOSTRAZIONE. Sia $\dim K = n$, altrimenti si ragiona su $\text{aff } K$ (vedi la discussione che segue il teorema 4.71). Sia allora senza perdita di generalità $0 \in \overset{\circ}{K}$. Quindi

$$K^* = \{w : w x_i \leq 1, \forall i = 1, \dots, h \text{ e } w d_j \leq 0, \forall j = 1, \dots, k\}$$

Per definizione quindi K^* è un poliedro. Dal teorema precedente i suoi punti estremi sono finiti e coincidono con i punti esposti, quindi, dal teorema 4.70, i piani tangenti a K sono in numero finito, e, dal teorema 4.71 discende che K è un poliedro. ■

La struttura delle facce di un poliedro è particolarmente importante. Il poliedro polare (supponendo di avere traslato l'origine all'interno del poliedro e di avere normalizzato in $Ax \leq \mathbf{1}$) è l'involuppo convesso dei vettori A^i . Quelli fra questi che risultano vertici del poliedro polare danno luogo alle facce di dimensione $n - 1$, dette *faccette*, corrispondenti ai piani tangenti, e sono ovviamente indispensabili per rappresentare il poliedro. Tutti gli altri vettori sono ridondanti e possono essere omessi nella rappresentazione.

4.82 TEOREMA. Sia $\dim P = n$. Se esiste un punto \hat{x} ed un indice k tali che $A^k \hat{x} = b_k$ e $A^i \hat{x} < b_i$, $i \neq k$, allora la disuguaglianza $A^k x \leq b_k$ definisce una faccetta di P .

DIMOSTRAZIONE. Supponendo di avere traslato l'origine all'interno del poliedro e di avere normalizzato come $Ax \leq \mathbf{1}$, si vede che nessuna combinazione convessa dei vettori A^i , $i \neq k$, può dare il vettore A^k . Quindi A^k è un vertice del poliedro polare ed una faccetta del poliedro. ■

Non necessariamente un poliedro è un corpo convesso. I seguenti risultati di immediata dimostrazione legano la dimensione di un poliedro alla sua rappresentazione.

4.83 TEOREMA. $\dim P = n$ se e solo se esiste $\hat{x} \in P$ tale che $A\hat{x} < b$. ■

4.84 TEOREMA. Sia $\dim P = n$. Allora $\hat{x} \in \overset{\circ}{P}$ se e solo se $A\hat{x} < b$. ■

4.85 TEOREMA. Sia $\dim P < n$. Allora esiste almeno una disuguaglianza che è soddisfatta come eguaglianza da ogni punto di P .

DIMOSTRAZIONE. Supponiamo che nessuna riga sia sempre soddisfatta come eguaglianza. Quindi esistono m punti $x^1, \dots, x^m \in P$ tali che $A^j x^j < b_j$ e $Ax^j \leq b$. Si prenda una combinazione convessa stretta di questi punti $y = \sum_{j=1}^m \alpha_j x^j$, $\alpha_j > 0$, $\sum_{j=1}^m \alpha_j = 1$. Si ottiene

$$A^i y = A^i \sum_{j=1}^m \alpha_j x^j = \sum_{j=1}^m \alpha_j A^i x^j = \sum_{\substack{j=1 \\ j \neq i}}^m \alpha_j A^i x^j + \alpha_i A^i x^i < \sum_{j=1}^m \alpha_j b_i = b_i$$

in contraddizione con il fatto che $\dim P < n$ per il teorema 4.83. ■

Se $\dim P < n$ si ripartiscano le disequaglianze come

$$Ax = \begin{pmatrix} A^0 \\ A^1 \end{pmatrix} x$$

dove A^0 contiene tutte le righe per cui $A^i x = b_i$ per ogni $x \in P$ e A^1 contiene le rimanenti righe.

4.86 TEOREMA. $\dim P = n - \text{rank } A^0$.

DIMOSTRAZIONE. Con lo stesso ragionamento della precedente dimostrazione esiste $y \in P$ tale che $A^1 y < b_1$. Se $\text{rank } A^0 = k$ esistono $n - k$ vettori linearmente indipendenti z_1, \dots, z_{n-k} tali che $A^0 z_i = 0$. Quindi esiste $\varepsilon > 0$ per cui i punti $x_i := y + \varepsilon z_i$ sono ammissibili. I punti $\{y, x_1, \dots, x_{n-k}\}$ formano un insieme di $n - k + 1$ vettori di P indipendenti in modo affine e quindi $\dim P \geq n - k$. D'altra parte $P \subset \{x : A^0 x = b_0\}$ e, dal teorema 4.23 $\dim \{x : A^0 x = b_0\} = n - k$ da cui $\dim P \leq n - k$. ■

Se un poliedro ha dimensione inferiore ad n si usa comunque il termine di faccetta per le facce di dimensione $(\dim P - 1)$ e il criterio espresso nel teorema 4.82 si riferisce all'insieme di disequaglianze A^1 .

4.87 ESEMPIO. Si abbia

$$\begin{aligned} x_1 + x_2 &\leq 0 \\ -x_1 + x_2 &\leq 0 \\ -x_2 &\leq 0 \\ x_3 &\leq 1 \\ -x_3 &\leq 0 \end{aligned}$$

Si vede subito che l'unico punto ammissibile per x_1 e x_2 è l'origine. Quindi P è semplicemente il segmento che congiunge i punti $(0, 0, 0)$ e $(0, 0, 1)$. Per ogni punto di P si ha

$$\begin{aligned} x_1 + x_2 &= 0 \\ -x_1 + x_2 &= 0 \\ -x_2 &= 0 \end{aligned}$$

e quindi A^0 corrisponde alle prime tre disequaglianze, solo due delle quali sono ovviamente linearmente indipendenti e quindi $\text{rank } A^0 = 2$ e $\dim P = 1$. ■

4.88 ESERCIZIO. Si determinino i vertici, gli spigoli e le faccette del poliedro definito da

$$\begin{aligned} x_1 + x_2 + x_3 &\leq 4 \\ x_1 + x_2 - x_3 &\leq 2 \\ x_1 \geq 0 \quad x_2 \geq 0 \quad x_3 \geq 0 \end{aligned}$$

4.89 ESEMPIO. In questo esempio si vuol far vedere il profondo legame che ci può essere fra poliedri e strutture combinatorie. Il poliedro definito qui sotto ha dimensione $n - 1$ e ogni sua faccia di dimensione k è in corrispondenza con una partizione ordinata (cioè i sottoinsiemi della partizione sono ordinati e la stessa partizione ordinata in modi diversi corrisponde a facce diverse) formata da $n - k$ sottoinsiemi. In particolare i vertici corrispondono alle permutazioni di n elementi e le faccette a sottoinsiemi propri. ■

Sia $N = \{1, \dots, n\}$ e sia definito il seguente poliedro P

$$\sum_{i \in J} x_i \geq \frac{|J|(|J| + 1)}{2} \quad \forall J \subset N, J \neq N$$

$$\sum_{i \in N} x_i = \frac{|N|(|N| + 1)}{2}$$

Il poliedro è definito da $2^n - 2$ disequaglianze (una per ogni sottoinsieme proprio di N) e una eguaglianza. La sua dimensione pertanto può essere al più $n - 1$. Consideriamo il punto con tutte le coordinate uguali a $(n + 1)/2$. Si vede che appartiene al poliedro e tutte le disequaglianze sono soddisfatte strettamente. Quindi A^0 è data da un'unica eguaglianza e applicando il teorema 4.86 si ha che la dimensione del poliedro è $(n - 1)$.

Facciamo ora vedere che ogni disequaglianza genera una faccetta. Si consideri un generico insieme S . Si consideri il punto di coordinate $\hat{x}_i := (|S| + 1)/2$ se $i \in S$ e $\hat{x}_i := (n + |S| + 1)/2$ se $i \notin S$. Ovviamente $\sum_{i \in S} \hat{x}_i = |S|(|S| + 1)/2$ e si può verificare che $\sum_{i \in J} \hat{x}_i > |J|(|J| + 1)/2$ se $J \neq S$. Allora, per il teorema 4.82, ad ogni sottoinsieme proprio di N corrisponde una faccetta.

Analizziamo ora la struttura dei vertici. Ogni vertice è dato dall'intersezione di n piani di cui uno è già fissato. Consideriamo due qualsiasi disequaglianze relative a due sottoinsiemi S e T e supponiamo che siano soddisfatte come eguaglianze. Sommandole si ha

$$\frac{|S|(|S| + 1)}{2} + \frac{|T|(|T| + 1)}{2} = \sum_{i \in S} x_i + \sum_{i \in T} x_i = \sum_{i \in S \cap T} x_i + \sum_{i \in S \cup T} x_i$$

Anche gli insiemi $S \cap T$ e $S \cup T$ corrispondono a due disequaglianze che devono essere soddisfatte per l'ammissibilità e quindi otteniamo

$$\frac{|S|(|S| + 1)}{2} + \frac{|T|(|T| + 1)}{2} \geq \frac{|S \cap T|(|S \cap T| + 1)}{2} + \frac{|S \cup T|(|S \cup T| + 1)}{2}$$

Non è difficile ora verificare che vale sempre la relazione

$$\frac{|S|(|S| + 1)}{2} + \frac{|T|(|T| + 1)}{2} \leq \frac{|S \cap T|(|S \cap T| + 1)}{2} + \frac{|S \cup T|(|S \cup T| + 1)}{2}$$

dove la disequaglianza è sempre stretta tranne quando $S \subset T$ o viceversa. Quindi n equaglianze sono soddisfatte come equazioni solo quando si riferiscono ad n insiemi ordinati per inclusione (ad esempio $\{4\}$, $\{4, 1\}$, $\{4, 1, 3\}$, $\{4, 1, 3, 2\}$ se $n = 4$). Ogni vertice è pertanto associato ad una permutazione di n elementi e un breve calcolo porta ad una permutazione delle coordinate $\{1, 2, \dots, n\}$ per ogni vertice.

Si lascia come esercizio la verifica che gli spigoli corrispondono a due vertici le cui permutazioni corrispondenti si ottengono una dall'altra scambiando due termini e, in generale la verifica che la struttura delle altre facce è quella indicata precedentemente. Inoltre, che relazione c'è con l'esercizio 1.35?

In figura 4.17 è raffigurato il caso $n = 3$ che risulta essere un esagono disegnato già in \mathbb{R}^2 (in realtà l'esagono si trova in \mathbb{R}^3). I numeri esterni indicano quali sottoinsiemi propri di $\{1, 2, 3\}$ danno luogo alle faccette (spigoli in questo caso). I numeri interni indicano la permutazione a cui si riferisce ogni vertice. La permutazione si ottiene guardando in che ordine sono inclusi i sottoinsiemi (cioè i piani) che generano il vertice. Ad esempio la permutazione $\{3, 1, 2\}$ è dovuta al fatto che il vertice è generato dai sottoinsiemi $\{3\}$,

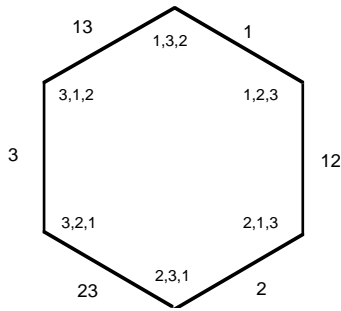


FIGURA 4.17

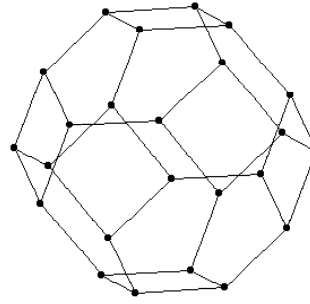


FIGURA 4.18

$\{1, 3\}$ e $\{1, 2, 3\}$ (quest'ultimo è sempre presente perché corrisponde all'unica uguaglianza che definisce il poliedro). Si notino le permutazioni agli estremi di ogni spigolo.

Più complesso è il caso $n = 4$. Il poliedro è raffigurato in figura 4.18. Le faccette sono di due tipi, esagoni e quadrati. Gli esagoni corrispondono a sottoinsiemi di un solo elemento oppure di tre elementi. I quadrati corrispondono a sottoinsiemi di due elementi. Ogni vertice si trova all'intersezione di due esagoni e un quadrato. ■

Spesso si deve proiettare un poliedro su un sottospazio. Necessariamente si ottiene un altro poliedro (perché?). Il problema è di trovare la rappresentazione del poliedro proiettato. Ci occupiamo ora del semplice caso in cui il sottospazio su cui si proietta è quello ottenuto ponendo uguali a zero alcune variabili. Sia quindi

$$P := \{(x, y) \in \mathbb{R}^{n+m} : Ax + By \leq b\}$$

e ci chiediamo quale sia la rappresentazione del poliedro P_x proiezione di P sul sottospazio $\{(x, y) : y = 0\}$. Sia, per ogni $x \in \mathbb{R}^n$,

$$P_y(x) := \{y \in \mathbb{R}^m : By \leq b - Ax\}$$

Il poliedro P_x è anche esprimibile come $P_x = \{x \in \mathbb{R}^n : P_y(x) \neq \emptyset\}$. Dobbiamo allora poter esprimere la condizione che $P_y(x) \neq \emptyset$. A questo scopo invochiamo il lemma di Farkas nella variante del corollario 4.77. Se $P_y(x) \neq \emptyset$ allora non è ammissibile l'insieme $F(x) := \{u : uB = 0, u(b - Ax) < 0, u \geq 0\}$. Inoltre $F(x)$ non è ammissibile se e solo se per tutti gli u tali che $u \geq 0$ e $uB = 0$ si ha $u(b - Ax) \geq 0$. L'insieme degli u che soddisfano le prime due condizioni è un cono poliedrico e quindi la terza condizione può essere espressa limitatamente alle sue direzioni estreme. Siano u^1, \dots, u^q tali direzioni estreme. Allora $P_y(x) \neq \emptyset$ se e solo se $u^i(b - Ax) \geq 0, i = 1, \dots, q$. Pertanto

$$P_x = \{x \in \mathbb{R}^n : u^i(b - Ax) \geq 0, i = 1, \dots, q\}$$

4.90 ESEMPIO. Sia dato il poliedro $P \subset \mathbb{R}^3$ definito da

$$\begin{pmatrix} -2 & -2 & 3 \\ 2 & 0 & -1 \\ 0 & 2 & -1 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \leq \begin{pmatrix} -2 \\ 2 \\ 2 \\ 0 \end{pmatrix} \tag{4.1}$$

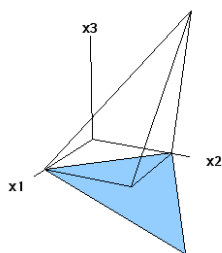


FIGURA 4.19

rappresentato in figura 4.19 insieme con la sua proiezione sul sottospazio delle variabili x_1 e x_2 .

In questo esempio B è data dalla terza colonna della matrice in (4.1). Quindi dobbiamo trovare le direzioni estreme di $u \geq 0$ e $3u_1 - u_2 - u_3 - u_4 = 0$. Le direzioni estreme hanno solo due componenti diverse da zero e quindi sono

$$U := \begin{pmatrix} u^1 \\ u^2 \\ u^3 \end{pmatrix} = \begin{pmatrix} 1 & 3 & 0 & 0 \\ 1 & 0 & 3 & 0 \\ 1 & 0 & 0 & 3 \end{pmatrix}$$

da cui

$$U b = \begin{pmatrix} 4 \\ 4 \\ -2 \end{pmatrix} \quad U A = \begin{pmatrix} 4 & -2 \\ -2 & 4 \\ -2 & -2 \end{pmatrix}$$

e quindi la proiezione è data da:

$$\begin{aligned} 4x_1 - 2x_2 &\leq 4 & 2x_1 - x_2 &\leq 2 \\ -2x_1 + 4x_2 &\leq 4 & \text{cioè} & -x_1 + 2x_2 &\leq 2 \\ -2x_1 - 2x_2 &\leq -2 & & x_1 + x_2 &\geq 1 \end{aligned}$$

■

4.91 ESERCIZIO. Sia dato il poliedro $P := \{x \in \mathbb{R}^n : Ax \leq b\}$ e si voglia trovare la rappresentazione in \mathbb{R}^n del poliedro P_H proiezione di P sul sottospazio $\{x \in \mathbb{R}^n : Hx = 0\}$. Si osservi che $x \in P_H$ se e solo se esiste una combinazione lineare delle righe di H , cioè $H^T y$, per cui si abbia $x + H^T y \in P$. A questo punto si possono applicare i risultati precedenti. ■

4.8. Funzioni convesse

Nello studio di funzioni convesse da \mathbb{R}^n in \mathbb{R} conviene estendere il codominio delle funzioni da \mathbb{R} a $\mathbb{R} \cup \pm\infty =: \bar{\mathbb{R}}$. Si definisce *epigrafo* di una funzione $f: \mathbb{R}^n \rightarrow \bar{\mathbb{R}}$ il seguente insieme:

$$\text{epi } f := \{(x_0, x) \in \mathbb{R} \times \mathbb{R}^n : x_0 \geq f(x)\}$$

(si noti che se $f(\bar{x}) = +\infty$ allora $(x_0, \bar{x}) \notin \text{epi } f, \forall x_0$, mentre se $f(\bar{x}) = -\infty$ allora $(x_0, \bar{x}) \in \text{epi } f, \forall x_0$).

4.92 DEFINIZIONE. Una funzione $f: \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ è convessa se $\text{epi } f$ è un insieme convesso. Inoltre una funzione f è concava se $-f$ è convessa ed è affine se è sia convessa che concava. ■

Il dominio effettivo di una funzione convessa è l'insieme dei punti su cui $f(x) < +\infty$ e viene indicato con $\text{dom } f$. Alternativamente $\text{dom } f$ è la proiezione su \mathbb{R}^n di $\text{epi } f$. Quest'ultima affermazione permette di dire che $\text{dom } f$ è un insieme convesso in quanto ogni trasformazione lineare di un insieme convesso dà luogo ad un insieme convesso.

4.93 ESERCIZIO. È convessa la seguente funzione $f: \mathbb{R}^2 \rightarrow \overline{\mathbb{R}}$:

$$f(x, y) := \begin{cases} +\infty & \text{se } x^2 + y^2 > 1 \\ -\infty & \text{se } x^2 + y^2 < 1 \\ \text{valori arbitrari} & \text{se } x^2 + y^2 = 1 \end{cases}$$

4.94 ESERCIZIO. Sia $f: \mathbb{R}^n \rightarrow \mathbb{R}$ una funzione convessa e $X \subset \mathbb{R}^n$ un insieme finito di punti. Si consideri la funzione:

$$\tilde{f}: \text{epi}(\tilde{f}) = \text{conv}\{(x, f(x) + R^+)\}_{x \in X}$$

Dimostrare che

$$\begin{aligned} \tilde{f}(y) = \min & \sum_{x \in X} \alpha_x f(x) \\ & \sum_{x \in X} \alpha_x x = y \\ & \sum_{x \in X} \alpha_x = 1 \\ & \alpha_x \geq 0 \quad \forall x \in X \end{aligned}$$

per qualsiasi $y \in \mathbb{R}^n$. Dimostrare inoltre, sfruttando la convessità di f che $\tilde{f}(x) = f(x)$ per qualsiasi $x \in X$. Far quindi vedere che $\min_{x \in X} f(x) = \min_{x \in \mathbb{R}^n} \tilde{f}(x)$. ■

Un'altra caratterizzazione di funzioni convesse è data dai seguenti teoremi di facile dimostrazione.

4.95 TEOREMA. $f: \mathbb{R}^n \rightarrow (-\infty, +\infty]$ è convessa se e solo se

$$f\left(\sum_1^m \alpha_k x_k\right) \leq \sum_1^m \alpha_k f(x_k), \quad \forall x_1, \dots, x_m, \quad \forall \alpha_1, \dots, \alpha_m \geq 0: \sum_1^m \alpha_k = 1$$

4.96 TEOREMA. $f: \mathbb{R}^n \rightarrow (-\infty, +\infty]$ è convessa se e solo se

$$f(\alpha x + (1 - \alpha)y) \leq \alpha f(x) + (1 - \alpha)f(y), \quad \forall x, y \quad \forall \alpha: 0 \leq \alpha \leq 1$$

L'importanza delle funzioni convesse risiede nelle loro molte proprietà, le cui due più importanti sono le seguenti:

4.97 TEOREMA. *Sia f convessa. Allora ogni minimo locale è anche globale.*

DIMOSTRAZIONE. Sia x minimo locale e y un punto qualsiasi di $\text{dom } f$. Quindi $f(x) \leq f((1-\alpha)x + \alpha y)$ per $\alpha > 0$ sufficientemente piccolo. Per la convessità si ha

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

cioè $\alpha f(x) \leq \alpha f(y)$ e, siccome $\alpha > 0$, ciò implica $f(x) \leq f(y)$. ■

Spesso gli insiemi ammissibili sono definiti da un vincolo su opportune funzioni. Se l'insieme ammissibile è convesso, il problema è di solito più facile. La seguente condizione, di facile dimostrazione, garantisce che gli insiemi ammissibili siano convessi:

4.98 TEOREMA. *Sia f convessa. Allora gli insiemi $\{x: f(x) \leq \beta\}$ sono convessi $\forall \beta$.* ■

Dai due teoremi discende immediatamente che:

4.99 COROLLARIO. *L'insieme dei minimi di una funzione convessa è convesso.* ■

Naturalmente se una funzione è concava i teoremi precedenti vanno opportunamente modificati, sostituendo 'massimi' al posto di 'minimi' e invertendo il senso della diseguaglianza nella definizione degli insiemi ammissibili. Si noti che i teoremi danno solo condizioni sufficienti.

La convessità di una funzione ha forti implicazioni sulla sua continuità, come vedremo subito. Consideriamo intanto la seguente definizione:

4.100 DEFINIZIONE. *Una funzione $f: \mathbb{R}^n \rightarrow \bar{\mathbb{R}}$ si dice chiusa se $\text{epi } f$ è un insieme chiuso.* ■

Si considerino ad esempio le funzioni:

$$f_1(x) = \begin{cases} 0 & \text{se } x < 0 \\ 1 & \text{se } x \geq 0 \end{cases} \quad f_2(x) = \begin{cases} 0 & \text{se } x \leq 0 \\ 1 & \text{se } x > 0 \end{cases}$$

f_1 non è chiusa, mentre f_2 lo è.

Ovviamente tutte le funzioni continue sono chiuse. Come si comportano le funzioni convesse rispetto alla chiusura e alla continuità? A tale proposito vale il seguente teorema:

4.101 TEOREMA. *Una funzione convessa è continua nell'interno relativo del suo dominio effettivo.*

DIMOSTRAZIONE. Sia (u_1, \dots, u_m) una base di H , generatore di $\text{aff dom } f$, con $\|u_i\|_\infty = 1 \quad \forall i$. Sia $x \in \text{dom } f$ e sia $\rho > 0$ tale che $x + h \in \text{dom } f, \forall h \in H: \|h\|_\infty \leq \rho$. La convessità di f implica le seguenti diseguaglianze ($0 \leq \eta \leq \rho$)

$$\begin{aligned} f(x + \eta u_i) - f(x) &\leq \frac{\eta}{\rho} [f(x + \rho u_i) - f(x)] & \forall i = 1, \dots, m \\ f(x - \eta u_i) - f(x) &\leq \frac{\eta}{\rho} [f(x - \rho u_i) - f(x)] & \forall i = 1, \dots, m \end{aligned}$$

Sia $M = \max_i \{f(x \pm \rho u_i) - f(x)\}$, e allora, se $|\eta| \leq \varepsilon \rho / M$, si ha $f(x \pm \eta u_i) - f(x) \leq \varepsilon$.

Ora $h \in \text{conv}(\pm u_i)_{i=1, \dots, m}$ a motivo della norma $\|\cdot\|_\infty$. Quindi $\forall h: \|h\|_\infty \leq 1$ e $\forall \eta: |\eta| \leq \varepsilon \rho / M$ si ha (con $\alpha_i^+ \geq 0, \alpha_i^- \geq 0, \sum \alpha_i^+ + \alpha_i^- = 1$):

$$\begin{aligned}
f(x + \eta h) - f(x) &= f\left(x + \eta \sum_i \alpha_i^+ u_i + \eta \sum_i -\alpha_i^- u_i\right) - f(x) = \\
&= f\left(\sum_i \alpha_i^+ (x + \eta u_i) + \sum_i \alpha_i^- (x - \eta u_i)\right) - f(x) \leq \\
&= \sum_i \alpha_i^+ f(x + \eta u_i) + \sum_i \alpha_i^- f(x - \eta u_i) - f(x) \leq \varepsilon
\end{aligned}$$

Inoltre dalla convessità di f :

$$2f(x) \leq f(x + \eta h) + f(x - \eta h) \quad \implies \quad f(x) - f(x - \eta h) \leq f(x + \eta h) - f(x)$$

e cioè
$$\forall \eta : |\eta| \leq \frac{\varepsilon \rho}{M} \quad |f(x + \eta h) - f(x)| \leq \varepsilon$$

Dal teorema discende che gli unici punti in cui f può non essere continua devono stare sulla frontiera relativa. Come si vede immediatamente le funzioni convesse chiuse sono quelle funzioni convesse i cui valori di frontiera coincidono con i valori ottenuti estendendo per continuità i valori di f a $\partial \text{dom } f$.

Alcune funzioni di uso frequente sono convesse. Ad esempio:

- 1) La funzione indicatrice di un insieme convesso K :

$$\delta(x, K) = \begin{cases} 0 & \text{se } x \in K \\ +\infty & \text{altrimenti} \end{cases}$$

- 2) Una qualsiasi norma $\|\cdot\|$;
 3) Una forma quadratica semidefinita positiva (cioè $x Q x \geq 0, \forall x$);
 4) Ogni funzione del tipo $f(x) = \sup_{\alpha \in A} \{g_\alpha(x)\}$, dove g_α è un'arbitraria famiglia di funzioni convesse. La dimostrazione della convessità di f è immediata ricordando che $\text{epi } f = \bigcap_{\alpha \in A} \text{epi } g_\alpha$. Naturalmente funzioni del tipo $f(x) = \inf_{\alpha \in A} \{g_\alpha(x)\}$ con g_α un'arbitraria famiglia di funzioni concave, sono concave.

Funzioni convesse (o concave) definite come al punto 4) con le funzioni g_α di tipo affine e insieme A finito sono lineari a tratti. Vedremo che tali funzioni intervengono frequentemente in problemi di ottimizzazione discreta. Per tali funzioni, i punti x per cui il sup (o l'inf) viene realizzato da più di un valore di α vengono detti *punti di rottura*. In tali punti infatti si passa da una regione di linearità ad un'altra con una pendenza diversa.

4.102 ESERCIZIO. Dalle proprietà della norma, ogni insieme $\{x : \|x\| \leq \alpha\}$ è un corpo convesso contenente l'origine nel proprio interno. Tuttavia è interessante vedere che a sua volta ogni corpo convesso contenente l'origine nel proprio interno può definire una norma. Sia dunque K un corpo convesso contenente l'origine nel proprio interno e si definisca la seguente funzione (detta *funzionale di Minkowski*)

$$K(x) := \inf \{\alpha \geq 0 : x \in \alpha K\}$$

Dimostrare che $K(x)$ è convessa e soddisfa le proprietà della norma. ■

4.9. Funzioni di supporto e funzioni coniugate

4.103 DEFINIZIONE. Si definisce funzione di supporto di un insieme K la funzione

$$\varphi(w; K) = \inf_{x \in K} w x$$

(scrivendo talvolta semplicemente $\varphi(w)$ se è chiaro quale sia K). ■

Dai risultati precedenti risulta che $\varphi(w; K)$ è una funzione concava per qualsiasi insieme K (non necessariamente convesso!). Ricordando la definizione di piano di supporto si ha che un piano di supporto è rappresentato dall'equazione $w x = \varphi(w; K)$ e un semispazio di supporto dalla disequazione $w x \geq \varphi(w; K)$.

Se K è chiuso e convesso c'è una corrispondenza biunivoca fra insiemi e funzioni di supporto, per cui si ha il seguente teorema, la cui facile dimostrazione viene lasciata come esercizio.

4.104 TEOREMA. Siano K e H due insiemi chiusi e convessi. Se $\varphi(w; K)$ è identicamente uguale a $\varphi(w; H)$ allora $K = H$. ■

Abbiamo visto che una funzione definita come $f(x) = \sup_{(w, \varphi) \in A} \{w x - \varphi\}$ è una funzione convessa, il cui epigrafo è chiuso (in quanto intersezione di insiemi chiusi), e propria, nel senso che $f(x) > -\infty, \forall x$. Possiamo porci la domanda se valga anche l'inverso, se cioè ogni funzione convessa, chiusa e propria sia rappresentabile come supremo puntuale di un'opportuna famiglia di funzioni affini.

La notevole risposta è che ciò è effettivamente possibile. Infatti epi f in quanto convesso e chiuso è rappresentabile come intersezione di semispazi di supporto. In \mathbb{R}^{n+1} un semispazio di supporto è dato da:

$$w_0 x_0 + w x \geq \varphi(w_0, w; \text{epi } f)$$

Se $w_0 = 0$ il semispazio viene detto verticale, mentre se $w_0 > 0$ il semispazio è l'epigrafo della funzione affine

$$h(x) = \frac{\varphi(w_0, w; \text{epi } f) - w x}{w_0} = \varphi\left(1, \frac{w}{w_0}; \text{epi } f\right) - \frac{w}{w_0} x = \varphi(1, u; \text{epi } f) - u x$$

dove si è posto $u = w/w_0$.

Siccome φ è funzione di supporto di un epigrafo vale sempre $w_0 \geq 0$. Quindi se riusciamo a dimostrare che epi f è dato dall'intersezione di semispazi non verticali è anche automaticamente dimostrato l'asserto iniziale.

Innanzitutto non tutti i semispazi di supporto sono verticali, perché in tal caso si avrebbe $f(x) = -\infty$ per qualche x . Sia quindi $x_0 + \hat{u} x \geq \varphi(1, \hat{u})$ un semispazio di supporto non verticale e sia $\bar{u} x \geq \varphi(0, \bar{u})$ un semispazio di supporto H^+ verticale. È sufficiente dimostrare che se $(x_0, x) \notin H^+$ allora esiste un semispazio di supporto non verticale H_1^+ tale che $(x_0, x) \notin H_1^+$. Sia quindi (\bar{x}_0, \bar{x}) tale che $\bar{u} \bar{x} < \varphi(0, \bar{u})$. Per un valore di α sufficientemente grande si ha

$$\bar{x}_0 + (\hat{u} + \alpha \bar{u}) \bar{x} < \varphi(1, \hat{u}) + \alpha \varphi(0, \bar{u})$$

Essendo φ concava e positivamente omogenea si ha

$$\varphi(1, \hat{u}) + \alpha \varphi(0, \bar{u}) \leq \varphi(1, \hat{u} + \alpha \bar{u})$$

da cui risulta che il piano di supporto

$$x_0 + (\hat{u} + \alpha \bar{u}) x = \varphi(1, \hat{u} + \alpha \bar{u})$$

determina il semispazio H_1^+ . Abbiamo così dimostrato il seguente:

4.105 TEOREMA. *Se f è chiusa, convessa e propria, allora*

$$f(x) = \sup_{u: (1,u) \in \text{dom } \varphi} \{ \varphi(1, u; \text{epi } f) - ux \} .$$

■

La precedente discussione ha anche messo in luce il seguente fatto, che risulta particolarmente utile quando $\text{epi } f$ sia un poliedro, in quanto permette di rappresentare la funzione convessa f come supremo puntuale di un numero comunque *finito* di funzioni affini, purché si limiti la rappresentazione a $\text{dom } f$:

4.106 COROLLARIO. *Sia $\text{epi } f = \bigcap_{\alpha \in A} H_{\alpha}^{+}$ dove H_{α}^{+} sono semispazi di supporto e A è finito. Allora $f(x) = \max_{\alpha \in \hat{A}} h_{\alpha}(x)$ se $x \in \text{dom } f$ per un opportuno insieme finito di funzioni affini h_{α} .*

DIMOSTRAZIONE. Siano H_{α}^{+} , $\alpha \in \hat{A} \subset A$, tutti i semispazi di supporto non verticali e sia $D := \mathbb{R} \times \text{dom } f \subset \mathbb{R}^{n+1}$. Allora, poiché $D \subset H_{\alpha}^{+}$ per ogni piano verticale H_{α}^{+} ,

$$\text{epi } f = \text{epi } f \cap D = \bigcap_{\alpha \in A} H_{\alpha}^{+} \cap D = \bigcap_{\alpha \in \hat{A}} H_{\alpha}^{+} \cap D.$$

Quindi se $h_{\alpha}(x)$ è la funzione affine tale che $\text{epi } h_{\alpha} = H_{\alpha}^{+}$ con $\alpha \in \hat{A}$, abbiamo la tesi. ■

Esaminiamo ora alcune delle notevoli conseguenze del teorema 4.105. Tenendo presente che

$$-\varphi(1, -u; \text{epi } f) = -\inf_x \{ f(x) - ux \} = \sup_x \{ ux - f(x) \} =: f^*(u)$$

dove la funzione f^* , definita come sopra, prende il nome di *funzione coniugata* di f , abbiamo

$$f(x) = \sup_u \{ \varphi(1, u; \text{epi } f) - ux \} = \sup_u \{ \varphi(1, -u; \text{epi } f) + ux \} = \sup_u \{ ux - f^*(u) \}$$

Quindi anche f è coniugata di f^* e fra f e f^* c'è una perfetta corrispondenza di dualità.

La relazione di coniugio lega in modo molto stretto, ad esempio, la funzione indicatrice $\delta(x, K)$ con la funzione di supporto $\varphi(u, K)$ di un insieme convesso K . Infatti è immediato stabilire che

$$\delta^*(u, K) = -\varphi(-u, K) \quad \text{e} \quad -\varphi^*(x, K) = \delta(-x, K)$$

Come conseguenza di questo fatto possiamo provare il seguente importante teorema:

4.107 TEOREMA. *Siano $f, g: \mathbb{R}^n \rightarrow \bar{\mathbb{R}}$ funzioni convesse tali che $\text{dom } f \cap \text{dom } g \neq \emptyset$. Allora*

$$\mu := \inf_x f(x) + g(x) = \sup_u -f^*(u) - g^*(-u)$$

DIMOSTRAZIONE. Per definizione si ha

$$f^*(u) \geq ux - f(x) \quad \forall u, x \quad \text{e} \quad g^*(-u) \geq -ux - g(x) \quad \forall u, x$$

da cui si ottiene $\mu \geq \sup_u -f^*(u) - g^*(-u)$.

Se $\mu = -\infty$ la tesi è provata dalla precedente disequaglianza. Si noti che $\mu \neq +\infty$ per ipotesi. Si supponga quindi μ finito.

Si indichi con $\text{epi}^- f$ l'insieme $\{(x_0, x) : (-x_0, x) \in \text{epi} f\}$. Si considerino ora i due insiemi $(\mu, 0) + \text{epi}^- f$ e $\text{epi} g$, che, essendo convessi e senza punti interni comuni, sono separabili. Si noti che il piano separatore non può essere verticale perché in tal caso separerebbe $\text{dom} f$ da $\text{dom} g$, cosa impossibile a causa dell'ipotesi fatta. Quindi la separazione si può esprimere come:

$$\exists u: \quad \mu - f(x) + ux \leq g(y) + uy \quad \forall x, y$$

Prendendo il sup a sinistra e l'inf a destra si ottiene $\mu + f^*(u) \leq -g^*(-u)$ che, combinato con la precedente disuguaglianza, fornisce la tesi. ■

Come immediata applicazione del teorema si può dimostrare un celebre risultato che verrà sfruttato nel seguito:

4.108 TEOREMA. (Von Neumann [1928]) *Siano A e B insiemi convessi in \mathbb{R}^n , con B limitato. Allora*

$$\inf_{x \in A} \sup_{u \in B} ux = \sup_{u \in B} \inf_{x \in A} ux$$

DIMOSTRAZIONE.

$$\inf_{x \in A} \sup_{u \in B} ux = \inf_{x \in A} -\varphi(-x, B) = \inf_x -\varphi(-x, B) + \delta(x, A) = \inf_x \delta^*(x, B) - \varphi^*(-x, A)$$

che, in base al precedente teorema, risulta essere uguale a:

$$\sup_u -\delta(u, B) + \varphi(u, A) = \sup_{u \in B} \inf_{x \in A} ux$$

■

4.10. Proprietà differenziali

Se f è differenziabile (una o più volte) in x indichiamo con $Df(x)$ il *gradiente* di f in x , cioè il vettore $\{\partial f/\partial x_1, \dots, \partial f/\partial x_n\}$, e con $D^2f(x)$ l'*Hessiano* di f in x , cioè la matrice di componenti $\partial^2 f/\partial x_i \partial x_j$. Ricordiamo che f è differenziabile in x se e solo se esiste un vettore y per cui

$$\lim_{\|h\| \rightarrow 0} \frac{|f(x+h) - f(x) - yh|}{\|h\|} = 0 \quad \forall h$$

Tale vettore è unico ed è il gradiente. Quindi, se f è derivabile, si ha:

$$\lim_{\alpha \downarrow 0} \frac{f(x + \alpha h) - f(x) - y(\alpha h)}{\|\alpha h\|} = \lim_{\alpha \downarrow 0} \frac{f(x + \alpha h) - f(x)}{\alpha \|h\|} - \frac{yh}{\|h\|} = 0$$

ovvero

$$\lim_{\alpha \downarrow 0} \frac{f(x + \alpha h) - f(x)}{\alpha} = Df(x)h$$

Funzioni convesse differenziabili si possono utilmente caratterizzare tramite i seguenti risultati:

4.109 TEOREMA. Sia f differenziabile due volte nell'interno di $\text{dom } f$. Allora f è convessa se e solo se il suo Hessiano è semidefinito positivo per ogni $x \in \text{dom } f$.

DIMOSTRAZIONE. Sia l'Hessiano semidefinito positivo; sia $x = \alpha y + (1 - \alpha)z$ con $0 \leq \alpha \leq 1$. Allora:

$$f(y) = f(x) + Df(x)(y - x) + \frac{1}{2} D^2 f(\xi)(y - x, y - x)$$

$$f(z) = f(x) + Df(x)(z - x) + \frac{1}{2} D^2 f(\zeta)(z - x, z - x)$$

dove ξ e ζ sono opportuni punti negli intervalli aperti $\text{conv}(y, x)$ e $\text{conv}(z, x)$ rispettivamente. Per l'ipotesi sull'Hessiano

$$f(y) \geq f(x) + Df(x)(y - x)$$

$$f(z) \geq f(x) + Df(x)(z - x)$$

Moltiplicando la prima disequazione per α , la seconda per $(1 - \alpha)$ e sommando si ottiene

$$\alpha f(y) + (1 - \alpha) f(z) \geq f(x) + Df(x)(\alpha y + (1 - \alpha)z - x) = f(x) = f(\alpha y + (1 - \alpha)z)$$

Viceversa sia f convessa. Allora dalla definizione si ottiene

$$f(x + \alpha h) - f(x) \leq \alpha (f(x + h) - f(x)) \quad \forall 0 \leq \alpha \leq 1 \quad \forall h$$

cioè

$$\frac{f(x + \alpha h) - f(x)}{\alpha} \leq f(x + h) - f(x)$$

e passando al limite per $\alpha \downarrow 0$ si ha $Df(x)h \leq f(x + h) - f(x)$. Siccome

$$f(x + h) - f(x) = Df(x)h + \frac{1}{2} D^2 f(x)(h, h) + o(\|h\|^3)$$

si ha $D^2 f(x)(h, h) + o(\|h\|^3) \geq 0$, che, essendo $o(\|h\|^3)$ infinitesimo di ordine superiore a $D^2 f(x)(h, h)$, implica $D^2 f(x)(h, h) \geq 0, \forall h$. ■

4.110 ESERCIZIO. Sia $f(x, y) := x^3 + ax^2y + ax^2y^2 + y^3$ definita sull'ortante positivo. Per quali valori di a la funzione è convessa (sull'ortante positivo)? ■

Un'altra caratterizzazione delle funzioni convesse, legata alle derivate prime, è data da:

4.111 TEOREMA. Sia f differenziabile in un aperto convesso. Allora f è convessa se e solo se

$$Df(x)h \leq f(x + h) - f(x) \quad \forall h$$

DIMOSTRAZIONE. La necessità è stata dimostrata nel teorema 4.109. Per la sufficienza si può procedere come nel precedente teorema con una espansione in serie fino al primo ordine anziché al secondo. ■

Se f non è differenziabile si può comunque introdurre, specialmente in presenza di convessità, un concetto più debole della derivata ma che permette comunque di ottenere risultati importanti.

4.112 DEFINIZIONE. La derivata direzionale $Df_x(h)$ di una funzione f nel punto x lungo la direzione h viene definita dal seguente limite (se esiste):

$$Df_x(h) = \lim_{\alpha \downarrow 0} \frac{f(x + \alpha h) - f(x)}{\alpha}$$

Il limite dipende ovviamente da x e da h e pertanto la derivata direzionale viene notata indicando in modo esplicito la dipendenza da x e h . In generale il limite può non esistere. Nel caso di funzioni convesse vale il seguente:

4.113 TEOREMA. Se f è convessa la derivata direzionale esiste (eventualmente uguale a $\pm\infty$) per ogni h e per ogni x . Inoltre risulta:

$$Df_x(h) = \inf_{\alpha > 0} \frac{f(x + \alpha h) - f(x)}{\alpha}$$

DIMOSTRAZIONE. Dimostriamo innanzitutto che il rapporto incrementale $(f(x + \alpha h) - f(x))/\alpha$, è, in quanto funzione di α soltanto, funzione monotona non decrescente (se f è convessa). Sia $0 \leq \alpha \leq \beta$. Allora

$$x + \alpha h = \frac{\beta - \alpha}{\beta} x + \frac{\alpha}{\beta} (x + \beta h)$$

cioè una combinazione convessa di x e $x + \beta h$. Applicando la definizione di funzione convessa si ottiene con semplici passaggi $(f(x + \alpha h) - f(x))/\alpha \leq (f(x + \beta h) - f(x))/\beta$. In questo caso

$$\lim_{\alpha \downarrow 0} \frac{f(x + \alpha h) - f(x)}{\alpha} = \inf_{\alpha > 0} \frac{f(x + \alpha h) - f(x)}{\alpha}$$

e quindi il limite esiste. ■

Si noti che dal teorema precedente segue che

$$Df_x(h) \leq \frac{f(x + \alpha h) - f(x)}{\alpha} \quad \forall h, \forall \alpha > 0 \quad (4.2)$$

La derivata direzionale $Df_x(h)$, in quanto funzione di h , gode delle seguenti proprietà:

- 1) è positivamente omogenea, cioè $Df_x(\alpha h) = \alpha Df_x(h)$, $\forall \alpha \geq 0$;
- 2) $Df_x(0) = 0$ (discende da 1), ma è utile dirlo in modo esplicito);
- 3) è convessa, infatti

$$\begin{aligned} Df_x(\beta h_1 + (1 - \beta) h_2) &= \inf_{\alpha > 0} \frac{f(\beta x + (1 - \beta) x + \alpha \beta h_1 + \alpha (1 - \beta) h_2) - f(x)}{\alpha} \leq \\ &= \inf_{\alpha > 0} \frac{\beta f(x + \alpha h_1) + (1 - \beta) f(x + \alpha h_2) - \beta f(x) - (1 - \beta) f(x)}{\alpha} = \\ &= \beta \inf_{\alpha > 0} \frac{f(x + \alpha h_1) - f(x)}{\alpha} + (1 - \beta) \inf_{\alpha > 0} \frac{f(x + \alpha h_2) - f(x)}{\alpha} = \beta Df_x(h_1) + (1 - \beta) Df_x(h_2); \end{aligned}$$

- 4) $\text{dom } Df_x = \text{cone}(\text{dom } f - x)$ (lo si dimostri per esercizio).

Inoltre Df_x è una funzione lineare (in h) se e soltanto se f è derivabile in x . Come è immediato constatare da (4.2), ponendo $\alpha = 1$, si ha:

$$Df_x(h) \leq f(x+h) - f(x) \quad \forall h$$

e quindi se f è derivabile in x si ritrova il teorema 4.111:

$$Df(x)h \leq f(x+h) - f(x) \quad \forall h$$

Sorge naturale la domanda se il gradiente sia l'unico vettore γ tale che $\gamma h \leq f(x+h) - f(x)$ per ogni h . Sia allora $\gamma(\alpha h) \leq f(x + \alpha h) - f(x), \forall \alpha, h$, ovvero

$$\gamma h \leq \frac{f(x + \alpha h) - f(x)}{\alpha} \quad \forall \alpha, h \quad \implies \quad \gamma h \leq Df_x(h) = Df(x)h \quad \forall h \quad (4.3)$$

e solo per $\gamma = Df(x)$ la disuguaglianza è vera per ogni h .

Quest'ultima proprietà, che definisce in modo univoco il gradiente quando f è derivabile, può essere sfruttata per definire un'estensione del concetto di gradiente, quando f non è derivabile.

4.114 DEFINIZIONE. Ogni $\gamma \in \mathbb{R}^n$ tale che :

$$\gamma h \leq f(x+h) - f(x) \quad \forall h \quad (4.4)$$

prende il nome di subgradiente di f in x . L'insieme dei subgradienti di f in x prende il nome di subdifferenziale di f in x e si indica con ∂f_x . ■

Le proprietà di ∂f_x si possono così riassumere:

4.115 TEOREMA.

(a) ∂f_x è chiuso e convesso

(b) $\gamma \in \partial f_x$ se e soltanto se $\gamma h \leq Df_x(h) \quad \forall h$

(c) sia f chiusa; $\gamma \in \partial f_x$ se e soltanto se il piano di equazione $\{(y_0, y): y_0 - \gamma y = f(x) - \gamma x\}$ è di supporto a epi f nel punto $(f(x), x)$.

DIMOSTRAZIONE. (a) ∂f_x è l'intersezione di infiniti semispazi chiusi e quindi è chiuso; (b) la sufficienza discende da (4.2) per $\alpha = 1$ e la necessità da (4.3); (c) per definizione di subgradiente si ha $f(y) - f(x) \geq \gamma(y-x)$ per ogni y , cioè $f(y) - \gamma y \geq f(x) - \gamma x, \forall y$, e quindi, per definizione di epigrafo $y_0 - \gamma y \geq f(x) - \gamma x, \forall (y_0, y) \in \text{epi } f$ da cui la tesi. Viceversa, se f è chiusa, per ogni $x \in \text{dom } f$ esiste un piano di supporto non verticale di coefficienti $(1, -\gamma)$ che separa epi f da $(f(x), x)$, cioè $y_0 - \gamma y \geq f(x) - \gamma x$ per ogni $(y_0, y) \in \text{epi } f$. Prendendo in particolare $y_0 = f(y)$ si vede che γ è un subgradiente. ■

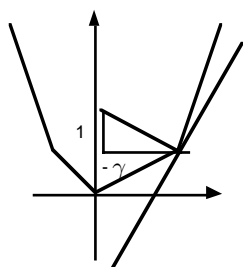


FIGURA 4.20

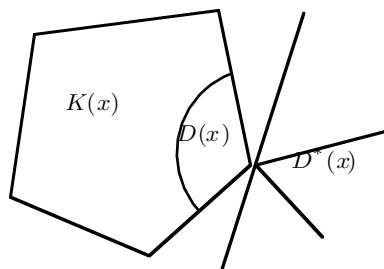


FIGURA 4.21

La proprietà (c) caratterizza geometricamente i subgradienti. Si consideri la figura 4.20 dove è raffigurata una funzione $f : \mathbb{R} \rightarrow \mathbb{R}$. Un piano di supporto all'epigrafo è in questo caso una retta di supporto al grafico della funzione. Si prenda la normale alla retta che abbia componente verticale uguale a 1 e la si proietti sul piano orizzontale. Il subgradiente corrispondente è il vettore opposto a questa proiezione. Si vede facilmente che è un valore compreso fra la derivata sinistra e la derivata destra nel punto considerato. Un'altra interpretazione geometrica la si ottiene pensando agli insiemi definiti dalle linee di livello di una funzione convessa, ovvero $K(x) := \{y \in \mathbb{R}^n : f(y) \leq f(x)\}$. Come conseguenza immediata del teorema (basta considerare la proprietà (c) ristretta all'insieme affine di $\{(y_0, y) \in \mathbb{R}^{n+1} : y_0 = f(x)\}$), dato un punto $x \in \mathbb{R}^n$, ogni subgradiente in ∂f_x è uguale ad un elemento di $D^*(x)$, cono polare del cono ammissibile $D(x)$. Quindi $\partial f_x \subset D^*(x)$ e anche ovviamente $\text{cone}(\partial f_x) \subset D^*(x)$. La seconda inclusione è stretta in generale, come mostra il seguente esempio $f(x) := (x-1)^2$ se $x \geq 1$, altrimenti $f(x) := 0$. Per $x = 1$ si ha $\partial f_x = \{0\}$ (con $\text{cone}(\partial f_x) = \{0\}$ quindi) e $D^*(x) = \mathbb{R}_+$. Tuttavia si può dimostrare che se \bar{x} non è minimo allora $\text{cone}(\partial f_x) = D^*(x)$, cioè ad ogni direzione in $D^*(x)$ corrisponde un subgradiente a meno di uno scalare positivo (l'ampiezza del subgradiente dipende dalla velocità con cui la funzione cresce, informazione che non è desumibile dall'insieme $K(x)$). Si veda la figura 4.21 dove si raffigura il caso di una funzione $f : \mathbb{R}^2 \rightarrow \mathbb{R}$.

È interessante applicare questi ragionamenti al subdifferenziale di una funzione indicatrice. Non è difficile dimostrare che per ogni insieme convesso K

$$\partial \delta_x(K) = D^*(K, x) \quad (4.5)$$

Facciamo ora vedere in che modo i subgradienti caratterizzano la derivata direzionale.

4.116 TEOREMA. *Sia f chiusa o alternativamente $x \in \overset{\circ}{\text{dom}} f$. Allora*

$$Df_x(h) = \sup_{\gamma \in \partial f_x} \gamma h$$

DIMOSTRAZIONE. Da $Df_x(h) - \gamma h \geq 0, \forall h$ discende che il piano $y_0 - \gamma y = 0$ è di supporto a $\text{epi } Df_x$ nell'origine. Siccome Df_x è positivamente omogenea, il suo epigrafo è un cono convesso. Allora tutti i piani di supporto di $\text{epi } Df_x$ passano per l'origine, cioè sono tutti del tipo $y_0 - \gamma y = 0$ con $\gamma \in \partial f_x$. In base alla rappresentabilità di una funzione convessa e chiusa come supremo degli epigrafi delle funzioni affini di supporto a $\text{epi } Df_x$ si ha la tesi. ■

Per quanto già osservato, se f è derivabile in x allora il subgradiente in x è unico e coincide con il gradiente. Il teorema precedente garantisce che l'unicità del subgradiente è anche condizione sufficiente alla differenziabilità.

4.117 COROLLARIO. *Se il subgradiente è unico in un punto $x \in \overset{\circ}{\text{dom}} f$, f è derivabile e il suo gradiente coincide con il subgradiente.* ■

La derivazione è un operatore lineare, per cui la derivata di una somma è uguale alla somma delle derivate. Facciamo ora vedere che questa proprietà si può estendere ai subdifferenziali:

4.118 TEOREMA. *Siano f e g funzioni convesse e sia $h := f + g$. Allora $\partial h_x = \partial f_x + \partial g_x$ per ogni x .*

DIMOSTRAZIONE. L'inclusione $\partial h_x \supset \partial f_x + \partial g_x$ è facilmente dimostrabile. Sia allora $\bar{\gamma} \in \partial h_x$ e $\bar{\gamma} \notin \partial f_x + \partial g_x$. Quindi $\bar{\gamma}$ è strettamente separabile da $\partial f_x + \partial g_x$ ed esiste \bar{h} tale che $\bar{\gamma} \bar{h} > \alpha \geq \gamma \bar{h}$ per ogni $\gamma \in \partial f_x + \partial g_x$, ovvero

$$\bar{\gamma} \bar{h} > \alpha \geq \gamma_f \bar{h} + \gamma_g \bar{h} \quad \forall \gamma_f \in \partial f_x, \forall \gamma_g \in \partial g_x$$

Possiamo pertanto scrivere

$$\bar{\gamma} \bar{h} > \alpha \geq \sup_{\substack{\gamma_f \in \partial f_x \\ \gamma_g \in \partial g_x}} (\gamma_f \bar{h} + \gamma_g \bar{h}) = \sup_{\gamma_f \in \partial f_x} \gamma_f \bar{h} + \sup_{\gamma_g \in \partial g_x} \gamma_g \bar{h} = Df(x, \bar{h}) + Dg(x, \bar{h}) = Dh(x, \bar{h})$$

per note proprietà dei limiti, e siccome $\bar{\gamma} \in \partial h_x$, si ha

$$\bar{\gamma} \bar{h} \leq \sup_{\gamma \in \partial h_x} \gamma \bar{h} = Dh(x, \bar{h})$$

contraddicendo la precedente disuguaglianza stretta. \blacksquare

Per funzioni definite da $f(x) = \sup_{i \in I} g^i(x)$ dove g^i sono funzioni convesse, la caratterizzazione di ∂f_x viene fatta tramite i ∂g_x^i . Sia infatti $f(x) = \sup_i g^i(x) = g^k(x)$ con $k \in I(x)$ dove appunto $I(x)$ è l'insieme di indici per i quali il supremo viene raggiunto. Allora si ha:

$$f(x+h) - f(x) = \sup_i g^i(x+h) - g^k(x) \geq g^k(x+h) - g^k(x) \geq \gamma^k h \quad \forall \gamma^k \in \partial g_x^k \quad k \in I(x)$$

e questo implica che $\text{conv}\{\partial g_x^k : k \in I(x)\} \subset \partial f_x$. L'inclusione può essere stretta in generale, come si vede dal seguente esempio dove $I = \{0\} \cup [1, +\infty)$:

$$g^i := \begin{cases} -x & \text{se } i = 0 \\ \frac{x}{i} - \frac{1}{2i^2} & \text{se } i \geq 1 \end{cases} \quad \text{e} \quad f(x) = \begin{cases} -x & \text{se } x \leq 0 \\ \frac{x^2}{2} & \text{se } 0 < x \leq 1 \\ x - \frac{1}{2} & \text{se } x > 1 \end{cases}$$

da cui $\partial f_0 = [-1, 0]$ mentre $I(0) = \{0\}$, quindi $\partial g_{0,0} = \{-1\}$. Si noti che si ottiene lo stesso risultato negativo anche se restringiamo l'insieme I agli interi non negativi. Tuttavia se I è un insieme finito vale l'eguaglianza come affermato nel seguente teorema:

4.119 TEOREMA. *Sia $f(x) := \max_{i \in I} g^i(x)$ con $\{g^i\}_{i \in I}$ un insieme finito di funzioni convesse, sia $I(x) := \{i \in I : f(x) = g^i(x)\}$ e sia $x \in \text{dom} f$. Allora*

$$\partial f_x = \text{conv}\{\partial g_x^i : i \in I(x)\}$$

DIMOSTRAZIONE. Per la finitezza di I e la continuità delle funzioni g^i esiste un intorno N di x per cui $I(x) \supset I(y)$ per ogni $y \in N$. Possiamo quindi supporre senza perdita di generalità che $I(x) = I$. Si indichi $G := \text{conv}\{\partial g_x^i : i \in I\}$. Se esistesse $\bar{\gamma} \in \partial f_x$ con $\bar{\gamma} \notin G$, esisterebbe un piano $\gamma \bar{h} = \bar{\gamma} \bar{h}$ che separa strettamente $\bar{\gamma}$ da G , cioè $\gamma \bar{h} \leq K < \bar{\gamma} \bar{h}$ per ogni $\gamma \in G$ e in particolare $\gamma \bar{h} \leq K < \bar{\gamma} \bar{h}$ per ogni $\gamma \in \partial g_x^i$ per ogni i . Quindi $Dg_x^i(\bar{h}) = \sup\{\gamma \bar{h} : \gamma \in \partial g_x^i\} \leq K < \bar{\gamma} \bar{h}$. Siccome $\bar{\gamma} \in \partial f_x$ si ha $Dg_x^i(\bar{h}) \leq K < Df_x(\bar{h})$ per ogni i .

Si consideri ora $f(x + \alpha \bar{h})$. Per ogni valore di α esiste un indice $k(\alpha)$ per cui $f(x + \alpha \bar{h}) = g^{k(\alpha)}(x + \alpha \bar{h})$. Per la finitezza di I esiste un indice k ed una successione $\alpha_1, \alpha_2, \dots$ tendente a 0, tale che $k = k(\alpha_j)$ per ogni α_j della successione. Quindi

$$Df_x(\bar{h}) = \lim_{\alpha \downarrow 0} \frac{f(x + \alpha \bar{h}) - f(x)}{\alpha} = \lim_{j \rightarrow \infty} \frac{f(x + \alpha_j \bar{h}) - f(x)}{\alpha_j} = \lim_{j \rightarrow \infty} \frac{g^k(x + \alpha_j \bar{h}) - g^k(x)}{\alpha_j} = Dg_x^k(\bar{h})$$

contraddicendo il precedente risultato. \blacksquare

4.120 ESEMPIO. Sia data

$$f(x) := \max \{-4x_1 + 8x_2; 8x_1 - 4x_2; -4x_1 - 4x_2; -x_1 - x_2 + 6\}$$

Siano g^1, g^2, g^3 e g^4 le funzioni affini che definiscono f . La funzione f è ovviamente lineare a tratti e il suo epigrafo è un poliedro. Può essere utile disegnare la proiezione sul dominio degli spigoli e dei vertici dell'epigrafo (figura 4.22). Le varie regioni rappresentano quella parte del dominio dove f è lineare. In queste regioni solo una delle funzioni g^i dà luogo al massimo e quindi si ha $f(x) = g^i(x)$ (in figura è indicato l'indice della g^i corrispondente). Sugli spigoli si ha lo stesso valore di massimo per due funzioni, ed esattamente per le due funzioni delle regioni adiacenti. Analogamente quando tre spigoli si incontrano in un vertice (vertici indicati con le lettere A, B e C), sono tre le funzioni che danno contemporaneamente luogo al massimo e cioè quelle tre le cui regioni circondano il vertice.

Nella figura 4.23 viene evidenziata la struttura dei subdifferenziali di f . Nelle regioni di linearità f è differenziabile e quindi il subdifferenziale si riduce ad un punto, negli spigoli diventa un segmento e nei vertici diventa un poligono con tanti vertici quante sono le funzioni che danno luogo al massimo. Si noti la relazione fra le due figure. Regioni di sinistra sono punti a destra e viceversa. Segmenti a sinistra corrispondono a segmenti ortogonali a destra (perché?). ■

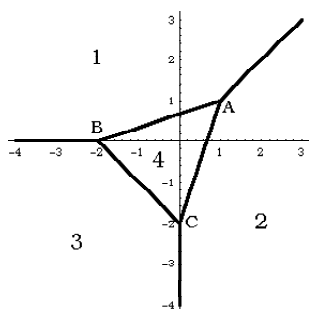


FIGURA 4.22

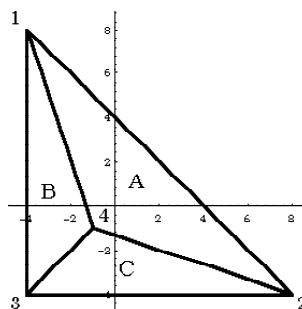


FIGURA 4.23

Spesso una funzione convessa viene definita imponendo il suo dominio effettivo con condizioni di vincolo. Sia quindi K insieme convesso e chiuso e $f : \mathbb{R}^n \rightarrow \mathbb{R}$ convessa con $\text{dom } f = \mathbb{R}^n$. Definiamo

$$\tilde{f}(x) = f(x) + \delta(x, K)$$

In base al teorema 4.118 si ha $\partial \tilde{f}_x = \partial f_x + \partial \delta_x(K)$, ed applicando (4.5) si ha

4.121 TEOREMA.

$$\partial \tilde{f}(x) = \partial f(x) + D^*(K, x)$$

4.11. Minimi di funzioni convesse

Le proprietà della derivata direzionale permettono di caratterizzare i minimi di una funzione convessa.

4.122 TEOREMA. Sia f chiusa o alternativamente $x \in \overset{\circ}{\text{dom}} f$. Allora $Df_x(h) \geq 0, \forall h$ se e solo se x è un punto di minimo.

DIMOSTRAZIONE. La sufficienza è ovvia. Per la necessità basta notare che

$$0 \leq Df_x(h) \leq \frac{f(x + \alpha h) - f(x)}{\alpha} \quad \forall \alpha > 0$$

e quindi anche per $\alpha = 1$, cioè $f(x + h) - f(x) \geq 0$. ■

4.123 TEOREMA. Sia f chiusa o alternativamente $x \in \overset{\circ}{\text{dom}}f$. Allora x è punto di minimo se e solo se $\mathbf{0} \in \partial f_x$.

DIMOSTRAZIONE. La sufficienza è immediata dalla definizione di subgradiente. Per la necessità si ha $f(x + h) - f(x) \geq 0, \forall h$, e, scrivendo $0 = \mathbf{0} \cdot h, \forall h$, si ha $\mathbf{0} \in \partial f_x$. ■

4.124 ESERCIZIO. Qual è il minimo della funzione dell'esempio 4.120? ■

4.125 ESERCIZIO. Siano dati n punti su una retta di coordinate a_i . Trovare il punto x che minimizza $\sum_{i=1}^n |x - a_i|$ usando la condizione del teorema 4.123. Generalizzando siano dati n punti nel piano di coordinate $a^i = (a_1^i, a_2^i)$. Trovare il punto x (o i punti x) che minimizza $\sum_{i=1}^n \|x - a^i\|_1$. ■

4.126 ESERCIZIO. Dimostrare che $\overline{\text{cone}(\partial f_x)} = D^*(x)$ se x non è di minimo per la funzione $f(x)$. ■

Se il minimo è vincolato ad appartenere ad un insieme convesso e chiuso K il teorema 4.123 si modifica in:

4.127 TEOREMA. Sia K convesso e chiuso e sia $D(x)$ il cono delle direzioni ammissibili in x . Un punto x è minimo su K di una funzione convessa f se e solo se $\partial f(x) \cap -D^*(x) \neq \emptyset$.

DIMOSTRAZIONE. Dal teorema 4.121

$$\mathbf{0} \in \partial \tilde{f}(x) \iff \mathbf{0} \in \partial f(x) + D^*(x) \iff \partial f(x) \cap -D^*(x) \neq \emptyset$$

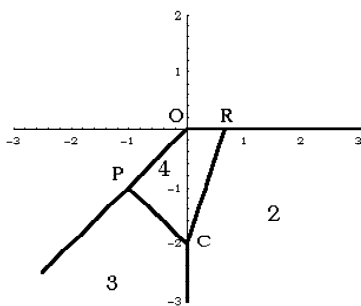


FIGURA 4.24

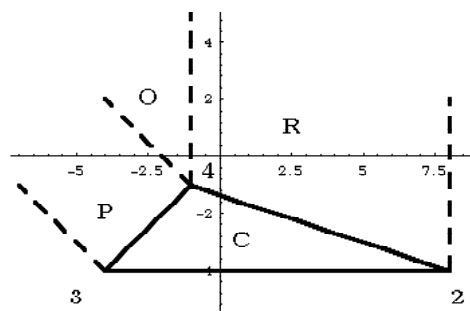


FIGURA 4.25

4.128 ESEMPIO. Si riconsideri l'esempio 4.120 e sia $K := \{x : x_2 \leq 0, x_2 \leq x_1\}$. Si vedano disegnati in figura 4.24 l'insieme ammissibile e i punti di rottura di f limitatamente a K . La frontiera di K può essere ripartita in tre insiemi:

$$\partial K_1 = \{x : x_2 = 0, x_1 > 0\}, \quad \partial K_2 = \{x : x_2 < 0, x_1 = x_2\}, \quad \partial K_3 = \{x : x_2 = 0, x_1 = 0\}$$

e si ha rispettivamente

$$D(x) = \{x : x_2 \leq 0\} =: D_1 \quad \text{se } x \in \partial K_1, \quad D(x) = \{x : x_1 - x_2 > 0\} =: D_2 \quad \text{se } x \in \partial K_2$$

$$D(x) = D_1 \cap D_2 =: D_3 \quad \text{se } x \in \partial K_3$$

i cui coni polari sono

$$D_1^* = \text{cone}(0, 1) \quad D_2^* = \text{cone}(-1, 1) \quad D_3^* = \text{cone}(D_1^*, D_2^*)$$

Dai risultati dell'esempio 4.120 nessun punto interno a K può essere minimo. Consideriamo pertanto i punti della frontiera. Con riferimento alla figura 4.24 nei punti di ∂K_1 alla destra del punto R, f è derivabile con gradiente il vettore $(8, -4)$. Tuttavia essendo punti di frontiera bisogna sommare al gradiente il cono D_1^* (linea tratteggiata in figura 4.25). Il nuovo insieme non contiene l'origine e quindi tali punti non sono minimi. Il punto R ha come subdifferenziale il segmento $2 - 4$ (figura 4.25). Sommando a tale segmento il cono D_1^* si ottiene l'insieme indicato come R in figura 4.25. Come si vede l'origine è inclusa nell'insieme e quindi R è minimo. I punti di ∂K_1 alla sinistra di R hanno come gradiente il punto indicato 2 in figura 4.25 e, sommando D_1^* si ottiene un insieme che non contiene l'origine. Si consideri ora il punto O nel quale f è derivabile con gradiente $(-1, -1)$. Sommando il cono D_3^* si ottiene l'insieme indicato O, che non contiene l'origine. Continuando nello stesso modo si vede che l'unico minimo è il punto R. ■

4.129 ESERCIZIO. Si dimostri in modo alternativo il teorema 4.76 (lemma di Farkas) sfruttando il precedente risultato. ■

Quando un punto x non è minimo è molto spesso necessario calcolare una direzione h che sia di discesa per f , cioè esiste $\alpha > 0$ tale che $f(x + \alpha h) < f(x)$. In questo modo tramite una serie di spostamenti si perviene ad un minimo. Ovviamente è vantaggioso cercare una direzione che risulti di massima discesa, cosa che può esser fatta minimizzando la derivata direzionale rispetto ad h , previa normalizzazione di h , cioè

$$w := \min_{\|h\|_p \leq 1} Df_x(h) \quad (4.6)$$

Se la funzione è differenziabile allora una direzione di massima discesa è immediatamente disponibile ed è data dall'antigradiente, cioè $h = -Df(x)$, valore che si ottiene da (4.6) usando la norma euclidea.

Nel caso generale bisogna invece usare l'espressione della derivata direzionale in funzione del subdifferenziale e applicare il teorema 4.108. Si ha

$$w = \min_{\|h\|_p \leq 1} Df_x(h) = \min_{\|h\|_p \leq 1} \sup_{\gamma \in \partial f_x} \gamma h = \sup_{\gamma \in \partial f_x} \min_{\|h\|_p \leq 1} \gamma h = \sup_{\gamma \in \partial f_x} -\|\gamma\|_q = - \inf_{\gamma \in \partial f_x} \|\gamma\|_q \quad (4.7)$$

dove, al solito, p e q sono due reali maggiori o uguali ad 1 (e incluso $+\infty$), tali che $1/p + 1/q = 1$.

È interessante dare una caratterizzazione geometrica della direzione di discesa che si calcola tramite (4.7). Siano quindi $\bar{\gamma}$ e \bar{h} le soluzioni di (4.7). Quindi

$$w = \bar{\gamma}\bar{h} = -\|\bar{\gamma}\|_q = \sup_{\gamma \in \partial f_x} \gamma\bar{h} \quad (4.8)$$

Da (4.8) si deduce

$$-\|\bar{\gamma}\|_q \geq \gamma\bar{h} \quad \forall \gamma \in \partial f_x \quad (4.9)$$

Inoltre dalle proprietà delle norme si ha

$$-\gamma\bar{h} \leq \|\gamma\|_q \|\bar{h}\|_p \leq \|\gamma\|_q \quad \forall \gamma$$

e quindi, indicando $\Gamma := \{\gamma: \|\gamma\|_q \leq \|\bar{\gamma}\|_q\}$, si ha

$$\gamma\bar{h} \geq -\|\bar{\gamma}\|_q \quad \forall \gamma \in \Gamma \quad (4.10)$$

Confrontando (4.9) e (4.10) si vede che il piano $\gamma\bar{h} = -\|\bar{\gamma}\|_q$ separa Γ da ∂f_x . La direzione di discesa è la normale al piano separatore. Si vede quindi che, se $\bar{\gamma} \in \partial f_x$ è un punto interno di una faccetta di ∂f_x , la direzione di discesa \bar{h} è invariante rispetto al tipo di norma usata nel calcolo di (4.7) (vedi figura 4.26 dove Γ_0 si ottiene usando la norma $\|\cdot\|_\infty$ e Γ_1 usando la norma $\|\cdot\|_1$).

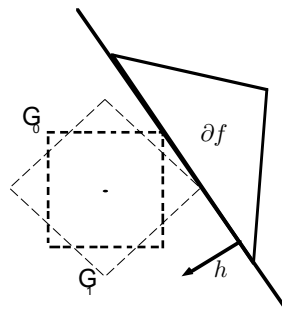


FIGURA 4.26

4.130 ESEMPIO. Si voglia minimizzare la funzione dell'esempio 4.120 con un metodo di massima discesa. Sia il punto iniziale ad esempio $x^0 := (1, -3)$. In x^0 la funzione è derivabile e il suo gradiente vale $\gamma^0 = (8, -4)$. La direzione di discesa è pertanto $h^0 := (-8, 4)$. Tipicamente un metodo di massima discesa prosegue lungo la direzione fino ad incontrare il primo punto di rottura. Questo è il punto $x^1 := (0, -2.5)$. In x^1 la funzione non è derivabile e il suo subdifferenziale è il segmento 2-3 in figura 4.23. Il punto di norma minima del subdifferenziale è il punto $\bar{\gamma} := (0, -4)$ sia che si usi la norma $\|\cdot\|_2$ oppure $\|\cdot\|_1$ (quali sono i punti a norma minima usando la norma $\|\cdot\|_\infty$?) e, in base alle considerazioni precedenti, la direzione di discesa è un vettore normale al segmento 2-3, cioè $h^1 := (0, 1)$. Si noti che, non a caso, questa direzione coincide esattamente con la linea di rottura. Proseguendo da x^1 lungo h^1 si incontra il punto di rottura indicato con C in figura 4.22 e il subdifferenziale in C è il triangolo indicato con C in figura 4.23. Il subdifferenziale di norma minima è un punto vicino al vertice 4 del triangolo e qualunque norma si usi, la direzione di discesa è normale al segmento 2-4 e quindi coincide con il segmento C-A di figura 4.22. Proseguendo lungo questo spigolo si arriva nel punto A il cui subdifferenziale contiene l'origine e quindi il punto è minimo. ■

Il metodo di massima discesa è alquanto complesso perché richiede la conoscenza del subdifferenziale e questo non sempre è immediato. Molto più semplice sarebbe scegliere come direzione una direzione opposta ad un generico subgradiente, dato che un subgradiente è di solito immediatamente disponibile. Purtroppo l'opposto di un subgradiente non è in generale una direzione di discesa come è facile vedere anche da esempi semplici, come in figura 4.27 dove \hat{x} è un punto di minimo, x è il punto corrente d'iterazione, l'area tratteggiata è l'insieme convesso $\{y : f(y) \leq f(x)\}$ e $-\gamma$ è la direzione di discesa data dall'opposto di un subgradiente.

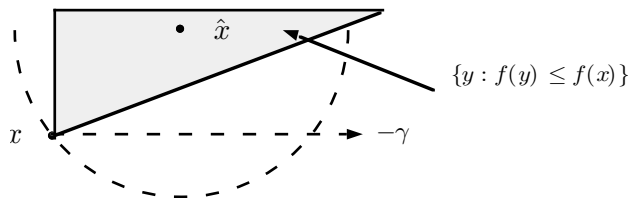


FIGURA 4.27

Tuttavia una tale direzione è una direzione di avvicinamento verso i minimi (si veda ancora la figura 4.27) e si può sfruttare questo fatto per ottenere un'iterazione convergente. Infatti sia $\gamma \in \partial f_x$ e \hat{x} un minimo e si calcolino i valori di $\bar{\alpha}$ per cui $\|x - \alpha \gamma - \hat{x}\|^2 = \|x - \hat{x}\|^2$. Si trova facilmente (oltre al valore $\bar{\alpha} = 0$ che non interessa)

$$\bar{\alpha} = 2 \frac{\gamma(x - \hat{x})}{\gamma \gamma} \tag{4.11}$$

Per ogni valore $0 < \alpha < \bar{\alpha}$ il punto $x - \alpha \gamma$ si trova dentro la sfera di raggio $\|x - \hat{x}\|$ e centro \hat{x} e per il valore $\bar{\alpha}/2$ si ottiene il massimo avvicinamento all'ottimo. Purtroppo $\bar{\alpha}$ non è calcolabile da (4.11) in quanto \hat{x} non è ovviamente noto. Si può però ottenere una limitazione inferiore di $\bar{\alpha}$ sfruttando la convessità di f , che implica $f(\hat{x}) - f(x) \geq \gamma(\hat{x} - x)$ e quindi

$$\bar{\alpha} = 2 \frac{\gamma(x - \hat{x})}{\gamma \gamma} \geq 2 \frac{f(x) - f(\hat{x})}{\gamma \gamma} =: 2\hat{\alpha} > 0 \tag{4.12}$$

Quindi per ogni $\alpha := \lambda \hat{\alpha}$, con $0 < \lambda < 2$ c'è la garanzia di un avvicinamento all'ottimo e si ha

$$\begin{aligned} \|x - \hat{x}\|^2 - \|x - \alpha \gamma - \hat{x}\|^2 &= 2\alpha \gamma(x - \hat{x}) - \alpha^2 \gamma \gamma \geq \\ &\geq 2\alpha (f(x) - f(\hat{x})) - \alpha^2 \gamma \gamma = \lambda(2 - \lambda) \frac{(f(x) - f(\hat{x}))^2}{\gamma \gamma} \end{aligned} \tag{4.13}$$

In (4.12) si richiede la conoscenza di $f(\hat{x})$, cioè il valore ottimo (e non l'ottimo stesso) che in qualche caso può essere noto a priori e in altri casi si può stimare (vedi più sotto). Se definiamo la seguente successione

$$x^{k+1} := x^k + \lambda \frac{f(\hat{x}) - f(x^k)}{\gamma^k \gamma^k} \gamma^k \tag{4.14}$$

con $0 < \lambda < 2$ si ottiene una successione di valori di f che tende verso il minimo valore. Per dimostrare la convergenza si noti che per la successione degli x^k , deve valere in base a (4.13)

$$\|\hat{x} - x^k\|^2 - \|\hat{x} - x^{k+1}\|^2 \geq \lambda(2 - \lambda) \frac{(f(x^k) - f(\hat{x}))^2}{\gamma^k \gamma^k}$$

cioè, sommando su tutte le disequazioni,

$$\|\hat{x} - x^1\|^2 \geq \|\hat{x} - x^1\|^2 - \|\hat{x} - x^k\|^2 \geq \lambda(2-\lambda) \sum_{i=1}^{k-1} \frac{(f(x^i) - f(\hat{x}))^2}{\gamma^i \gamma^i},$$

che, supponendo i subgradienti limitati superiormente in norma, implica $\lim f(x^k) = f(\hat{x})$. La successione generata da (4.14) prende il nome di *quasi discesa*. La difficoltà pratica di usare tale formula consiste, come già osservato, nel fatto che $f(\hat{x})$ non è noto, tranne in rari casi. Si tratta allora di usare al posto di $f(\hat{x})$ una stima \hat{f} di tale valore. Cosa succede se $\hat{f} \neq f(\hat{x})$? Consideriamo i due casi.

Sia $\hat{f} > f(\hat{x})$, cioè il valore ottimo è sottostimato (nel senso che è migliore della stima). In questo caso è come se l'iterazione fosse applicata alla funzione convessa $\tilde{f}(x) := \max\{f(x), \hat{f}\}$ il cui valore minimo è noto e vale \hat{f} . Allora l'iterazione converge verso i minimi di \tilde{f} cioè l'insieme $\{x : f(x) \leq \hat{f}\}$ e quindi la successione x^k 'si ferma' prima di raggiungere il minimo vero. L'iterazione è inoltre più lenta di quel che sarebbe stato usando il vero valore ottimo, in quanto il parametro $\hat{\alpha}$ risulta più piccolo.

Sia $\hat{f} < f(\hat{x})$, cioè il valore ottimo è sovrastimato (nel senso che è peggiore della stima). Il valore $\hat{\alpha}$ è più grande di quel che dovrebbe essere e, a seconda della stima, può anche essere superiore al valore critico $\bar{\alpha}$ per cui il nuovo valore x^{k+1} è più distante dal minimo di x^k . Sia $\Delta := f(\hat{x}) - \hat{f}$ e $\varepsilon_k := f(x^k) - f(\hat{x})$. Finché

$$\lambda \frac{f(x^k) - \hat{f}}{\gamma^k \gamma^k} < 2 \frac{f(x^k) - f(\hat{x})}{\gamma^k \gamma^k} \implies \varepsilon_k \geq \frac{\lambda}{2-\lambda} \Delta$$

l'iterazione si avvicina comunque al minimo. Non appena l'iterazione entra nella fascia al di sotto dell'errore $\lambda/(2-\lambda) \Delta$, l'andamento dell'iterazione comincia a divenire erratico e quindi si può interrompere l'iterazione cercando di correggere la stima. È bene notare che metodi di quasi discesa si usano normalmente quando interessa solamente un valore approssimato del minimo. Il valore più conveniente di λ dipende da vari fattori. Il valore $\lambda = 1$ massimizza la limitazione inferiore in (4.13) e quindi garantisce una rapida convergenza. Tenendo conto che il massimo avvicinamento all'ottimo si ottiene per $\alpha := \bar{\alpha} \geq \hat{\alpha}$ un valore di λ leggermente superiore a 1 può essere conveniente. Questa scelta però può rivelarsi controproducente se l'errore Δ della stima dell'ottimo è elevato.

4.131 ESEMPIO. Sia dato un sistema lineare $Ax \leq b$. Il sistema ammette soluzione ammissibile se e solo se il valore minimo v della funzione

$$f(x) = \max\{A_i x - b_i : i \in \{1, \dots, m\}\}$$

è minore o uguale a zero. Possiamo sempre pensare di introdurre un'ulteriore disequazione $0 \cdot x \leq 0$, che sarà sempre soddisfatta, e il cui scopo è di far sì che il minimo è uguale a 0 se e solo se il sistema è ammissibile. Pertanto, supponendo il sistema ammissibile, possiamo impiegare la seguente iterazione (ponendo $\lambda := 1$)

$$x := x + \frac{b_h - A_h x}{A_h A_h} A_h$$

dove h è l'indice della disequazione maggiormente violata per il valore x .

Si considerino ad esempio i due sistemi

$$\begin{array}{ll} -x_1 - x_2 \leq 0 & -x_1 - x_2 \leq 0 \\ x_1 - x_2 \leq 0 & x_1 - x_2 \leq 0 \\ -x_2 \leq -1 & x_2 \leq -1 \end{array}$$

Supponiamo di iniziare le iterazioni con il punto $(0, 0)$ in ambedue i casi. Per il primo sistema si ottiene

$$\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} + \frac{-1}{1} \begin{pmatrix} 0 \\ -1 \end{pmatrix} \quad \text{con } h = 3$$

e quindi si ottiene immediatamente una soluzione ammissibile. Nel secondo caso si ha invece

$$\begin{aligned} \begin{pmatrix} 0 \\ -1 \end{pmatrix} &= \begin{pmatrix} 0 \\ 0 \end{pmatrix} + \frac{-1}{1} \begin{pmatrix} 0 \\ 1 \end{pmatrix} && \text{con } h = 3 \\ \begin{pmatrix} 1/2 \\ -1/2 \end{pmatrix} &= \begin{pmatrix} 0 \\ -1 \end{pmatrix} + \frac{-1}{2} \begin{pmatrix} -1 \\ -1 \end{pmatrix} && \text{con } h = 1 \\ \begin{pmatrix} 0 \\ 0 \end{pmatrix} &= \begin{pmatrix} 1/2 \\ -1/2 \end{pmatrix} + \frac{-1}{2} \begin{pmatrix} 1 \\ -1 \end{pmatrix} && \text{con } h = 2 \end{aligned}$$

e a questo punto è evidente che l'iterazione è entrata in ciclo. Infatti in questo caso il sistema è non ammissibile. ■

4.132 ESEMPIO. Si riconsideri l'esempio 4.120. Si inizi l'iterazione a partire dal punto $x^0 := (1, -3)$ con valore $\lambda := 1$ e si usi nell'iterazione il valore minimo vero $f(\hat{x}) = 4$. L'iterazione che si ottiene è rappresentata in figura 4.28 e in figura 4.29 (dettaglio vicino al minimo). Se si usa una sottostima $\hat{f} = 5 > f(\hat{x})$, si ottiene l'iterazione di figura 4.30, che converge appunto verso un valore x tale che $f(x) = 5$. Se si usa una sovrastima $\hat{f} = 3 < f(\hat{x})$, si ottiene l'iterazione di figura 4.31, che nell'intorno del minimo comincia ad avere un andamento irregolare (si veda il dettaglio vicino al minimo in figura 4.32). In figura 4.33 è raffigurato il grafico di $f(x^k)$ in funzione di k . Come si vede il metodo non è di discesa e si vede chiaramente inoltre che da una certa iterazione in poi i valori di f hanno un andamento non prevedibile. ■

4.12. Funzioni concave

Le definizioni e le proprietà date valgono per funzioni convesse. Proprietà e definizioni simili valgono ovviamente anche per funzioni concave con opportune modifiche. Ci limitiamo semplicemente ad elencare la versione per funzioni concave dei risultati dati. Nel seguito si intende che f sia sempre una funzione concava:

— $\text{dom } f = \{x : f(x) > -\infty\}$

— (teorema 4.96) f è concava se e solo se

$$f(\alpha x + (1 - \alpha)y) \geq \alpha f(x) + (1 - \alpha)f(y) \quad \forall x, y \quad \forall \alpha : 0 \leq \alpha \leq 1$$

— (teorema 4.97) Ogni massimo locale di f è anche globale.

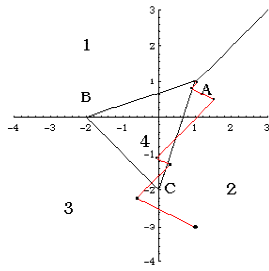


FIGURA 4.28

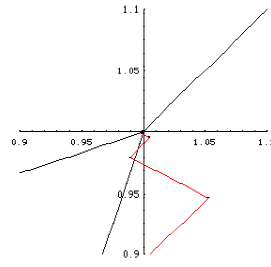


FIGURA 4.29

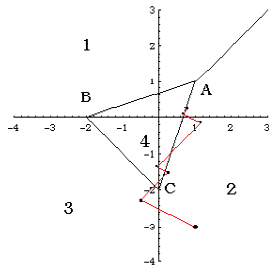


FIGURA 4.30

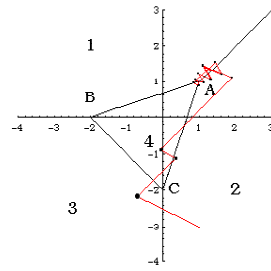


FIGURA 4.31

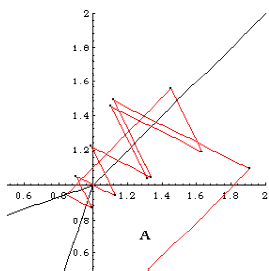


FIGURA 4.32

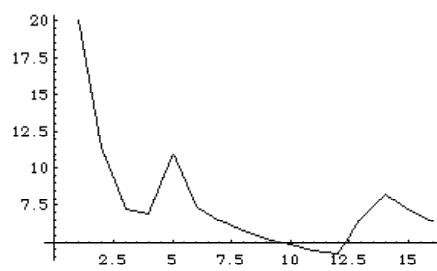


FIGURA 4.33

- (teorema 4.98) Gli insiemi $\{x: f(x) \geq \beta\}$ sono convessi $\forall \beta$.
- f è propria se $f(x) < +\infty, \forall x$.
- (teorema 4.105) Se f è chiusa, concava e propria, allora

$$f(x) = \inf_{u:(1,u) \in \text{dom } \varphi} \{ux - \varphi(1, u; \text{epi} - f)\}$$

- (teorema 4.109) Sia f differenziabile due volte nell'interno di $\text{dom } f$. Allora f è concava se e solo se il suo Hessiano è semidefinito negativo per ogni $x \in \text{dom } f$.
- (teorema 4.111) Sia f differenziabile in un aperto convesso, allora f è concava se e solo se

$$Df(x)h \geq f(x+h) - f(x) \quad \forall h$$

- $Df_x(h)$ è concava in h .

— (definizione 4.114) Sono subgradienti di f in x tutti i $\gamma \in \mathbb{R}^n$ tali che :

$$\gamma h \geq f(x+h) - f(x) \quad \forall h$$

— Quindi $\partial f_x = -\partial(-f)_x$ dove i due subdifferenziali sono definiti uno per le funzioni concave e l'altro per le funzioni convesse.

— (teorema 4.115) $\gamma \in \partial f_x$ se e soltanto se $\gamma h \geq Df_x(h) \quad \forall h$

— (teorema 4.116) Sia f chiusa o alternativamente $x \in \mathring{\text{dom}}f$. Allora

$$Df_x(h) = \inf_{\gamma \in \partial f_x} \gamma h$$

— (teorema 4.119) Sia $f(x) := \min_{i \in I} g^i(x)$, con $\{g^i\}_{i \in I}$ un insieme finito di funzioni concave, sia $x \in \mathring{\text{dom}}f$ e sia $I(x) := \{i \in I : f(x) = g^i(x)\}$. Allora

$$\partial f_x = \text{conv}\{\partial g_x^i : i \in I(x)\}$$

— (teorema 4.121) Sia

$$\tilde{f}(x) = \begin{cases} f(x) & \text{se } x \in K \\ -\infty & \text{altrimenti} \end{cases}$$

allora

$$\partial \tilde{f}(x) = \partial f(x) - D^*(x)$$

— (teorema 4.122) $Df_x(h) \leq 0, \forall h$ se e solo se x è un punto di massimo.

— (teorema 4.123) Sia f chiusa o alternativamente $x \in \mathring{\text{dom}}f$. Allora x è punto di massimo se e solo se $0 \in \partial f_x$.

— (teorema 4.127) Sia K convesso e chiuso e sia $D(x)$ il cono delle direzioni ammissibili in x . Un punto x è massimo su K di una funzione concava f se e solo se $\partial f(x) \cap D^*(x) \neq \emptyset$.

— (equazione (4.7)) Direzioni h di massima ascesa soddisfano

$$w = \max_{\|h\|_p \leq 1} Df_x(h) = \max_{\|h\|_p \leq 1} \inf_{\gamma \in \partial f_x} \gamma h = \inf_{\gamma \in \partial f_x} \max_{\|h\|_p \leq 1} \gamma h = \inf_{\gamma \in \partial f_x} \|\gamma\|_q$$

— (equazione (4.14)) L'iterazione di quasi ascesa è definita da (uguale al caso convesso):

$$x^{k+1} := x^k + \lambda \frac{f(\hat{x}) - f(x^k)}{\gamma^k \gamma^k} \gamma^k$$