Capitolo 6

Ottimizzazione vincolata

6.1 Introduzione

Con riferimento al problema di ottimizzazione

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

con $S \subset R^n$, illustriamo nel seguito alcune condizioni di ottimalità in relazione alle classi di problemi di nostro interesse. In termini molto generali, una condizione di ottimalità è una condizione (necessaria, sufficiente, necessaria e sufficiente) perché un punto x^* risulti una soluzione ottima (locale o globale) del problema. Ovviamente, una condizione di ottimalità sarà significativa se la verifica della condizione risulta più "semplice" o più "vantaggiosa" (da qualche punto di vista) rispetto all'applicazione diretta della definizione. Le condizioni di ottimalità si esprimono infatti, tipicamente, attraverso sistemi di equazioni, sistemi di disequazioni, condizioni sugli autovalori di matrici opportune.

Lo studio delle condizioni di ottimalità ha sia motivazioni di natura teorica, sia motivazioni di natura algoritmica. Dal punto di vista teorico, una condizione di ottimalità può servire a caratterizzare analiticamente le soluzioni di un problema di ottimo e quindi consentire di svolgere analisi *qualitative*, anche in assenza di soluzioni numeriche esplicite; un esempio è l'analisi della sensibilità delle soluzioni di un problema di ottimo rispetto a variazioni parametriche. Dal punto di vista algoritmico, una condizione *necessaria* può servire a restringere l'insieme in cui ricercare le soluzioni del problema originario e a costruire algoritmi finalizzati al soddisfacimento di tale condizione; una condizione *sufficiente* può servire a dimostrare che un punto ottenuto per via numerica sia una soluzione ottima del problema e quindi a definire criteri di arresto del procedimento risolutivo.

In particolare in questo paragrafo ci riferiremo a problemi i cui

- S è l'intero spazio $\mathbb{R}^n$ (problemi non vincolati)
- S è un generico insieme convesso,
- S è un poliedro.

Al fine di individuare le condizioni di ottimo, introduciamo i concetti di direzione di discesa e direzione ammissibile che consentiranno di caratterizzare i punti di minimo locale.

6.2 Direzioni di discesa

Allo scopo di definire condizioni di ottimalità è necessario introdurre la seguente definizione.

Definizione 6.2.1 (Direzione di discesa) *Sia $f : R^n \to R$ e $x \in R^n$. Si dice che un vettore $d \in R^n$, $d \neq 0$ è una direzione di discesa per f in x se esiste $\tilde{t} > 0$ tale che*

$$f(x + td) < f(x), \quad \textit{per ogni} \quad t \in (0, \tilde{t}\,].$$

Il concetto di direzione di discesa ci consentirà di caratterizzare i minimi locali di una funzione. Nel caso di funzione lineare $f(x) = c^T x$, si ottiene facilmente applicando la definizione

$$c^T(x + td) = c^T x + tc^T d < c^T x \qquad \text{per ogni } t > 0$$

che una direzione è di discesa se e solo risulta $c^T d < 0$ per ogni $t > 0$. Osserviamo che, in questo caso particolare, la caratterizzazione di direzione di discesa non dipende dal punto x in cui ci troviamo.

Se f è una funzione lineare, ossia se $f(x) \equiv c^T x$ allora d è una direzione di discesa (in un qualsiasi punto $x \in \mathbb{R}^n$) se e solo se $c^T d < 0$.

In generale, per funzioni continuamente differenziabili, è possibile caratterizzare la direzione di discesa utilizzando informazioni sulle derivate prime della funzione. tali caratterizzazioni saranno discusse nel paragrafo 6.5

6.3 Direzione ammissibile

Una direzione è "ammissibile" in un punto $\bar{x} \in S$ se è possibile muoversi lungo tale direzione mantenendosi in S. Più formalmente possiamo dare la seguente definizione riferita ad un generico insieme $S \subseteq \mathbb{R}^n$.

Definizione 6.3.1 (Direzione ammissibile) *Sia S un sottoinsieme di R^n e $\bar{x} \in S$. Si dice che un vettore $d \in R^n$, $d \neq 0$ è una direzione ammissibile per S in $\bar{x}$ se esiste $t_{\max} > 0$ tale che*

$$\bar{x} + td \in S, \quad \textit{per ogni} \quad t \in [0, t_{\max}\,].$$

Esempio 6.3.2 Sia dato l'insieme (poliedro)

$$S = \{x \in R^2 : \ x_1 + x_2 \leq 3, \ x_1 \geq 1, \ x_2 \geq 0\}$$

rappresentato in Figura 6.12 e consideriamo il punto ammissibile $\bar{x} = (1, \ 1)$ (indicato con un puntino rosso in figura).

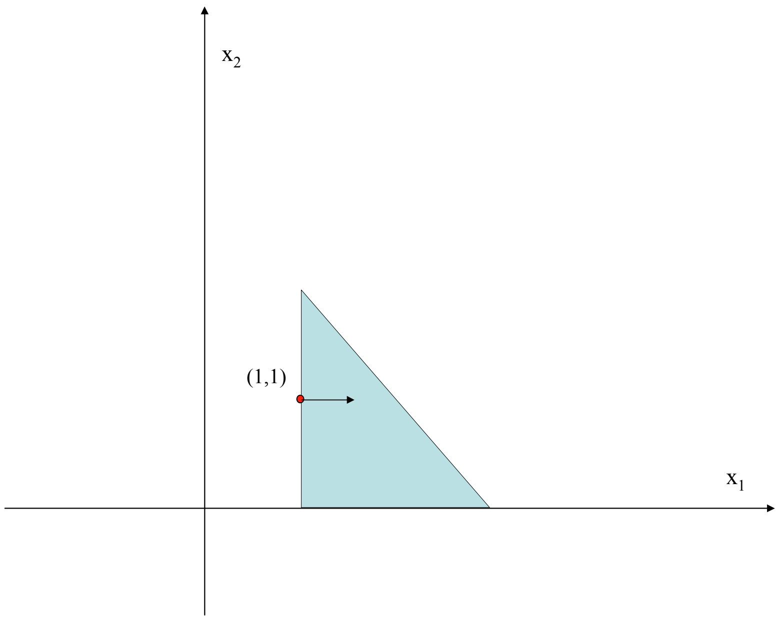

Figura 6.1: Poliedro Esempio 6.3.2.

È facile convincersi che la direzione $d = (1,\ 0)$ è ammissibile in $\bar{x}$. Risulta infatti

$$\bar{x} + td = \begin{pmatrix} 1 \\ 1 \end{pmatrix} + t \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1+t \\ 1 \end{pmatrix}$$

Sostituendo nei vincoli si ha:

1. $x_1 + x_2 = (1 + t) + 1 = 2 + t$ che risulta ≤ 3 per valori di $t \leq 1$.
2. $x_1 = 1 + t$ che risulta ≥ 1 per ogni $t \geq 0$;
3. $x_2 = 1 > 0$.

Quindi per ogni $t \in [0, 1]$ ($t_{\max} = 1$) risulta $\bar{x} + td \in S$.

È altrettanto facile convincersi che la direzione opposta $-d = (-1,\ 0)$ non è ammissibile in $\bar{x}$. Risulta infatti

$$y = \bar{x} + t(-d) = \begin{pmatrix} 1 \\ 1 \end{pmatrix} + t \begin{pmatrix} -1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1-t \\ 1 \end{pmatrix}$$

Sostituendo nei vincoli si ha:

1. $x_1 + x_2 = (1 - t) + 1 = 2 - t$ che risulta ≤ 3 per valori di $t \geq -1$.

2. $x_1 = 1 - t$ che risulta ≥ 1 solo per $t \leq 0$;
3. $x_2 = 1 > 0$ per ogni t.

Quindi NON esiste un $t_{\max} > 0$ per cui il punto y appartiene ad S. La direzione non è dunque ammissibile. □

Il concetto di direzione ammissibile è riferito ad un generico insieme S non necessariamente poliedrale né con particolari proprietà. Vedremo nel seguito come caratterizzare le direzioni ammissibili nel caso di particolari insiemi ammissibili.

6.4 Condizioni di ottimo vincolate

In questo paragrafo ricaviamo le prima condizioni di ottimalità per il generico problema

$$\begin{array}{ll} \min & f(x) \\ & x \in S \end{array} \qquad (P)$$

Una conseguenza immediata della Definizione 6.2.1 di direzione di discesa e della Definizione 6.3.1 di direzione ammissibile è la condizione necessaria di minimo locale enunciata nel teorema successivo.

Teorema 6.4.1 (Condizione necessaria di minimo locale) *Sia $x^* \in S$ un punto di minimo locale del problema (P) allora non può esistere una direzione ammissibile in x^* che sia anche di discesa per f.*

Dimostrazione. Sia x^* un minimo locale. Per assurdo, se esistesse una direzione d al tempo stesso ammissibile e di discesa in x^*, allora in ogni intorno di x^* sarebbe possibile trovare, per $t > 0$ abbastanza piccolo, un punto $x^* + td \in S$ tale che $f(x^* + td) < f(x^*)$, il che contraddice l'ipotesi che x^* sia un punto di minimo locale. □

La condizione necessaria può consentire di selezionare tra tutti i punti ammissibili i potenziali candidati ad essere punti di minimo locale.
Il risultato espresso nel Teorema 6.4.1 può essere poco significativo qualora vi siano regioni (di frontiera) dell'insieme ammissibile, tali che nei punti di queste regioni non esistano affatto direzioni ammissibili. In tal caso, la condizione del Teorema 6.4.1 non è in grado di discriminare le soluzioni ottime dagli altri punti ammissibili. Una situazione del genere si verifica, tipicamente, quando siano presenti vincoli di eguaglianza non lineari.

Consideriamo, a titolo esemplificativo, un problema in due variabili.

Esempio 6.4.2 Sia dato il problema

$$\begin{array}{ll} \min & x_2 \\ & x_1^2 + x_2^2 = 1 \end{array}$$

L'unico punto di minimo è il punto $(0, \ -1)^T$. Se consideriamo un qualsiasi punto $(x_1, x_2) \in S$ l'insieme delle direzioni ammissibili è vuoto e quindi la condizione del Teorema 6.4.1 è soddisfatta in ogni punto di S e non dà alcuna informazione aggiuntiva per la selezione di punti candidati ad essere ottimi. Tale situazione è illustrata nella figura 6.2.

□

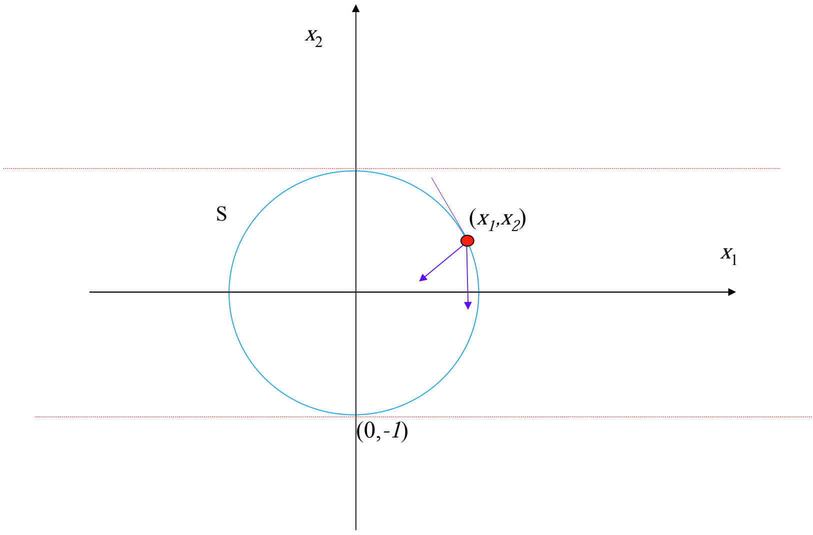

Figura 6.2: Esempio 6.4.2 in cui in qualunque punto ammissibile non esiste direzione ammissibile.

Per ricavare delle condizioni di ottimalità dalla condizione enunciata nel Teorema 6.4.1 occorre utilizzare una caratterizzazione analitica delle direzioni di discesa e delle direzioni ammissibili. In effetti la verifica della condizione espressa dal Teorema 6.4.1 è essenzialmente equivalente all'applicazione della definizione di minimo locale.

6.5 Caratterizzazione direzione di discesa

Vale in particolare la seguente condizioni sufficiente:

Teorema 6.5.1 (Condizione di discesa) *Supponiamo che f sia continuamente differenziabile nell'intorno di un punto $x \in R^n$ e sia $d \in R^n$ un vettore non nullo. Se risulta*

$$\nabla f(x)^T d < 0,$$

allora la direzione d è una direzione di discesa per f in x.

Dimostrazione. È noto (vedi Appendice A) che per una qualunque funzione continuamente differenziabile possiamo scrivere

$$f(x + td) = f(x) + t\nabla f(x)^T d + \alpha(x, t) \tag{6.1}$$

dove $\alpha(x,t)$ soddisfa $\lim_{t\to 0} \frac{\alpha(x,t)}{t} = 0$. Per valori sufficientemente piccoli di t possiamo dunque scrivere

$$f(x+td) - f(x) \simeq t\nabla f(x)^T d;$$

se $\nabla f(x)^T d < 0$ risulta allora $f(x+td) - f(x) < 0$ che quindi dimostra che d è una direzione di discesa. $\square$

Ricordando che $\nabla f(x)^T d = \|\nabla f(x)\| \|d\| \cos\theta$ dove θ è l'angolo compreso tra $\nabla f(x)$ e d, dal punto di vista geometrico la condizione $\nabla f(x)^T d < 0$ esprime il fatto che se una direzione d forma un angolo ottuso con il gradiente di f in x, allora d è una direzione di discesa per f in x.

Sia f continuamente differenziabile e sia $d \in R^n$ un vettore non nullo. Se l'angolo θ tra $\nabla f(x)$ e d soddisfa

$$\theta > 90^o$$

allora d è di discesa per f in x.

Tra le direzioni di discesa, un ruolo particolarmente importante è svolto dalla direzione dell'*antigradiente* $d = -\nabla f(x)$. Se $\nabla f(x) \neq 0$, la direzione dell'antigradiente $d = -\nabla f(x)$ è sempre una direzione di discesa, infatti risulta

$$\nabla f(x)^T d = -\nabla f(x)^T \nabla f(x) = -\|\nabla f(x)\|^2 < 0.$$

La definizione è illustrata nella figura 6.3, con riferimento agli insiemi di livello di una funzione in R^2. La direzione dell'antigradiente è utilizzata per la definizione di un semplice algoritmo

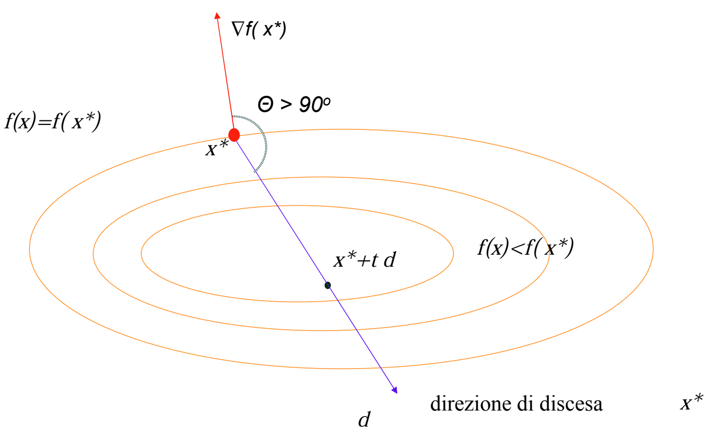

Figura 6.3: Esempio di direzione d di discesa in $I\!R^2$.

per la determinazione di punti di minimo (vedi il paragrafo 6.8).

Naturalmente è possibile dare una condizione sufficiente analoga al teorema 6.5.1 affinché una direzione sia di salita.

Sia f sia continuamente differenziabile nell'intorno di un punto $x \in R^n$ e sia $d \in R^n$ un vettore non nullo. Se risulta

$$\nabla f(x)^T d > 0,$$

allora la direzione d è una direzione di salita per f in x.

Naturalmente se $\nabla f(x) \neq 0$, la direzione del gradiente $d = \nabla f(x)$ è sempre una direzione di salita in x.

Osserviamo che nel caso di funzione lineare $f = c^T x$, risulta $\nabla f(x) = c$, quindi la condizione espressa dal Teorema 6.5.1 coincide con la caratterizzazione ottenuta in base della definizione di direzione di discesa. Nel caso lineare, la relazione (6.1) vale con $\alpha(x, t) \equiv 0$ e dunque la condizione del Teorema 6.5.1 caratterizza tutte e sole le direzioni di discesa. Si tratta cioè di una condizione necessaria e sufficiente, mentre nel caso generale è solo una condizione sufficiente. Nel caso di funzione lineare vale quindi la seguente caratterizzazione:

Sia f è una funzione lineare, ossia $f(x) \equiv c^T x$ allora

1. d è una direzione di discesa in x se e solo se $\nabla f(x)^T d = c^T d < 0$;
2. d è una direzione di salita in x se e solo se $\nabla f(x)^T d = c^T d > 0$;
3. d è una direzione lungo cui la funzione si mantiene costante rispetto al valore in x se e solo se $\nabla f(x)^T d = c^T d = 0$.

Nel caso generale, la condizione di discesa enunciata è solo sufficiente, in quanto possono esistere direzioni di discesa tali che $\nabla f(x)^T d = 0$. Illustriamo le varie possibilità nel seguente esempio.

Esempio 6.5.2 Consideriamo la funzione

$$f(x) = x_1^3 + x_2^2 - 3x_1.$$

sia dato il punto $\bar{x} = (-1,\ 0)^T$ e la direzione $d = (-1,\ 0)^T$. Il gradiente della funzione vale

$$\nabla f(x) = \begin{pmatrix} 3x_1^2 - 3 \\ 2x_2 \end{pmatrix}.$$

Si verifica facilmente che il gradiente della funzione si annulla nel punto $\bar{x}$, quindi la condizione sufficiente espressa dal teorema 6.5.1 non è soddisfatta in quanto $\nabla f(\bar{x})^T d = 0$. La direzione d è però di discesa in $\bar{x}$. Infatti risulta $\bar{x} + td = (-1 - t,\ 0)^T$ e la funzione vale:

$$f(\bar{x} + td) = (-1 - t)^3 - 3(-1 - t) = -1 - t^3 - 3t - 3t^2 + 3 + 3t = -t^3 + 3t^2 + 2.$$

Mostriamo che esiste un valore di $\bar{t}$ tale che $-t^3 - 3t^2 + 2 < 2 = f(\bar{x})$ per ogni $0 < t \leq \bar{t}$. La condizione che si ottiene è $-t^3 - 3t^2 < 0$ che risulta verificata per qualunque valore di $t > 0$, quindi effettivamente la direzione è di discesa nel punto.

si consideri ora il punto $\tilde{x} = (1,\ 0)^T$ e la stessa direzione $d = (-1,\ 0)^T$. Si verifica facilmente che il gradiente della funzione si annulla nel punto $\tilde{x}$, e quindi la condizione sufficiente espressa dal teorema 6.5.1 anche in questo caso non è soddisfatta in quanto $\nabla f(\tilde{x})^T d = 0$. In questo caso però la direzione d non è di discesa, risulta infatti $\tilde{x} + td = (1 - t,\ 0)^T$ e la funzione vale:

$$f(\bar{x} + td) = (1-t)^3 - 3(1-t) = 1 - t^3 - 3t + 3t^2 - 3 + 3t = -t^3 + 3t^2 - 2$$

Dobbiamo quindi verificare se esiste un $\bar{t} > 0$ tale che per ogni $t \in (0, \bar{t}]$ risulta $-t^3 + 3t^2 = t^2(-t+3) < 0$. Si verifica facilmente che la condizione è verificata solo per valori di t abbastanza grandi ($t > 3$) e quindi d non è di discesa. Si tratta invece di una direzione di salita, infatti scegliendo $\bar{t} < 3$, si ottiene che $f(\tilde{x} + td) > f(\tilde{x})$ per ogni $t \in (0, \bar{t}]$.
Consideriamo ora il punto $\hat{x} = (0,\ 0)^T$ in cui il gradiente di f vale $\nabla f(\hat{x}) = (-3,\ 0)^T$. Risulta che la direzione $d = (-1,\ 0)^T$ è di salita in $\hat{x}$, mentre la direzione opposta $-d = (1,\ 0)^T$ è di discesa.

Dall'esempio e dalle considerazioni precedenti risulta evidente che dato in un punto x ed una direzione d, sono possibili solo questi casi:

Sia f sia continuamente differenziabile nell'intorno di un punto $x \in R^n$ e sia $d \in R^n$ un vettore non nullo.

1. se $\nabla f(x)^T d < 0$ allora la direzione è di discesa in x;
2. se $\nabla f(x)^T d > 0$ allora la direzione è di salita in x;
3. se $\nabla f(x)^T d = 0$ allora non si può concludere se la direzione è di discesa o di salita in x o lungo la quale la funzione si mantiene costante.

La condizione del teorema 6.5.1 diventa anche necessaria nel caso di convessità della funzione obiettivo.

Teorema 6.5.3 *Se f è convessa allora una direzione d è di discesa in un punto x se e solo se*

$$\nabla f(x)^T d < 0.$$

Dimostrazione. Sia f convessa, dobbiamo solo dimostrare che la condizione $\nabla f(x)^T d < 0$ è necessaria (la sufficienza è data dal teorema 6.5.1). Sia $x \in I\!R^n$ e sia d una direzione di discesa per f in x. Possiamo scrivere per la convessità della funzione

$$f(x + \alpha d) \geq f(x) + \alpha \nabla f(x)^T d.$$

Se per assurdo fosse $\nabla f(x)^T d \geq 0$ allora per $\alpha > 0$ sufficientemente piccolo, si avrebbe $f(x + \alpha d) \geq f(x)$ contraddicendo l'ipotesi che d sia di discesa. □

Se f è differenziabile due volte è possibile caratterizzare l'andamento di f lungo una direzione assegnata tenendo conto anche delle derivate seconde e di ciò si può tener conto, come si vedrà in seguito, per stabilire condizioni di ottimo del secondo ordine.

Introduciamo la definizione seguente.

Definizione 6.5.4 (Direzione a curvatura negativa)

Sia $f : R^n \to R$ due volte continuamente differenziabile nell'intorno di un punto $x \in R^n$. Si dice che un vettore $d \in R^n$, $d \neq 0$ è una direzione a curvatura negativa per f in x se risulta

$$d^T \nabla^2 f(x) d < 0. \qquad \square$$

Una direzione a curvatura negativa è quindi tale che la derivata direzionale seconda è negativa in x, per cui diminuisce localmente la derivata direzionale del primo ordine. In particolare vale il risultato seguente.

Teorema 6.5.5 (Condizione di discesa del secondo ordine)

Sia $f : R^n \to R$ due volte continuamente differenziabile nell'intorno di un punto $x \in R^n$ e sia $d \in R^n$ un vettore non nullo. Supponiamo che risulti

$$\nabla f(x)^T d = 0,$$

e che d sia una direzione a curvatura negativa in x, ossia che

$$d^T \nabla^2 f(x) d < 0.$$

Allora la direzione d è una direzione di discesa per f in x.

Dimostrazione. Poiché f è differenziabile due volte, si ha:

$$f(x + td) = f(x) + t\nabla f(x)^T d + \frac{1}{2} t^2 d^T \nabla^2 f(x) d + \beta(x, td)$$

in cui $\beta(x, td)/t^2 \to 0$. Essendo per ipotesi $\nabla f(x)^T d = 0$, si può scrivere:

$$\frac{f(x + td) - f(x)}{t^2} = \frac{1}{2} d^T \nabla^2 f(x) d + \frac{\beta(x, td)}{t^2}$$

e quindi, poiché

$$\lim_{t \to 0} \frac{\beta(x, td)}{t^2} = 0,$$

per valori sufficientemente piccoli di t si ha $f(x + td) - f(x) < 0$, per cui d è una direzione di discesa. $\square$

Esempio 6.5.6 Consideriamo la funzione del'esempio 6.5.2,

$$f(x) = x_1^3 + x_2^2 - 3x_1.$$

nel punto $\bar{x} = (-1,\ 0)^T$. La direzione $d = (-1,\ 0)^T$ è tale che $\nabla f(\bar{x})^T d = 0$, quindi la condizione sufficiente espressa dal teorema 6.5.1 non è soddisfatta. Verifichiamo la condizione del Teorema 6.5.5. La matrice hessiana è

$$\nabla^2 f(x) = \begin{pmatrix} 6x_1 & 0 \\ 0 & 2 \end{pmatrix} \overset{\bar{x}}{=} \begin{pmatrix} -6 & 0 \\ 0 & 2 \end{pmatrix}.$$

Risulta allora $d^T \nabla^2 f(\bar{x}) d = -6$ e quindi la direzione soddisfa la condizione sufficiente del Teorema 6.5.5, per cui si può concludere che è di discesa.

6.6 Condizioni di ottimo del 1^o ordine

Si può però specificare ulteriormente la condizione necessaria utilizzando la caratterizzazione di direzione di discesa del teorema 6.5.1.

Si ottiene allora la seguente condizione necessaria.

Se x^* è un punto di minimo locale del problema (P) non può esistere una direzione ammissibile $d \in R^n$ in x^* tale che $\nabla f(x^*)^T d < 0$.

Tale condizione può essere equivalentemente enunciata come segue.

Teorema 6.6.1 (Condizione necessaria di minimo locale) *Se x^* è un punto di minimo locale del del problema (P), allora risulta*

$$\nabla f(x^*)^T d \geq 0 \qquad \textit{per ogni direzione ammissibile } d \in R^n \textit{ in } \; x^*.$$

Ricordando che $\nabla f(x^*)^T d = \|\nabla f(x^*)\| \|d\| \cos\theta$ dove θ è l'angolo compreso tra $\nabla f(x^*)$ e d, dal punto di vista geometrico la condizione $\nabla f(x^*)^T d \geq 0$ richiede che $0^o \leq \theta \leq 90^o$ per ogni d ammissibile.

Se x^* è un punto di minimo locale del del problema (P), allora per ogni d ammissibile l'angolo θ formato con $\nabla f(x^*)$ è

$$0^o \leq \theta \leq 90^o.$$

La condizione espressa dal Teorema 6.6.1 è solo necessaria di minimo locale. Illustriamo la condizione con un esempio.

Esempio 6.6.2 Sia dato il problema di ottimizzazione

$$\begin{array}{ll} \min & -x_1^2 - x_2^2 \\ & -1 \leq x_1 \leq 1 \\ & -1 \leq x_2 \leq 1 \end{array}$$

Si tratta della minimizzazione di una funzione quadratica concava su un poliedro (la regione ammissibile e le curve di livello della funzione obiettivo sono rappresentate in Figura 6.4.) Geometricamente si individuano i minimi globali sui vertici del cubo ammnissibile. Consideriamo il punto di minimo $x^* = (1,\ 1)^T$. Il gradiente della funzione obiettivo è

$$\nabla f(x) = \begin{pmatrix} -2x_1 \\ -2x_2 \end{pmatrix}$$

e in x^* vale $\nabla f(x^*) = (-2,\ -2)^T$ (freccia rossa in Figura 6.4). Si osservi che il gradiente della funzione obiettivo NON si annulla in x^*. Le direzioni ammissibili d nel punto x^* sono in base

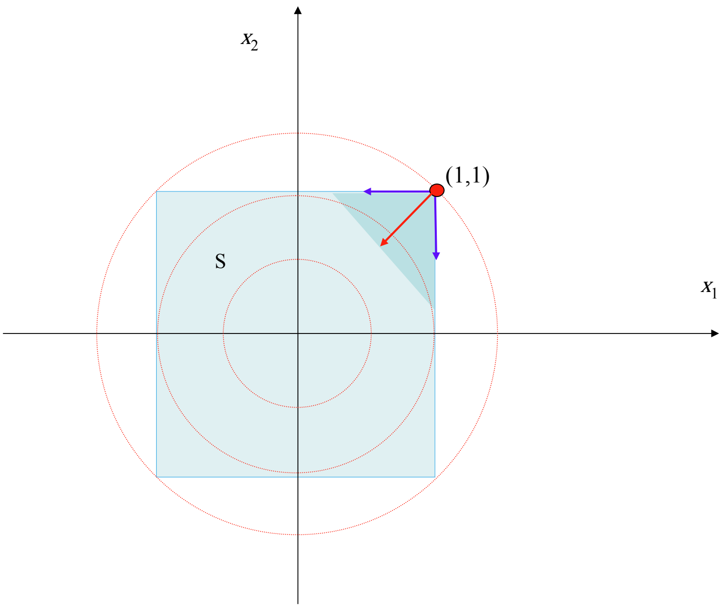

Figura 6.4: Illustrazione Esempio 6.6.2.

alla definizione le soluzioni del sistema definito da $x^* + td$ con $t \geq 0$ (in Figura 6.4 sono i vettori nella regione compresa tra le frecce blu), cioè:

$$\begin{array}{l} x_1^* + td_1 = 1 + td_1 \geq -1 \\ x_1^* + td_1 = 1 + td_1 \leq 1 \\ x_2^* + td_2 = 1 + td_2 \geq -1 \\ x_2^* + td_2 = 1 + td_2 \leq 1 \end{array}$$

ovvero le direzioni ammissibile sono quelle che soddisfano

$$\begin{array}{l} 2 + td_1 \geq 0 \\ td_1 \leq 0 \\ 2 + td_2 \geq 0 \\ td_2 \leq 0. \end{array}$$

La seconda e quarta condizione impongono $d_1, d_2 \leq 0$, mentre la prima e terza impongono solo delle condizioni sullo spostamento t e richiedono rispettivamente che $0 \leq t \leq 2/|d_1|$ e $0 \leq t \leq 2/|d_2|$. Dunque risulta $\nabla f(x^*)^T d = (-2 \ \ -2) \begin{pmatrix} d_1 \\ d_2 \end{pmatrix} = -2(d_1 + d_2) \geq 0$ per ogni d ammissibile (ovvero tale che $d_1, d_2 \leq 0$). Quindi la condizione necessaria è soddisfatta.
Sia dato ora il punto $\bar{x} = (0, \ 0)^T$ interno alla regione ammissibile (nessun vincolo attivo). Ogni vettore $d \in I\!R^2$ è ammissibile in $\bar{x}$. Inoltre risulta $\nabla f(\bar{x})^T d = 0$ per ogni d, in quanto

$\nabla f(\bar{x}) = 0$, quindi è verificata la condizione necessaria. Ciononostante il punto $\bar{x}$ non è un minimo locale, anzi è un massimo. Infatti per ogni $d \in I\!R^2$ risulta $\bar{x} + td = (td_1,\ td_2)^T$ e la funzione vale $f(\bar{x} + td) = -t^2(d_1^2 + d_2^2)$. Per ogni $t \neq 0$ e $d \not\equiv 0$ risulta

$$f(\bar{x} + td) = -t^2(d_1^2 + d_2^2) < f(\bar{x}) = 0,$$

ovvero qualunque d non identicamente nulla è di discesa in $\bar{x}$. □

Possiamo dare delle condizioni che utilizzano la caratterizzazione del secondo ordine di direzione di discesa espressa al Teorema 6.5.5. In particolare quindi abbiamo:

Teorema 6.6.3 (Condizione necessaria del 2^o ordine di minimo locale) *Se x^* è un punto di minimo locale del del problema (P), allora per ogni d ammissibile in x^* tale che $\nabla f(x^*)^T d = 0$ risulta*

$$d^T \nabla^2 f(x^*)^T d \geq 0 \qquad \textit{per ogni direzione ammissibile } d \in R^n \textit{ in } x^*.$$

A partire dalle condizioni espresse dai teoremi 6.6.1 e 6.10.3 è possibile ricavare condizioni più specifiche in base alle caratterizzazioni delle direzioni ammissibili, e quindi in base alla struttura e alle proprietà delle funzioni che definiscono i vincoli del problema.
Considereremo nel seguito solo i casi di nostro interesse.

6.7 Ottimizzazione non vincolata

Abbiamo già osservato nel capitolo 2 che i problemi non vincolati sono una particolare classe di problemi non lineari, in cui $S = R^n$. Il problema in oggetto è quindi

$$\min_{x \in I\!R^n} f(x), \qquad (P - NV)$$

con $f : I\!R^n \to I\!R$ non lineare e continuamente differenziabile. Osserviamo preventivamente che in questo caso non si può applicare in modo diretto il teorema di Weierstrass per l'esistenza di un minimo. Condizioni di esistenza possono essere comunque ottenute con riferimento ad un insieme ammissibile "fittizio". In particolare, dato un qualunque punto $\hat{x}$ in cui la funzione vale $f(\hat{x})$, e si considera l'insieme non vuoto (contiene almeno $\hat{x}$)

$$\mathcal{L}(\hat{x}) = \{x \in I\!R^n :\ f(x) \leq f(\hat{x})\}.$$

Possiamo allora considerare il problema "vincolato"

$$\begin{array}{ll} \min & f(x) \\ & x \in \mathcal{L}(\hat{x}) \end{array} \qquad (P - NV\mathcal{L})$$

Se $\hat{x}$ non è il minimo globale di (P-NV), allora il minimo globale x^*, se esiste, sicuramente si trova in $\mathcal{L}(\hat{x})$. Infatti per definizione $f(x^*) \leq f(x)$ per ogni x e quindi in particolare vale anche $f(x^*) \leq f(\hat{x})$. Quindi risolvere (P-NV) é equivalente a risolvere (P-NV$\mathcal{L}$).
Per assicurare l'esistenza di un minimo globale di (P-NV$\mathcal{L}$) (e quindi di (P-NV), possiamo utilizzare il teorema di Weierstrass. È sufficiente richiedere che esista un $\hat{x}$ per cui l'insieme $\mathcal{L}(\hat{x})$ sia compatto.

Condizione di esistenza di una soluzione.
Se esiste un $\hat{x} \in I\!R^n$ per cui l'insieme $\mathcal{L}(\hat{x})$ é compatto, allora il problema (P-NV) ammette un minimo globale.

Deriviamo ora le condizioni di ottimo per (P-NV). Una conseguenza immediata della Definizione 6.2.1 di direzione di discesa è la condizione necessaria di minimo locale enunciata nel teorema successivo.

Teorema 6.7.1 (Condizione necessaria di minimo locale) *Sia $x^* \in \mathbb{R}^n$ un punto di minimo locale del problema (P-NV) allora non può esistere in x^* una direzione di discesa per f.*

La condizione necessaria può consentire di selezionare tra tutti i punti ammissibili i potenziali candidati ad essere punti di minimo locale. Per ricavare delle condizioni di ottimalità dalla condizione enunciata nel Teorema 6.7.1 occorre utilizzare una caratterizzazione analitica delle direzioni di discesa. In effetti la verifica della condizione espressa dal Teorema 6.7.1 è essenzialmente equivalente all'applicazione della definizione di minimo locale.
Si può però specificare ulteriormente la condizione necessaria utilizzando la caratterizzazione di direzione di discesa. Si ottiene allora la seguente condizione necessaria.

Se x^* è un punto di minimo locale del del problema (P) non può esistere una direzione $d \in R^n$ in x^* tale che $\nabla f(x^*)^T d < 0$.

Tale condizione può essere equivalentemente enunciata come segue.

Teorema 6.7.2 (Condizione necessaria di minimo locale) *Se x^* è un punto di minimo locale del del problema (P), allora risulta*

$$\nabla f(x^*)^T d \geq 0 \qquad \textit{per ogni direzione } d \in R^n.$$

Ricordando che $\nabla f(x^*)^T d = \|\nabla f(x^*)\| \|d\| \cos\theta$ dove θ è l'angolo compreso tra $\nabla f(x^*)$ e d, dal punto di vista geometrico la condizione $\nabla f(x^*)^T d \geq 0$ richiede che $0^o \leq \theta \leq 90^o$ per ogni d.

Se x^* è un punto di minimo locale del del problema (P), allora per ogni d l'angolo θ formato con $\nabla f(x^*)$ è

$$0^o \leq \theta \leq 90^o.$$

Si ottiene quindi la seguente condizione necessaria.

Teorema 6.7.3 (Condizione necessaria di minimo locale non vincolato) *Sia $f : R^n \to R$ continuamente differenziabile su R^n e sia $x^\star \in R^n$. Se x^* è un punto di minimo locale non vincolato di f in R^n allora si ha*

$$\nabla f(x^\star) = 0.$$

Dimostrazione. Se $x^* \in R^n$ è un punto di minimo locale non possono esistere direzioni di discesa in x^*. Se $\nabla f(x^*) \neq 0$, esisterebbe una direzione $d = -\nabla f(x^*) \neq 0$ di discesa in x^*. Ma questo è assurdo. □

Definizione 6.7.4 (Punto stazionario) *Sia $f : R^n \to R$ continuamente differenziabile su R^n. Un punto $\bar{x}$ tale che $\nabla f(\bar{x}) = 0$ si dice* **punto stazionario** *di f.*

I possibili candidati ad essere punti di minimo locale del problema (P-NV) si determinano determinando le soluzioni di $\nabla f(x) = 0$.

Esempio 6.7.5 Sia dato il problema

$$f(x_1, x_2) = {x_1}^4 + {x_2}^4 - 3x_1x_2,$$

sapendo che esiste un punto di minimo, determinarlo utilizzando le condizioni necessarie.
Il gradiente della funzione è dato da:

$$\nabla f = \begin{pmatrix} 4x_1^3 - 3x_2 \\ 4x_2^3 - 3x_1 \end{pmatrix}.$$

Dall'annullamento del gradiente si ottiene il sistema

$$\begin{cases} 4x_1^3 - 3x_2 = 0 \\ 4x_2^3 - 3x_1 = 0 \end{cases} \qquad \begin{cases} x_2\left(\frac{256}{27}x_2^8 - 3\right) = 0 \\ x_1 = \frac{4}{3}x_2^3 \end{cases}$$

Si ottengono le soluzioni: $A(0,0)$, $B\left(\frac{\sqrt{3}}{2}, \frac{\sqrt{3}}{2}\right)$, $C\left(-\frac{\sqrt{3}}{2}, -\frac{\sqrt{3}}{2}\right)$.
Le curve di livello della funzione sono nella Figura 6.5 da cui si individuano i tre punti stazionari. Il punto $(0,\ 0)^T$ è un punto di sella con valore della funzione $f(0,0) = 0$.

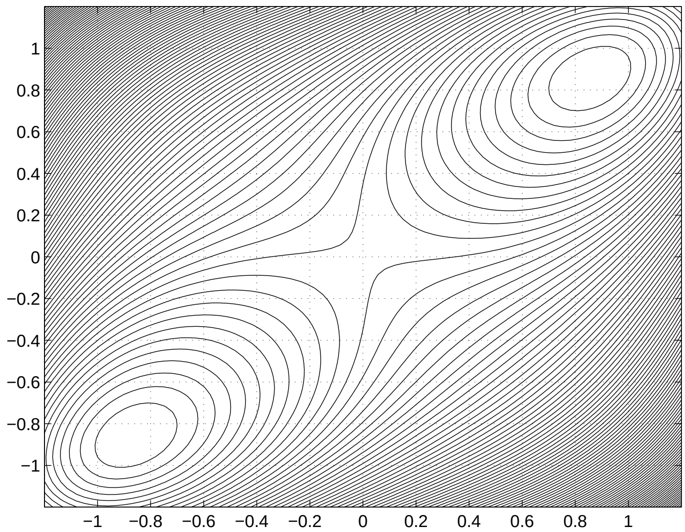

Figura 6.5: Curve di livello della funzione $x_1^4 + x_2^4 - 3x_1x_2$.

I punti $B, C = \left(\pm\frac{\sqrt{3}}{2}, \pm\frac{\sqrt{3}}{2}\right)$ abbiamo

$$\nabla^2 f = 3\begin{pmatrix} 3 & -1 \\ -1 & 3 \end{pmatrix}$$

sono punti di minimo globale di valore della funzione obiettivo $f = -\frac{9}{8}$.

$\square$

Senza ulteriori ipotesi sulla funzione obiettivo, la condizione è solo necessaria. Ricordando che se $\nabla f(x^*) \neq 0$ la direzione $d = \nabla f(x^*)$ è di salita nel punto x^*, si ottiene che la condizione $\nabla f(x^*) = 0$ è anche necessaria di massimo locale. In generale si hanno questi possibili casi.

Sia $f : R^n \to R$ continuamente differenziabile su R^n e sia $x^\star \in R^n$ tale che $\nabla f(x^\star) = 0$ allora è vera una delle seguenti affermazioni:

1. $x^\star$ è un minimo locale (non esistono in $x^\star$ direzioni di discesa);
2. $x^\star$ è un massimo locale (non esistono in $x^\star$ direzioni di salita);
3. $x^\star$ è un punto di sella (esistono in $x^\star$ sia direzioni di discesa che di salita).

Un esempio di punto di sella è in Figura 6.6.

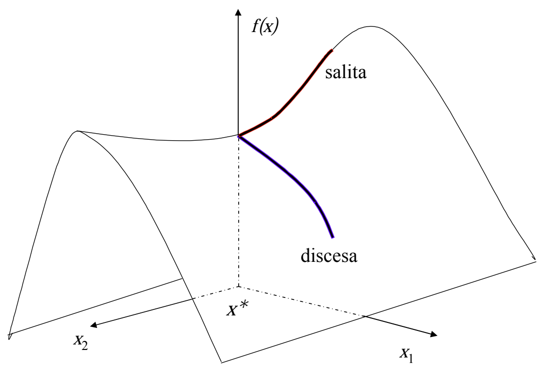

Figura 6.6: Esempio di punto di sella in $I\!R^2$.

Teorema 6.7.6 (Condizione sufficiente di minimo globale non vincolato) *Sia $f : R^n \to R$ continuamente differenziabile su R^n e sia $x^\star \in R^n$. Si supponga che f sia convessa. Se $\nabla f(x^\star) = 0$, allora $x^\star$ è un punto di minimo globale di f su R^n. Inoltre, se f è strettamente convessa su R^n, allora $x^\star$ è l'unico punto di minimo globale di f su R^n.*

Dai teoremi 6.7.3 e 6.7.6 si ottiene quindi una condizione necessaria e sufficiente di minimo globale non vincolato.

Teorema 6.7.7 (Condizione necessaria e sufficiente di minimo globale non vincolato) *Sia $f : R^n \to R$, con ∇f continuo su R^n e si supponga che f sia convessa. Allora $x^\star$ è un punto di minimo globale di f su R^n se e solo se $\nabla f(x^\star) = 0$. Inoltre, se f è strettamente convessa su R^n e se in $x^\star$ si ha $\nabla f(x^\star) = 0$, allora $x^\star$ è l'unico punto stazionario di f e costituisce anche l'unico punto di minimo globale della funzione.*

Possiamo dare delle condizioni che utilizzano la caratterizzazione del secondo ordine di direzione di discesa espressa al Teorema 6.5.5. In particolare quindi abbiamo:

Teorema 6.7.8 (Condizione necessaria del 2^o ordine di minimo locale) *Se x^* è un punto di minimo locale del del problema (P), allora per ogni d tale che $\nabla f(x^*)^T d = 0$ risulta*

$$d^T \nabla^2 f(x^*)^T d \geq 0 \qquad \textit{per ogni direzione } d \in R^n$$

Se si considera la matrice Hessiana si ottiene una condizione necessaria del secondo ordine.

Teorema 6.7.9 (Condizione necessaria del secondo ordine)

Sia $f : R^n \to R$ due volte continuamente differenziabile in un intorno di $x^\star \in R^n$. Allora, se $x^\star$ è un punto di minimo locale non vincolato di f si ha:

(a) $\nabla f(x^\star) = 0$;

(b) $\nabla^2 f(x^*)$ *è semidefinita positiva, ossia* $y^T \nabla^2 f(x^\star) y \geq 0$, per ogni $y \in R^n$.

Dimostrazione. La (a) segue dalla caratterizzazione di direzioni di discesa del secondo ordine data dal Teorema 6.5.5. □

Se tuttavia $\nabla^2 f(x^\star)$ è definita positiva, si può stabilire un risultato più forte, che riportiamo nel teorema successivo.

Teorema 6.7.10 (Condizione sufficiente del secondo ordine)

Sia $f : R^n \to R$ due volte continuamente differenziabile in un intorno di $x^* \in R^n$.

Supponiamo che valgano le condizioni:

(a) $\nabla f(x^\star) = 0$

(b) *la matrice Hessiana è definita positiva in $x^\star$, ossia:*

$$y^T \nabla^2 f(x^\star) y > 0 \quad \text{per ogni} \quad y \in R^n, \quad y \neq 0.$$

Allora $x^\star$ è un punto di minimo locale stretto.

Dimostrazione. Utilizzando il Teorema di Taylor, e tenendo conto del fatto che $\nabla f(x^\star) = 0$, possiamo scrivere:

$$f(x^\star + \alpha d) = f(x^\star) + \frac{1}{2}\alpha^2 d^T \nabla^2 f(x^\star) d + o(\alpha^2),$$

e quindi si ottiene:

$$\frac{f(x^\star + \alpha d) - f(x^\star)}{\alpha^2} = \frac{1}{2} d^T \nabla^2 f(x^\star) d + \frac{o(\alpha^2)}{\alpha^2}$$

Quindi per valori di α piccoli, l'ultimo termine è trascurabile ($o(\alpha^2)/\alpha^2 \to 0$ per $\alpha \to 0$), per cui $f(x^\star + \alpha d) - f(x^\star) > 0$ per ogni d, il che prova l'enunciato. □

Dai risultati precedenti si possono dedurre facilmente condizioni necessarie e condizioni sufficienti di massimo locale (basta infatti imporre condizioni di minimo su $-f$).

In particolare, la condizione che $x^\star$ sia *un punto stazionario*, ossia tale che $\nabla f(x^\star) = 0$ è condizione necessaria sia perché $x^\star$ sia un punto di minimo, sia perchè $x^\star$ sia un punto di massimo.

Una condizione necessaria del secondo ordine per un massimo locale è che $\nabla f(x^\star) = 0$ e $\nabla^2 f(x^\star)$ sia *semidefinita negativa*. Condizione sufficiente perchè $x^\star$ sia un punto di massimo locale stretto è che $\nabla f(x^\star) = 0$ e $\nabla^2 f(x^\star)$ sia *definita negativa*.

Se risulta $\nabla f(x^\star) = 0$ e la matrice Hessiana $\nabla^2 f(x^\star)$ è *indefinita* (ossia, esistono vettori d per cui $d^T \nabla^2 f(x^\star) d > 0$ e altri per cui $d^T \nabla^2 f(x^\star) d < 0$) allora si può escludere che $x^\star$ sia un punto di minimo o di massimo locale e $x^\star$ viene usualmente denominato *punto di sella*.

Un punto di sella è un punto stazionario in corrispondenza al quale esistono sia direzioni *di discesa* (quando $d^T \nabla^2 f(x^\star) d < 0$) sia direzioni *di salita* (quando $d^T \nabla^2 f(x^\star) d > 0$).

o Se $\nabla f(x^\star) = 0$ e $\nabla^2 f(x^\star)$ è *semidefinita* (negativa o positiva) non si può determinare la natura di $x^\star$ in assenza di altre informazioni. Tuttavia se $\nabla^2 f(x^\star)$ *non* è semidefinita positiva (negativa) si può escludere che x^* sia un punto di minimo (massimo).

Una classe di particolare interesse di funzioni non lineari è quella delle funzioni quadratiche, ovvero del tipo

$$f(x) = \frac{1}{2} x^T Q x + c^T x$$

con Q matrice simmetrica $n \times n$ e $c \in \mathbb{R}^n$. Poiché la matrice hessiana di una funzione quadratica è costante è facile verificare la convessità della funzione (vedi paragrafo 5.2 del Capitolo 5). Nel caso quadratico è anche possibile formalizzare i risultati di esistenza di un minimo globale.

Vale in particolare la seguente caratterizzazione.

Teorema 6.7.11 (Minimizzazione di una funzione quadratica) *Sia* $q(x) = \frac{1}{2} x^T Q x + c^T x$, *con* Q *simmetrica e* $c \in R^n$. *Allora:*

1. $q(x)$ *ammette un punto di minimo se e solo se* Q *è semidefinita positiva ed esiste* $x^\star$ *tale che* $Qx^\star + c = 0$;
2. $q(x)$ *ammette un unico punto di minimo globale se e solo se* Q *è definita positiva;*
3. *se* Q *è semidefinita positiva ogni punto* $x^\star$ *tale che* $Qx^\star + c = 0$ *è un punto di minimo globale di* $q(x)$.

Dimostrazione. Ricordiamo che risulta $\nabla q(x) = Qx + c$ e $\nabla^2 f(x) = Q$. Inoltre se risulta Q semidefinita positiva, $q(x)$ è convessa, se risulta Q definita positiva, $q(x)$ è strettamente convessa.
Assegnati $x^\star, x \in R^n$ e posto $x = x^\star + s$ si può scrivere:

$$q(x) = q(x^\star + s) = q(x^\star) + (Qx^\star + c)^T s + \frac{1}{2} s^T Q s. \tag{6.2}$$

Supponiamo ora che $Qx^\star + c = 0$ e che Q sia semidefinita positiva; in tal caso dalla (6.2) segue $q(x) \geq q(x^\star)$ per ogni $x \in R^n$.
Inversamente, se q ammette un punto di minimo $x^\star$, per il Teorema 6.7.3 deve essere $\nabla q(x^\star) = 0$ e risulta $q(x) \geq q(x^\star)$ per ogni $x \in I\!R^n$. Dalla (6.2) segue $0 \leq q(x) - q(x^\star) = \frac{1}{2} s^T Q s$ con $s = x - x^\star$, ovvero $Q \succeq 0$. Ciò prova la (1).
La (2) segue dalla (1) e dal fatto che, $s^T Q s > 0$ per ogni s se e solo se Q è definita positiva.
Infine, la (3) segue ancora dalla (6.2) perchè per ogni $x^\star$ tale che $\nabla q(x^\star) = 0$ si ha $q(x) \geq q(x^\star)$.
□

Illustriamo questo teorema con un esempio.

Esempio 6.7.12 Sia data la funzione

$$q(x_1, x_2) = \frac{1}{2}(\alpha x_1^2 + \beta x_2^2) - x_1.$$

Studiamo le esistenza e la natura dei punti estremali al variare dei parametri α e β.

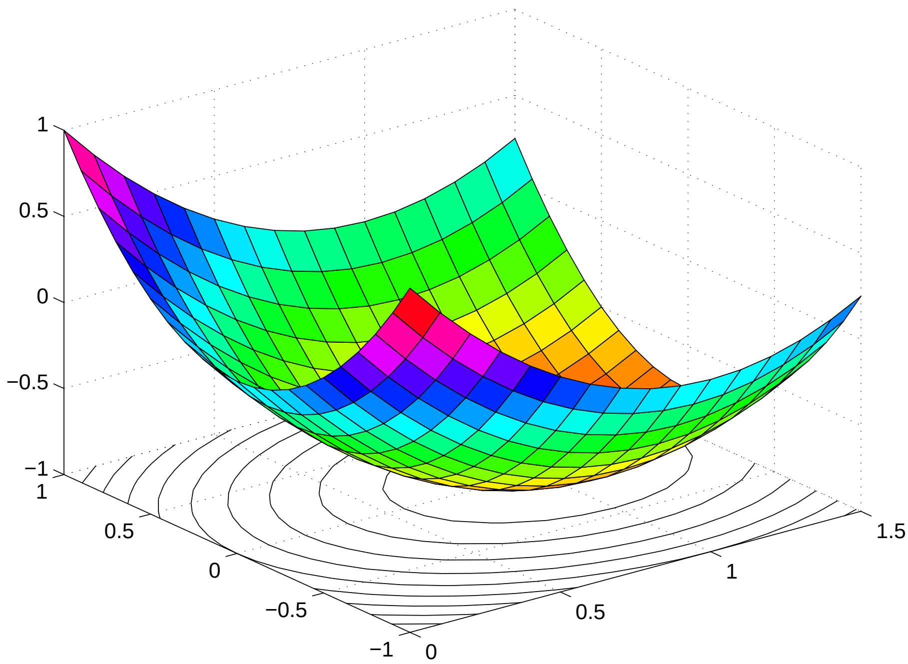

Figura 6.7: Grafico di $q(x_1, x_2)$ per $\alpha = \beta = 1$.

Scriviamo il gradiente e la matrice hessiana di q. Si ha

$$\nabla q = \begin{pmatrix} \alpha x_1 - 1 \\ \beta x_2 \end{pmatrix} \qquad \nabla^2 q = \begin{pmatrix} \alpha & 0 \\ 0 & \beta \end{pmatrix}$$

Se $\alpha > 0$ e $\beta > 0$, allora esiste un unica soluzione al sistema $\nabla q = 0$ ed è $\left(\dfrac{1}{\alpha}, 0\right)^T$. Inoltre la matrice $\nabla^2 q$ è definita positiva; si tratta quindi dell'unico punto di minimo globale.

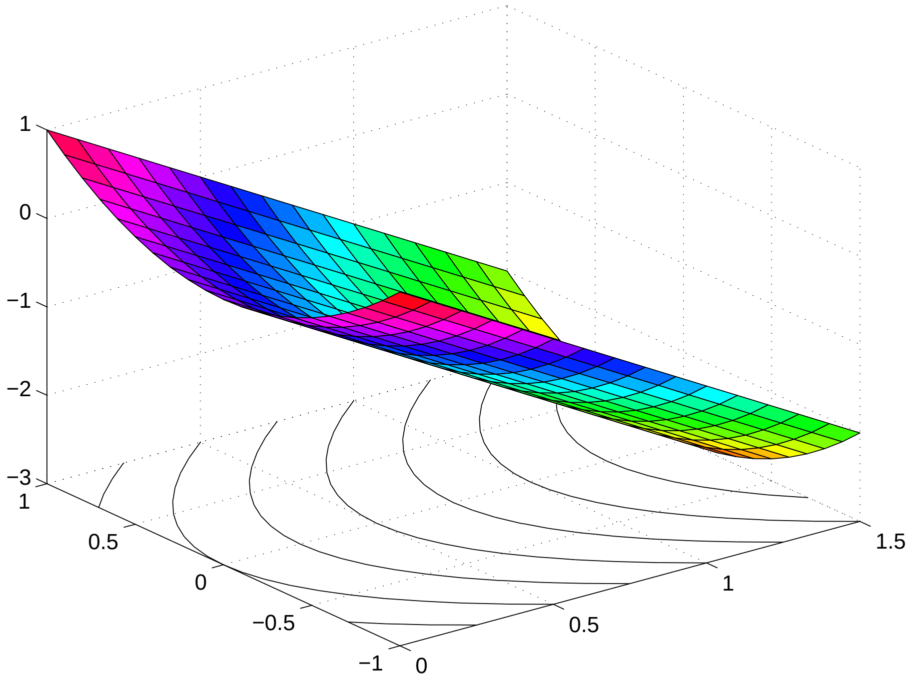

Figura 6.8: Grafico di $q(x_1, x_2)$ per $\alpha = 0$ $\beta = 1$.

Se $\alpha = 0$ e β è qualsiasi, non esiste soluzione al sistema $\nabla q = 0$. Notare che se $\beta \geq 0$ la matrice è semidefinita positiva, ma questo non assicura l'esistenza del minimo globale.

Se $\alpha > 0$ e $\beta = 0$, esistono infinite soluzioni al sistema $\nabla q = 0$ ed è $\left(\dfrac{1}{\alpha}, \xi\right)^T$ con ξ qualsiasi. Inoltre la matrice $\nabla^2 q$ è semidefinita positiva; si tratta quindi di infiniti punti di minimo globale.

Se $\alpha < 0$ e $\beta > 0$ si ha un'unica soluzione $(\dfrac{1}{\alpha}, 0)$. Ma la matrice hessiana è indefinita; si tratta quindi di un punto di sella.

Nel caso di $\alpha < 0$ e $\beta < 0$, allora esiste un unica soluzione al sistema $\nabla q = 0$ ed è $\left(\dfrac{1}{\alpha}, 0\right)^T$. Inoltre la matrice $\nabla^2 q$ è definita negativa; si tratta quindi dell'unico punto di massimo globale.
□

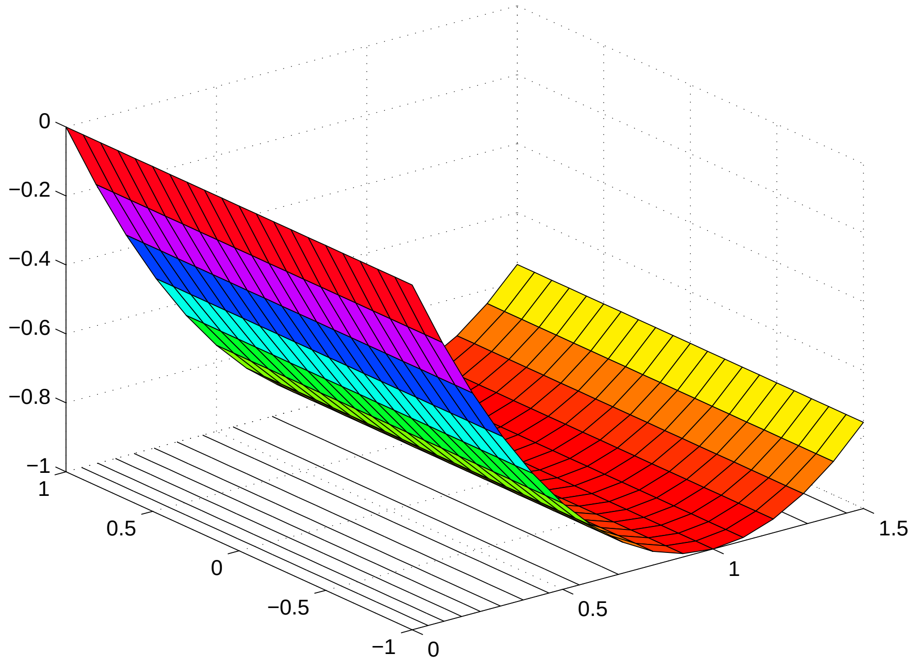

Figura 6.9: Grafico di $q(x_1, x_2)$ per $\alpha = 1$ $\beta = 0$.

6.8 Utilizzo algoritmico delle condizioni di ottimo non vincolate

Consideriamo il problema di ottimizzazione non vincolata:

$$\min_{x \in R^n} f(x) \tag{6.3}$$

in cui $f : R^n \to R$ e supponiamo verificata la seguente ipotesi che garantisce l'esistenza di un punto di minimo x^*:

Ipotesi 1 *La funzione $f : R^n \to R$ è una funzione continuamente differenziabile ed esiste un $x^0 \in R^n$ tale che l'insieme di livello $\mathcal{L}_{x^0}$ sia compatto.*

Gli algoritmi per la soluzione del problema (6.3) consentono, in generale, soltanto la determinazione di punti stazionari di f, ossia di punti dell'insieme

$$\Omega := \{\omega \in R^n : \nabla f(\omega) = 0\}.$$

In generale si riesce anche a garantire che, se x^0 non è già un punto stazionario, vengano comunque ottenuti punti in cui la funzione obiettivo assume un valore inferiore al valore assunto nel punto iniziale x^0 e ciò consente di ottenere soluzioni soddisfacenti in molte applicazioni.
Se la funzione obiettivo è convessa, la determinazione di un punto stazionario risolve completamente il problema, poiché, come è noto, ogni punto stazionario di una funzione convessa è un punto di minimo globale.

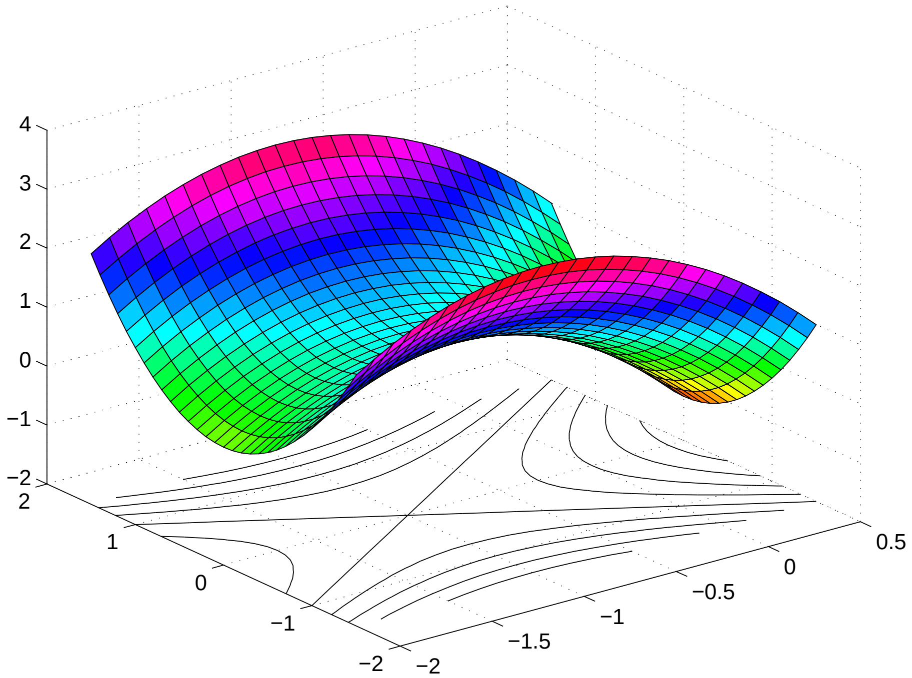

Figura 6.10: Grafico di $q(x_1, x_2)$ per $\alpha = -1$ $\beta = 1$.

Gli algoritmi che ci proponiamo di studiare possono essere descritti per mezzo dello schema concettuale seguente:

Schema generico di algoritmo di ottimizzazione non vincolata

1. Si fissa un *punto iniziale* $x^0 \in R^n$ e si pone $k = 0$.
2. Se $x^k \in \Omega$ stop.
3. Si calcola una *direzione di discesa* $d^k \in R^n$.
4. Si calcola un *passo* $\alpha^k \in R$ lungo d^k;
5. Si determina un nuovo punto $x^{k+1} = x^k + \alpha^k d^k$. Si pone $k = k + 1$ e si ritorna al Passo 2.

Commentiamo brevemente lo schema considerato.

1. **Scelta del punto iniziale.** Il punto iniziale dell'algoritmo è un *dato* del problema e deve essere fornito in relazione alla particolare funzione che si intende minimizzare. Il punto x^0 dovrebbe essere scelto come la migliore stima disponibile della soluzione ottima, eventualmente facendo riferimento a un modello semplificato della funzione obiettivo. Nella maggior parte dei casi, tuttavia, non esistono criteri generali per effettuare una buona scelta di x^0.

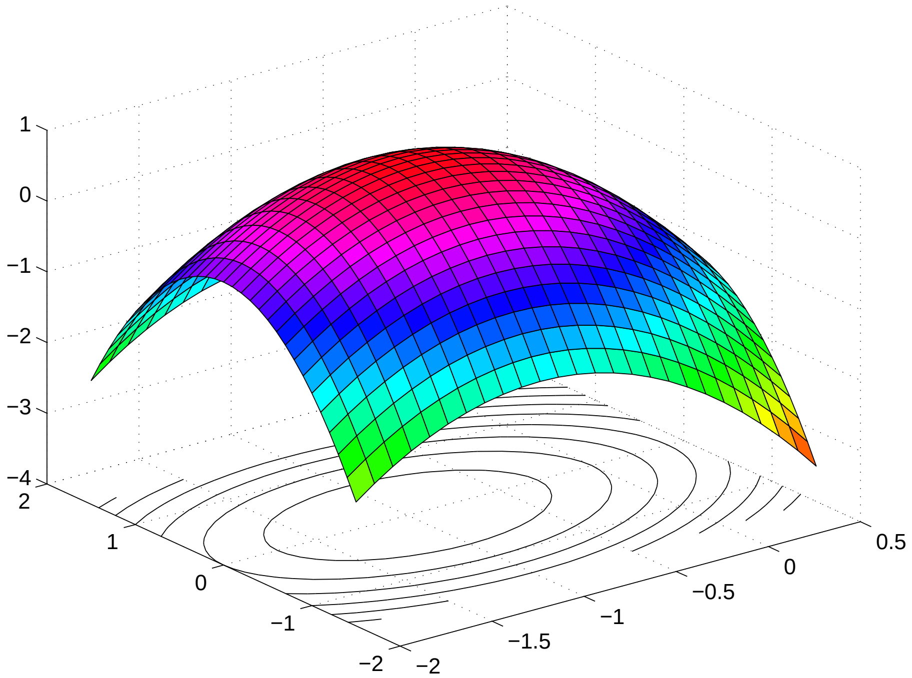

Figura 6.11: Grafico di $q(x_1, x_2)$ per $\alpha = \beta = -1$.

Nella soluzione di problemi applicativi (non convessi) può essere conveniente ripetere la ricerca a partire da punti iniziali differenti, ad esempio generati casualmente, e scegliere poi il punto stazionario migliore tra quelli così determinati.

2. **Criterio di arresto.** La verifica effettuata al Passo 2 sull'appartenenza di x^k all'insieme Ω equivale a controllare se $\nabla f(x^k) = 0$. In pratica, per l'utilizzo su calcolatore con precisione finita, occorre specificare un *criterio di arresto*.

   Una possibilità consiste nell'arrestare l'algoritmo quando

$$\|\nabla f(x^k)\| \leq \varepsilon \tag{6.4}$$

   in cui $\varepsilon > 0$ è un valore sufficientemente piccolo (valori standard richiedono $\varepsilon \in [10^{-8}, 10^{-3}]$).

   Per quanto riguarda la scelta della norma, tra le più utilizzate ci sono $\|\cdot\|_2$ e $\|\cdot\|_\infty$. Ricordiamo che $\|x\|_2 = \sqrt{\sum_{i=1}^n x_i^2}$ e $\|x\|_\infty = \max_{1 \leq i \leq n}\{|x_i|\}$.

3. **Scelta della direzione.** I criteri seguiti nella scelta della direzione di ricerca d^k individuano il particolare metodo di ottimizzazione utilizzato.

   Tra i metodi che utilizzano informazioni del primo ordine riportiamo il *metodo del gradiente* che consiste nello scegliere come direzione d^k in x^k quella che minimizza la derivata direzionale

$$\min_{\|d\|=1} \nabla f(x^k)^d$$

Poiché $\nabla f(x^k)^d = \|\nabla f(x^k)\| \|d\| \cos\theta$ dove θ ̀l'angolo tra $\nabla f(x^k)$ e d, segue banalmente che la soluzione del problema è $d^k = -\nabla f(x^k)$; per questo motivo è detto anche *metodo della discesa più ripida*. La generica iterazione del metodo del gradiente è dunque

$$x^{k+1} = x^k - \alpha^k \nabla f(x^k).$$

4. **Calcolo del passo.** Il calcolo dello scalare $\alpha^k > 0$ costituisce la cosiddetta *ricerca unidimensionale* o *ricerca di linea* e viene effettuato valutando la funzione obiettivo lungo la direzione d^k. Per la ricerca unidimensionale sono disponibili diversi algoritmi. Si può scegliere il valore di α^k che risolve il problema unidimensionale

$$\min_{\alpha} \; f(x^k + \alpha d^k).$$

In questo caso si parla di *ricerca di linea esatta* (limitata ad alcuni casi particolari, ad es. per funzioni quadratiche convesse). Oppure si possono utilizzare regole di *ricerca di linea inesatta* che consentono di l'individuazione di un intervallo di valori accettabili per α^k ai fini dell'ottenimento della convergenza. In questo corso non entreremo nel dettagli delle Ricerche di Linea.

6.9 Esempi di modelli di ottimizzazione non vincolata

Esempio 6.9.1 *Il problema di discriminazione del prezzo descritto nell'Esempio 2.4.5 è un problema di programmazione quadratica strettamente convessa del tipo*

$$\begin{aligned} min \quad & x^T M x - (a-c)^T x \\ & x \geq 0. \end{aligned}$$

Si osservi che il minimo non vincolato x^*, ovvero il punto che annulla il gradiente della funzione obiettivo $2Mx^* - (a-c) = 0$, ha componenti $x_i^* = \dfrac{a_i - c_i}{2m_i}$. Dunque se $a_i - c_i \geq 0$ pe ogni i, il punto x^* ha tutte le componenti non negative ed é ottimo (unico) anche del problema vincolato. Nell'esempio numerico riportato in Es. 2.4.5 si ha

$$\nabla f = \begin{pmatrix} 10x_A - 30 \\ 20x_B - 80 \end{pmatrix} \qquad \nabla^2 f = \begin{pmatrix} 10 & 0 \\ 0 & 20 \end{pmatrix}$$

La matrice hessiana è definita positiva, ed esiste un unico punto di che annulla il gradiente $x^* = (3,4)^T$ che ha componenti positive, dunque é soluzione del problema vincolato originario. □

Esempio 6.9.2 (Minimi quadrati)
Consideriamo il caso 1 definito nel Esempio 2.4.6.

$$\min_x \left[x_0^2 + (x_1 + x_0 - 1)^2 + (2x_1 + x_0 - 4)^2 + (3x_1 + x_0 - 9)^2\right].$$

Si tratta di un problema quadratico del tipo

$$\min \begin{pmatrix} x_1 \\ x_0 \end{pmatrix}^T Q \begin{pmatrix} x_1 \\ x_0 \end{pmatrix} + c^T \begin{pmatrix} x_1 \\ x_0 \end{pmatrix}$$

con

$$Q = \begin{pmatrix} 14 & 6 \\ 6 & 4 \end{pmatrix} \qquad c = \begin{pmatrix} -72 \\ -28 \end{pmatrix}$$

Il problema è strettamente convesso. Esiste un unico punto di minimo globale che si ottiene come soluzione di $2Qx + c = 0$ e vale $(-1,\ 3)^T$. Dunque la funzione lineare che minimizza l'errore quadratico è $y = 3t - 1$. $\square$

6.10 Ottimizzazione su insieme convesso generico

Consideriamo un problema di ottimo in cui in cui supponiamo che l'insieme ammissibile S sia un *insieme convesso*, ovvero

$$\begin{array}{ll} min & f(x) \\ & x \in S. \end{array} \qquad (P - CONV)$$

Possiamo caratterizzare le direzioni ammissibili di questo insieme.

Teorema 6.10.1 (Direzioni ammissibili insieme convesso) *Sia $S \subseteq I\!R^n$ un insieme convesso e sia $\bar{x}$ un qualsiasi punto di S. Allora comunque si fissi $x \in S$ tale che $x \neq \bar{x}$, la direzione $d = x - \bar{x}$ è una direzione ammissibile per S in $\bar{x}$.*

Dimostrazione. Sia $\bar{x} \in S$. Comunque preso $x \in S$, con $x \neq \bar{x}$, per la convessità di S, si ha che $(1 - \beta)\bar{x} + \beta x \in S$ per ogni $\beta \in [0, 1]$ e quindi $\bar{x} + \beta(x - \bar{x}) \in S$ per ogni $\beta \in [0, 1]$. Da cui segue che la direzione $d = x - \bar{x}$ è una direzione ammissibile per S in $\bar{x}$. $\square$

È facile verificare che, inversamente, se $d \neq 0$ è una direzione ammissibile per S in $\bar{x}$ allora esiste un punto $x \in S$ e uno scalare t tale che $d = t(x - \bar{x})$.

Quindi utilizzando direttamente il teorema 6.6.1, otteniamo la seguente condizione necessaria.

Teorema 6.10.2 [Condizioni necessarie di minimo locale con insieme ammissibile convesso]

Sia $x^* \in S$ un punto di minimo locale del problema (P-CONV) e supponiamo che f sia continuamente differenziabile su $I\!R^n$. Allora si ha:

$$\nabla f(x^*)^T (x - x^*) \geq 0, \quad \text{per ogni} \quad x \in S. \qquad (6.5)$$

Dim. Osserviamo innanzitutto che in base al teorema 6.10.1 la direzione $d = x - x^*$ è una direzione ammissibile per f in S per ogni $x \in S$. Se esistesse $d = x - x^*$ tale che $\nabla f(x^*)^T d < 0$ per il Teorema 6.5.1 la direzione ammissibile d sarebbe di discesa ma questo è assurdo perché in base al Teorema 6.4.1 sappiamo che non può esistere una direzione ammissibile di discesa in x^*. $\square$

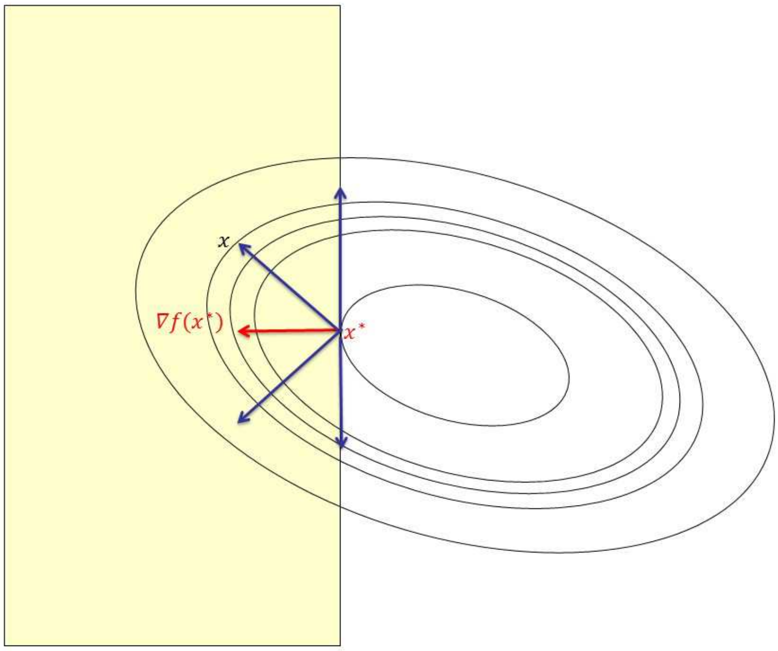

Figura 6.12: Rappresentazione grafica delle condizioni di ottimo su insieme convesso.

Nel caso che f sia convessa, si può dimostrare che la condizione è anche sufficiente. Vale in effetti il seguente risultato.

Teorema 6.10.3 [Condizioni necessarie e sufficienti di minimo globale nel caso convesso]

Sia S un sottoinsieme convesso di R^n e supponiamo che f sia una funzione convessa continuamente differenziabile su $I\!R^n$. Allora $x^* \in S$ è un punto di minimo globale del problema (P-CONV) se e solo se

$$\nabla f(x^*)^T(x - x^*) \geq 0, \quad \text{per ogni} \quad x \in S. \tag{6.6}$$

Dimostrazione. La necessità segue dal Teorema 6.10.2. La sufficienza segue dal Teorema 5.4.1. Infatti, se f è convessa e $x^* \in S$ deve essere, per ogni $x \in S$:

$$f(x) \geq f(x^*) + \nabla f(x^*)^T(x - x^*),$$

e quindi se vale la (6.6) si ha che $f(x) \geq f(x^*)$ per ogni $x \in S$, il che prova che x^* è un punto di minimo globale. □

Teorema 6.10.4 (Condizioni necessarie del secondo ordine di minimo locale con insieme ammissibile convesso)

Sia $x^* \in S$ un punto di minimo locale del problema (P-CONV). Se f è due volte continuamente differenziabile su un insieme aperto contenente S, risulta

$$(x - x^*)^T \nabla^2 f(x^*)(x - x^*) \geq 0, \quad \text{per ogni} \quad x \in S \quad \text{tale che} \quad \nabla f(x^*)'(x - x^*) = 0. \tag{6.7}$$

Dimostrazione. In base al Teorema 6.4.1 sappiamo che non può esistere una direzione ammissibile di discesa. Se f è due volte continuamente differenziabile ed esiste $x \in S$ tale che la direzione ammissibile $d = x - x^*$ soddisfi $\nabla f(x^*)^T d = 0$. Applicando il teorema 6.5.5 si ottiene allora la (6.7). □

Consideriamo il seguente esempio in cui i vincoli sono solo limitazioni superiori ed inferiori sulle variabili.

Esempio 6.10.5 Consideriamo l'esempio 6.6.2

$$\begin{array}{ll} \min & -x_1^2 - x_2^2 \\ & -1 \leq x_1 \leq 1 \\ & -1 \leq x_2 \leq 1 \end{array}$$

L'insieme ammissibile è convesso, quindi le condizioni necessarie richiedono che in un punto x^* risulti $\nabla f(x^*)^T(x - x^*) \geq 0$ per ogni $x \in S$. In particolare in questo caso

$$\nabla f = \begin{pmatrix} -2x_1 \\ -2x_2 \end{pmatrix}$$

e deve risultare quindi

$$-2\bar{x}_1(x_1 - \bar{x}_1) - 2\bar{x}_2(x_2 - \bar{x}_2) \geq 0 \quad \text{per ogni} \quad -1 \leq x_1 \leq 1 \quad -1 \leq x_2 \leq 1.$$

Verifichiamo che le condizioni sono soddisfatte in $\bar{x} = (1,\ 1)^T$. Deve risultare

$$-2(x_1 - 1) - 2(x_2 - 1) \geq 0 \ \ \text{per ogni} \quad -1 \leq x_1 \leq 1 \quad -1 \leq x_2 \leq 1.$$

Il termine $x_1 - 1 \leq 0$ per punti $x \in S$ e analogamente $x_2 - 1 \leq 0$ per $x \in S$, quindi la condizione è effettivamente soddisfatta. □

6.11 Utilizzo algoritmico delle condizioni di ottimo per problemi con vincoli convessi

Consideriamo il problema di ottimizzazione vincolata:

$$\min_{x \in S} \ f(x) \tag{6.8}$$

in cui $f : R^n \to R$ con S convesso e supponiamo verificata la seguente ipotesi (semplificativa) che garantisce l'esistenza di un punto di minimo $x^* \in S$:

Ipotesi 2 *La funzione $f : R^n \to R$ è una funzione continuamente differenziabile ed S sia compatto.*

Gli algoritmi per la soluzione del problema (6.8) consentono, in generale, soltanto la determinazione di punti dell'insieme

$$\Omega := \{\omega \in R^n : \nabla f(\omega)^T (x - \omega) \geq 0, \text{per ogni } x \in S\}.$$

Un possibile schema concettuale di algoritmo è il seguente:

Schema generico di algoritmo di ottimizzazione vincolata per S convessi

1. Si fissa un *punto iniziale* $x^0 \in S$ e si pone $k = 0$.
2. Se $x^k \in \Omega$ stop.
3. Si calcola una *direzione ammissibile e di discesa* $d^k \in R^n$.
4. Si calcola un *passo* $\alpha^k \in R$ lungo d^k;
5. Si determina un nuovo punto $x^{k+1} = x^k + \alpha^k d^k \in S$. Si pone $k = k + 1$ e si ritorna al Passo 2.

Per i commenti generali rimandiamo al paragrafo 6.8. Qui ci limitiamo a commentare i seguenti punti: Scelta della direzione e Criterio di arresto.
Assegnato un punto x^k, possiamo considerare il problema vincolato

$$\min_{x \in S} \nabla f(x^k)^T (x - x^k). \tag{6.9}$$

Osserviamo che la funzione obiettivo è lineare e poiché si è supposto che S sia compatto e convesso il problema precedente è convesso e ammette una soluzione x^{k*}; inoltre per la convessità del problema non esistono minimi locali non globali. Per definizione di minimo risulta

$$\nabla f(x^k)^T (x - x^k) \geq \nabla f(x^k)^T (x^{k*} - x^k) \quad \forall\, x \in S$$

Inoltre per la convessit di S la direzione $d^k = x^{k*} - x^k$ è una direzione ammissibile in x^k. Se risulta $\nabla f(x^k)^T d^k = \nabla f(x^k)^T (x^{k*} - x^k) \geq 0$ allora il punto x^k soddisfa le condizioni di ottimo per il problema (6.8) (e dunque risulta soddisfata la verifica effettuata al Passo 2 sull'appartenenza di x^k all'insieme Ω) e l'algoritmo si ferma. Viceversa se $\nabla f(x^k)^T (x^{k*} - x^k) < 0$ la direzione $d^k = x^{k*} - x^k$ risulta essere di discesa. È dunque possibile determinare con opportuni metodi di ricerca unidimensinoale (esatta o inesatta) un passo $\alpha^k > 0$ lungo d^k, tale che $x^{k+1} \in S$ e $f(x^{k+1}) < f(x^k)$. L'algoritmo cosìdefinito è noto come *Metodo di Frank-Wolfe* o *Conditional gradient method*.

Osserviamo che se l'insieme ammissibile S è un poliedro, ovvero è definito da vincoli lineari, il sottoproblema (6.9) relativo al calcolo della direzione diviene un problema di programmazione lineare la cui soluzione è in generale "semplice".

6.12 Ottimizzazione su un poliedro

Una classe significativa di problemi di programmazione matematica con insieme ammissibile convesso è quella in cui l'insieme ammissibile è un poliedro. Valgono naturalmente per questa categoria di problemi le stesse condizioni di ottimo descritte del paragrafo 6.10; tuttavia la linearità dei vincoli consente di ottenere condizioni più specifiche.

Consideriamo, senza perdita di generalità (vedi considerazione del capitolo 3), un poliedro definito come

$$S = \{x \in I\!R^n : \; Ax \geq b\}$$

in cui A è una matrice reale $m \times n$ e $b \in I\!R^m$. Il problema in esame è quindi

$$\begin{aligned} \min \quad & f(x) \\ & Ax \geq b \end{aligned} \qquad (P - POL)$$

in cui f è una funzione continuamente differenziabile.

6.12.1 Direzioni ammissibili di un poliedro

Per caratterizzare le direzioni ammissibili di un poliedro è necessario introdurre ulteriori definizione. A tal scopo analizziamo nuovamente l'esempio 6.3.2.

Esempio 6.12.1 Consideriamo nuovamente il poliedro

$$S = \{x \in R^2 : \; x_1 + x_2 \leq 3, \; x_1 \geq 1, \; x - 2 \geq 0\}$$

in Figura 6.12 e il punto ammissibile $\bar{x} = (1, \; 1)$.

In Figura 6.13 sono rappresentati i due semispazi definiti singolarmente dai due vincoli $x_1 + x_2 \leq 3$ e $x_1 \geq 1$ che definiscono il poliedro S. Se valutiamo i vincoli in $\bar{x} = (1, \; 1)^T$ (in rosso nella figura 6.13) abbiamo:

$$\begin{aligned} &\bar{x}_1 + \bar{x}_2 < 3, \\ &\bar{x}_1 = 1 \\ &\bar{x}_2 > 0 \end{aligned}$$

ovvero solo il vincolo $x_1 \leq 1$ è soddisfatto all'uguaglianza.

Dalla figura 6.13 è evidente che nel punto $\bar{x}$ ogni direzione $d \in I\!R^n$ è ammissibile rispetto al vincolo $x_1 + x_2 \leq 3$ a patto di scegliere adeguatamente il valore dello spostamento t. Analogamente rispetto al vincolo $x_2 \geq 0$ (non rappresentato in figura 6.13) che individua il primo e secondo quadrante, nel punto $\bar{x}$ ogni direzione è ammissibile. Se invece consideriamo il semispazio individuato dal vincolo $x_1 \geq 1$ è evidente che non tutte le direzioni sono ammissibili. Ad esempio la direzione $d = (-1 \; 0)^T$ (considerata nell'esempio 6.3.2 e disegnata in rosso in figura 6.13) non consente di mantenere l'ammissibilità del vincolo per nessun valore di $t > 0$.

Da queste semplici osservazioni appare evidente che in un punto ammissibile, le direzioni ammissibili sono "determinate" da quei vincoli che sono soddisfatti all'uguaglianza. Mentre il valore di $t_{\max}$ dipende dal valore degli altri vincoli.
Introduciamo allora la seguente definizione:

Definizione 6.12.2 (Vincoli attivi) *Sia $S = \{x \in R^n : \; Ax \geq b\}$ e sia $\bar{x} \in S$. Se $\bar{x}$ soddisfa il $a_i^T \bar{x} = b_i$ si dice che il vincolo i-esimo è* attivo *in $\bar{x}$.*
Dato $\bar{x} \in S$ si indica con $I(\bar{x})$ l'insieme degli indici di tutti i vincoli attivi n $\bar{x}$, ovvero:

$$I(\bar{x}) = \{i \in \{1, \ldots, m\} : \; a_i^T \bar{x} = b_i\}$$

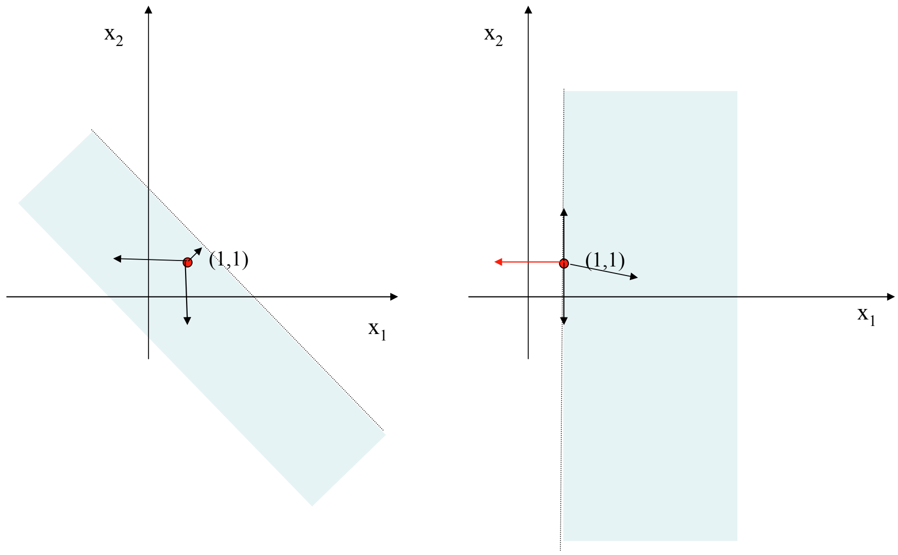

Figura 6.13: Semispazi individuati rispettivamente dai vincoli $x_1 + x_2 \leq 3$ e $x_1 \geq 1$ dell'Esempio 6.12.1.

Nell'esempio 6.3.2, in $\bar{x}$ è attivo il secondo vincolo e cioè $I(\bar{x}) = \{2\}$.

La definizione 6.12.2 di vincolo attivo non è legata alla particolare forma del poliedro, ma può essere estesa al caso generale. In particolare, notiamo che i vincoli di uguaglianza sono, per definizione, sempre attivi in un punto ammissibile. Possiamo quindi affermare

Dato un poliedro definito come $S = \{x \in I\!R^n : \ Ax = b\}$, risulta $I(\bar{x}) = \{1, \dots, m\}$.

Dato un poliedro in forma standard $S = \{x \in I\!R^n : \ Ax = b, \ x \geq 0\}$, i vincoli attivi sono quelli di uguaglianza più gli eventuali componenti di $\bar{x}_j = 0$, ovvero risulta $I(\bar{x}) = \{1, \dots, m\} \cup \{j \in \{1, \dots, n\} : \ \bar{x}_j = 0\}$.

Possiamo ora caratterizzare le direzioni ammissibili di un poliedro.

Teorema 6.12.3 *Dato il poliedro $S = \{x \in R^n : \ Ax \geq b\}$. Sia $\bar{x}$ un punto ammissibile e sia $I(\bar{x}) = \{i \in \{1, \dots, m\} : \ a_i^T \bar{x} = b_i\}$ l'insieme degli indici attivi in $\bar{x}$. Una direzione d è ammissibile in $\bar{x}$ se e solo se risulta*

$$a_i^T d \geq 0, \qquad \textit{per ogni} \ \ i \in I(\bar{x}). \tag{6.10}$$

Inoltre per ogni direzione ammissibile d, il punto $\bar{x} + td$ è ammissibile per ogni valore di t che soddisfa la condizione

$$0 < t \leq t_{\max} = \min_{j \notin I(\bar{x}):\ a_j^T d < 0} \frac{a_j^T \bar{x} - b_j}{|a_j^T d|} \tag{6.11}$$

dove si intende che, se l'insieme $\{j \notin I(\bar{x}) :\ a_j^T d < 0\}$ su cui si effettua il minimo è vuoto, il valore $t_{\max} = \infty$.

Dimostrazione. Sia $\bar{x}$ tale che $a_i^T \bar{x} \geq b_i$ per ogni $i = 1, \ldots, m$. Definiamo un punto $x = \bar{x} + td$ con $t > 0$.
Verifichiamo che le direzione ammissibili sono tutte e sole quelle per cui vale la (6.10). Si tratta quindi di verificare le condizioni su d per cui risulta

$$a_i^T(\bar{x} + td) = a_i^T \bar{x} + t a_i^T d \geq b_i \quad \text{ per ogni } t > 0.$$

Per ogni $i \in I(\bar{x})$, vale $a_i^T \bar{x} = b_i$ da cui $a_i^T \bar{x} + t a_i^T d = b_i + t a_i^T d$. Quindi si ottiene per $i \in I(\bar{x})$:

$$a_i^T \bar{x} + t a_i^T d \geq b_i \ \text{ per ogni } t > 0 \qquad \text{se e solo se} \quad a_i^T d \geq 0.$$

Per ogni $j \notin I(\bar{x})$, vale $a_j^T \bar{x} > b_i$ da cui $a_j^T \bar{x} + t a_j^T d > b_j + t a_j^T d$. Abbiamo due possibili casi: 1.) $a_j^T d \geq 0$ oppure 2.) $a_j^T d < 0$.
Se $a_j^T d \geq 0$, allora per ogni $t > 0$ risulta $b_j + t a_j^T d \geq b_j$ e il punto è sempre ammissibile rispetto al vincolo $j-$esimo. Se invece $a_j^T d < 0$, è necessario scegliere t abbastanza piccolo in modo che $a_j^T \bar{x} - t|a_j^T d|$ continui ad essere $\geq b_j$ per $t > 0$. In ogni caso (1. o 2.) la direzione risulta ammissibile.

Verifichiamo che il valore di $t_{\max}$ deve soddisfare la (6.11).
Sia quindi data d ammissibile. Dalle considerazioni precedenti si ha che per ogni $i \in I(\bar{x})$, vale $a_i^T(\bar{x} + td) \geq b_i$ per ogni $t > 0$.
Consideriamo allora gli indici $j \notin I(\bar{x})$. Abbiamo già visto che nel caso $a_j^T d \geq 0$, possiamo scrivere per ogni $t > 0$ $a_j^T(\bar{x} + td) = a_j^T \bar{x} + t a_j^T d \geq a_j^T \bar{x} > b_j$ per ogni $t > 0$.
Consideriamo allora il caso $a_j^T d < 0$ in cui sappiamo che t non può assumere tutti i valori positivi. Dobbiamo trovare il valore di t che soddisfa la condizione di ammissibilità espressa da $a_j^T(\bar{x} + td) = a_j^T \bar{x} + t a_j^T d = a_j^T \bar{x} - t|a_j^T d| \geq b_j$. Risolvendo la condizione $a_i^T \bar{x} - t|a_j^T d| \geq b_j$ si ottiene che

$$t \leq \frac{a_j^T \bar{x} - b_j}{|a_i^T d|}.$$

Poiché questo deve valere per ogni $j \notin I(\bar{x})$ e tale che $a_j^T d < 0$, si ottiene la limitazione data dalla (6.11). □

Introduciamo la seguente notazione matriciale: indichiamo con $A_{I(\bar{x})}$ la matrice $|I(\bar{x})| \times n$ costituita dalle righe di A con indice in $I(\bar{x})$, ovvero, se $A = (a_i^T)_{i=1,\ldots,m}$, si ha

$$A_{I(\bar{x})} = (a_i^T)_{i \in I(\bar{x})}.$$

Sia $\bar{x} \in S$, le direzioni ammissibili in $\bar{x}$ sono tutte e sole le soluzioni del sistema lineare

$$A_{I(\bar{x})} d \geq 0$$

dove $A_{I(\bar{x})}$ è la matrice $A_{I(\bar{x})} = (a_i^T)_{i \in I(\bar{x})}$ e $I(\bar{x}) = \{i :\ a_i^T \bar{x} = b_i\}$.

Illustriamo con un esempio.

Esempio 6.12.4 Sia dato il poliedro

$$\begin{array}{l} 3x_1 - 2x_2 \geq -30 \\ 2x_1 - x_2 \geq -12 \\ x_1 \geq 0 \; x_2 \geq 0 \end{array}$$

ed il punto ammissibile $\bar{x} = (0,\ 12)^T$. Risulta

$$\begin{array}{l} 3\bar{x}_1 - 2\bar{x}_2 = -12 > -30 \\ 2\bar{x}_1 - \bar{x}_2 = -12 \\ \bar{x}_1 = 0 \; \bar{x}_2 = 12 > 0 \end{array}$$

Quindi $I(\bar{x}) = \{2, 3\}$ e le direzioni ammissibili $d = (d_1,\ d_2)$ devono soddisfare il sistema

$$\begin{array}{l} 2d_1 - d_2 \geq 0 \\ d_1 \geq 0 \end{array}$$

Il valore di $t_{\max}$ si ottiene considerando i vincoli non attivi e dipende ovviamente dalla particolare direzione ammissibile considerata. Sia, ad esempio $\bar{d} = (1,\ 2)^T$. Risulta $a_1^T\bar{d} = 3 - 4 < 0$ e $a_4^T\bar{d} = 1 > 0$. Quindi il valore di $t_{\max}$ è individuato dal primo vincolo e si ha in particolare

$$t_{\max} = \frac{a_1^T\bar{x} - b_1}{|a_1^T d|} = \frac{3\bar{x}_1 - 2\bar{x}_2 + 30}{|3\bar{d}_1 - 2\bar{d}_2|} = \frac{6}{1} = 6.$$

In figura 6.14 è rappresentato il poliedro, il punto $\bar{x}$, l'insieme delle direzioni ammissibili (insieme compreso tra le due frecce blu tratteggiate), i punto ottenibili spostandosi da $\bar{x}$ lungo le direzioni ammissibili (regione compresa tra le due frecce blu a tratto intero) e il valore del passo $t_{\max}$ relativo alla direzione ammissibile $\bar{d} = (1,\ 2)^T$. Il punto $\bar{x} + t_{\max}\bar{d} = (6,\ 24)^T$ è indicato in rosso in figura 6.14. □

Corollario 6.12.5 *Dato il poliedro* $S = \{x \in \mathbb{R}^n : \ Ax \geq b\}$. *Sia* $\bar{x}$ *un punto ammissibile e sia* $I(\bar{x}) = \{i \in \{1, \dots, m\} : \ a_i^T\bar{x} = b_i\}$ *l'insieme degli indici attivi in* $\bar{x}$ *e supponiamo che esista un vettore* d *non nullo tale che*

$$a_i^T d = 0, \qquad \textit{per ogni} \;\; i \in I(\bar{x}).$$

Allora le direzioni d *e* $-d$ *sono ammissibili in* $\bar{x}$ *ed esistono valori* $t^+_{\max} > 0$ *e* $t^-_{\max} > 0$ *tali che*

$$\begin{array}{l} \bar{x} + td \in S \;\textit{per ogni}\; t \in [0, t^+_{\max}) \\ \bar{x} - td \in S \;\textit{per ogni}\; t \in [0, t^-_{\max}) \end{array}$$

con

$$t^+_{\max} = \min_{j \notin I(\bar{x}):\ a_j^T d < 0} \frac{a_j^T\bar{x} - b_j}{|a_j^T d|} \qquad t^-_{\max} = \min_{j \notin I(\bar{x}):\ a_j^T d > 0} \frac{a_j^T\bar{x} - b_j}{a_j^T d} \tag{6.12}$$

Si intende che se l'insieme degli indici su cui viene calcolato il minimo è vuoto, il corrispondente valore $t^+_{\max}$ *e/o* $t^-_{\max}$ *vale* ∞.

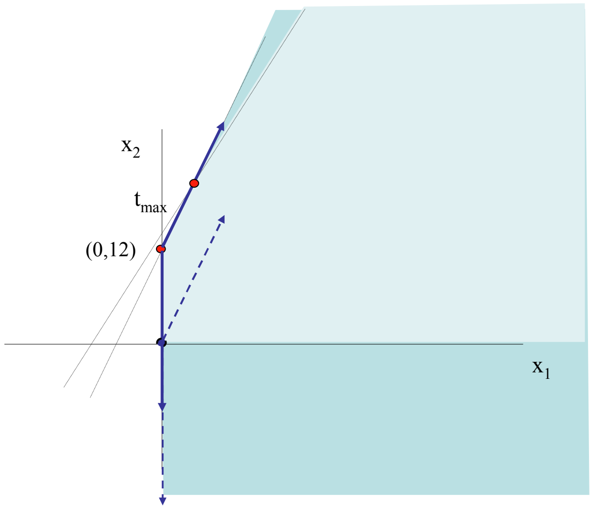

Figura 6.14: Poliedro Esempio 6.12.4.

Dimostrazione. La dimostrazione segue facilmente dal Teorema 6.12.3, applicandolo separatamente alla due direzioni d e $-d$. $\square$

Dato un poliedro $S = \{x \in I\!R^n : \; Ax \geq b\}$ e un punto $\bar{x} \in S$. Supponiamo che esista una direzione $\bar{d}$ tale che

$$A_{I(\bar{x})}\bar{d} = 0 \qquad I(\bar{x}) = \{i \in \{1, \ldots, m\} : \; a_i^T\bar{x} = b_i\}.$$

Allora il punto $y = \bar{x} \pm t\bar{d}$ con $t \in [0, t_{max}^{\pm}]$ è ammissibile e risulta

$$I(y) \supseteq I(\bar{x}) \quad \forall\, t \in [0, t_{max}^{\pm}].$$

Naturalmente è possibile caratterizzare anche le direzioni ammissibili nel caso di poliedri in forma standard $S = \{x \in I\!R^n : \; Ax = b, \; x \geq 0\}$.

Preventivamente consideriamo il caso di soli vincoli di uguaglianza $Ax = b$. Sia $\bar{x}$ un punto ammissibile, ovvero tale che $A\bar{x} = b$. La definizione di direzione ammissibile d richiede che $A(\bar{x} + td) = b$ per t sufficientemente piccolo. Sviluppando si ottiene

$$A\bar{x} + tAd = b + tAd = b \quad \Longleftrightarrow \quad Ad = 0 \quad \text{per ogni } t.$$

Si osservi che in questo caso, t può assumere qualunque valore non nullo (anche negativo). Abbiamo quindi dimostrato il seguente risultato.

Teorema 6.12.6 *Dato un poliedro definito da un sistema di equazioni lineari del tipo $Ax = b$, e un punto ammissibile $\bar{x}$. Il vettore $d \in I\!R^n$ é una direzione ammissibile in $\bar{x}$ se e solo se:*

$$Ad = 0. \tag{6.13}$$

Osservazione 6.12.7 *Notiamo che nel caso lineare una direzione d che soddisfa la (6.13) è ammissibile qualunque sia il punto $\bar{x}$ ammissibile, e l'ammissibilità si conserva per ogni valore di $t \neq 0$.*

Ad esempio, consideriamo il sistema di equazioni lineari

$$\begin{aligned} x_1 + x_2 - x_3 &= 1 \\ x_1 - x_2 &= 2 \end{aligned} \tag{6.14}$$

Il sistema $Ad = 0$ si scrive

$$\begin{pmatrix} 1 & 1 & -1 \\ 1 & -1 & 0 \end{pmatrix} \begin{pmatrix} d_1 \\ d_2 \\ d_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

e ha infinite soluzione del tipo $d = (t,\ t,\ 2t)^T = t(1,\ 1,\ 2)^T$ con $t \in I\!R$.

Riportiamo senza dimostrazione il seguente risultato relativo a poliedri in forma standard.

Teorema 6.12.8 (Direzioni ammissibili di un poliedro in forma standard) *Sia dato un poliedro del tipo $\{x \in I\!R^n :\ Ax = b,\ x \geq 0\}$ e sia $\bar{x}$ un punto ammissibile e sia $J(\bar{x}) = \{j \in \{1, \ldots, n\} :\ \bar{x}_j = 0\}$. Una direzione d è ammissibile in $\bar{x}$ se e solo se risulta*

$$\begin{aligned} & Ad = 0, \\ & d_j \geq 0 \quad \text{per ogni} \ \ j \in J(\bar{x}). \end{aligned} \tag{6.15}$$

Inoltre per ogni direzione ammissibile d, un punto $\bar{x} + td$ è ammissibile per ogni valore di t che soddisfa la condizione

$$0 \leq t \leq t_{\max} = \min_{j:\ \bar{x}_j \neq 0:\ d_j < 0} \frac{\bar{x}_j}{|d_j|}$$

dove si intende che se l'insieme su cui si effettua il minimo $\{j :\ \bar{x}_j \neq 0 :\ d_j < 0\}$ è vuoto, il valore $t_{\max} = \infty$.

6.13 Condizioni di ottimo per soli vincoli di uguaglianza lineari

Consideriamo un problema del tipo

$$\begin{aligned} \min \quad & f(x) \\ & Ax = b \end{aligned} \tag{$P-EQ$}$$

ove A è la matrice $m \times n$ le cui righe sono a_j^T, e b è il vettore di dimensione m con componenti b_j.

Possiamo caratterizzare le direzioni ammissibili utilizzando la proposizione 6.12.6 e ottenere la seguente condizione necessaria di ottimo.

Teorema 6.13.1 *Condizione necessaria affinchè x^* sia una soluzione locale del Problema ((P-EQ)) è che risulti*

$$\nabla f(x^*)^T d \geq 0, \quad \text{per ogni } d \in I\!R^n \;: Ad = 0.$$

Pertanto, mentre nel caso non vincolato la condizione $\nabla f(x^*)^T d \geq 0$ deve valere per ogni $d \in I\!R^n$, in quanto ogni direzione è ammissibile, nel presente caso la stessa condizione deve valere solo per le direzioni ammissibili, che sono quelle individuate mediante la (6.13).

Per ottenere una condizione di più pratico impiego, possiamo approfondire l'analisi del vincolo lineare.
Supponiamo a questo scopo e a titolo puramente semplificativo che *la matrice A abbia rango m.*

A tale scopo, indichiamo con $A_1, A_2, \ldots, A_n$ le colonne della matrice A, e scriviamo il vincolo mettendole in evidenza:

$$A_1 x_1 + A_2 x_2 + \ldots + A_n x_n - b = 0. \tag{6.16}$$

Osserviamo poi che risulta:

$$\nabla(Ax - b) = A^T;$$

Possiamo allora estrarre dalla matrice A un numero m di colonne linearmente indipendenti, e possiamo riordinare le colonne $A_1, A_2, \ldots, A_n$ e di conseguenza le variabili $x_1, x_2, \ldots, x_n$ in modo tale che quelle linearmente indipendenti siano le prime m. Possiamo cioè riscrivere il sistema (6.16), con l'intesa che le colonne $A_1, A_2, \ldots, A_m$ siano linearmente indipendenti.

Con questa intesa, introduciamo le notazioni:

$$B = [A_1, A_2, \ldots, A_m], \quad N = [A_{m+1}, A_{m+2} \ldots, A_n], \quad A = (B \; N)$$

$$x_B = \begin{bmatrix} x_1 \\ x_2 \\ \ldots \\ x_m \end{bmatrix}, \quad x_N = \begin{bmatrix} x_{m+1} \\ x_{m+2} \\ \ldots \\ x_n \end{bmatrix}, \quad d_B = \begin{bmatrix} d_1 \\ d_2 \\ \ldots \\ d_m \end{bmatrix}, \quad d_N = \begin{bmatrix} d_{m+1} \\ d_{m+2} \\ \ldots \\ d_n \end{bmatrix},$$

ove B è una matrice $m \times m$ non singolare.

Utilizzando una terminologia adottata nella Programmazione Lineare, diremo che B è una *matrice di base* per la matrice A, mentre N è una *matrice non di base*; che x_B è il vettore delle *variabili di base*, mentre x_N è il vettore delle *variabili non di base*; che d_B è il vettore delle *direzioni di base*, mentre d_N è il vettore delle *direzioni non di base*.
Possiamo riscrivere il vincolo (6.16) nella forma:

$$Bx_B + Nx_N - b = 0, \tag{6.17}$$

e riscrivere la (6.13) nella forma:

$$Bd_B + Nd_N = 0. \tag{6.18}$$

Concludiamo questa analisi preliminare osservando che, poichè B è non singolare, dalla (6.18) si ottiene:

$$d_B = -B^{-1} N d_N; \tag{6.19}$$

quindi, se si fa riferimento ad una base B di A, ogni direzione ammissibile si può esprimere nella forma

$$d = \begin{bmatrix} -B^{-1} N d_N \\ d_N \end{bmatrix}, \tag{6.20}$$

ove $d_N \in I\!R^{(n-m)}$ è arbitrario.
Ad esempio, con riferimento al sistema (6.14), possiamo considerare la partizione

$$B = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \quad N = \begin{pmatrix} -1 \\ 0 \end{pmatrix}, \quad x_B = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \quad x_N = x_3, \quad d_B = \begin{pmatrix} d_1 \\ d_2 \end{pmatrix}, \quad d_N = d_3.$$

Possiamo quindi scrivere

$$d_B = \begin{pmatrix} d_1 \\ d_2 \end{pmatrix} = -\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}^{-1} \begin{pmatrix} -1 \\ 0 \end{pmatrix} d_3 = \begin{pmatrix} 1/2 \\ 1/2 \end{pmatrix} d_3.$$

Partiziondo di conseguenza anche il gradiente di f nella forma seguente:

$$\nabla f(x^*) = \begin{bmatrix} \nabla_B f(x^*) \\ \nabla_N f(x^*) \end{bmatrix}, \tag{6.21}$$

ove il vettore $\nabla_B f(x)$ è dato dalle derivate di f rispetto alle variabili di base x_B e il vettore $\nabla_N f(x)$ è dato dalle derivate di f rispetto alle variabili non di base x_N. La condizione necessaria della Proposizione (6.13.1) può essere riscritta nella forma:

$$\nabla_B f(x^*)^T d_B + \nabla_N f(x^*)^T d_N \geq 0, \text{ per ogni } \; d_B \in I\!R^m, \text{e } \; d_N \in I\!R^{(n-m)} : \; B d_B + N d_N = 0.$$

Usando (6.19) si ottiene

$$-\nabla_B f(x^*)^T B^{-1} N d_N + \nabla_N f(x^*)^T d_N \geq 0, \text{per ogni } \; d_N \in I\!R^{n-m}.$$

Mettendo in evidenza d_N si ottiene:

$$(-\nabla_B f(x^*)^T B^{-1} N + \nabla_N f(x^*)^T) d_N \geq 0, \text{ per ogni } \; d_N \in I\!R^{n-m}. \tag{6.22}$$

ovvero

$$(\nabla_N f(x^*) - N^T (B^{-1})^T \nabla_B f(x^*))^T d_N \geq 0, \text{ per ogni } \; d_N \in I\!R^{n-m}.$$

Analogamente a come abbiamo ragionato nel caso non vincolato (vedi Teorema 6.7.3), affinchè la disuguaglianza precedente valga per ogni $d_N \in I\!R^{n-m}$, deve risultare nullo il vettore che premoltiplica d_N, deve cioè risultare:

$$\nabla_N f(x^*) - N^T (B^{-1})^T \nabla_B f(x^*) = 0; \tag{6.23}$$

altrimenti, prendendo $d_N = -[\nabla_N f(x^*) - N^T (B^{-1})^T \nabla_B f(x^*)]$, la disuguaglianza non sarebbe soddisfatta.

Osserviamo che l'equazione (6.23) non è altro che il gradiente della funzione ristretta allo spazio $d_N \in I\!R^{n-m}$. Infatti usando il fatto che

$$x_B = B^{-1}b - B^{-1}Nx_N,$$

il problema (P-EQ) puó essere riscritto come

$$\min_{x_N \in I\!R^{n-m}} f(B^{-1}b - B^{-1}Nx_N, x_N)$$

ovvero si tratta della minimizzazione non vincolata di una funzione composta $f(x_B(x_N), x_N)$ in cui $x_B : I\!R^{n-m} \to I\!R^m$ e il cui gradiente é $\nabla_N x_B = (B^{-1}N)^T$ (matrice $(n-m) \times m$). Il gradiente della funzione $\nabla_N f(x_B(x_N), x_N)$ si ottiene utilizzando le regola di derivazione di una funzione composta ed é:

$$\nabla_N f(x_B, x_N) = \nabla_N x_B \nabla_B f(x_B(x_N), x_N) + \nabla_N f(x_B, x_N) = (B^{-1}N)^T \nabla_B f(x_B(x_N), x_N) + \nabla_N f(x_B, x_N)$$

Questa condizione puó essere riscritta utilizzando la funzione Lagrangiana per il Problema (P-EQ) definita come

$$L(x, \mu) = f(x) + \sum_{j=1}^{m} \mu_j (a_j^T x - b_j),$$

o, in forma vettoriale,

$$L(x, \mu) = f(x) + \mu^T (Ax - b).$$

Possiamo ora enunciare la seguente condizione necessaria di ottimalità per il Problema (P-EQ).

Teorema 6.13.2 *Condizione necessaria affinchè x^* sia una soluzione locale del Problema (P-EQ) è che esistano dei moltiplicatori $\mu_1^*, \mu_2^*, \dots \mu_m^*$ tali che:*

$$\nabla_x L(x, \mu) = \nabla f(x^*) + \sum_{j=1}^{m} \mu_j^* a_j = 0, \tag{6.24}$$

o, in forma matriciale:

$$\nabla_x L(x^*, \mu^*) = \nabla f(x^*) + A^T \mu^* = 0. \tag{6.25}$$

Dimostrazione. La condizione (6.25) puó essere riscritta

$$\begin{pmatrix} \nabla_B f(x^*) \\ \nabla_N f(x^*) \end{pmatrix} + \begin{pmatrix} B^T \\ N^T \end{pmatrix} \mu^* = 0.$$

ovvero

$$\begin{aligned} \nabla_B f(x^*) + B^T \mu^* &= 0, \\ \nabla_N f(x^*) + N^T \mu^* &= 0. \end{aligned}$$

Ponendo μ^* soluzione della prima condizione cioè

$$\mu^* = -(B^{-1})^T \nabla_B f(x^*), \tag{6.26}$$

e sostituendo nella seconda, si ottiene la (6.23) □

Dunque analogamente a quanto accade nel caso non vincolato, i punti che soddifano le condizioni di ottimio del 1 ordine, ovvero i candidati ad essere minimo locale, possono essere individuati risolvendo un sistema di equazioni non lineare.

IN un problema con soli vincoli di uguaglianza lineari $Ax = b$, i possibili candidati a punto di minimo sono punti stazionari della funzione Lagrangiana, ciov soddisfano

$$\begin{aligned} \nabla_x L(x, \mu) &= \nabla f(x) + A^T \mu = 0. \\ \nabla_\mu L(x^*, \mu^*) &= Ax - b = 0 \end{aligned}$$

6.14 Condizioni di ottimo su un poliedro

Utilizzando la caratterizzazione di direzioni ammissibili per un poliedro, possiamo applicare il teorema 6.4.1 ed ottenere la seguente condizione necessaria.

Teorema 6.14.1 (Condizione necessaria di minimo locale su un poliedro) *Se x^* è un punto di minimo locale del problema (P-POL) allora risulta*

$$\nabla f(x^*)^T d \geq 0 \quad \textit{per ogni direzione } d \in R^n : \ a_i^T d \geq 0, \quad \forall i \in I(x^*) = \{i : \ a_i^T x^* = b_i\}.$$

Questa condizione può essere formulata come *non esistenza* di soluzione di una sistema lineare. Utilizzando la notazione matriciale introdotta precedentemente, possiamo scrivere la condizione necessaria espressa dal teorema 6.14.1 come segue.

Se x^* è un punto di minimo locale di (P-POL) allora *non esiste* una soluzione $d \in R^n$ al sistema di disequazioni lineari

$$\begin{aligned} &A_{I(x^*)} d \geq 0, \\ &\nabla f(x^*)^T d < 0. \end{aligned} \tag{6.27}$$

Illustriamo questa condizione necessaria con un esempio.

Esempio 6.14.2 Consideriamo il problema

$$\begin{array}{ll} \min & (x_1 + 10)^2 + (x_2 - 12)^2 \\ & 3x_1 - 2x_2 \geq -30 \\ & 2x_1 - x_2 \geq -12 \\ & x_1 \geq 0, \\ & x_2 \geq 0 \end{array}$$

il cui poliedro è stato descritto nell'Esempio 6.12.4. Consideriamo il punto ammissibile $\bar{x} = (0,\ 12)^T$ che risulta essere minimo (vedi Figura 6.15 per le curve di livello della funzione obiettivo e la regione ammissibile).
Il sistema delle direzione ammissibili è già stato definito nell'Esempio 6.12.4. Il gradiente della funzione obiettivo

$$\nabla f(x) = \begin{pmatrix} 2(x_1 + 10) \\ 2(x_2 - 12) \end{pmatrix}$$

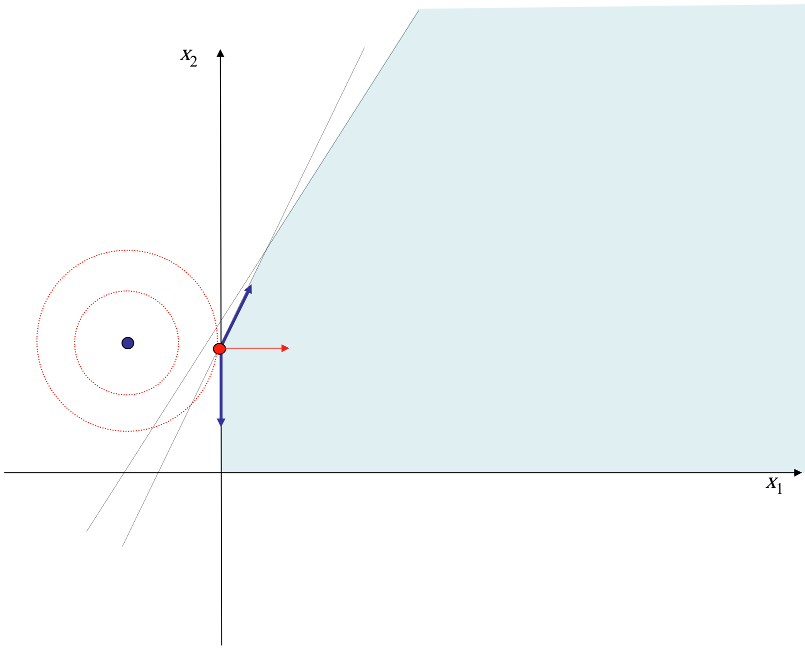

Figura 6.15: Poliedro Esempio 6.14.2.

che in $\bar{x}$ vale $\nabla f(\bar{x}) = (20,\ 0)^T$. Il sistema (6.27) diventa

$$\begin{array}{l} 2d_1 - d_2 \geq 0 \\ d_1 \geq 0 \\ 20d_1 < 0 \end{array} \tag{6.28}$$

che non ha ovviamente soluzione (le ultime due disequazioni incompatibili). □

Ricordando che un poliedro è un insieme convesso e quindi vale la caratterizzazione delle direzione ammissibili data nel teorema 6.10.1 è possibile applicare il Teorema 6.10.3 e ottenere il seguente risultato.

Teorema 6.14.3 (Condizione necessaria e sufficiente di minimo globale su un poliedro) *Sia $f(x)$ convessa. Un punto $x^* \in S$ è un minimo globale per problema (P-POL) se e solo se*

$$\nabla f(x^*)^T d \geq 0 \quad \textit{per ogni direzione } d \in R^n : \ a_i^T d \geq 0, \quad \forall i \in I(x^*) = \{i : \ a_i^T x^* = b_i\}.$$

Anche in questo caso si può riscrivere la condizione come:

Sia f una funzione convessa. Un punto $x^* \in S$ è minimo globale di (P-POL) se e solo se *non esiste* una soluzione $d \in R^n$ al sistema di disequazioni lineare

$$\begin{aligned} &A_I d \geq 0, \\ &\nabla f(x^*)^T d < 0. \end{aligned} \tag{6.29}$$

Consideriamo nuovamente un esempio.

Esempio 6.14.4 Consideriamo nuovamente il problema dell'Esempio 6.14.2. La funzione obiettivo è strettamente convessa. Quindi la condizione (6.29) è necessaria e sufficiente di minimo globale. Quindi il punto $(0,\ 12)^T$ che la verifica è un minimo globale.
Verifichiamo cosa succede in un punto diverso, ad esempio $\hat{x} = (0,\ 0)^T$. In $\hat{x}$ risulta $I(\hat{x}) = \{3, 4\}$ e le direzioni ammissibili soddisfano: $d_1 \geq 0,\ d_2 \geq 0$. Inoltre si ha $\nabla f(\hat{x}) = (20,\ -24)^T$, quindi il sistema (6.29) è

$$\begin{aligned} &20d_1 - 24d_2 < 0 \\ &d_1 \geq 0 \\ &d_2 \geq 0 \end{aligned}$$

In Figura 6.16 sono rappresentate le direzioni ammissibili (tutti i vettori tra le due frecce blu) e il gradiente della funzione obiettivo (freccia rossa). È facile verificare che esistono soluzioni al sistema; in particolare sono soluzione tutti i vettori $d = (d_1,\ d_2)^T$ tali che $0 \leq d_1 < \frac{6}{5} d_2$. Quindi, come già sapevamo, il punto $\hat{x}$ non è minimo globale in quanto non è soddisfatta la condizione necessaria e sufficiente. □

Possiamo inoltre dare una condizione necessaria del secondo ordine utilizzando il teorema 6.6.1.

Teorema 6.14.5 (Condizioni necessarie del 2^o ordine di minimo locale sul poliedro) *Sia x^* un punto di minimo del Problema (P-POL). Allora risulta*

$$d^T \nabla^2_x f(x^*) d \geq 0 \quad \textit{per ogni} \;\; d \in I\!R^n : \;\; a_i^T d \geq 0, \;\; i \in I(x^*) \;\; \textit{e} \; \nabla f(x^*)^T d = 0.$$

Le considerazioni svolte nel caso di poliedro descritto da soli vincoli di disuguaglianza $Ax \geq b$, possono essere applicate anche a poliedri definiti da vincoli di uguaglianza e disuguaglianza. Possiamo infatti utilizzare la caratterizzazione delle direzioni ammissibili data nel paragrafo 6.12.1. Consideriamo allora il caso piú generale di problema

$$\begin{aligned} \min \quad & f(x) \\ & a_i^T x \geq b_i \quad i = 1, \dots, m \\ & a_j^T x = b_j \quad j = 1, \dots, p \end{aligned} \tag{6.30}$$

Se x^* è minimo globale del problema (6.30) allora *non esiste* una soluzione $d \in R^n$ al sistema di disequazioni lineare

$$\begin{aligned} &a_i^T d \geq 0 \qquad && i \in I(x^*) \subseteq \{1, \dots, m\} \\ &a_j^T d = 0 && j = 1, \dots, p \\ &\nabla f(x^*)^T d < 0 && . \end{aligned}$$

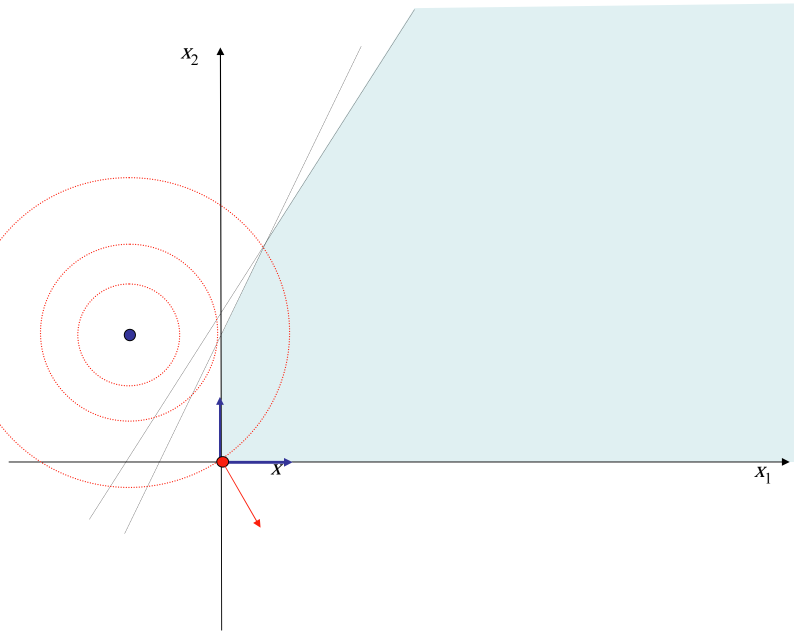

Figura 6.16: Poliedro Esempio 6.14.4.

Un esempio importante è un problema con vincoli in forma standard del tipo:

$$
\begin{array}{ll}
\min & f(x) \\
 & Ax = b \\
 & x \geq 0
\end{array}
\qquad (P - POL - ST)
$$

In questo caso, utilizzando la condizione espressa dal Teorema 6.12.8, possiamo affermare:

Se x^* è minimo globale del problema (P-POL-ST) allora *non esiste* una soluzione $d \in R^n$ al sistema di disequazioni lineare

$$
\begin{array}{l}
Ad = 0, \\
d_{J(x^*)} \geq 0 \\
\nabla f(x^*)^T d < 0.
\end{array}
\qquad (6.31)
$$

con $J(x^*) = \{j \in \{1, \ldots, n\} : \ x_j^* = 0\}$.

Naturalmente anche in questo caso le condizioni diventano necessarie e sufficienti nel caso in cui la funzione obiettivo sia convessa.

Tra i problemi di minimizzazione su poliedro particolare interesse rivestono i problemi di Programmazione Lineare del tipo

$$\begin{aligned} \min \quad & c^T x \\ & Ax \geq b. \end{aligned} \tag{PL}$$

Utilizzando il teorema 6.14.3, possiamo enunciare la seguente condizione che caratterizza le soluzioni ottime di un problema di Programmazione Lineare.

Sia dato il problema (PL). Un punto x^* è minimo globale se e solo se *non esiste* una soluzione $d \in R^n$ al sistema di disequazioni lineare

$$\begin{aligned} & A_{I(x^*)} d \geq 0, \\ & c^T d < 0. \end{aligned} \tag{6.32}$$

Analogamente nel caso di problemi di Programmazione Lineare in forma standard del tipo

$$\begin{aligned} \min \quad & c^T x \\ & Ax = b \\ & x \geq 0, \end{aligned} \tag{$PL-ST$}$$

possiamo affermare:

Un punto x^* è minimo globale del problema di Programmazione Lineare (PL-ST) se e solo se *non esiste* una soluzione $d \in R^n$ al sistema di equazioni e disequazioni lineari

$$\begin{aligned} & Ad = 0, \\ & d_{J(x^*)} \geq 0 \\ & c^T d < 0. \end{aligned} \tag{6.33}$$

con $J(x^*) = \{j \in \{1, \ldots, n\} : \ x_j^* = 0\}$.

Le condizioni di "non esistenza" di soluzioni di un sistema lineare ottenute in questo paragrafo, possono essere formulate in modo equivalente come condizioni di esistenza di un diverso sistema di equazioni e disequazioni non lineari. Questa nuova formulazione costituisce la forma classica delle condizioni di ottimo per problemi di programmazione matematica in presenza di vincoli di disuguaglianza. Sarà descritta nel Capitolo 8.